Advanced Machine Learning Empowering Microwave Sensors with Generative/Predictive Capabilities

by

Nazli Kazemi

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Software Engineering and Intelligent Systems

Department of Electrical and Computer Engineering

University of Alberta

Abstract

This dissertation, addressing the critical shortage of microwave planar sensors, embarks on a journey to advance noncontact sensing applications, with a particular emphasis on the application of machine learning techniques to address existing problems. It pioneers the introduction of a unique planar reflective sensor, employing complementary split ring resonators (CSRRs) as the primary sensing element, complemented by an additional loss-compensating negative resistance. Experimental results validate that this innovative design substantially amplifies the sensor's resolution without any detriment to sensitivity.

The thesis further unveils a novel design that utilizes coupled CSRRs, resulting in a sensor with heightened sensitivity. The sensor response in passive mode undergoes processing using CycleGAN, a machine learning algorithm typically harnessed for image-to-image translation. This groundbreaking application of machine learning effectively transmutes low-quality factor sensor profiles in the passive domain into their high-quality factor counterparts in the active domain. The resolution achieved by this CycleGAN-enhanced response rivals that of an active sensor, marking a significant leap in the resolution of passive sensors. This is a clear demonstration of how machine learning can be used to address the limitations of traditional sensor designs.

In the biomedical sensing arena, the thesis emphasizes the paramount importance of continuous glucose monitoring systems that obviate the need for invasive finger pricking, thereby ameliorating the comfort and lifestyle of diabetic patients. It advocates for a compact planar resonator-based sensor for noncontact glucose sensing, delivering real-time response and a strong linear correlation between sensor readings and blood glucose levels. The integration of the Long Short Term Memory (LSTM) algorithm, another

machine learning technique, enables the prediction of glucose level fluctuations in both non-diabetic and diabetic patients. This facilitates early interventions in case of abnormal glucose trends and aids in identifying sensor anomalies. This is another example of how machine learning can be used to address real-world problems, in this case, improving the management of diabetes.

This research makes a monumental contribution to the field of microwave planar sensors by bolstering their robustness and resolution using machine learning algorithms of CycleGAN and LSTM. This paves the way for an expansion of potential applications in material characterization and biomedical analysis. The insights extracted from this thesis are instrumental in propelling the development of advanced sensor technologies, with a focus on crafting miniaturized devices. The central role of machine learning in these advancements cannot be overstated.

Preface

The research presented in this thesis was conducted at the ENergy digiTizAtIon Lab (ENTAIL) under the guidance of Dr. Petr Musilek. This thesis includes three main publications.

[1] Kazemi, Nazli, and Petr Musilek. "A review on application of machine learning algorithms in microwave sensors" will be submitted to Engineering Applications of Artificial Intelligence (2023)

[2] Kazemi, Nazli, and Petr Musilek. "Enhancing Microwave Sensor Performance With Ultrahigh Q Features Using CycleGAN." IEEE Transactions on Microwave Theory and Techniques 70, no. 12 (2022): 5369-5382.

[3] Kazemi, Nazli, and Petr Musilek. "In–human testing of a non-invasive continuous low–energy microwave glucose sensor with advanced machine learning capabilities", Biosensors and Bioelectronics (2023) : 115668

Chapter 3, starting from page (38), is based on a journal paper that will be submitted to the Engineering Applications of Artificial Intelligence Journal. In our review paper, my responsibility was to organize the steps for conducting a systematic literature review on the applications of machine learning in microwave sensors. Dr. Petr Musilek, as a supervisory author, provided valuable insights and input to improve the quality of the review paper, remaining involved throughout the entire review process.

Chapter 4, starting from page (75), is derived from our Transactions paper [1]. In this research study, I took on the responsibility of designing the research methodology and conducting measurements and analysis of the results. Dr. Petr Musilek, acting as a supervisory author, provided significant assistance with concept formation and manuscript editing.

Chapter 5, starting from page (94), is based on our paper published in Biosensors and Bioelectronics [2]. In this chapter, I was responsible for various tasks, including analytical modeling, numerical simulations, sensor fabrication, measurements, development of the machine learning model, and manuscript writing. Mohammad Abdolrazzaghi provided valuable assistance in the theoretical analysis and validation of the simulation and measurement processes. Dr. Peter E. Light contributed to the preparation of the ISF solution with biological analytes and manuscript editing. Dr. Petr Musilek served as the supervisory author and participated in concept formation, coordination, writing review, and editing.

Acknowledgements

I would like to express my sincere gratitude and appreciation to all those who have contributed to the completion of this doctoral thesis. Without their support, guidance, and encouragement, this journey would not have been possible.

First and foremost, I am deeply grateful to my supervisor, Prof. Petr Musilek, for his unwavering commitment, expert guidance, and invaluable mentorship throughout this research endeavor. Prof. Musilek's profound knowledge, patience, and dedication have been instrumental in shaping this thesis. His insightful feedback, constructive criticism, and constant encouragement have significantly enhanced the quality of my work. I am truly fortunate to have had the privilege of working under his supervision.

I would also like to extend my heartfelt appreciation to the members of my supervisory committee, Prof. Marek Reformat, Prof. Kambiz Moez. Their expertise, insightful comments, and valuable suggestions have greatly enriched this research work. I am grateful for their time, effort, and contributions to the improvement of my thesis.

I am indebted to my colleagues at ENTAIL lab and fellow researchers, Prof. Peter. E. Light, Mohammad Abdolrazzaghi, Kalvin Schofield, Mikael Sabouhi, Nasatan Gholizadeh, Daniel May, Akhtar Hussain, and Zahra Nazari, for their support and fruitful discussions throughout the course of my doctoral studies. Their diverse perspectives, constructive criticism, and shared knowledge have broadened my understanding and helped me to overcome various challenges.

I am grateful to my family and friends for their unwavering love, encouragement, and understanding throughout this demanding journey. Their belief in me, constant support, and motivation have been my pillars of strength, inspiring me to persist even during the most challenging times.

I extend my heartfelt gratitude to my mom and dad for their unwavering support, encouragement, and endless patience throughout my journey in life; their love and belief in me have been my greatest motivation.

Thank you all for being part of this incredible journey.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Microwave sensors have several advantages when compared to their optical and electrochemical counterparts. They include portability, small form factor, label-free detection and most importantly, noncontact sensing. A high level of accuracy can be achieved with inexpensive and low-profile designs. Over the past two decades, applications of microwave sensors have been growing in various fields including biosensing [4,5], food quality monitoring [6], material characterization [7–17], environmental monitoring [18–25], and monitoring of mechanical motion and strain [26,27], measurement of distance, movement, shape, and particle size. But they are mainly used in applications related to measurement of material properties.

The principle of operation for microwave sensors is based on the interaction between the materials and electromagnetic waves. This interaction could be in different ways including reflection, refraction, emission, scattering, absorption, or change of speed and phase. Microwave sensors are divided into groups such as resonators, transmission sensors, reflection, radar sensors, etc, based on the phenomenon used in designing the sensor and the arrangement of the measurement. The interaction between electromagnetic waves (microwaves) and the medium of propagation is determined by the relative permittivity of the medium. In material characterization this permittivity is highly dependent on the composition of the material under test and, in the case of measurement of dimensions and distance, the contrasts in the permittivity at interfaces cause reflection. There are several advantages the microwave sensors for material measurements are listed below:

- Microwave sensors perform a non-contact mode of sensing. This allows for the interaction of materials in the proximity of the sensor to be interrogated without direct contact. Therefore, it enables the sensor to operate in harsh environments and inaccessible areas. This eventually prolongs sensor life and significantly decreases the maintenance cost.
- Using microwaves not only the surface but also the volume of material can be reached since they can penetrate all materials except metals.
- Microwave sensors are well suited for water content measurements; the contrast between the dielectric constant of water and other materials makes these sensors capable of such sensing.

- Microwave resonator sensors are less sensitive to high temperatures (in contrast to semiconductor sensors) and to surrounding conditions including moisture, dust, etc. (in comparison to infrared sensors).
- Microwaves do not affect the material under test in any way.
- Experiments using microwaves are safe and fast; they use non-ionized radiation instead of ionized (radioactive) radiation.
- Resonator-based sensors can be easily integrated with CMOS technology.

The measurement methods and sensor design are highly dependent on the nature of MUT (both physically and electrically), the frequency of operation, and the accuracy needed to detect the material. Considering the frequency of operation, at low and medium ranges of frequency, resonant methods have been widely used for material characterization. However, at higher frequencies, resonant cavities, transmission lines, and free space techniques are preferred. In general, material characterization methods are divided into two main categories including reflection and transmission techniques in which resonant and non-resonant-based sensors with open and closed structures play the main role.

1.2 Objective

Although microwave sensors have proven to be effective in various applications, they are not without their challenges. The primary objective of my PhD study is to address these inherent challenges associated with microwave sensors:

- Planar structures incur high loss due to undesired propagation through air rather than MUT resulting in a low resonance profile (low Q-factor) and low resolution.
- The planar nature of resonator-based sensors leaves them vulnerable to ambient changes with a potential impact on the perception of MUT. Temperature, relative humidity, proximity of other materials, displacement of MUT, and electromagnetic interference are the most important factors that affect sensor response and are considered undesired environmental impacts.
- The contactless sensing scheme becomes problematic when the material under test (MUT) resides within an unknown background material. Similarly, when an interfering medium/material is present in the sensing range of a microwave planar sensor, measurement results are affected.
- Microwave sensors are often designed for sensing a specific material; for example, if the sensor is designed to measure the level of moisture in soil it cannot be used to measure the moisture in other materials with the same calibration.

These issues significantly limit the potential applications and effectiveness of these sensors in various fields of engineering and bioengineering. In general, this thesis proposes the integration of machine learning approaches, which have been successful in addressing similar issues in other areas of engineering and bioengineering. The overarching goal is to leverage the power of artificial intelligence (AI) to enhance the functionality of microwave sensors. AI, with its ability to discern linear and nonlinear patterns in complex systems, offers a promising avenue for improving the performance of these sensors. The thesis

will explore the application of AI in learning the patterns of sensor responses in correlation with numerous parameters. By interpreting large datasets and resolving the relationship between the final outcome and features of sensor response, we aim to optimize the output of microwave sensors. Several algorithms will be examined in this process, with a comprehensive study conducted to compare them, assess their properties, and determine the best performers for given applications.

One of the challenges related to microwave sensors is the low resolution of their passive versions for certain material concentrations. This issue significantly hampers the effectiveness of these sensors, particularly in applications that require high-quality sensing. The reliance on power-intensive and error-prone active sensors to achieve high-resolution responses is not a sustainable or efficient solution. To address this issue, we aim to explore the potential of generative algorithms. These algorithms, which are a branch of machine learning, have shown promise in various fields for their ability to generate new data that mimic the distribution of the training set. In the context of microwave sensors, generative algorithms could potentially be used to enhance the resolution of sensor responses, transforming lower-quality profiles obtained from passive sensors into equivalent high-resolution profiles typically associated with active sensors. Basically, the goal of this work is to harness the power of generative algorithms to improve the quality of sensor responses, thereby addressing the limitations of current microwave sensor technology. This approach could pave the way for the development of more efficient, high-resolution sensing technologies, contributing to advancements in various fields that rely on these sensors.

The other challenges assigned to microwave sensors include their vulnerability to environmental impacts and the discrepancies caused by sensor delays and malfunctions, which can significantly affect the continuous analysis of sensor responses. These issues are particularly problematic in applications that require real-time, accurate sensing. To address these challenges, our focus turns towards the potential of recurrent neural networks, specifically Long Short-Term Memory (LSTM) networks. These networks are designed to remember patterns over time, making them particularly suited for tasks that involve continuous data, such as the analysis of sensor responses. Ultimately, the aim of this work is to leverage the capabilities of LSTM networks to enhance the robustness and accuracy of microwave sensors, thereby mitigating their vulnerability to environmental impacts. This innovative approach could lead to the development of more resilient, high-resolution sensing technologies, fostering advancements in various fields that rely on these sensors.

The last two challenges are not elaborated explicitly in this thesis, however, they are explored during my PhD studies.

In essence, this thesis aims to bridge the gap between microwave sensor technology and AI, with the ultimate goal of enhancing sensor performance, reducing costs, and increasing resilience against environmental impacts. This endeavor aligns with the broader vision of harmonizing various devices and platforms within the expansive realm of the Internet of Things, thereby contributing to the advancement of this critical field.

1.3 Contributions

This thesis aims to enhance the capabilities of microwave sensors by employing machine learning algorithms. To accomplish this, a series of specific investigations are conducted to introduce fundamental algorithms like classification and clustering to the realm of microwave sensors. These investigations ultimately aim to incorporate complex models similar to generative models into the field of microwave sensors.

1. The necessity for high-quality sensing in various applications, coupled with the growing interest in noninvasive wearable devices, underscores the need for innovative solutions. A novel system is introduced that leverages machine learning to achieve high-resolution responses, circumventing the need for power-intensive and error-prone active sensors. This system is grounded in the CycleGAN technique, which facilitates the transformation between passive and active sensor responses using unpaired datasets. It captures the characteristics of low-to-moderate quality factor profiles from the passive response and generates equivalent high-resolution profiles in the active domain. This transformation is achieved without the need for paired training examples, a significant advancement in the field. Despite the passive and active domains having identical distributions, there can be multiple mappings that produce the same distribution. To ensure consistency, the system employs a cycle consistency loss, which ensures that the mapping remains consistent with a reverse mapping. This, combined with adversarial loss, leads to the unpaired passive-to-active mapping. By conceptualizing the sensor response as an image, the system can utilize image processing techniques to enhance the accuracy of passive sensors. This approach paves the way for the development of an intelligent sensory system, marking a significant stride in the field [1].
2. Microwave sensors designed for continuous, real-time measurements in various configurations and practical applications are crucial in many fields. However, several challenges persist, including sensor delays, malfunctions, and erroneous environmental effects, leading to discrepancies between the measured and actual values. To address these challenges, the integration of continuous prediction mechanisms, facilitated by machine learning algorithms, is proposed. This approach aims to reduce natural latency and fill in missing measurements. A Long Short-Term Memory (LSTM) network is notably utilized to forecast values 30 and 60 minutes into the future, aiding in anomaly detection arising from omitted readings or unintentional changes in the measurement setup due to environmental impacts. The effectiveness of these techniques is thoroughly validated through proof-of-concept experiments and practical applications, suggesting a promising avenue for more reliable and accurate sensor technologies.

In addition to the publications included in this thesis, the author has conducted other studies that significantly contribute to the development of machine learning for microwave sensors. Specifically, the following prior works provide a broader context for understanding the relevance and potential impact of integrating machine learning into microwave sensor applications.

3. A comparative analysis of various machine learning techniques was conducted to compensate for temperature variations in microwave sensors. Classification and regression models were utilized to

identify material types and extract temperature information from the same dataset. Techniques such as Multilayer Perceptron (MLP), Support Vector Machine (SVM), Decision Tree (DT), Linear Discriminant Analysis (LDA), Random Forest, and K-Nearest Neighbors (KNN) were found to demonstrate superior performance in classifying sensor data into corresponding classes. Additionally, the potential of machine learning systems to reduce the limit of detection (LOD) in sensing was explored, with a focus on distinguishing lower concentrations of methanol in water. It was concluded that through the optimization of selected classifiers and regressors via hyperparameter optimization, temperature-variant sensor responses for low concentrations could be accurately interpreted [14].

4. A wide-frequency range multi-resonator sensor using metamaterial transmission line zeroth order resonators is developed to detect binary mixtures of water/methanol. The sensor's response is converted into heatmaps, and through the utilization of Style-GAN, a large number of images representing intermediate concentrations are generated. These images are then fed to a convolutional neural network (CNN) classifier, resulting in the classification of predefined mixtures. This approach showcases the potential of combining Style-GAN with a CNN classifier to develop microwave sensors with selectivity for specific materials [28].
5. A machine learning algorithm is introduced to enhance the resolution of a planar microwave sensor, which utilizes complementary split-ring resonators to increase sensitivity. By exploiting the transmission zero response and employing a μ Fluidic channel, the sensor achieves improved performance in characterizing various concentrations of glucose in the interstitial fluid. The resolution is further enhanced by incorporating a super-resolution generative adversarial network (SRGAN), allowing the classification of glucose samples into 10 different classes and demonstrating the potential for future clinical trials [29].

1.4 Outline

This thesis is structured as follows: A brief overview of sensing techniques is given in chapter 2 in order to provide a thorough outlook on various microwave sensing techniques. Next chapter 3 presents a number of algorithms that can help resolve some challenges of microwave sensors. This chapter ends with fundamentals of recurrent neural networks (RNN) and generative adversarial networks (GAN) that are used in this study as potential solutions for anomaly detection in sensor response, making sensor selective to specific material, and having higher resolution in terms of sensor response and readout circuitry.

In chapter 5, a compact planar resonator-based sensor is introduced for noncontact sensing of glucose. Continuous glucose monitoring schemes that avoid finger pricking are of utmost importance to enhance the comfort and lifestyle of diabetic patients. To this aim, we propose a microwave planar sensing platform as a potent and fast-growing sensing technology that extends its applications to biomedical analytes. Furthermore, *in-vivo* and *in-vitro* tests using a microfluidic channel system and in clinical trial settings demonstrate its reliable operation. The proposed sensor offers real-time response and a high linear correlation ($R^2 \sim 0.913$) between the measured sensor response and the blood glucose level (GL). The sensor is also enhanced with machine learning to predict the variation of body glucose levels for

various types of non-diabetic and diabetic patients. This addition is instrumental in triggering preemptive measures in cases of unusual glucose level trends. In addition, it allows for the detection of common dispositional artifacts of the sensor as anomalies so that they can be removed from the measured data. The proposed system is designed to monitor a human signature, introducing the opportunity to create a custom, wearable apparatus with the ability to learn.

In chapter 4, we introduce a CycleGAN-boosting algorithm as a learnable technique to improve the sensing performance of low-quality resonance-based sensor response. In particular, an ultra-high sensitive microwave sensor that leverages a coupled CSRR design is used as a core of a loss-compensated active sensor to work at ≈ 3.5 GHz. The mutual coupling between the resonators provides a region with strongly enhanced sensitivity. The sensor performance in the loss-compensated mode is significantly improved, with $\approx$ 10-fold increase of the quality factor compared to the passive mode response. To avoid the complexity of circuit design and the need for constant powering, a novel method is introduced that involves a learning network called CycleGAN. It maps the measured passive mode sensor transmission response of the sensor to an equivalent active mode response while increasing the sensor transmission response quality factor. It is shown that while a low concentration of water in the mixture of water in ethanol results in intermingled transmission profiles, applying CycleGAN on the passive mode sensor response not only improves their quality factor by an order of magnitude but also helps differentiate them with higher resolution. Following the same procedure for noninvasive glucose detection, the CycleGAN-boosted sensor response resulting in very low measurement errors, as confirmed by Clarke error grid analysis. This technique is potentially applicable to a wide range of applications, where high-end sensing can be conducted with low-profile sensors amended with machine learning methods [1].

The thesis is concluded in chapter 6, with a summary of results and suggestions for future work. Also, supplementary information on the evaluation metrics of algorithms as well as human trial results is given as an Appendix A.

Chapter 2

Background on Dielectric Measurement Techniques

2.1 Material Properties

Sensors come in various forms and cover an extensive range. They serve as instruments that identify the characteristics of a measured subject and transform the measured data into a perceivable signal presented on a device. Sensing technology, at its core, employs sensors to gather information by recognizing physical, chemical, or biological properties and converting them into understandable signals. Every sensory system has three main components, including measurement objective, sensor, and output information. Taking into account this sequence, sensors are designed based on the characteristics and properties of the MUT, and consequently, the features of the output signal will be determined based on the sensor.

In the case of electromagnetic (EM) sensors, the interaction of electromagnetic fields with MUT changes the output signal. Among them, microwave sensors operate at various frequencies ranging from 3 GHz to 30 GHz, as shown in Fig. 2.1 (a). EM sensors have different applications, such as proximity, position, temperature, and material sensing. For Material characterization, EM sensors rely on the permittivity of the material (ε), a complex parameter that characterizes how the material interacts with the applied electric field. To derive these parameters, let us start with Maxwell's equations in phasor form [30, 31]:

$$\nabla \times \bar{\mathbf{E}} = -j\omega\bar{\mathbf{B}} - \bar{\mathbf{M}}, \tag{2.1}$$

$$\nabla \times \bar{\mathbf{H}} = j\omega\bar{\mathbf{D}} + \bar{\mathbf{J}}, \tag{2.2}$$

$$\nabla \cdot \bar{\mathbf{D}} = \rho, \tag{2.3}$$

$$\nabla \cdot \bar{\mathbf{B}} = 0, \tag{2.4}$$

where $\bar{\mathbf{E}}$ is the electric field (V/m), $\bar{\mathbf{H}}$ is the magnetic field (A/m), $\bar{\mathbf{D}}$ is the electric flux density (Coul/m^2), $\bar{\mathbf{B}}$ is the magnetic flux density (Wb/m^2), $\bar{\mathbf{M}}$ is the (fictitious) magnetic current density (V/m^2), $\bar{\mathbf{J}}$ is the electric current density (A/m^2), ρ is the electric charge density (Coul/m^3). Electric and magnetic current

sources $\bar{\mathbf{J}}$ (2.1) and $\bar{\mathbf{M}}$ (2.2) are volume current densities with units of $\mathrm{A/m^2}$ and $\mathrm{V/m^2}$, respectively.

When a dielectric material is exposed to an electric field $\bar{\mathbf{E}}$, the atoms or molecules of the material shift to create electric dipole moments that increase the flux of total displacement $\bar{\mathbf{D}}$. This extra polarization vector is called $\bar{\mathbf{P_e}}$, the electric polarization, where:

$$\bar{\mathbf{D}} = \varepsilon_0 \bar{\mathbf{E}} + \bar{\mathbf{P_e}} \tag{2.5}$$

The relationship between electric polarization and applied electric field in a linear medium is as follows:

$$\bar{\mathbf{P_e}} = \varepsilon_0 \chi_e \bar{\mathbf{E}}, \tag{2.6}$$

where χ_e, which may be complex, is called the *electric susceptibility*. Then

$$\bar{\mathbf{D}} = \varepsilon_0 \bar{\mathbf{E}} + \bar{\mathbf{P_e}} = \varepsilon_0 (1 + \chi_e) \bar{\mathbf{E}} = \varepsilon \bar{\mathbf{E}}, \tag{2.7}$$

where:

$$\varepsilon = \varepsilon^{'} - j\varepsilon^{"} = \varepsilon_0 (1 + \chi_e). \tag{2.8}$$

ε refers to the complex permittivity of a given medium. The imaginary component of a material's dielectric constant signifies the loss within that medium, which occurs due to the damping of oscillating dipole moments and the subsequent production of heat. In particular, free space exhibits real permittivity (ε) and therefore does not incur any loss. The principles of energy conservation dictate that the imaginary component of ε must be negative. Moreover, the loss within a dielectric material may be equated to a comparable conductor loss. When such a material has conductivity (σ), a conduction current density which is *Ohm's law* from an electromagnetic field point of view will be as follows [30]:

$$\bar{\mathbf{J}} = \sigma \bar{\mathbf{E}} \tag{2.9}$$

Considering Eq. (2.9) Maxwell's curl equation for $\bar{\mathbf{H}}$ in (2.2) becomes:

$$\begin{aligned} \nabla \times \bar{\mathbf{H}} &= j\omega \bar{\mathbf{D}} + \bar{\mathbf{J}}, && (2.10)\\ &= j\omega\varepsilon \bar{\mathbf{E}} + \sigma \bar{\mathbf{E}}, && (2.11)\\ &= j\omega\varepsilon^{'} \bar{\mathbf{E}} + (\omega\varepsilon^{"} + \sigma) \bar{\mathbf{E}}, && (2.12)\\ &= j\omega \left(\varepsilon^{'} - j\varepsilon^{"} - j\frac{\sigma}{\omega} \right) \bar{\mathbf{E}}. && (2.13) \end{aligned}$$

It should be noted that the loss attributed to dielectric damping ($\omega\varepsilon^{"}$) is practically indistinguishable from the conductivity loss (σ). Consequently, the combined value of ($\omega\varepsilon^{"} + \sigma$) can be viewed as the total effective conductivity. Additionally, a related parameter of significance is the loss tangent, which is defined as:

$$\tan(\delta) = \frac{\omega\varepsilon^{"} + \sigma}{\omega\varepsilon^{'}} \tag{2.14}$$

The loss tangent can be interpreted as the quotient of the actual and imaginary components of the

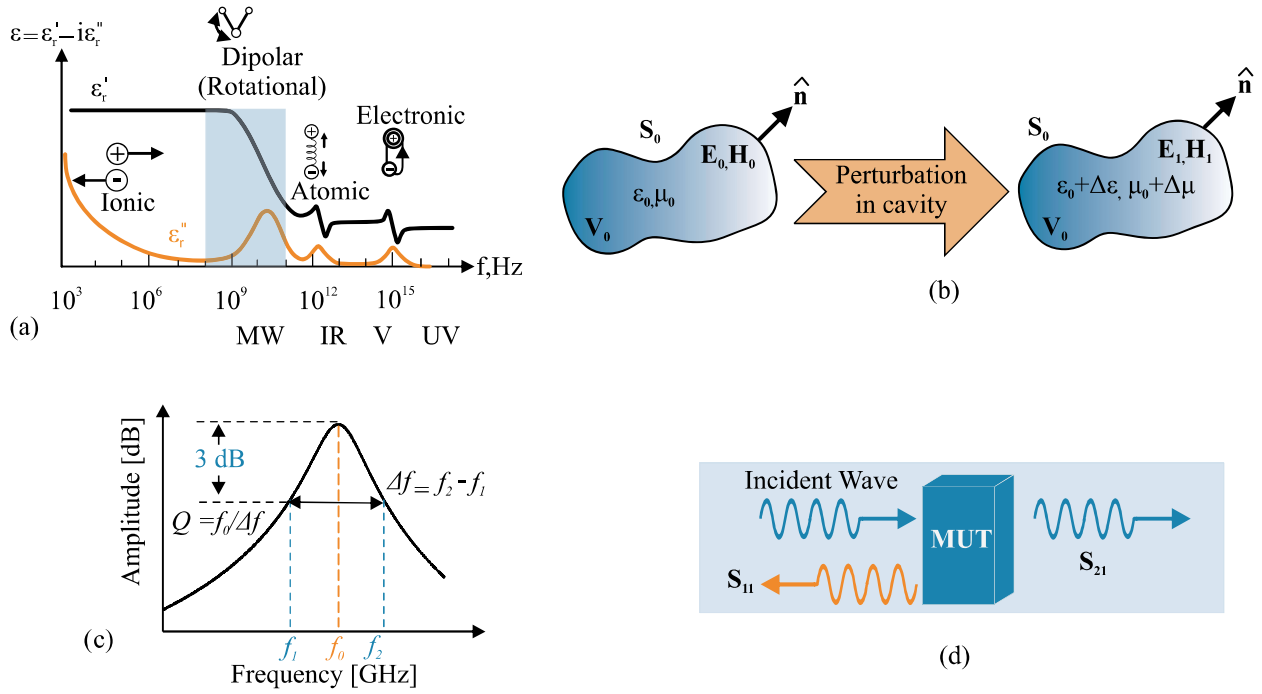


Figure 2.1: (a) Dispersive dielectric properties of materials (b) Principle of operation in perturbation theory (c) Features of sensor response (d) Diagram of reflected and transmitted waves from/through the material.

total displacement current. Microwave materials are mainly distinguished by defining the actual relative permittivity (that is, *the dielectric constant*), ε_r, together with the loss tangent $tan(\delta)$, which reflects the performance of the material at a specific frequency.

Knowing the electrical properties of dielectric materials helps to understand the sensor response theory that can be derived from a fundamental origin called *perturbation theory*.

Microwave sensors utilizing material perturbation theory and field interactions rely on the systematic analysis of electromagnetic field variations induced by the presence of objects or materials with distinct dielectric and magnetic properties.

The field perturbation shown in Fig. 2.1 (b) encompasses the variations induced in the amplitude, quality factor, phase, and frequency of the microwave signals by the material-induced perturbations in the electromagnetic fields (see Fig. 2.1 (c)). The magnitude of these variations is dependent on the material properties and the frequency of the microwave signal. Consequently, these perturbations lead to the transmission/reflection of a portion of the microwave signals through/back toward the sensor (see Fig. 2.1 (d)). The transmitted/reflected signals encode information about the field perturbations, which are intrinsically linked to the material properties of the object.

As depicted in Fig. 2.1 (b), a cavity experiences perturbation due to alterations in the permittivity or permeability (μ) of the material occupying the cavity, either partially or entirely. Considering $\bar{\mathbf{E_0}}$ and $\bar{\mathbf{H_0}}$ as the original cavity's electromagnetic fields and $\bar{\mathbf{E}}$ and $\bar{\mathbf{H}}$ as the perturbed cavity's fields, Maxwell's

curl equations can be formulated for both scenarios accordingly [30].

$$\nabla \times \bar{\mathbf{E}}_0 = -j\omega_0\mu\bar{\mathbf{H}}_0, \tag{2.15}$$

$$\nabla \times \bar{\mathbf{H}}_0 = j\omega_0\varepsilon\bar{\mathbf{E}}_0, \tag{2.16}$$

$$\nabla \times \bar{\mathbf{E}} = -j\omega(\mu + \Delta\mu)\bar{\mathbf{H}}, \tag{2.17}$$

$$\nabla \times \bar{\mathbf{H}} = j\omega(\varepsilon + \Delta\varepsilon)\bar{\mathbf{E}}. \tag{2.18}$$

In the given context, ω_0 denotes the resonant frequency of the unperturbed cavity, while ω signifies the resonant frequency associated with the perturbed cavity. It is evident from these equations that the change in the material property ($\Delta\varepsilon$, $\Delta\mu$) leads to a change in the electric and magnetic fields. (meaning: $\bar{\mathbf{E}}_0 \mapsto \bar{\mathbf{E}}$ and $\bar{\mathbf{H}}_0 \mapsto \bar{\mathbf{H}}$), which is manifested in the measured transmission/reflection parameters of the sensor.

Microwave sensors are designed to analyze alterations occurring in fields following their interaction with a MUT. Techniques for measuring dielectric properties can be categorized into various types, including reflection and transmission methods, resonant and non-resonant systems, as well as open and closed structures.

When selecting a method to characterize a material, three important factors should be taken into account.

- First, the frequency range utilized must be considered.
- Secondl the physical properties of the MUT must be evaluated, including whether it is magnetic or non-magnetic, lossy or low-loss, isotropic or anisotropic, homogeneous or heterogeneous, and whether it is dispersive or not.
- Last, the shape and nature of the available samples should be considered, such as whether they are plate or thin films, liquid or solid, elastomeric or granular.

It is essential to evaluate all these factors to choose the most appropriate characterization method and obtain accurate results. Using a network analyzer, one can measure the experimental device's reflection and transmission coefficients, which have a direct relationship to the electromagnetic characteristics of the subject material (i.e. sample's permittivity and permeability).

2.2 Non-Resonant Methods

Non-resonant methods are a common approach used to determine the properties of materials based on their impedance and wave velocities. These methods involve measuring the reflection and transmission of electromagnetic waves as they propagate from one material to another, resulting in changes in the characteristic wave impedance and wave velocity. Reflection methods involve measuring the magnitude and phase of the reflected signals to deduce the material properties, while reflection/transmission measurements require measuring the magnitude and phase of both reflected and transmitted signals. To improve the accuracy of the results, it is recommended to use both reflection and transmission measurements. To

perform these measurements, any type of transmission line can be used, such as a coaxial line, a hollow metallic waveguide, or a planar transmission line [32].

The properties of an electromagnetic wave, as it moves in the positive z-direction, can be determined by solving the wave equation, which is described by the equations for the electric and magnetic fields [30]:

$$E(t,z) = E_0 e^{-j\omega t - \gamma z}, \tag{2.19}$$

$$H(t,z) = H_0 e^{-j\omega t - \gamma z} = \frac{E_0}{\eta} e^{-j\omega t - \gamma z}, \tag{2.20}$$

where γ is the propagation constant, given by $\gamma = \alpha + j\beta$. α is the attenuation constant, β is the phase constant, and η is the intrinsic impedance of the medium given by $\sqrt{\frac{\mu}{\varepsilon}}$. Therefore, it is easy to understand how the electrical properties of materials are a great impact factor on the behavior of waves that come to interact with their structures.

An analytical method known as the Nicholson Ross Weir (NRW) [33–35] derivation is used to calculate a material's permittivity and permeability using the material's measured S_{11} and S_{21} properties. This method considers the multiple wave reflections that occur at the interfaces between the air and the sample, whether the sample is in free space or within a waveguide. It also takes into consideration how the scattering parameters affect the characteristics of the material. An abbreviated form of the NRW derivation for scattering parameters observed in free space is provided by the following equations [31, 36–38]:

$$S_{11} = \frac{\Gamma \cdot (1 - T^2)}{(1 - \Gamma^2 \cdot T^2)}, \tag{2.21}$$

$$S_{21} = \frac{T \cdot (1 - \Gamma^2)}{(1 - \Gamma^2 \cdot T^2)}, \tag{2.22}$$

$$z^2 = e^{-2 \cdot j \cdot \gamma \cdot d}, \tag{2.23}$$

$$\Gamma = \frac{(Z_1 - Z_0)}{(Z_1 + Z_0)} = \frac{(Z - 1)}{(Z + 1)}, \tag{2.24}$$

$$Z = \frac{\mu}{\varepsilon}. \tag{2.25}$$

where d is the sample thickness, Z_1 and Z_0 are the transmission line characteristic impedances in the region with and without sample, respectively. Γ is the first partial reflection at the air-sample interface, Z is the characteristic impedance of the material and depends on any of the material as stated in Eq. (2.25), and T is the propagation factor, i.e.

$$T = e^{-\gamma d} = e^{-jkd}. \tag{2.26}$$

Relation of the propagation constant γ with ε and μ of the material is given by Eq. (2.27).

$$\gamma = \frac{j \cdot 2 \cdot \pi \cdot \sqrt{\varepsilon \cdot \mu}\cdot}{\lambda}, \tag{2.27}$$

$$\lambda = \frac{c}{f}. \tag{2.28}$$

where λ is a free space wavelength, c is speed of light in vacuum and f is the frequency.

$$\gamma = -\frac{1}{d} \cdot \log \frac{1}{T} \tag{2.29}$$

From the Eq. (2.23), γ is expressed as a function of z and d (Eq. (2.27)). d is material slab thickness and z is calculated from Eq. (2.21) and Eq. (2.22). N_m is the material's refractive index and can be expressed in terms of permittivity and permeability as:

$$N_m = \sqrt{\mu \cdot \varepsilon} \tag{2.30}$$

By combining Eq. (2.27) and Eq. (2.28) we obtain an expression for the material's refractive index (Eq. (2.31)). G1 is expressed as a function of S_{11} and S_{21} and Z is expressed as a function of G_1. Considering two equations, two unknown parameters can extract easily. Finally, ε and μ can be calculated as a functions of N_m and Z [Eqs. (2.32) and (2.33)].

$$N_m = -j \cdot \frac{\lambda \cdot \gamma}{2 \cdot \pi} \tag{2.31}$$

$$\varepsilon = \frac{N_m}{Z} \tag{2.32}$$

$$\mu = N_m \cdot Z \tag{2.33}$$

Although the NRW method is a popular technique used for measuring the complex permittivity (dielectric constant) of materials over a wide frequency range, there are some limitations and issues associated with this method, which include:

1. Vulnerability to noisy measured data: The NRW method uses S_{11} and S_{21} measurements to calculate the complex permittivity (ε) and permeability (μ) for each frequency point. However, even small amounts of noise in the data can result in significant errors in the calculated values of ε and μ.
2. Inaccurate measurements at high frequencies: The NRW method assumes that the material being measured is non-magnetic, and the measurement results can be inaccurate at high frequencies when the material exhibits magnetic properties.
3. Sensitivity to sample preparation: The accuracy of the NRW method is highly dependent on the sample preparation, including the size, shape, and surface condition of the sample.
4. Reflection measurement limitations: The NRW method requires the use of a reflection measurement, which can introduce errors due to impedance mismatches, cable losses, and other measurement errors.

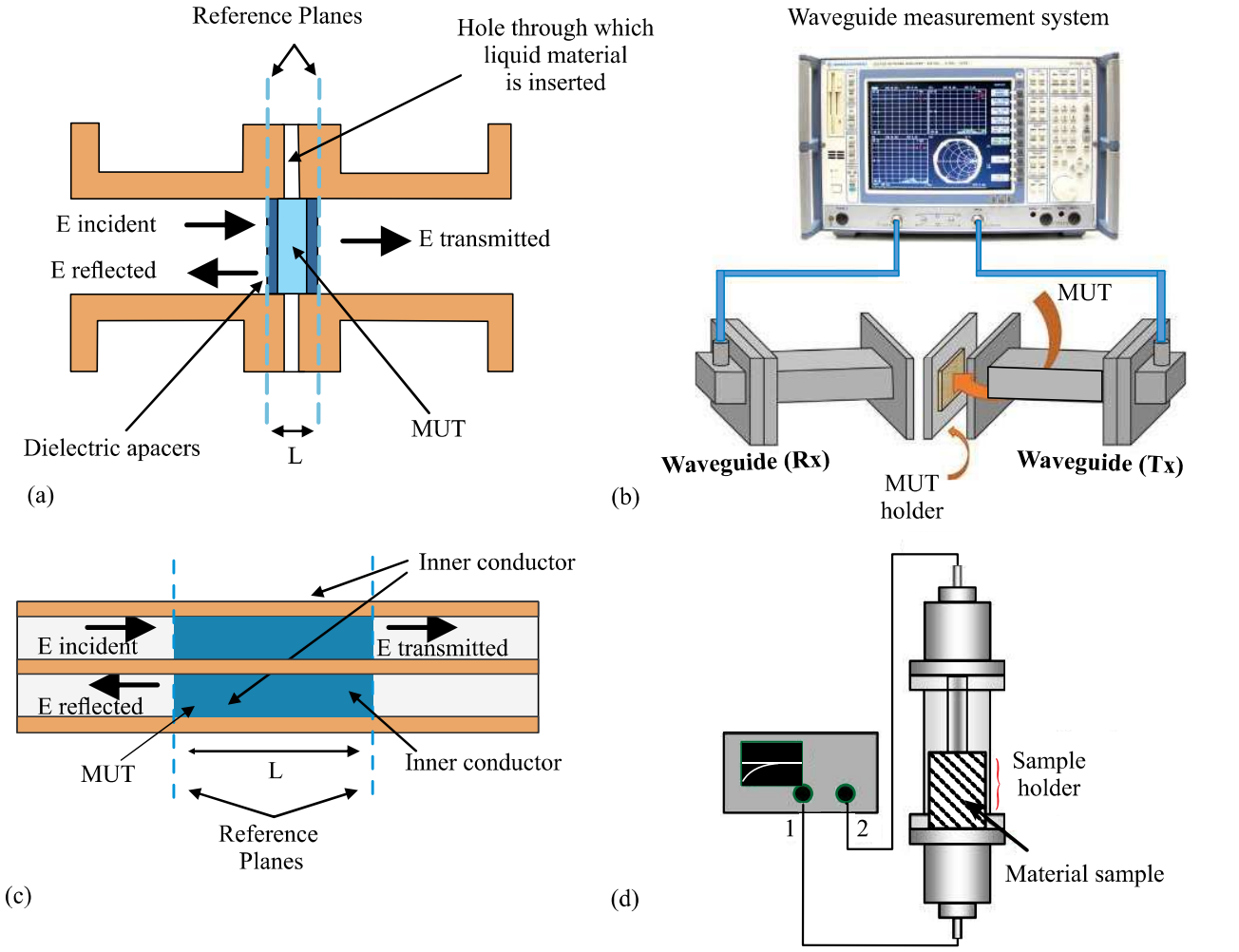


Figure 2.2: (a) Waveguide fixture and (b) Waveguide measurement setup (c) Coaxial fixture and (d) Coaxial setup for transmission measurements.

2.2.1 Waveguide and Coaxial Transmission Line Techniques

Waveguide Fig. 2.2 (a-b) and coaxial transmission line Fig. 2.2 (c-d) cells for complex permittivity measurements were initially introduced by Nicholson-Ross [39] and subsequently by Weir [40], who analyzed the structure in both time and frequency domains. Following these works, Baker-Jarvis [41] developed iterative solution methods [40]. The investigations conducted by Nicholson, Ross, and Weir resulted in the formulation of explicit equations for calculating permittivity and permeability, commonly known as the NRW algorithm.

This method involves filling a specific length of a transmission line with the sample to be measured (Fig. 2.2 (a)). Changes in the propagation constant γ and the characteristic impedance Z_0 cause partial reflections of the wave at the interfaces. The propagation constant is associated with the attenuation coefficient α and propagation coefficient β through the equation $\gamma = \alpha + j\beta$. The relation between the

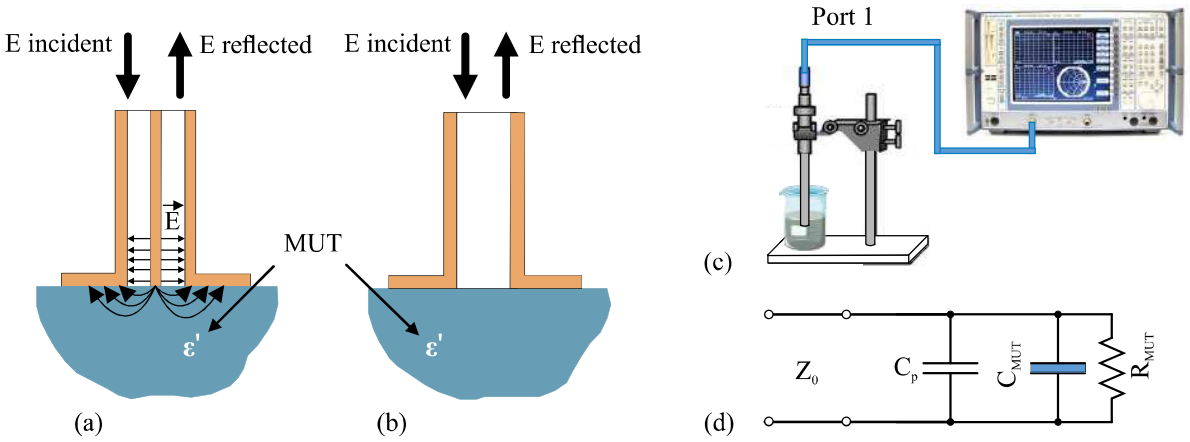


Figure 2.3: (a) Open-ended coaxial and (b) Waveguide transmission lines. (c) Dielectric probe test setup, (d) Equivalent circuit for the dielectric probe.

complex permittivity of the filling material and the wave propagation coefficient within a transmission line for a dielectric material is as follows [42–45]:

$$\gamma = j\sqrt{\frac{\omega^2 \mu_r \varepsilon_r}{c^2} - (\frac{2\pi}{\lambda_c})^2} \tag{2.34}$$

where ω is the angular frequency, μ_r is the permeability of the material which is equal to 1, c is the speed of light, and λ_c is the cutoff wavelength of the transmission line. The cutoff wavelength in the case of a coaxial transmission line supporting TEM propagation is assumed to be infinite. From Eq. (2.34), it becomes possible to extract the complex permittivity of the material filling the transmission line ($\varepsilon_r = \varepsilon_r' - j\varepsilon_r''$).

$$\varepsilon_r' = \frac{(\frac{2\pi}{\lambda_c})^2 - (\alpha + \beta)^2}{\omega^2} \tag{2.35}$$

$$\varepsilon_r'' = \frac{2\alpha\beta}{\omega} \tag{2.36}$$

The method of deriving the complex propagation coefficient in the line segment containing the material involves using S-parameter measurements taken by a network analyzer. Calibration techniques such as Thru-Reflect-Line (TRL) or Short-Open-Load-Thru (SOLT) [46] are utilized to define reference planes. The simplicity of this method has led to its broad adoption and continuous improvements over time. These developments have facilitated the estimation of permittivity for samples with high and moderate losses, yielding a significant degree of precision.

While a variety of transmission lines can be used with this method, waveguides, and coaxial lines are often favored for frequencies below 30 GHz. The transmission mode can change depending on whether the waveguide transmission cells are rectangular or circular. Waveguides, however, might not be the ideal option for frequencies below 2.45 GHz because of the substantial test sample volume they call for. Because of their small size and wide bandwidth, coaxial transmission lines are frequently used in these circumstances.

The frequency range to be used while sensing proximate structures will depend on a number of variables, including the desired resolution, the depth of penetration, and the characteristics of the substance being investigated. Higher frequencies could offer better clarity whereas lower frequencies might encounter more attenuation in particular materials. Lower frequencies might offer deeper penetration but might sacrifice resolution.

Higher-order TE and TM modes must be prevented from propagating in order to avoid errors in the measured permittivity. As a result, common cell sizes like 3.5 mm (0 - 34.5 GHz), 7 mm (0 - 18.2 GHz), and 14 mm (0 - 8.6 GHz) are used. When the cell size becomes excessively small, transmission line cells may experience problems with air gaps, resulting in severe measurement inaccuracies. Transmission line cell designs have seen various improvements and modifications over the past three decades, all of which are tuned to particular dielectric classes. Cell dimensions get too small for frequencies above 30 GHz, which forces the employment of free-space measurement methods.

2.2.2 Open-ended transmission line Techniques

Open-ended transmission line techniques (Fig. 2.3 (a-b)) offer a non-invasive and highly convenient fixture for the assessment of dielectric permittivity of semi-solids and liquids without the need for sample preparation. The technique was initially developed by Stuchly and Athey [47, 48], who employed it to evaluate the dielectric properties of biological materials.

This technique has experienced considerable refinement over the years. In this method, the sample to be tested is placed against the terminated section of the transmission line, and the phase and magnitude of the reflected signal are measured (see Fig. 2.3). Numerous models have been proposed in the last decade to correlate the material's permittivity with the reflection coefficient measured at the probe's aperture. The equivalent circuit of a measurement setup (Fig. 2.3 (c)) is shown in Fig. 2.3 (d). The capacitance C_p represents the concentration of the electric field inside the dielectric-filled part of the coaxial cable. The capacitance C_{MUT} corresponds to the fringing field concentration in the dielectric, while the conductance $G(\varepsilon)$ signifies radiation into the dielectric surrounding the cable. The value of the conductance $G = 1/R_{MUT}$ depends on the frequency f. The equivalent admittance of the open-ended coaxial cable Y can be expressed as [49, 50]

$$Y = j\omega C_p + j\omega C_{MUT} + G = j\omega C_p + j\omega C_0 \varepsilon + G_0 \varepsilon^{5/2} \tag{2.37}$$

where C_0 symbolises the capacitance of the air-filled parallel plate capacitance, ε signifies the permittivity of the object under test, and G_0 stands for the external radiation conductance of a coaxial cable. Based on the transmission line theory [51], the terminal admittance Y in terms of the reflection coefficient Γ is defined as

$$Y = \frac{1-\Gamma}{1+\Gamma}\frac{1}{Z_0} \tag{2.38}$$

Here, Z_0 represents the characteristic impedance of the coaxial line (50 Ω). The external radiation conductance of a coaxial cable is minimal, and hence, it can be neglected ($G_0 \approx 0$). Therefore, the relationship

between the measured complex permittivity of the object under test and the reflection coefficient Γ is

$$Y = \frac{1-\Gamma}{1+\Gamma}\frac{1}{Z_0} = j\omega C_p + j\omega C_0 \varepsilon \tag{2.39}$$

The relative dielectric constant ε' and the loss factor ε'' are calculated using Eq. (2.40) and (2.41):

$$\varepsilon' = \frac{1}{2\pi f Z_0 C_0} \times \frac{-2|\Gamma|\sin(\phi)}{1+2|\Gamma|\cos(\phi)+|\Gamma|^2} - \frac{C_p}{C_0} \tag{2.40}$$

$$\varepsilon'' = \frac{1}{2\pi f Z_0 C_0} \times \frac{1-|\Gamma|^2}{1+2|\Gamma|\cos(\phi)+|\Gamma|^2} \tag{2.41}$$

where Γ and ϕ are the modulus and phase of the input reflection coefficient, respectively.

Coaxial transmission lines are often chosen due to their small size and wide frequency range capabilities. This technique is particularly useful for analyzing lossy solvent materials because it can effectively measure samples with high loss or high dielectric constants. The method requires establishing calibration reference planes at the probe's aperture, a task that could be challenging with a coaxial probe. Calibration is typically performed using commonly known reference materials.

2.2.3 Free Space Transmission Techniques

Free-space methodologies [52, 53], primarily employed for higher frequencies often above 10 GHz, serve as non-invasive and non-contact measurement approaches [54]. These techniques do not mandate any particular preparation of samples and are especially advantageous for examining materials at elevated temperatures or with heterogeneous dielectrics. They are extensively adopted in industrial scenarios for ongoing monitoring and management. In these approaches, a sample is situated between a transmission and a reception antenna. The attenuation and phase alteration of the signal is then analyzed to infer the dielectric attributes of the sample. These techniques are capable of precisely evaluating permittivity over a broad frequency spectrum. However, the precision of the measurements is contingent upon the measurement system and the accuracy of the model utilized for the computations. The techniques assume the material under examination to be homogeneous, the incidence of a uniform plane wave on the sample surface to be normal, and the planar sample to extend laterally to infinity. Predominant sources of errors such as multiple reflections, mismatches, and diffraction impacts at the edges of the sample, must be duly acknowledged and addressed. The accuracy of measurement can be improved by careful selection of radiating elements, design of the sample holder, and consideration of the sample's geometry and location.

The performance of free-space techniques is greatly enhanced when the specific characteristics of the materials under test are considered, including their thermal properties, structural consistency, and electromagnetic behavior. Further, these techniques can be enriched by employing advanced computational models and algorithms that can factor in potential anomalies and complexities in the material's behavior. Moreover, integrating the free-space techniques with advanced artificial intelligence methodologies for data analysis can also contribute to the enhancement of measurement accuracy, thereby widening the scope and applicability of these techniques in various scientific and industrial domains. A typical

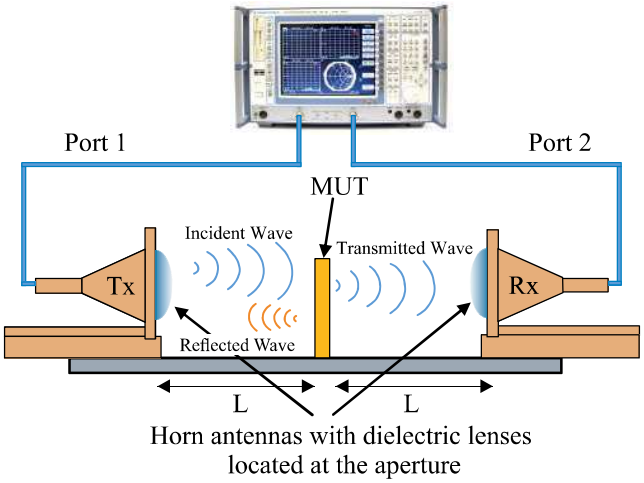


Figure 2.4: Free space dielectric measurement setup.

arrangement of a free-space measurement setup is shown in Fig 2.4.

The NRW method stands as the predominant approach to simultaneously derive the complex relative permittivity ($\varepsilon_r = \varepsilon_r^{'} - j\varepsilon_r^{"}$) and permeability ($\mu_r = \mu_r^{'} - j\mu_r^{"}$), excluding cases involving non-magnetic materials, where $\mu_r = 1$. Implementation of the NRW method necessitates specific knowledge concerning both the thickness of the material, denoted as d, and the separation between the Material Under Test (MUT) and the antennas, represented by L. A schematic diagram of the measurement setup is shown in Fig 2.4, which consists of a vector network analyzer (VNA), two horn antennas, and accessories. The setup is used for wave transmission and reception in free space, with the required dimensions of L and d indicated.

Figure 2.4 shows how a portion of the electromagnetic (EM) wave that is incident on MUT is reflected and measured by the reflection coefficient (S_{11}), while another part passes through and is transmitted and measured by the transmission coefficient (S_{21}). However, a small portion of the EM wave refracts and is transmitted with lag, which can affect the S_{21} measurement unless microwave absorbers are used to minimize this effect. To determine the electrical characteristics of the MUT sample, the measured scattering parameters (S_{11} and S_{21}), the thickness of the MUT sample (d), and its distance to the transmitting (Tx) and receiving (Rx) antennas (L) are used in the expressions of the Nicolson-Ross-Weir (NRW) method. The main equations of the NRW method for determining the relative permittivity and magnetic permeability are shown in Eq. (2.44) and Eq. (2.43) [52, 55–57].

$$k = \frac{1}{d}[j \ln |T| - \phi] \tag{2.42}$$

$$\mu_r = \frac{k}{k_0}\frac{1+\Gamma}{1-\Gamma} \tag{2.43}$$

$$\epsilon_r = \frac{k}{k_0}\frac{1-\Gamma}{1+\Gamma} \tag{2.44}$$

where $k_0 = \omega\sqrt{\mu_0\epsilon_0}$ is the free space wavenumber. Equations (2.45) to (2.47), which are presented in [26, 42], are used to determine the transmission (T) and reflection (Γ) coefficients of the MUT sample using

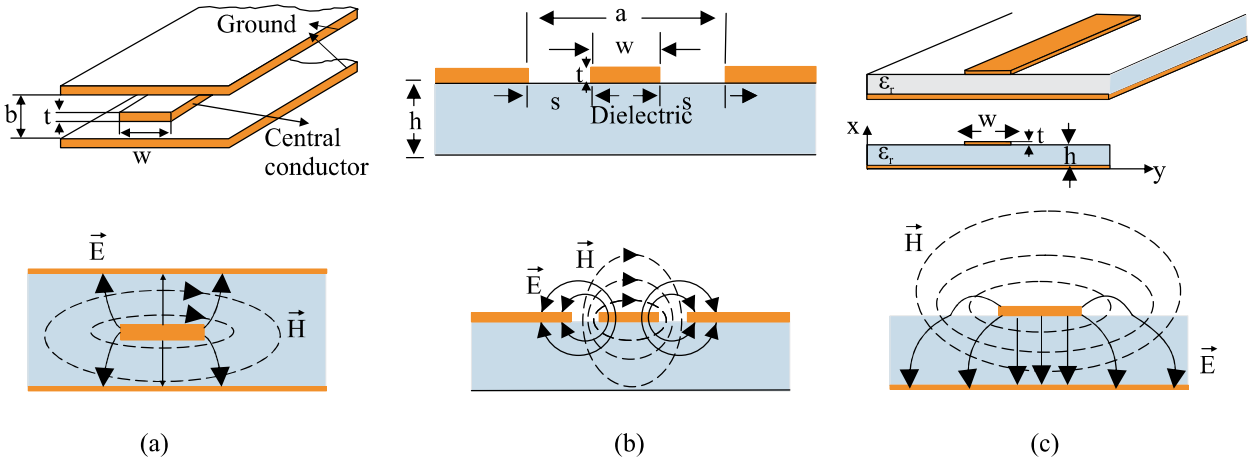


Figure 2.5: Structure of (a) Stripline TL, (b) Coplanar waveguide TL, and (c) Microstrip line TL.

the measured scattering parameters (S_{11} and S_{21} [37,38]).

$$K = \frac{S_{11}^2 - S_{21}^2 + 1}{2S_{21}} \tag{2.45}$$

$$\Gamma = K \pm \sqrt{K^2 - 1}, \quad |\Gamma| \leq 1 \tag{2.46}$$

$$T = \frac{(S_{11} - S_{21}) + \Gamma}{1 - (S_{11} + S_{21})\Gamma} \tag{2.47}$$

[58] modeled the propagation channel as a two-port network (quadripole) using an equivalent transmission line and the scattering matrix, [S]. Thus Eq. (2.48) can be used for the relative permittivity (ε_r) for non-magnetic materials.

$$\varepsilon_r = \left[\left(j\frac{c}{\omega L}\right)\ln(T)\right]^2 \tag{2.48}$$

The scattering parameters S_{11} and S_{21} are represented in phasor or complex number form, with indications of module, phase, and real and imaginary parts. Similarly, the transmission (T) and reflection (Γ) coefficients of the MUT sample can be expressed in these forms ($T = |T|\, e^{j\phi}$). The use of complex numbers in Eq. (2.42) to (2.44) enables the real and imaginary parts of the electrical permittivity and magnetic permeability to be obtained. For non-magnetic materials, such as most building materials, the relative magnetic permeability (μ_r) is equal to 1. These equations use the free space velocity (c) and angular frequency (ω) as parameters.

2.2.4 Planar Transmission Line Techniques [Non-Resonant]

There are several different transmission lines as shown in Fig. 2.5 including stripline TL, coplanar waveguide (CPW) TL, and microstrip TL with each type is being used primarily in a specific application. For one, stripline is mostly used for signal integrity, whereas CPW and microstrip lines can be deployed in sensing due to their openness to the surrounding medium. Since CPW includes ground planes around the signal line, it has been found to be much more robust in many applications, which is elaborated as

follows for complex permittivity characterization. A CPW consists of a central conductor and two ground planes on the same side of a dielectric substrate. The MUT is placed on top or inside the CPW, and its permittivity affects the propagation constant and the characteristic impedance of the CPW. By measuring the reflection and/or transmission coefficients of the CPW, the permittivity of the MUT can be extracted using an appropriate algorithm. A CPW microwave sensor can provide broadband, accurate, and non-destructive measurements of various materials, especially liquids. Its unique characteristics, such as low dispersion, planar configuration, and easy fabrication, make it suitable for these sensing applications.

To calculate the dielectric constant (ε_r) of the MUT using the CPW transmission line method there are several ways in the literature. Among them one practical and easy way is briefly explained as follows:

Let us consider a sensing structure implemented by means of an ordinary transmission line with physical length l and phase constant β at a certain frequency f_0 (the operating frequency). The electrical length, or phase, of the line at that frequency is thus $\phi = \beta l$, which in turn can be expressed as [30]

$$\phi = \beta l = \frac{\omega_0 l}{v_p} \tag{2.49}$$

where $\omega_0 = 2\pi f_0$ is the angular frequency and v_p is the phase velocity of the line. According to Eq. (2.49), any change in the electrical length of the line should be entirely due to a variation in the phase velocity, since the frequency is set to a given value and the physical length is constant (i.e. it does not depend on the input variable, typically the dielectric constant of the MUT, ε_{MUT}).

The phase velocity can be expressed as [30]

$$v_p = \frac{c}{\sqrt{\varepsilon_{\text{eff}}}} \tag{2.50}$$

where c represents the speed of light in vacuum and ε_{eff} is the effective dielectric constant of the line. Hence, an increase in the dielectric constant of the substrate, ε_r, which is closely associated with ε_{eff}, results in a reduction of the line's phase velocity.

For microstrip lines, the approximation considered yields the effective dielectric constant [30]:

$$\varepsilon_{\text{eff}} = \frac{\varepsilon_r + \varepsilon_{\text{MUT}}}{2} + \frac{\varepsilon_r - \varepsilon_{\text{MUT}}}{2} F \tag{2.51}$$

where F is a geometric factor that is determined by the width of the line, W, and the thickness of the substrate, h. The geometric factor F can be defined as $F = (1 + \frac{12h}{W})^{-\frac{1}{2}}$, assuming $W > h$. If $W < h$, the geometric factor is given by:

$$F = 1 + (\frac{12h}{W})^{-\frac{1}{2}} + 0.04\left(1 - \frac{W}{h}\right)^2 \tag{2.52}$$

Equation (2.51) is valid under the assumption that $t < h$, which is generally the case, where t denotes the thickness of the metallic layer. For Coplanar Waveguide (CPW) transmission lines where $t < h$, the effective dielectric constant can be approximated by $\varepsilon_{\text{eff}} = \frac{\varepsilon_r + \varepsilon_{\text{MUT}}}{2}$.

A comparison table for different non-resonant structures used for material characterization is given in

Table 2.1: Comparison of Nonresonant Microwave Sensing Structures

Parameter	Coaxial TL	Waveguide TL	CPW TL	Free Space
Operating Frequency Range	Broad (DC to 110 GHz)	Moderate (1 GHz to 100 GHz)	Moderate (10 MHz to 100 GHz)	Broad (DC to 300 GHz)
Size	Small	Larger than coaxial	Thin and planar	N/A
Transmission Mode	TEM	TE and TM modes	Quasi-TEM	TEM
Losses	Low	Low to moderate	Moderate	Low
Isolation	High	High	Low to moderate	Moderate to low
Fabrication and Integration	Simple	Moderate	Easy integration with PCBs	No fabrication needed
Flexibility	Flexible	Rigid	Rigid	N/A
Material Compatibility	Wide range	Limited by waveguide size	Wide range	N/A
Sensing in Complex Structures	Suitable	Less suitable	Suitable	Suitable for open structures

Table. 2.1.

2.3 Resonant Methods

Microwave components have historically employed planar transmission lines, such as microstrip and coplanar waveguides. These lines provide advantages such as ease of fabrication, low cost, and compactness, making them ideal for industrial applications involving dielectric permittivity measurements. In the planar transmission line approach, the test material frequently functions as a superstrate, substrate, or component of either. Solid dielectric samples can serve as both a substrate and a superstrate, although liquids and semi-solids are more practical. Several studies on the use of planar circuits for sophisticated liquid permittivity measurements have been undertaken [36, 59–63]. Figures 5 and 6 are typical planar cells for determining the dielectric permittivity of liquid materials. The liquid under examination completely covers the planar circuit and is confined within a low-loss container that is securely attached or affixed to the top of the board. This enclosure produces readily calibrated mismatches.

The liquid dielectric alters the effective permittivity (ε_{eff}) and characteristic impedance (Z_0) of the line. The dielectric properties can then be determined from the change in effective permittivity using applicable calculations from the literature [62].

2.3.1 Cavity Resonators

Resonant cavities are enclosed metallic constructs, typically made by using a short-circuited waveguide segment (either rectangular or circular) at both ends [180, 181]. These cavities have the ability to trap electromagnetic energy, which corresponds to various resonant modes. The frequencies of these modes are contingent upon the cavity's geometry and the dielectric constant of the material inside the cavity. For a rectangular waveguide cavity resonator that has a width of 'a', height of 'b', and length of 'd', the

frequencies for the different modes, whether TE_{mnl} or TM_{mnl}, are specified by [30].

$$f_{mnl} = \frac{c}{2\sqrt{\varepsilon_r \mu_r}} \sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2 + \left(\frac{l}{d}\right)^2} \tag{2.53}$$

Here, c denotes the speed of light in a vacuum, while ε_r and μ_r represent the relative permittivity and permeability of the material within the cavity, respectively. The mode numbers are represented by m, n, and l. For a circular waveguide cavity resonator, also known as a cylindrical cavity, the resonance frequencies of various modes are derived from formulas similar to, but distinct from, equation 2.53. These formulas differ for TE_{mnl} and TM_{mnl} modes [30]. According to Eq. (2.53), it is clear that the dielectric constant of the material filling the waveguide can be determined from the resonance frequency of a specific mode. For a dielectric material, $\mu_r = 1$. Therefore, whether rectangular or cylindrical, the resonant cavity functions as a sensor for the dielectric constant.

The main drawback of the resonant cavity technique, as outlined, relates to the necessity for the MUT to occupy the entire volume of the cavity, which is impractical in many situations. Nonetheless, resonant cavity sensors can still be used for dielectric characterization of material samples that only fill a small part of the cavity volume. This approach, typically known as the perturbation method (as referenced in [30, 64]), relies on the changes in both the resonance frequency and the quality factor of the cavity when it is loaded or disrupted with the sample being tested. This method assumes that the fields in the cavity, when loaded with the sample, are only slightly disturbed compared to those in the unloaded or unperturbed cavity. This is a reasonable assumption in many practical scenarios (such as small samples, or samples with small dielectric constants). An approximate formula for the resonance frequency of the disturbed cavity was derived in [30]. If we name the (angular) frequency of the perturbed cavity as ω_p, and the resonance frequency of the unperturbed cavity asω_0, the fractional change in the resonance frequency due to the cavity's disruption can be calculated using the following formula [30]:

$$\frac{\omega_p - \omega_0}{\omega_0} = -\frac{\int_{V_0} (\Delta\varepsilon \left|\vec{E}_0\right|^2 + \Delta\mu \left|\vec{H}_0\right|^2) dv}{\int_{V_0} (\varepsilon \left|\vec{E}_0\right|^2 + \mu \left|\vec{H}_0\right|^2) dv} \tag{2.54}$$

where V_0 represents the volume of the cavity, while E_0 and H_0 are the field values in the unperturbed cavity. The permittivity and permeability of the undisturbed cavity are signified by ε and μ respectively (for cavities filled with air, $\varepsilon = \varepsilon_0$ and $\mu = \mu_0$). The terms $\Delta\varepsilon$ and $\Delta\mu$ illustrate the changes in permittivity and permeability respectively, which are the results of the disturbance. The general formula derived in [30] contemplates potential changes in both permittivity and permeability of the cavity. However, for dielectric materials, $\Delta\mu$ equals zero, so the formula can be simplified further. The integrals in equation 2.54 cover the entire volume of the cavity. Yet, it is apparent that the reduction in the resonance frequency of the disturbed cavity is more pronounced if the perturbation in permittivity happens in areas where the electric field intensity is high.

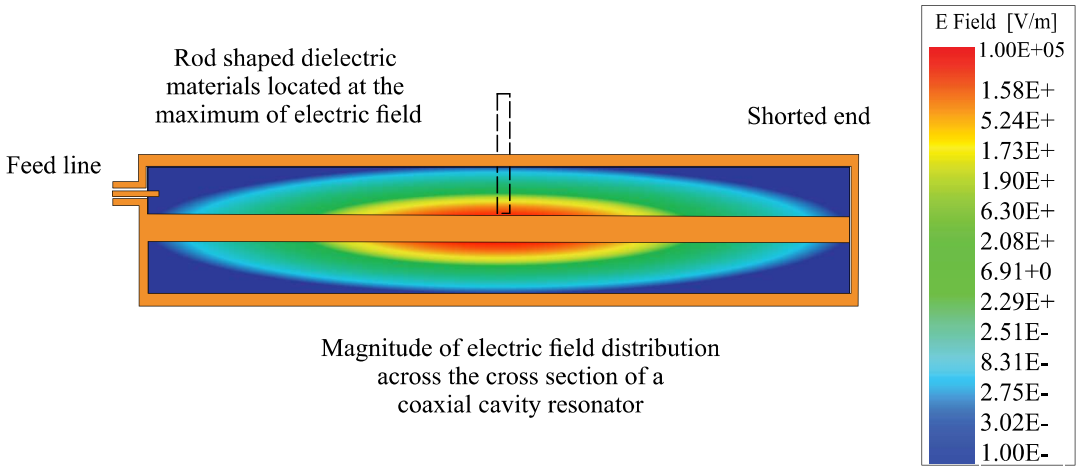


Figure 2.6: Coaxial cavity resonators for material characterization.

2.3.2 Coaxial Cavity Resonators

The coaxial cavity illustrated in Fig. 2.6 is additionally capable of determining the complex permittivity of samples with medium and high losses. This method, first demonstrated in [65], can be examined through the lens of perturbation theory. Its key advantage is its high-quality factors, making it a reliable tool for evaluating samples with high loss. The coaxial cavity resonator is fundamentally a single coaxial transmission line connected to another coaxial line at one end, while the other end is either closed or left open. The cavity presents a fundamental resonance with a single peak of the electric field, along with several higher-order resonances, which can also be exploited for frequency-specific measurements. A rod-shaped sample is introduced into the cavity via a slot carved into the outer conductor, enabling it to be moved along the cavity's length. As coaxial lines possess a broad frequency range, resonance measurements can span a significantly wide bandwidth [66].

2.3.3 Open-Ended Resonant Lines

The resonant open-ended line method, although not commonly employed, is a crucial technique for measuring complex permittivity. Initially introduced in [67], this method involves using a sensor, which is a long coaxial line. The input feed line of this sensor is positioned at a distance of $L_1 = \lambda/4$ from the shorted end, and $L_2 = \lambda/2$ from the open-ended section, as depicted in Fig. 2.7. At resonance, the fields are concentrated at the aperture of the open end of the resonator, where the test material is placed. This method is especially well-suited for analyzing high-loss samples and doesn't necessitate any sample preparation. The sensor is first calibrated in the air prior to interfacing with the test material. The material's presence at the sensor's aperture causes a detuning effect on the response. The response can be returned to its initial frequency by modifying the length L_1. This alteration in length correlates directly with the material's permittivity. After tuning, any decrease in the resonator's quality factor delivers data on the sample's loss [67].

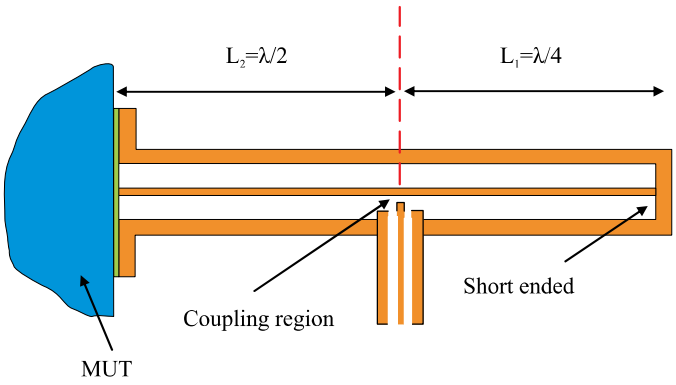


Figure 2.7: Open-ended resonant line for material characterization.

2.3.4 Substrate Integrated Waveguide

The Substrate Integrated Waveguide (SIW) technology enables the integration of a traditional waveguide into a planar substrate, making it suitable for high-frequency applications where low loss and high Q-factor performance are required [68]. In microwave sensing, SIW structures can be used as resonant sensors, where the resonance frequency or quality factor of the SIW resonator changes in response to variations in the properties of the MUT [69].

Due to their high Q-factor and shielding from external interference, SIW resonators can offer high sensitivity, making them ideal for precise material characterization. The resonance of the SIW can be made sensitive to the dielectric constant or permeability, allowing for the measurement of these properties based on the detected changes in the resonance. As an example of SIW-based resonant sensors, Fig 2.8 (a) represents the top view of the sensor and how the material is accommodated. The exploded view of this SIW sensor as shown in Fig 2.8 (b), demonstrates vias surrounding the cavity as well as a tube that passes through the substrate and contains the MUT. It is important to note that the material size is significantly smaller than the cavity volume as it is evident from the cross-section view as shown in Fig. 2.8 (c).

However, designing and fabricating SIW structures for material characterization can be challenging due to the need to control the size, shape, and alignment of the vias that form the waveguide walls. Furthermore, while SIW technology offers benefits such as high Q-factor performance and shielding from external interference, these advantages decrease as the frequency decreases, making SIW less effective for lower-frequency applications.

2.3.5 Metamaterial

Metamaterials are artificially engineered materials designed to have properties not typically found in natural substances, often achieved by structuring them at a scale smaller than the wavelength of the waves they interact with. Figure 2.9 illustrates the unique attributes of TL constructed from metamaterials, characterized by their permittivity (ε_r) and permeability (μ_r). Besides the standard right-handed (RH) TL, there exist three varieties of metamaterial-based TL: 1) the combined right-/left-handed (CRLH)

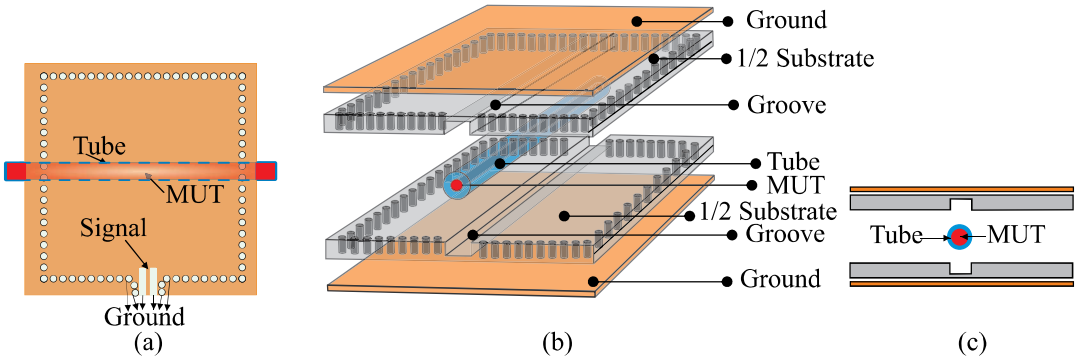


Figure 2.8: SIW structure with the tube passing through it in (a) Top view, (b) Exploded view, and (c) Cross-section.

TL; 2) the transmission line integrated with split ring resonators (SRR); and 3) the transmission line incorporating complementary split ring resonators (CSRR) [70].

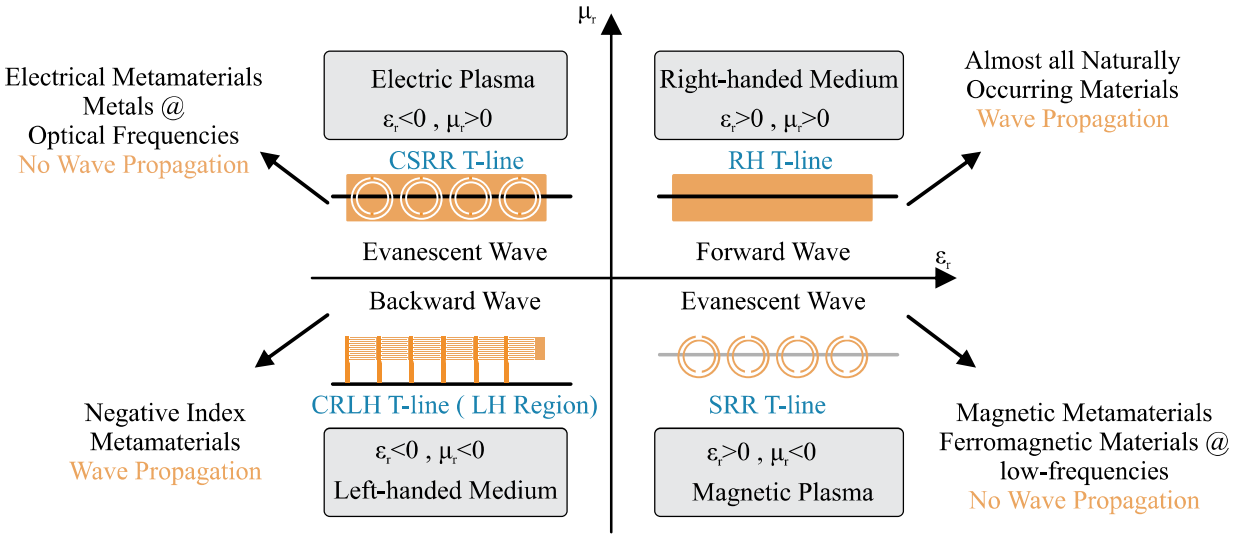


Figure 2.9: Characterization of four types of Transmission line: RH TL, CRLH TL, SRR TL, and CSRR TL.

SRR, CSRR, and CRLH transmission lines, which are common inclusions in metamaterial design, are widely used for material characterization and sensing applications due to their unique and tunable electromagnetic properties [71].

The design of CRLH TL, illustrated in Fig. 2.10, involves various components. C_{LH}, derived from interdigital capacitors, and L_{LH}, obtained from a lengthy metal strip connected to the ground via vias, are included. The conventional TL model is represented as series L_L and parallel C_L. Parasitic and lossy elements of TLs are ignored for simplicity. The interdigital capacitor, utilizing the smallest possible metal pinch as per design rules, fabricates the capacitor C_{LH} to provide maximum capacitance with minimal layout overhead. The inductor L_{LH} is sourced from microstrip lines grounded through via bars (or a collection of combined via bars), which helps decrease ohmic losses. The bottom-most metal layer, acting

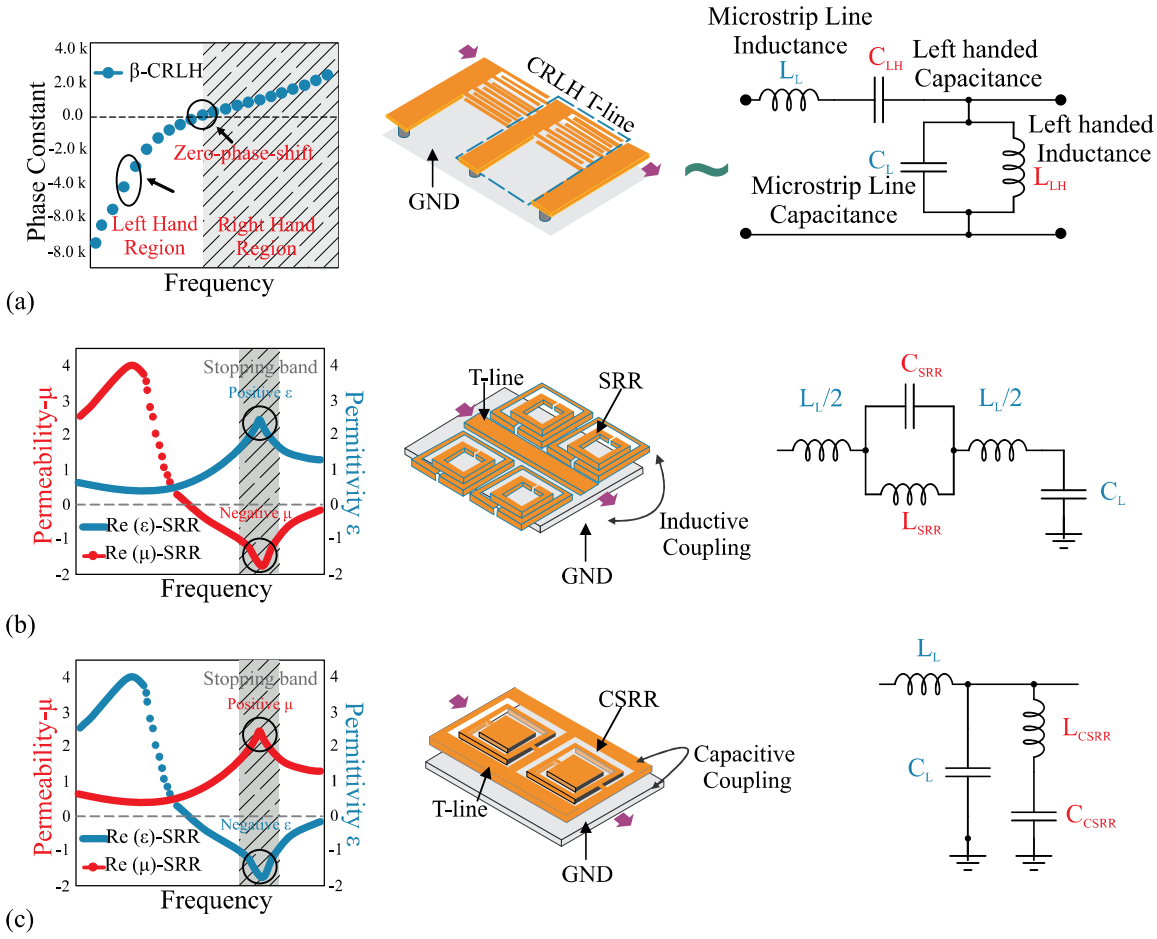


Figure 2.10: Three types of metamaterial T-lines unit-cells including (a)CRLH TL, (b) SRR, (c) CSRR with the corresponding dispersion diagrams (left) and equivalent circuit models (right).

as a ground layer, serves as shielding to minimize substrate loss. A significant characteristic of the CRLH TL is that its phase constant transitions from negative to positive as the operational frequency increases.

In the SRR where the effective permeability becomes negative $\varepsilon > 0$ and $\mu < 0$ (see Fig. 2.10 (b), phase constant $k = \alpha + j\beta = \omega/c = \omega\sqrt{\varepsilon\mu}$ of wave equation $Ae^{i(\omega t - kx)}$ is complex, which results in an evanescent wave followed by reflections from SRR. This results in a resonance profile in a narrow band with a high-quality factor.

In the CSRR case, on the other hand, the effective permittivity is negative $\varepsilon < 0$ and $\mu > 0$ (see Fig. 2.10 (c), where the phase constant $k = \alpha + j\beta = \omega/c = \omega\sqrt{\varepsilon\mu}$ becomes complex, which results in an evanescent wave followed by reflections from CSRR. This again results in a resonance profile in a narrow band with a high-quality factor. In the following section, some examples of sensing elements are elaborated.

Microstrip Resonators

A microstrip resonator is a type of TL resonator, which forms a fundamental component in microwave engineering and RF applications. It consists of a conductive strip, generally made of copper, positioned on a dielectric substrate, which itself is backed by a conductive ground plane. The structure creates a planar resonator capable of guiding microwave signals. The resonant frequency, which defines the frequency at which the structure will efficiently trap and oscillate energy, is primarily a function of the physical dimensions of the conductive strip and the relative permittivity (dielectric constant) of the substrate. some common examples of microstrip resonators are comprised of microstrip configurations with specific strip lengths of λ, $\lambda/2$, $\lambda/4$. These structures then develop into more complex metamaterial inclusions such as SRR and CSRR, which are explained as follows [32, 72].

λ: In a full-wavelength resonator [73], the length of the conductive strip is approximately equal to the wavelength of the resonant frequency in the dielectric medium (see Fig. 2.11(a)). This resonator configuration is less commonly used than its quarter- or half-wavelength counterparts due to its larger physical size.

A standing waveforms along the length of the strip, with voltage nodes at both ends and a voltage antinode at the center. The electric field is therefore maximum at the center and minimum at both ends.

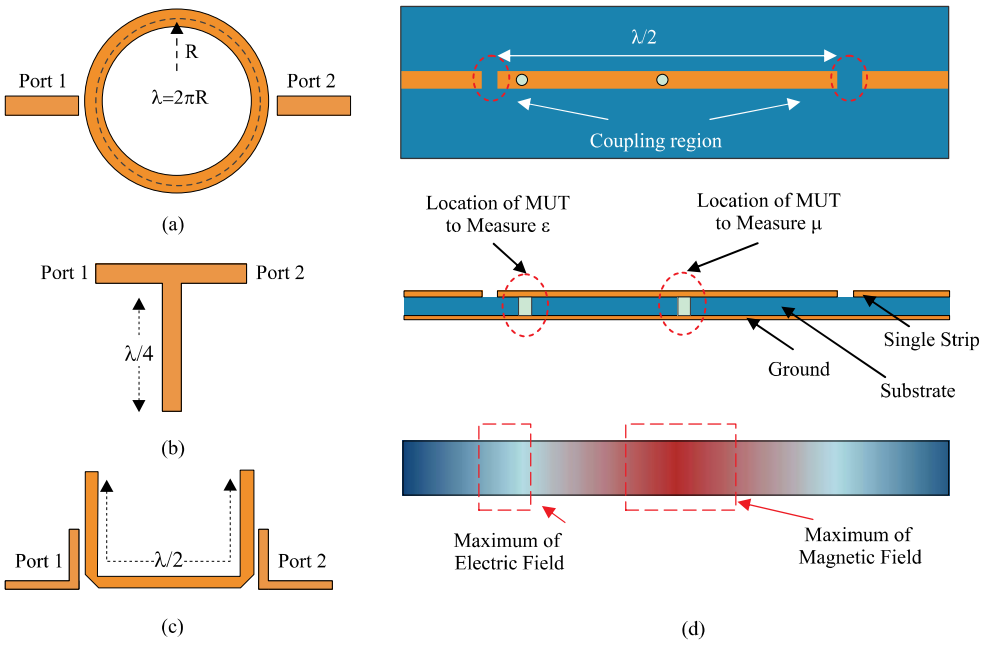


Figure 2.11: Topology of microstrip resonators including (a) Full wavelength λ, (b) Quarter wavelength $\lambda/4$, and (c) half wavelength $\lambda/2$. (d) Representation of maximum electric/magnetic field in a half wavelength resonator.

$\lambda/4$: In a quarter-wavelength resonator [74, 75], the length of the conductive strip is approximately one-quarter of the wavelength of the resonant frequency in the dielectric medium (see Fig. 2.11(b)). A standing waveform along the length of the strip, with a voltage node at the open end and a voltage antinode at the shorted (or grounded) end. The electric field is therefore maximum at the open end and

minimum at the shorted end.

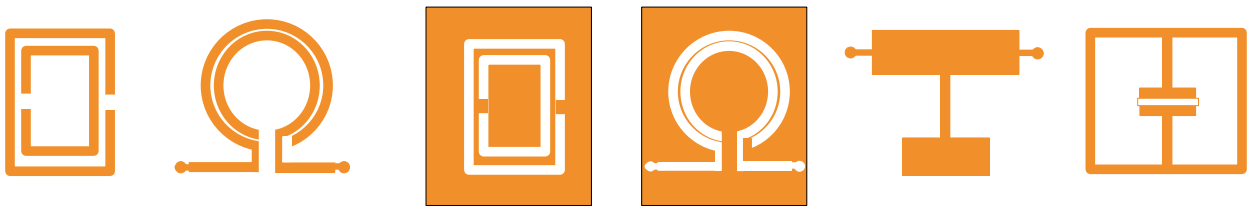
$\lambda/2$: In a half-wavelength resonator [76], the conductive strip is approximately one-half wavelength of the resonant frequency in the dielectric medium (see Fig. 2.11(c)). Here, standing waves form along the length of the strip with two voltage antinodes and a voltage node in the middle. The electric field is therefore maximum at the open ends and minimum at the center.

When the material is introduced into these resonators, it will interact with the microwave energy, causing changes in the resonator's characteristics. This will result in a shift in the resonator's resonant frequency and possibly also its quality factor (Q-factor), both of which can be measured. By comparing these changes with those observed when the resonator is empty or filled with a known reference material, the electromagnetic properties of the material (such as its complex permittivity and permeability) can be deduced.

Figure 2.11(d) illustrates a flat, linear microstrip resonator device as an example of a half-wavelength resonator sensor for material characterization. This configuration is notably effective for conducting resonance disturbance evaluations on liquid samples. The sample is inserted into a minuscule aperture created in the substrate. Due to the higher electric field concentration at the ends of the resonator, the dielectric sample can be situated close to these areas. The interaction between these electric fields and the dielectric triggers a capacitive fluctuation in the response. The main advantage of this approach is that it requires only a tiny quantity of the sample (several microlitres) for analysis. This is particularly beneficial in sectors like pharmaceuticals, where accurate characterization of materials in minute amounts is of paramount importance. If the permeability of the MUT is desired to be measured, it should be located where the magnetic field is maximum.

SRRs and CSRRs are more recent developments that provide additional capabilities for material characterization, particularly for meta-materials and complex electromagnetic properties. SRRs consist of a pair of concentric annular rings with a slit in each, while CSRRs are essentially the 'negative' of the SRRs - they have conductive areas where SRRs have gaps and vice versa.

SRRs and CSRRs possess distinct resonant frequencies determined by their shape, size, and the permittivity and permeability of the material inside them. The presence of the splits in the rings introduces an effective capacitance, which combines with the self-inductance of the rings to form a resonant LC (Inductor-Capacitor) circuit. By observing changes in these resonant frequencies when a material is introduced, the material's properties can be characterized. In particular, they have found substantial use in the study of negative index materials, where unique electromagnetic responses are sought.



(a) (b) (c)

Figure 2.12: Typical resonant elements including (a) SRRs, (b) CSRRs, (c) LC resonators

SRR: In microwave sensing, SRRs are often used as the building blocks of metamaterial-based sensors as shown in Fig. 2.12 (a). The resonant frequency of an SRR structure is sensitive to the properties of the MUT, such as its dielectric constant and permeability. Therefore, by placing the MUT in close proximity to the SRR and measuring the change in the resonant frequency or quality factor, the properties of the MUT can be determined.

SRRs offer the benefit of high sensitivity due to their resonant nature, as well as the ability to engineer their response by modifying their shape and size. However, accurate material characterization using SRRs can be challenging due to their sensitivity to the angle and polarization of the incident wave, as well as to the position and orientation of the MUT. Thus, the use of SRRs in microwave sensing requires careful design and calibration.

CSRR: Complementary Split Ring Resonators (CSRRs), the 'negative image' counterparts of SRRs, offer similar advantages and face similar challenges as SRRs in microwave sensing as shown in Fig. 2.12 (b). The resonant frequency of a CSRR structure is sensitive to the properties of the MUT, allowing for the measurement of these properties based on the detected changes in the resonance.

In addition, CSRRs offer the ability to induce strong electric responses from incident magnetic fields, providing a complementary response to SRRs and expanding the range of materials that can be characterized. However, like SRRs, CSRRs are sensitive to the angle and polarization of the incident wave, as well as to the position and orientation of the MUT, requiring careful design and calibration for accurate material characterization.

Both SRR and CSRR are crucial elements in the field of metamaterials, used for tailoring the electromagnetic properties of a medium. While they are closely related, their operational principles and effects on electromagnetic waves are different. SRRs are designed to provide a magnetic response, i.e., when an external oscillating magnetic field interacts with the structure, it induces a magnetic resonance. Thus, an SRR can essentially be seen as an LC (inductor-capacitor) resonant circuit, where the gaps act as capacitors and the rings as inductors. SRRs are used to achieve negative permeability in metamaterials. CSRRs, on the other hand, are designed to provide an electric response. They react to the electric field component of an electromagnetic wave, rather than the magnetic field component like SRRs. Therefore, they are used to achieve negative permittivity in metamaterials.

LC Resonator: In the corresponding circuit model (see Fig. 2.12 (c)), The equivalent circuit model represents the capacitive square patch, which is etched out of the ground plane, as the C capacitor. The L inductor is embodied by a thin metallic terrace of a specified width w, present on both the top surface and the ground plane, and they are interconnected through a metallic via hole. The sensor's microstrip feed line is likened to a 50 Ω transmission line segment in the circuit model. A 50 Ω lumped resistor is arranged in series with the inductor. During resonance, the LC tank becomes a short circuit, and the microstrip transmission line encounters a corresponding 50 Ω load. This results in a dip in the reflection coefficient (S_{11}) at the resonance frequency.

Regarding SRR-based microwave sensors as shown in Fig. 2.13 (a), material load can be modeled with a parallel RC circuit as shown in Fig. 2.13 (b), where C represents the capacitive effects and R accounts for the losses incurred due to the material. The capillary containing the MUT in a microstrip line sensor

can be positioned either atop the copper trace (see Fig. 2.13 (c)) or within the substrate (see Fig. 2.13 (d)), and the location significantly influences sensitivity due to the degree of field penetration into the material. It is observed that field lines are more concentrated within the substrate than in the air, which implies that positioning the material inside the substrate results in higher sensitivity. However, inserting a capillary into the substrate may be challenging with thin substrates, but tests show a clear increase in sensitivity when the MUT is placed inside the substrate as opposed to on top of the copper track.

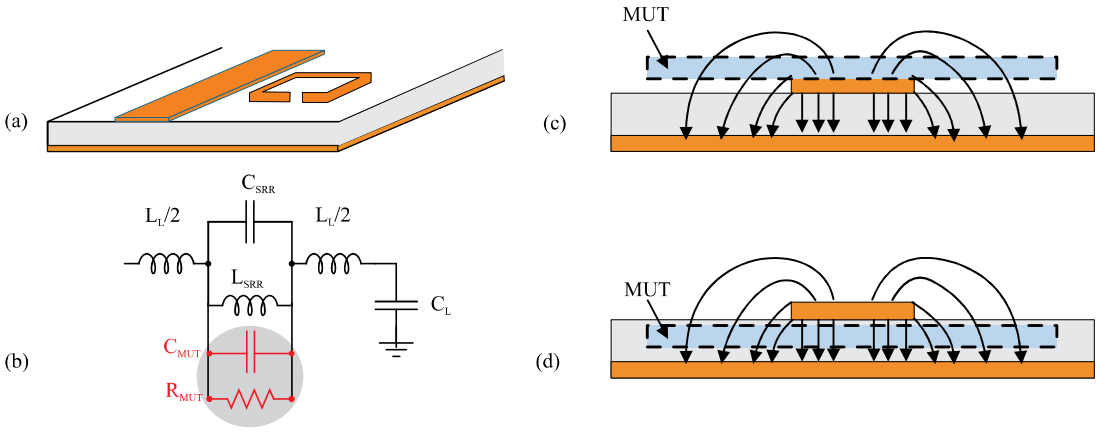


Figure 2.13: (a) Geometry of a microstrip (b) Equivalent circuit of the sensor considering the presence of MUT (c) Capillary located on top of a microstrip line and (d) inside the substrate.

A comparison table for different resonant structures used for material characterization is given in Table.2.2.

Table 2.2: Comparison of resonant microwave sensing structures

Parameter	**Cavity Resonator**	**Dielectric Resonator**	**SRR[1]/CSRR[2]**	**LC Resonator**
Resonant Mode	TE, TM, and Hybrid	TE and TM modes	TE and TM modes	TE and TM modes
Operating Frequency	1-100 GHz	1-100 GHz	0.1-100 GHz	0.1-100 GHz
Losses	Low to moderate	Low to moderate	Low to moderate	Low to moderate
Isolation	High	High	High	High
Sensing Volume	Enclosed cavity	Enclosed cavity	Open/Exposed region	Open/Exposed region
Size	Variable	Variable	Variable	Variable
Measured Property	ε_r, μ_r	ε_r	ε_r, μ_r	ε_r, μ_r
Measurement Accuracy	High	High	High	High

2.3.6 Differential Sensors

Sensors operating on frequency variation principles rely on changes in the resonance frequency (and possibly quality factor or peak/notch magnitude) of a resonant element, instigated by the subject being measured. A typical configuration for these sensors often involves a transmission line loaded with the resonant element, either in direct contact or coupled to it in a SRR or CSRR configuration (Figure 2.14 (a-b)). These sensors are straightforward in their construction, but they can be prone to cross-sensitivities due to alterations in environmental conditions, such as temperature and humidity. Consequently, they require calibration before deployment [77].

Symmetry-based sensors, including coupling modulation sensors, are designed to minimize the impact of environmental cross-sensitivities due to their invariance to environmental changes. These sensors are constructed by symmetrically loading a transmission line with a symmetric resonator (Figure 2.14 (c)); however, for effective operation, the symmetry plane of the resonator and the line must be of different electromagnetic types, one electric and the other magnetic. When symmetry is preserved, coupling between the line and resonator is prevented, resulting in a transparent line without resonance; but when symmetry is disrupted (by asymmetrical dielectric loading or a relative displacement between the line and resonator), coupling arises and resonance appears, with the magnitude of the resonance notch reflecting the degree of asymmetry. The most common applications for these sensors are in angular and linear measurements, but their use is limited by the small dimensions of the sensing resonators, and by the fact that measuring the notch magnitude is more sensitive to noise than frequency measurements [78].

In differential-mode sensors (as shown in Figure 2.14 (d)), One of the sensors is sensitive to the reference (REF) sample (or measurand in general) while the other is sensitive to the MUT sample. If the differential sensor is based on a pair of independent and uncoupled sensing elements, the differential transmission coefficient will be as follows [72]:

$$S_{21} = \frac{1}{2}(S_{21,MUT} - S_{21,Ref}) \tag{2.55}$$

The sensors under consideration exhibit resilience to cross-sensitivities, and their output is fundamentally defined as the discrepancy between the individual sensing elements' outputs. Methods such as substituting meander lines with artificial constructs can be employed to enhance sensitivity in these sensors, all the while sustaining reasonable dimensions. This makes them especially pertinent for applications demanding high sensitivity, such as the detection of minute volume fractions or the quantification of electrolyte concentrations in fluids. Even though differential-mode sensors usually operate with four-port measurements, they have the flexibility to be modified to function with two-port measurements by incorporating supplementary circuit components. This adaptability adds to their extensive applicability [79].

2.3.7 Passive Microwave Sensors

Planar microwave sensors are engineered to operate by resonating at predetermined frequencies, thereby exhibiting sensitivity to the characteristics of nearby materials. Any alterations in these material properties may cause corresponding changes in the resonant frequency or quality factor (Q-factor) of the resonator, allowing for precise material measurement and characterization. Despite this, realizing optimal functionality through the planar structural design is fraught with difficulties. Challenges encompass the tendency for a reduced quality factor (Q-factor), along with increased measurement inaccuracies that might restrict their practicality across various applications. The materials being assessed, especially when possessing lossy attributes such as biological tissues in biomedical contexts, can negatively impact the sensor's performance. This particularly manifests in the degradation of their Q-factor and resolution, a problem that is accentuated in transmission-based sensors and in the impedance matching process for reflection-based sensors.

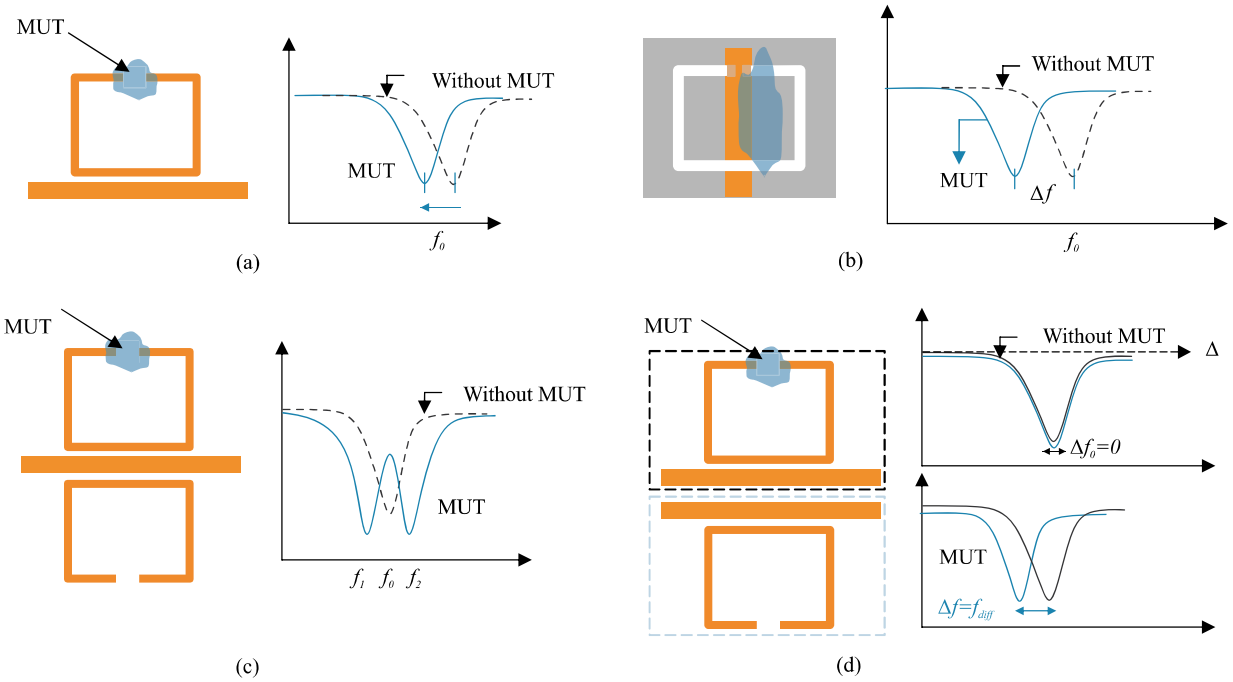


Figure 2.14: Common layout and operational concept of the contemplated planar microwave resonant sensors: (a) Frequency variation, (b) coupling modulation, (c) frequency splitting, and (d) differential.

There has been a thrust to devise novel design approaches to overcome these drawbacks Researchers, with some progress being reported in the development of SIW sensors. With advances in SIW technology, it is possible to develop planar sensors with levels of functionality that were previously only associated with conventional waveguide structures. These sensors are more compact, simpler, have lower manufacturing costs, and are easier to fabricate than their conventional counterparts. These sensors also offer higher quality factors compared to other planar devices, which can help to minimize measurement error detection [80].

In terms of cost, while SIW resonators are more cost-effective than traditional microwave sensor designs, their complexity can limit their use. Existing sensor designs are often costly due to their complex measurement setup and construction. Certain sensors require a large volume of liquid to be completely submerged to gain any interaction between the electromagnetic wave signal and the dielectric properties of the sample, which can also drive up costs [80]. This method of improving the resolution in the passive mode of sensing is associated with a contact form of operation, which is not preferred in most applications when access to material is limited. In order to resolve the low-to-moderate quality factor concern in passive sensors, a considerably redesigned form of sensing concerning the loss compensation is proposed and developed that are discussed as follows.

2.3.8 Active Microwave Sensors

Utilizing active components within resonators, such as transistors or operational amplifiers, can facilitate gain, empowering the resonator to counterbalance losses, and present a superior quality factor (Q-factor).

This is a distinguishing aspect of active resonators, setting them apart from their passive counterparts, which lack this compensatory capability, often resulting in increased losses.

There are two primary strategies that can be leveraged to obtain this gain: negative resistance and regenerative feedback amplifiers.

Regenerative feedback amplifiers use a segment of their output signal to feedback into their input. This feedback can be positive (regenerative) or negative (degenerative). Positive feedback can augment the gain of the amplifier, enabling it to sustain oscillation and counteract losses within the resonator as shown in Fig. 2.15 (a). Within the framework of a microwave planar sensor, it can be tailored to amplify the microwave signal post-interaction with the material and prior to reaching the detector. This amplification serves to cover the losses encountered during signal propagation, ensuring the presence of a detectable signal for analysis. The design of the regenerative amplifier is key to its effectiveness in loss compensation. It must be fashioned to deliver the required gain without introducing extraneous noise or distortion into the signal, necessitating careful component selection, arrangement, and precise parameter tuning. The capabilities of regenerative feedback resonant circuits have been exploited in creating multi-modal sensors that can identify specific actions or changes in the area being monitored, with potential applications ranging from proximity sensing to chemical monitoring.

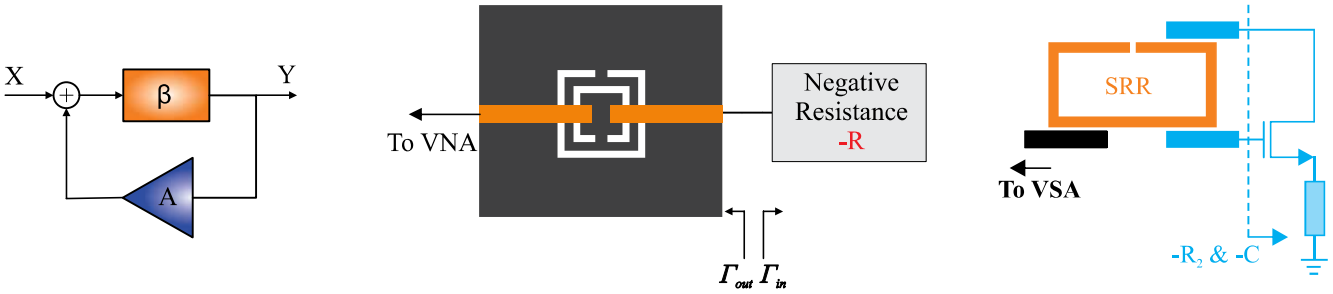


Figure 2.15: (a) Equivalent building block of two port active sensor, (b) One-port active resonator, (c) One-port oscillator with a loss-compensated resonator.

Negative resistance, on the other hand, is a characteristic where an increase in voltage across an element leads to a decrease in current through it. This phenomenon can be used to provide amplification and compensate for losses within the resonator (see Fig. 2.15 (b)). In the negative resistance configuration, the microwave signal, after interacting with the MUT, navigates through the negative resistance device before arriving at the detector. The device effectively amplifies the attenuated signal, compensating for signal losses, and ensuring a detectable signal is available for analysis. The implementation of a negative resistance configuration holds an advantage over conventional amplifiers as it can offer high gain without the necessity for external power, although caution is needed to circumvent instability or oscillation.

Active resonators, in particular, have been deployed in microwave planar sensors, fostering their use in material characterization including liquid/gas [81–84]. The advent of SRRs, subwavelength elements inspired by a metamaterial concept, has expedited this trend due to their high sensitivity, high-quality factor, and compact size. This has led to numerous material characterizations on solids, liquids, and

gases mainly due to high-quality factor, subwavelength dimension, and a localized sensitive region to tiny analytes under test at resonance. The newly developed loss-compensated planar sensors using negative resistance have alleviated concerns regarding signal transmission in lossy mediums.

The following sections will provide a detailed analysis of these sensor types with examples, focusing on their application in material characterization in the microwave domain. Their specific roles in enhancing resolution through the active components and methods utilized for loss compensation will be explored, providing an in-depth understanding of the principles and benefits of active resonators in microwave sensors.

2.4 Resolution Enhancement

Microwave sensors can be of transmission or reflection types based on the application and the number of ports required. The resonators in these sensors can be represented as series or parallel resonators [84, 85]. The loss of MUT has a significant impact on the sensor profile (S_{21}) or (S_{11}), with sharper profiles being easier to distinguish and therefore preferable for high-end applications. However, material loss can degrade sensor performance by lowering the quality factor of transmission-based sensors and affecting the matching of reflection-based ones.

The sensor resolution is a key performance metric that indicates how separate sensor responses become as a result of a given change in the MUT. A high-resolution sensor would result in low-to-no overlap between the original and shifted (S_{21}) profiles when a change in the environment causes a shift in the (S_{21}) profile of a sensor. Various methodologies can be applied to a given resonator to improve its resolution or profile sharpness.

2.4.1 Active Two-port Sensors

Microwave sensors, which are commonly deployed in a two-port setup, are compatible with their inclusion as components in amplifiers or oscillators. The subsequent discussion will primarily focus on high-resolution sensors that function based on transmission response.

Loss in the sensor can originate from several factors such as MUT loss, interference from nearby objects, loss from dielectrics or conductors, and issues related to electromagnetic propagation. Regardless of the specific cause, these losses typically result in a decrease in the sharpness of the transmission profile. The main focus in this context, however, is material loss. One technique to extract information from a lossy MUT is to counterbalance the sensor's loss. A detailed explanation of this approach can be found in [4] and other related works [82, 86–88] have followed it. The proposition of these authors is to use an active sensor as a means to mitigate losses within the system. An active sensor incorporates electronics to offset the resonator's loss. In Fig. 2.15 (a), a simplified version of the proposed system is shown where the resonator is linked to an amplifier within the feedback path to produce a regenerative amplifier. The transfer function of a feedback system with an amplifier gain A and a resonator transfer function β is given by:

$$\frac{Y}{X} = \frac{\beta}{1 + A\beta} \tag{2.56}$$

This function is dependent not just on the forward path β, but is also influenced by the amplification coefficient $(1 + A\beta)$. The ideal conditions for the sensor response involve both the gain and phase of the resonator and the amplifier as follows. The gain $|A| \gg 1$ is used to counterbalance the resonator loss β; i.e., $|A| < 1/|\beta|$. It should be noted that this criterion should be considered with much care since a slightly higher gain can lead to oscillations in the sensor response, an undesired state of the active resonator that can be harmful to the connected devices. Moreover, the wrapped phase of resonator β needs to match the amplifier phase; i.e., $\angle A = \angle \beta$.

Through this method, the resonator's loss can be managed, and the primary problem of system loss, mostly due to MUT loss, can be effectively addressed. It should be noted that only partial compensation is required. If the compensation exceeds this, the sensor will enter into an oscillation state and the mode of operation will be altered.

2.4.2 Active One-port Sensors

Aside from sensors based on transmission, resonators with a single port have been enhanced with loss compensation to function as reflection sensors, utilizing the reflection scattering parameter (S_{11}), as presented in [15]. These types of sensors are ideal for operations necessitating a single port, thereby minimizing computational load during data interpretation as shown in Fig. 2.15 (b).

In the example of a CSRR (subtracted from the ground plane), which provides a highly responsive area for the characterization of dielectric materials, the reflection S_{11} is devised to be at its lowest value at the resonance frequency. The resonance's magnitude is affected by the MUT's loss, and the S_{11} curve becomes less distinct. Using this resonator in a circuit that compensates for losses aids in the recovery of the lost power. The strategy for this compensation is borrowed from the concept of negative resistance, usually applied in oscillator design but utilized here for the improvement of the resonator response. The circuit's negative resistance hinges on the Barkhausen criterion, in which oscillations begin when the reflection parameter at the circuit's input (Γ_{in}) and the CSRR's output (Γ_{out}) fulfill the condition [89]:

$$|\Gamma_{in}| \times |\Gamma_{out}| > 1 \tag{2.57}$$

It is crucial to note that only a portion of the sensor loss is compensated to keep the sensor operating in resonance mode without slipping into the oscillation regime, which means $|\Gamma_{in}| \times |\Gamma_{out}| < 1$. Reclaiming the resonator loss amounts to recouping the power stored in the resonator by adjusting the input impedance of the active circuit to match the complex conjugate of the resonator's output impedance, consequently reducing the sensor's reflections. The degree of compensation can be modulated to control the depth of reflection response at the resonance frequency.

A significant form of loss compensating pertains to complete compensation, enabling the retrieval of sensor response as well as the use of the sensor in single-port applications without external excitation. This method allows the transformation of resonance behavior into oscillation once the resonator's lost

power at the resonance frequency (f_{res}) is completely recovered. This approach offers the advantage of producing a signal at f_{res}, which remains tightly associated with the parameters of the resonator, such as resistance, capacitance, and inductance. The first instance of a planar oscillator sensor was revealed in [90], where the use of a transistor (source degenerated MOSFET) aids in providing the negative resistance needed to alleviate the SRR loss. This process allows for a high-resolution sensing paradigm as the whole transmission/reflection from a resonator is reduced to a single frequency.

In the study of an oscillator Fig. 2.15 (c), the operational frequency is prone to fluctuations in the transistor current that can be originated from white noise, flicker noise, etc. When there is no apparent variation in the MUT, these fluctuations determine the sensor's bit resolution, which is typically conducted through phase noise analysis [91]. Phase noise significantly impacts the stability of microwave oscillators, influencing their utility in sensing applications. Specifically, these phase instabilities lead to noise that radiates from the carrier of an RF signal in the frequency domain. The phase noise $L(f)$ of a system is given by:

$$L(f) = \frac{kT}{2Q^2P\left(\frac{f_0}{f}\right)^2} \tag{2.58}$$

where k represents Boltzmann's constant, T is the temperature, Q denotes the quality factor, P is the output power, f_0 is the center frequency, and f is the frequency offset from the center frequency. Higher phase noise can overshadow the phase changes due to the quantity under measurement, introducing measurement errors and reducing sensor sensitivity and precision.

In the context of SRR-based biosensors, stability is ensured by carefully designing and optimizing the oscillator circuit for maximum output power and load independence on the input and output port. The sensor's bit resolution, *i.e.*, its limit of detection, is determined by fluctuations in the transistor current stemming from white noise and flicker noise.

2.5 RFID Sensors

Chipless RFID microwave planar sensors operate slightly differently from conventional Radio Frequency Identification (RFID) systems. Instead of storing and transmitting data via an integrated circuit or "chip," chipless RFID systems encode data directly onto the RFID tag itself, often in the form of specific structural features that can modify the tag's electromagnetic response in distinct ways [92–98].

Here's a simplified explanation of how an electromagnetic chipless RFID microwave planar sensor works:

RFID Tag Construction: The tag is made from a material that is capable of resonating or reflecting incoming microwave radio waves in a specific way based on its physical characteristics. It does not have a traditional integrated circuit. Instead, the data is encoded directly onto the tag itself in the form of distinctive physical features, such as specific patterns, sizes, or materials that can affect how it interacts with incoming radio waves.

RFID Reader: The reader emits microwave radio waves toward the tag. These waves interact with

the tag, and their properties are modified based on the physical characteristics of the tag [72].

Signal Reflection and Detection: The modified radio waves are reflected back to the reader. The reader then interprets the reflected waves, with the specific changes to the wave properties corresponding to the encoded data on the tag. For example, the tag might be designed such that its physical characteristics produce distinctive resonances at different frequencies, which can be interpreted as a unique identification code or sensor data.

In the context of a sensor [99], the physical features of the tag could be designed to change in response to specific environmental conditions (e.g., temperature, humidity, pressure), thus altering the reflected signal in a measurable way. This allows the tag to function as a sensor, wirelessly transmitting information about its environment back to the reader.

This chipless approach to RFID technology has several benefits. It can be more cost-effective to produce, as it eliminates the need for a microchip. It also allows for potentially more robust and flexible designs, as the tags do not require electrical power and can be made from a wider range of materials. However, the technology is also generally more limited in terms of the complexity and amount of data it can store and transmit compared to conventional chip-based RFID systems [100, 101].

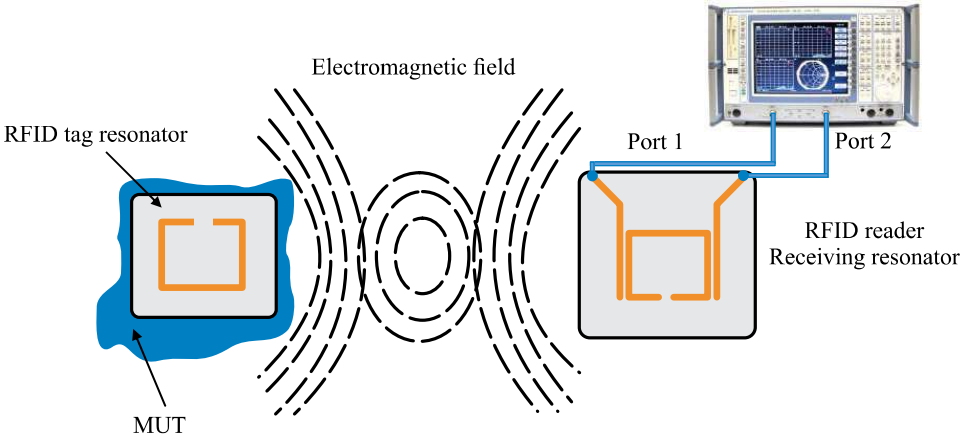


Figure 2.16: Schematic chipless RFID sensing system for material characterization.

RFID sensing technology is becoming increasingly integral in various sectors due to its low-cost, wireless, and resilient features. In healthcare, RFID sensors, whether wearable, implanted, or environmental, can provide essential medical information for healthcare professionals, contributing to the "smart health" concept and bolstering the Internet of Medical Things. In the food industry, these sensors offer critical information about product freshness and cleanliness, aiding quality control. They also have potential applications in smart agriculture for monitoring environmental variables and crop quality. Structural health monitoring in civil engineering can benefit from RFID sensors to detect damage in infrastructure like buildings, bridges, and railways. They are being used in the automotive industry for tire pressure monitoring, intelligent parking, and license plate identification, while in smart cities, they are employed for localization and activity monitoring. Space platforms, where weight reduction and energy saving are crucial, can also use RFID sensors. Despite some challenges related to cost and performance balance, future advancements are expected to further enhance the potential of RFID sensing technology, particularly

benefiting sectors like healthcare, agriculture, and IoT-focused smart city applications [101, 102].

Chapter 3

ML Techniques in MW Sensing

3.1 Machine Learning in Science and Engineering

Scientific machine learning (SciML) [103], a fundamental pillar of artificial intelligence (AI), is a computing tool capable of learning through scientific data to supplement or potentially take over tasks typically performed by humans [104]. ML has found extensive applications in various scientific fields, driving significant advancements and discoveries. In biomedical research, for example, ML is used to predict disease progression [105, 106], analyze genetic data for patterns linked to diseases [107, 108], and aid in the discovery of new drugs [109]. In the field of astrophysics, ML algorithms are utilized to sift through vast amounts of data from telescopes to classify galaxies [110] and detect phenomena such as exoplanets or supernovae [111]. Environmental scientists employ MLg to model complex systems such as climate change [112], where algorithms can help understand and predict patterns related to weather [113–116], greenhouse gas emissions [117], and biodiversity [118].

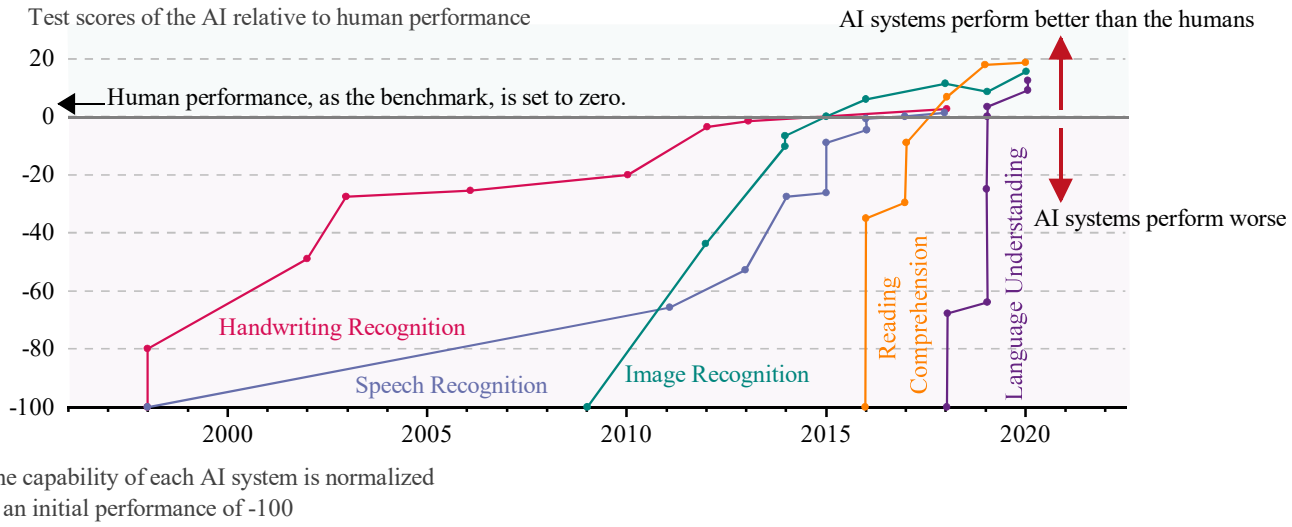


Figure 3.1: The history of language and image recognition capabilities of AI systems, adopted from [3].

AI has made tremendous strides in areas such as handwriting recognition, speech recognition, image

recognition, reading comprehension, and natural language understanding. Fig. 3.1 [119] illustrates the evolution of AI performance in the last 20 years in five domains, ranging from handwriting recognition to language understanding, using tests where both human and AI capabilities were assessed. Performance metrics are normalized such that initial AI performance is marked as -100, and human performance as zero, thus any point where the model's performance surpasses zero signifies an instance where the AI outperformed humans in the corresponding test. Beginning in the 1990s and early 2000s, handwriting, and speech recognition saw substantial improvements due to Hidden Markov Models and, later, deep learning methods. In the 2010s, the advent of deep learning, particularly CNNs, revolutionized image recognition, leading to technologies such as facial recognition and self-driving cars. Reading comprehension and natural language understanding, long-standing challenges in AI, saw significant progress with sequence-to-sequence models and transformer architectures such as Bidirectional Encoder Representations from Transformers (BERT) [120] and Generative Pretraining Transformer (GPT) [121]. These models, trained on large text corpora, have shown remarkable abilities in tasks such as answering questions about text passages, translating languages, and even generating coherent, human-such as text, although true human-level understanding remains an open research question.

3.1.1 History of AI

The history of AI can be categorized into several eras (see Fig. 3.1), including two notable periods known as "AI winters," where progress seemed to stall and interest and funding in the field waned.

1. The Birth of AI (1950s - early 1960s): This era saw the establishment of the theoretical groundwork for AI, with figures such as Alan Turing, John McCarthy, and Marvin Minsky. The focus was on creating systems that could mimic human intelligence, with the development of rule-based and search algorithms, such as the Logic Theorist [122] and the General Problem Solver [123].

2. First AI Winter (mid-1970s - early 1980s): The initial enthusiasm faded due to the high computational cost and limitations of the symbolic reasoning and rule-based approaches, which failed to deliver on earlier promises of creating general AI. The limited success of these early systems led to reduced funding and interest, marking the first AI winter.

3. The Rise of Expert Systems (early 1980s - late 1980s): This period saw the development of expert systems, which encoded knowledge from human experts into rule-based systems. These systems, such as MYCIN for medical diagnoses [124], were successful in narrow domains but struggled to scale or generalize, leading to disillusionment and the onset of another AI winter.

4. Second AI Winter (late 1980s - mid-1990s): The limitations of expert systems, combined with the high costs and limited computational power, led to the second AI winter, with reduced funding and skepticism about the feasibility of AI.

5. The Emergence of ML (mid-1990s - 2010): With the growth of the Internet and the availability of large data sets, as well as increased computational power, the focus shifted towards data-driven AI.

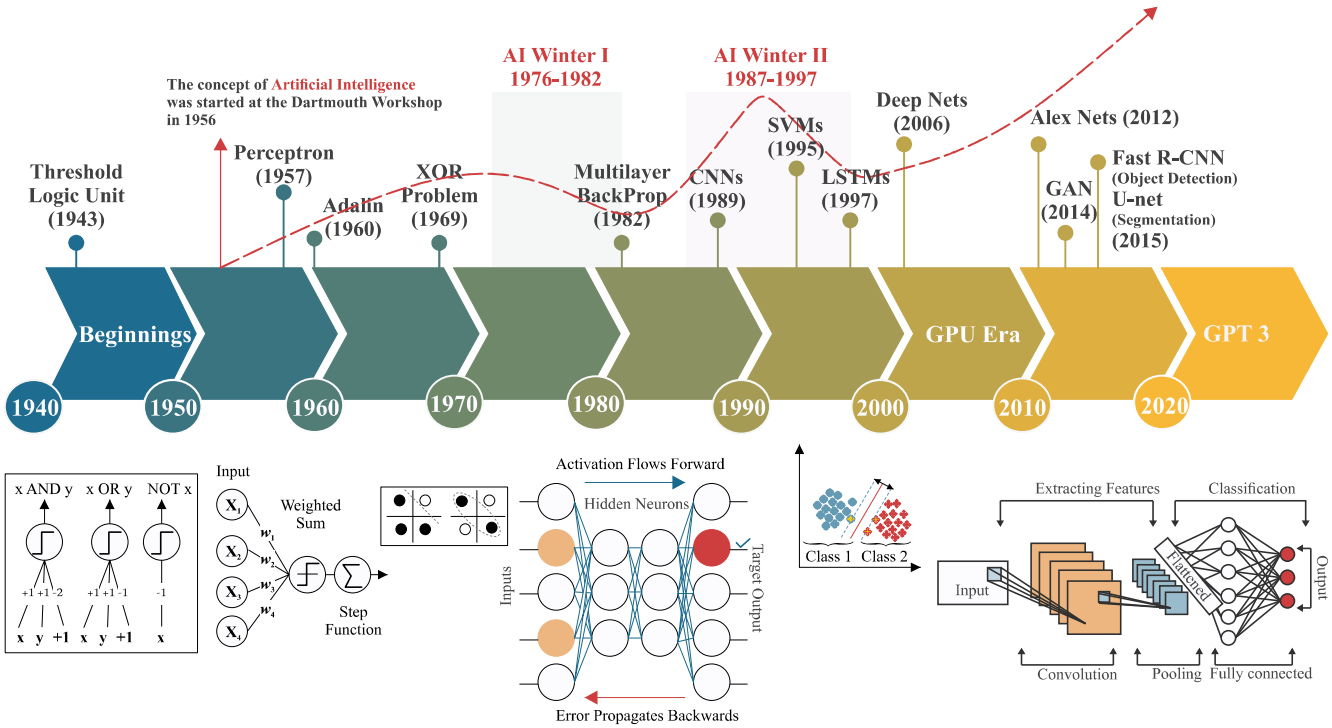


Figure 3.2: Journey through time: The evolution of AI - from early foundations to the era of deep learning.

Algorithms such as SVM [125], decision trees [126], and ensemble methods such as boosting and bagging became popular [127].

6. The Deep Learning Revolution (2010 onwards): The modern era of AI is characterized by the application of deep learning. Neural networks, which had been around for decades but had limited success due to computational constraints and a lack of data, came to the forefront again. Techniques such as CNNs [128] for image recognition and recurrent neural networks (RNNs) [129] for sequence data, along with advances in hardware and the availability of large data sets, led to significant progress in various fields. The success of deep learning continues into the present, with transformer models revolutionizing natural language processing and RL [130] showing promise in areas such as gameplay.

The relationship between artificial intelligence, ML, neural networks, and deep learning can be explained as follows (see Fig. 3.3):

- AI encompasses the capacity of machines or systems to carry out functions traditionally necessitating human intellect. This includes abilities such as critical thinking, making choices, and resolving problems, among others.
- ML, a subset of AI, refers to the unique ability of a system to enhance its performance by learning from data, without needing specific programming for each task. Through ML algorithms, systems can automate the creation of analytical models, and address a wide range of tasks, such as

classification, predicting numerical outcomes (regression), and grouping data based on similarities (clustering).

I trust this revision keeps the essence of the original content while ensuring its originality. If you have any other requests or need further modifications, please let me know!

- A neural network is a type of ML model that mimics the structure and function of the human brain through a set of interconnected nodes or neurons. NNs can learn complex patterns and relationships from data by adjusting the weights and biases of the connections. NNs can have different architectures and layers depending on the problem and the data.
- DL is a subfield of ML that uses ANNs with multiple hidden layers to learn from large amounts of data. DL models can achieve higher accuracy and performance than shallow ML models and traditional data analysis approaches for many applications, such as computer vision, natural language processing, speech recognition, etc.

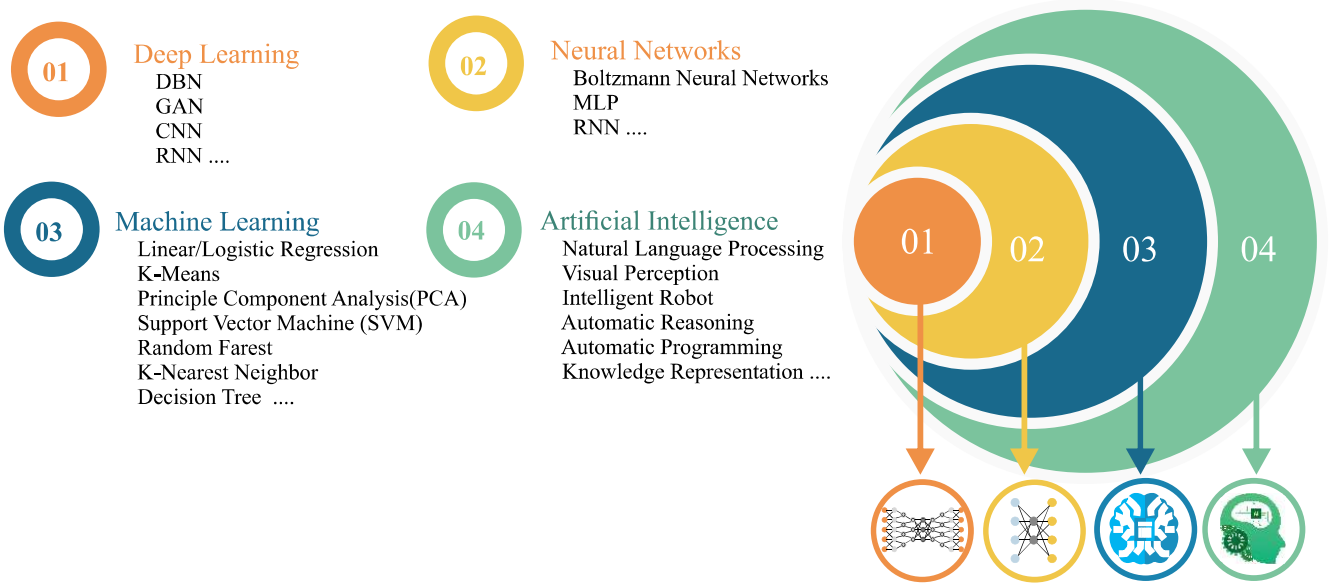


Figure 3.3: Relation of DL with NN, ML, and AI.

3.1.2 Artificial Neural Networks

ANNs are computational structures that draw inspiration from the biological neural networks found in animal brains. Unlike traditional computing systems, ANNs have the ability to 'learn' various tasks through the examination of examples rather than relying on explicit, task-specific programming. In the human brain, a neuron receives input signals through its dendrites, processes the signals, and then sends output signals along its axon, Fig. 3.4 (a). Similarly, an artificial neuron (often called a node or unit) takes a set of input values, performs operations on them, and then produces an output value.

A conventional NN is built upon a specific architecture that includes an input layer, one or several hidden layers, and an output layer. Within these layers, multiple nodes exist, and every node in one layer maintains connections to every node in the neighboring layers. These connections come with "weights," representing the significance or strength of the inputs. During the learning phase, these weights are fine-tuned to alter the inputs' importance in the calculations. Each node also includes a "bias" value, permitting a shift in the activation function, thus enhancing the model's adaptability. Activation functions play a key role in determining a node's output based on its inputs, transforming them into an output signal. Popular activation functions encompass sigmoid, ReLU, and tanh. Finally, the loss or cost function is essential for the training process, as it measures the discrepancy between predicted and actual values. Different tasks may employ various loss functions, such as mean squared error for regression, and cross-entropy for classification.

During the training phase, a neural network learns to correlate inputs with corresponding outputs by modifying the weights and biases. This adjustment is in reaction to the loss computed from the training data and is facilitated by optimization algorithms like gradient descent. The network's performance is subsequently assessed using unseen data known as the test set. This set is distinct from the training set, thereby providing an evaluation of how effectively the model can generalize to new data, as illustrated in Fig. 3.4 (b). Over a series of training iterations, referred to as "epochs," the model is trained to minimize error. With a well-structured network and sufficient data, it can even learn to generalize its predictions to new data in the test set. Striking the right equilibrium between the complexity of the network and the duration of training is crucial. This balance helps to avoid a common pitfall known as overfitting and leads to efficient neural network training.

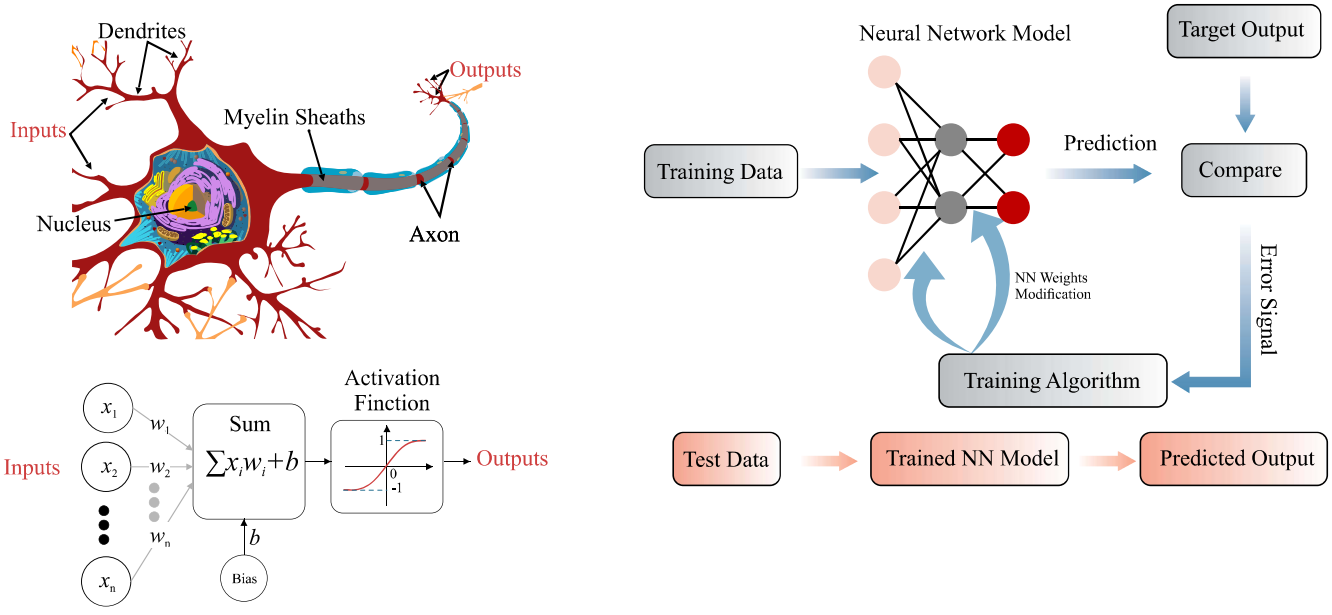


Figure 3.4: (a) Neural network model inspired by a biological neuron, (b) NN in its operational mode.

3.2 Microwave Sensing and Data Curation

Microwave sensing technology, a cornerstone of modern material characterization, has seen significant advancements in recent years [72]. This technology, which leverages the interaction between microwave signals and materials, provides a wealth of information about the physical and chemical properties [131–133] of the materials under test. As we delve into the specifics of material characterization techniques and the role of neural analysis in processing sensor data, it is important to note the transformative potential of integrating ML with these sensors. This integration promises to enhance the capabilities of microwave sensing, leading to more accurate and efficient material characterization [134–136].

Recently, the potential of incorporating ML to empower microwave sensors has emerged as a promising avenue, revolutionizing the way we approach material characterization. ML algorithms, with their ability to analyze vast amounts of data and detect intricate patterns, can greatly enhance the capabilities of microwave sensors. This goes beyond mere automation, offering intelligent data analysis, interpretation, and actionable insights. By integrating these intelligent algorithms, researchers can achieve real-time analysis, predictive modeling, and adaptive sensing, paving the way for more dynamic and responsive systems. For example, in the field of healthcare, ML-enabled microwave sensors can provide continuous monitoring of vital signs, adapting to individual patient needs, and predicting potential health issues before they become critical. In the automotive industry, predictive maintenance systems foresee potential failures before they occur, improving safety and reliability. Furthermore, in environmental monitoring, self-adjusting sensors that adapt to environmental changes can provide more accurate and timely data, essential for understanding and combating climate change. This synergy between microwave sensing technology and ML not only represents a significant technological advancement but also embodies a paradigm shift in our approach to material characterization. The integration of these two technologies leads to the development of intelligent, adaptive systems that can optimize themselves, learn from new information, and respond to changes with unprecedented speed and accuracy. From healthcare and transportation to environmental studies and industrial automation, this combination promises a new era of innovation, efficiency, and impact, positioning itself as a critical technology for the future of human-centric solutions and applications.

3.2.1 1D and 2D Representations of Microwave Sensor Data

There are several possibilities for feeding sensor data into a neural network, each with its unique characteristics and applications. In the pursuit of material characterization and system analysis using MW sensor responses, one-dimensional data can be instrumental. These data are extracted from parameters such as S_{11} or S_{21} profiles, amplitude, phase, frequency, quality factor, and input impedance. They form inputs to NN to extract properties such as the dielectric constant or loss tangent of the MUT, or to implement applications such as classification of different concentrations, time series analysis, and other domain-specific tasks. Using one-dimensional data as input, the NN benefits from a reduced computational burden and a streamlined architecture, making it more manageable to process and analyze. The efficiency of using one-dimensional data works well with the complex demands of MW sensors, allowing for

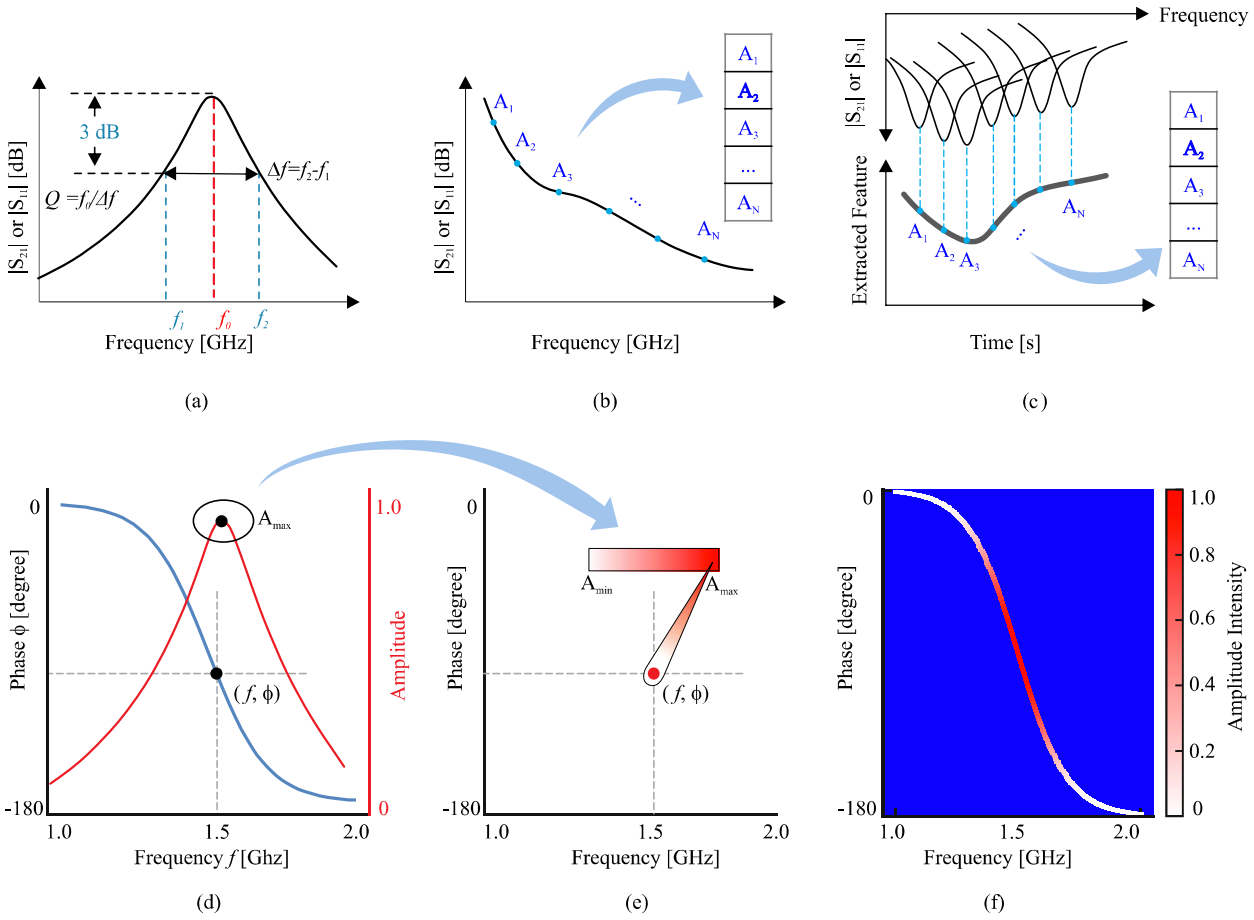


Figure 3.5: (a) Transmission/reflection profile of the sensor, with key features such as resonance frequency (f), amplitude (A), and quality factor (Q) extracted as 1D data for network input. (b) For non-resonant methods, the entire response over a wide frequency range is converted into a 1D vector for network input. (c) Time-based analysis of the data, where the most important feature (in this case, the amplitude at resonance) is extracted as 1D data for network analysis. (d) Symbolic sensor response in magnitude and phase. (e) Process of converting a sensor response into an image. (f) The final heatmap (image version) is generated from a combination of 1D sensor responses (phase and amplitude). This heatmap serves as a 2D input to the neural network, capturing complex spatial relationships within the data. The use of 1D data reduces the computational burden and streamlines the network architecture, while the 2D heatmap provides a comprehensive visualization of sensor response.

quicker training and often leading to accurate and robust results. Whether determining intricate material properties, categorizing various substances, or analyzing temporal patterns, the use of one-dimensional sensor response data offers a versatile and effective approach. When the specific requirements of the MW field are met with the powerful capabilities of NN, one-dimensional data serves as a bridge, enabling advanced analysis techniques and fostering the development of cutting-edge solutions in materials science and engineering. However, this approach might be applicable only when the focus is solely on a specific aspect of the signal.

A more comprehensive visualization can be achieved by combining the most important features of the sensor response, the amplitude and phase information, into a single plot. This combination can be effectively represented through a heatmap, where the phase vs. frequency plot is color-coded based on the magnitude values. The generated heatmap is then treated as an input image, where each pixel can accept a combination of colors. Using a heatmap or other image representation of the microwave sensor output

allows the data to be treated as a 2D structure. This approach can capture complex relationships within the data. The choice of data representation largely depends on the specific algorithm and application at hand. Some algorithms are inherently designed to work with image data, leveraging spatial correlations and patterns. This leads to the observation that working with image representations may sometimes be more understandable or easier, especially considering the vast array of image processing and computer vision techniques available. Transformation of microwave sensor data into various formats opens the doors to a myriad of analysis techniques and neural network architectures. Whether one chooses a simple 1D representation or a more complex image-based format such as a heatmap, the decision should align with the problem's nature, the algorithm's requirements, and the desired insights.

The trend towards image-based data processing, particularly in the microwave sensor domain, capitalizes on the rich set of tools developed in the fields of image processing and computer vision. It not only offers a visually intuitive understanding of the data, but it also leverages advanced algorithms designed to work with complex spatial relationships. This adaptability illustrates the flexibility and potential of modern data analysis in handling intricate and multidimensional data, fostering a deeper understanding of the underlying physical processes.

3.3 Application of Machine Learning in Sensors

As planar microwave sensors are increasingly utilized for material characterization, their performance can be further enhanced by ML. These sensors are typically used to detect and measure the properties of materials by analyzing the interaction between the material and the microwave signal. One of the key applications of ML in this context is the estimation of the dielectric properties of materials. ML models can be trained on data collected from materials with known properties. These trained models can then predict the properties of unknown materials based on the interaction of the microwave signal with the material.

Another important application of ML with planar microwave sensors is in manufacturing for defect detection and quality control. ML can be employed to detect defects or inconsistencies in materials. This is achieved by learning the normal pattern of a microwave signal when interacting with a standard or 'healthy' material. The system can then identify anomalies when the signal pattern deviates, which can indicate a defect or quality issue.

ML holds the capability to construct predictive models for determining material characteristics or identifying internal flaws, specifically utilizing the signals captured by planar microwave sensors. Such an approach paves the way for non-invasive or non-destructive testing—an essential procedure across various domains. For instance, in the field of civil engineering, this technique can be applied to evaluating structures such as concrete or pavement, ensuring integrity without causing damage.

Moreover, ML can be used to classify materials based on their microwave signatures. Different materials interact with microwave signals in different ways, and ML algorithms can be trained to recognize these patterns and classify the material accordingly.

In all these applications, various ML techniques can be used, from classic methods such as SVMs or

Decision Trees to more advanced deep learning methods, depending on the complexity of the task and the amount and quality of the available data. To set the ground, it is helpful to review a list of key terms in ML.

The features or attributes in the context of MW sensors are the individual measurable characteristics or properties of the data instances or samples, including variables such as the resonance frequency, amplitude, and the quality factor from the sensor response. These features represent critical input to make predictions or classifications, reflecting the interplay between microwave sensing and ML. They can be individual values (such as the magnitude and phase of a microwave signal) or characteristics (S21 profile), providing valuable information for training ML models to learn patterns and relationships within the data. In the domain of microwave sensing, feature selection, and engineering become vital, allowing the model to better understand the underlying patterns of microwave signals and make accurate predictions or classifications. Entropy in machine learning can also be related to uncertainty in the sensor response, where a lower entropy indicates a greater purity in the signal and a greater predictability of the data. Information gain refers to the reduction in entropy or uncertainty achieved by splitting a data set based on a particular attribute, such as a specific frequency component, indicating the usefulness of that attribute in making accurate predictions or classifications within the microwave spectrum. Supervised learning is applied when the model learns from labeled training data (associating microwave sensor input features with corresponding target labels) to make predictions or classifications on new, unseen data. Unsupervised learning operates when the model extracts insights from unlabeled data, such as raw readings from sensors. In this context, it seeks to identify underlying patterns, structures, or connections without reliance on specific target labels, allowing for the discovery of inherent formations or clusters within microwave data. In contrast, semi-supervised learning merges both labeled and unlabeled information. It capitalizes on the limited quantity of labeled data coupled with a more extensive set of unlabeled data. This combination enables the creation of predictions or classifications concerning new, previously unobserved microwave signals. Active learning in the context of microwave sensors might involve the model interacting with a human or an oracle to selectively and iteratively query for additional labeled data, aiming to improve its performance by actively seeking the most informative samples for training. RL is a type of ML that could be uniquely applied to the domain of microwave sensors, particularly in designing sensor configurations and structures that are not easily attainable through human intuition or traditional methods. Through RL, an agent learns to make decisions by interacting with a real or simulated environment, receiving rewards or penalties for its actions in the sensor design space. The iterative learning process helps to explore innovative design pathways, attempting to maximize the total reward over time, which may correspond to sensor performance metrics such as sensitivity. It involves learning a policy, which is a mapping from states to actions that leads to optimal sensor designs, using methods such as value iteration, policy iteration, or Q-learning. This approach opens new horizons for sensor design, allowing for the discovery and implementation of novel and complex configurations that can elevate the field of microwave sensing to new heights. This integration of RL into the sensor design process represents a cutting-edge intersection of ML and microwave engineering, providing a promising avenue for future research and development. A detailed and comprehensive view of the various ML paradigms — supervised, unsupervised, and RL —

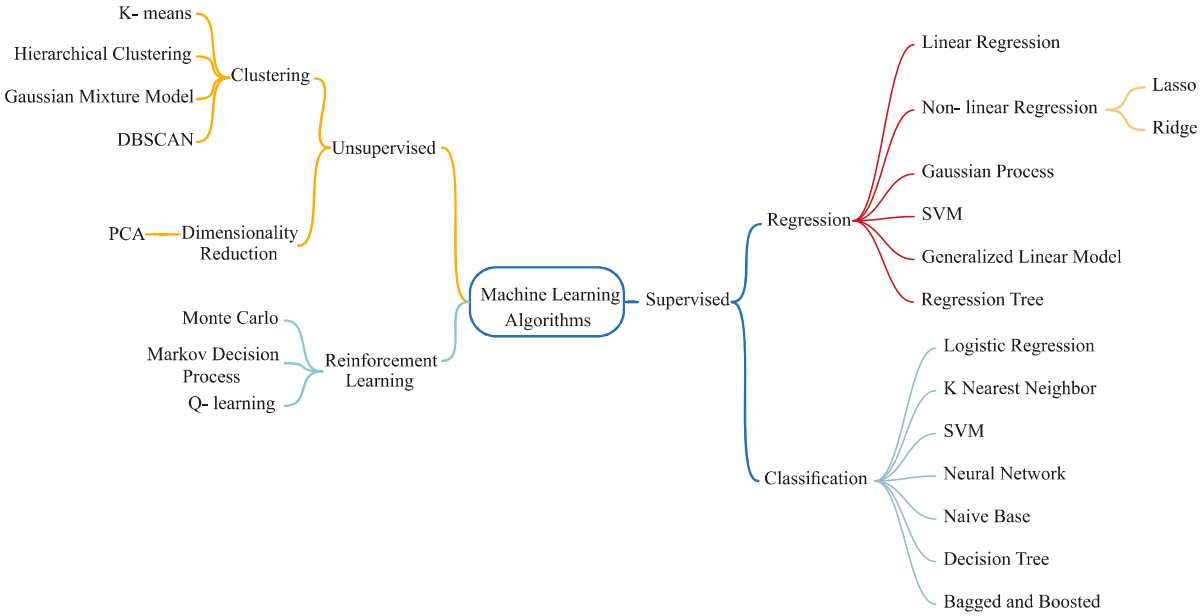


Figure 3.6: Exploring machine learning landscapes: a visual guide to supervised, unsupervised, and reinforcement learning paradigms and their respective algorithms

and the specific algorithms associated with each are showcased in Fig. 3.6. Although there are a plethora of applicable ML techniques, the following sections explore the ML approaches most commonly used in MW sensing.

3.3.1 Classification

Classification represents a distinct form of supervised learning where an algorithm is meticulously trained on labeled training data to create a model. Once this model has been tuned, it is put into action to predict the class or category of unseen data, as depicted in Fig. 3.7. The classification procedure comprises a sequence of stages, each playing a critical role in comprehending and utilizing the model effectively. The typical steps involved in the classification process include:

Preprocessing: The preprocessing stage is a foundational step in preparing raw data for machine learning. It includes a series of crucial actions such as managing missing data points, filtering out anomalies, converting categorical variables into a numerical format, normalizing, and extracting features.

Within this context, normalization sometimes referred to as standardization, stands out as a pivotal preprocessing task. It is a specialized scaling process that alters all variables in the dataset so they have a mean of zero and a standard deviation of one. This transformation is carried out without changing the differences in the value ranges or losing any vital information. Normalization's significance in machine learning extends across various aspects, not least of which is its role in ensuring a consistent approach. ML algorithms, particularly those that rely on the Euclidean distance between data points in their computations, tend to perform optimally when the numerical input variables are scaled to a standard range. Without normalization, features on larger scales can dominate this distance, skewing the algorithm's performance. Second, normalization can accelerate the learning process. By enhancing the numerical

stability of learning algorithms, it reduces the falsehood of encountering problems during computation, leading to quicker convergence. This efficiency can be particularly beneficial in large-scale applications or when computational resources are limited. Third, normalization helps to prevent bias in the model. If the scales of the features vary significantly, a feature could disproportionately influence the model, even if it is not more informative or predictive. This could lead to the model being unduly biased towards that feature. Normalization ensures that all features contribute on a level playing field, thereby avoiding such bias. Last, normalization aids in interpretability. When features are on the same scale, particularly in linear models, it becomes easier to compare coefficients. This can provide valuable information on the relative importance of different features, improving our understanding of the model and the relationships it has learned. However, it is important to remember that normalization might not always be necessary or beneficial for all types of ML algorithms. For instance, tree-based models are not affected by the scale of input features. Understanding your data and the model's requirements is crucial when deciding whether to use normalization.

Train-Validation-Test Split: In the context of developing a machine learning model, the available data is commonly partitioned into three distinct subsets: a training set, a validation set, and a test set. The training set serves as the foundation upon which the model builds its understanding, adjusting its internal parameters to minimize errors. The validation set is then employed to fine-tune the hyperparameters, ensuring that the model is neither overfitting nor underfitting. Finally, the test set provides an unbiased evaluation ground, allowing for an assessment of how the fully-tuned model performs on unseen data, ensuring its applicability to real-world scenarios.

Model Training: Involves feeding the training data into the classifier and allowing it to learn from the data. The learning process involves optimizing a loss function, which measures the discrepancy between the model predictions and the actual labels. The choice of classifiers depends on the data type, size, etc. Several common classifiers are elaborated in Table 3.1 with parameters to be tuned.

Hyperparameter Tuning: Hyperparameters are parameters that are not learned from the data but set prior to training. They include learning rate, regularization parameter, etc. Tuning them involves using techniques such as grid search or random search to find the values that give the best performance on the validation set.

Model Evaluation: This step involves assessing the performance of the model using various metrics. For classification tasks, these include the following:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.1}$$

$$Precision = \frac{TP}{TP + FP} \tag{3.2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3.3}$$

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{3.4}$$

where TP, TN, FP, FN are the numbers of true positives, true negatives, false positives, and false

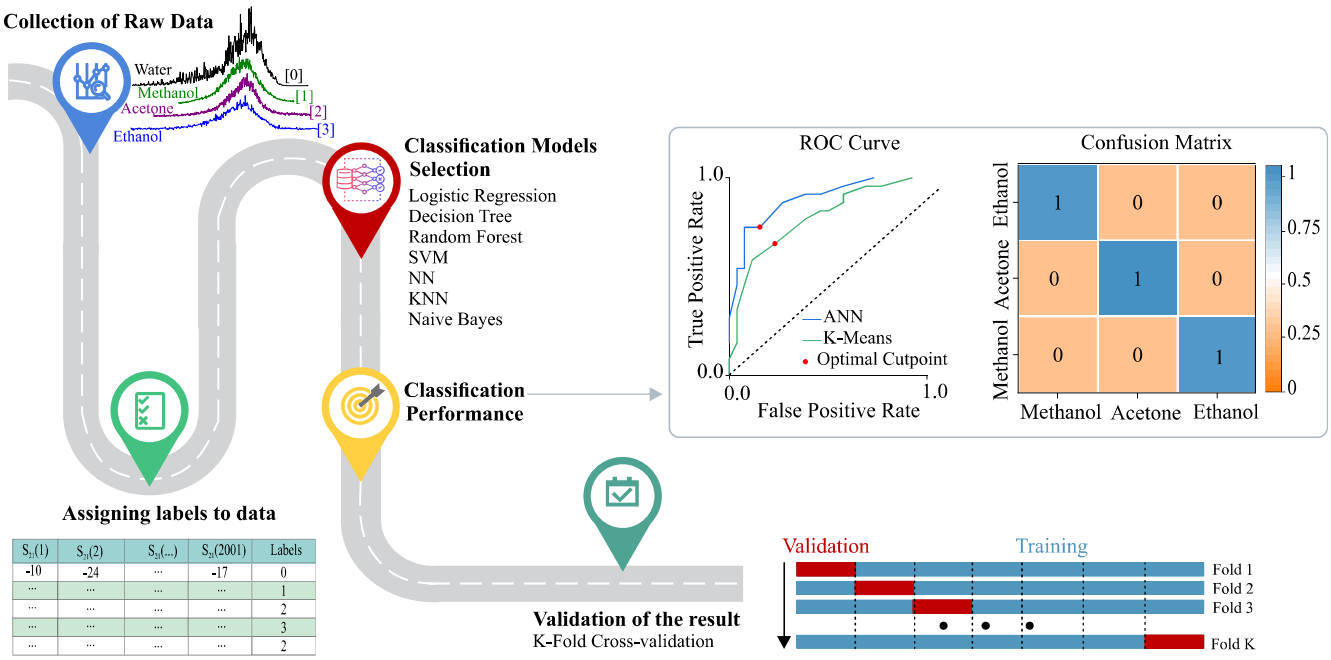


Figure 3.7: A roadmap to classification algorithm in ML.

negatives, respectively.

Performance Metrics and Their Application: The confusion matrix, delineated in Fig. 3.7, is a fundamental table that articulates the efficacy of a classification model on data with known true values. Comprising components such as *TP* (True Positives), *FP* (False Positives), *TN* (True Negatives), and *FN* (False Negatives), it serves as the cornerstone for deriving classification attributes such as *Accuracy*, *Precision*, *Recall*, and the $F1$-score. These components collectively furnish an all-encompassing snapshot of the model's capabilities. Furthermore, the model's performance is enriched by the utilization of the ROC curve (receiver operating characteristic) and the AUC (area under the curve), which delineate the true positive rate (TPR) against the false positive rate (FPR) at differing threshold levels. AUC acts as an encapsulated representation of the classifier's efficiency in discerning between classes. The judicious choice of metrics is imperative, tailored to the distinct problem, potential ramifications of incorrect predictions, and the intrinsic attributes of the data set.

Accuracy is a measure that reflects the ratio of correctly predicted instances by the classification model. The metrics of *Precision* and *Recall* take a different angle, focusing on the correct positive predictions and the proportion of actual positives recognized, respectively. The F_1 score, acting as a harmonic mean between *Precision* and *Recall*, serves as a more nuanced gauge for wrongly classified instances compared to *Accuracy* alone. Selecting the most suitable metric is a vital task and relies heavily on the balance of the dataset, and the significance of false positives and negatives in the context. These choices frequently tie back to the specific field of application. For instance, when dealing with anomaly detection within sensory data (where anomalies such as abrupt pressure changes or sudden environmental variations are infrequent compared to typical data), the $F1$ score becomes an essential tool. By equalizing *Precision* and *Recall*, this measure forms a robust evaluation criterion for the model, ensuring that the rare but

Table 3.1: Comparison table of different classifiers.

Classifier	**Advantages**	**Disadvantages**	**Hyperparameters for Optimization**
Logistic Reg	Easy to implement, interpret, and efficient to train. Good for binary classification. Provides probability score for outcomes.	Assumes linear decision boundary which may not be suitable for complex data with many features.	Learning rate, regularization parameter
Decision Tree	Easy to understand and visualize. Handles categorical and numerical data.	Can easily overfit or underfit the data. Sensitive to small changes in the data.	Maximum depth, minimum samples split, maximum features
Random Forest	Handles overfitting problem of decision trees by averaging multiple trees. Less affected by noise and outliers.	Complexity, slower prediction time.	Number of trees, maximum depth, minimum samples split
Naive Bayes	Simple and fast. Works well with high dimensional data.	Assumes all features are independent, which is not always the case in real-world data.	Prior probabilities, smoothing parameter
SVM	Effective in high-dimensional spaces and with complex boundaries.	May be inefficient with large datasets. Requires good choice of kernel.	C (Error term), kernel type, gamma (Kernel Coefficient)
K-NN	Simple, no training time. Naturally handles multi-class cases.	Computationally expensive during testing, sensitive to irrelevant features and the scale of the data.	Number of neighbors, distance metric
NN	Highly flexible, can model complex, nonlinear relationships.	Require large amounts of data and computational resources. Can be difficult to interpret.	Learning rate, number of layers, number of neurons per layer, activation function, batch size, epochs, regularization

critical events are identified without distortion by class imbalance.

In ML, overfitting and underfitting stand as two prevalent issues linked with the model's complexity and its capability to adapt to new, unseen data. Overfitting occurs when a model, due to its excessive complexity in relation to the available data and noise, performs exceptionally well on the training set but fails to yield good results on new data. Underfitting, in contrast, is the situation where the model is too simplistic to understand the intrinsic patterns within the data. To address overfitting, strategies such as acquiring more training data, simplifying the model, applying regularization methods, and utilizing ensemble techniques may be employed. For combating underfitting, methods such as enhancing the model's complexity and reducing the regularization factors are found to be beneficial.

Validation techniques, such as cross-validation, are vital in making sure that the model works correctly. This is about finding a good middle ground between how complicated the model is and what it can learn from the data. This idea is sometimes called the bias-variance tradeoff. Ways that classifiers can be used to help with processing data from sensors and figuring out how to classify materials are shown in Table 3.2. The best model performance in terms of a confusion matrix occurs when all instances are correctly identified. The presented overview emphasizes different aspects of classification modeling, shedding light on vital concepts, challenges, and field techniques, enriching the understanding and application of ML models.

Table 3.2 provides a detailed overview of various analyses performed using classification in the field of microwave sensors for material characterization. It encompasses a broad range of sensor types, including SRR, CSRR arrays, time domain reflectometry, and interdigital capacitors. Some prominent examples include the characterization of methanol/acetone using an SRR sensor with an MLP (multilayer percep-

tron) NN model, achieving an accuracy of 97.9%, and olive oil adulteration analysis using CSRR with NN, resulting in 100% accuracy. The table emphasizes the diversity of applications and highlights the efficacy of various ML algorithms in microwave sensing.

Table 3.2: Some examples of classification-based analysis in microwave sensors

Ref.	Sensor Type	f_{res} [GHz]	MUT	Best Algorithm	Metric
[14]	SRR	1.19	Methanol/Acetone	MLP	Acc = 97.9%
[137]	CSRRs array	1.3, 3, 5, 6.8, 8.9	Wood+Plastic/Cardboard	DSVM	Acc = 86.4%
[138]	CSRR	2.16	Honey	BNN	Acc = 100%
[139]	SRRs array	3.05	Ice	ANN	Acc = 94.6%
[140]	CSRR	1.1	olive oil adulteration	NN	Acc = 100%
[141]	CSRRs array	3.4	chloroform/ cyclohexane	BRANN	Acc = 99.6 %
[142]	4-port antenna	3-3.15	Ethanol/Water	ANN	Acc = 93.5 %
[143]	TDF[1]	26.5	Defects in Dielectric	ANN	Acc = 96%
[144]	Interdigital capacitor	1.5-2.9	Oxygen/Nitrogen	ANN	MAE < 0.15%
[145]	SIW+ CRLH	2.5	plastic/input: $[f_s,Q_s]$	k-NN	Acc = 90.2%
			plastic/input: $[f_s,\|S_11\|]$	SVM	Acc = 98.8%

[1]Time Domain Reflectometry

Further examination of Table 3.2 reveals the choice of best ML algorithms and the corresponding metrics used in each study. Algorithms range from basic NN to more specialized models such as Dual Support Vector Machines (DSVM) with SKF cross-validation and Bilayered Neural Networks (BNNs). The preferred metric in most cases is *Accuracy*, reflecting the proportion of correct predictions. The mean absolute error (MAE) is used for NN learning in characterizing oxygen/nitrogen using an interdigital capacitor. The diversity in algorithms and metrics shows the adaptability and tailored approach needed for different material characterizations using microwave sensors, reinforcing the field's complexity and the importance of selecting appropriate methodologies.

3.3.2 Regression

Regression is a supervised ML technique used to predict continuous numerical values based on input variables. The principle behind regression involves fitting a mathematical model to the training data. The resulting model captures the relationship between the input variables and the target variable. The goal is to find the best-fitting function that minimizes the difference between the predicted values and the actual values in the training data set. This is achieved by adjusting the model's parameters using optimization algorithms such as gradient descent, which iteratively updates the parameter values to reduce prediction errors.

During the training process, the regression model learns the patterns and trends present in the training data. The model identifies the optimal combination of input features and assigns the corresponding weights to each feature, determining their influence on the predicted output. The choice of regression algorithm depends on the characteristics of the data and the underlying assumptions about the relationship between the variables. Linear regression assumes a linear relationship, while other algorithms (such as polynomial regression, support vector regression, decision tree regression, or neural network-based regression) can capture more complex relationships.

Here are the metrics that can be considered to evalute the performance of the regressors.

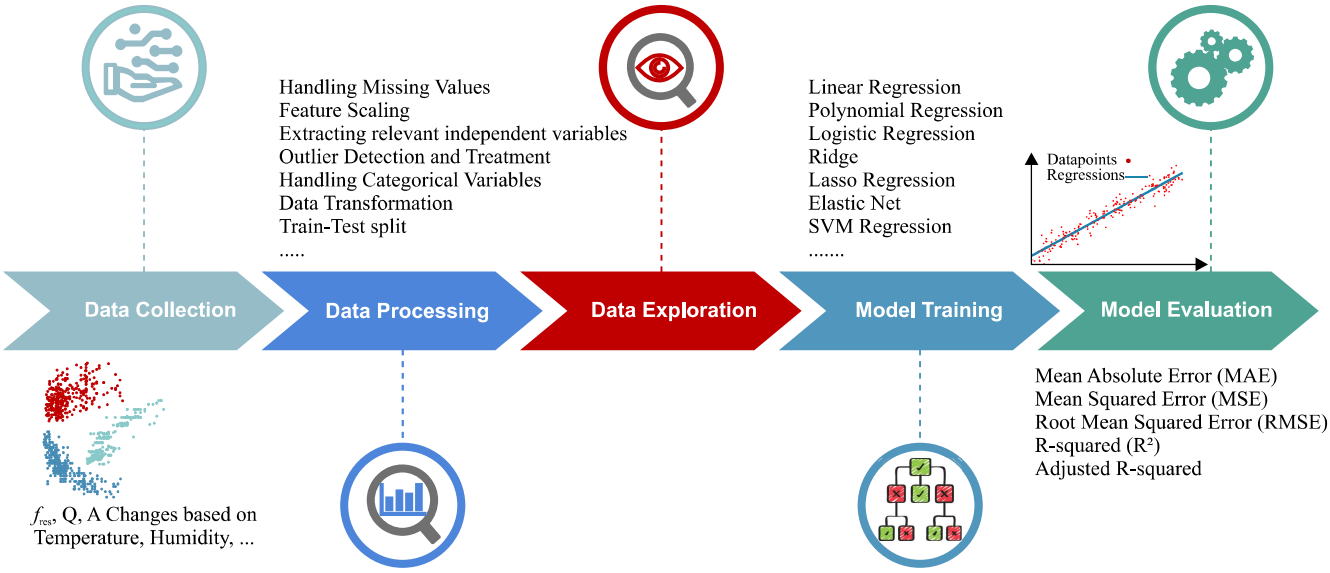


Figure 3.8: A roadmap to Regression algorithm in ML.

Here are the metrics can be considered to avalute the performance of the regressors.

- **Mean Squared Error (MSE)**: This is the average of the squared differences between the predicted and actual values. The formula is:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \tag{3.5}$$

  where Y_i are the actual values, $\hat{Y}_i$ are the predicted values, and n is the number of observations.
- **Root Mean Squared Error (RMSE)**: This is the square root of the MSE. The formula is:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2} \tag{3.6}$$

- **Mean Absolute Error (MAE)**: This is the average of the absolute differences between the predicted and actual values. The formula is:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - \hat{Y}_i| \tag{3.7}$$

- **R-Squared (R^2)**: This is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model.

Table 3.3: Comparison of different regression algorithms in machine learning.

Regressor	Advantages	Disadvantages	Key Parameters
Linear Reg.	Simple, easy to interpret	Limited complexity	
Ridge Reg.	Regularization helps prevent overfitting	Parameter tuning needed	Regularization strength (alpha)
Lasso Reg.	Can perform feature selection	Can be too aggressive, excluding relevant features	Regularization strength (alpha)
Elastic Net	Combines Ridge and Lasso	Requires tuning of two parameters	L1 ratio, Regularization strength (alpha)
Polynomial Reg.	Can fit complex, non-linear data	Can easily overfit	Polynomial degree
SVM	Can model non-linear relationships	Sensitive to hyperparameters	C (error penalty), kernel, gamma (for RBF kernel)
Decision Tree	Interpretable, handles non-linear data	Can overfit easily	Max depth, min samples split
Random Forest	Robust to overfitting, handles non-linear data	Less interpretable, long training time	Number of trees, max depth, min samples split
Gradient Boosting	Can handle non-linear data, robust	Sensitive to hyperparameters, long training time	Learning rate, number of estimators, max depth
NN	Can model complex functions	Requires much tuning, needs large data	Layer sizes, learning rate, regularization, number of epochs

The formula is:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \tag{3.8}$$

where $\bar{Y}$ is the mean of the actual values.

- **Adjusted R-Squared**: This is a modified version of R-squared that has been adjusted for the number of predictors in the model. The formula is:

$$\text{Adjusted } R^2 = 1 - (1 - R^2)\frac{n-1}{n-p-1} \tag{3.9}$$

where n is the number of observations and p is the number of predictors.

- **Mean Squared Logarithmic Error (MSLE)**: This is a variant of MSE. The formula is:

$$MSLE = \frac{1}{n}\sum_{i=1}^{n}(\log(Y_i + 1) - \log(\hat{Y}_i + 1))^2 \tag{3.10}$$

- **Mean Absolute Percentage Error (MAPE)**: This measures the size of the error in percentage terms. It is calculated as the average of the unsigned percentage error. The formula is:

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right| \tag{3.11}$$

Once the regression model is trained, it can be used to make predictions on new, unseen data. The model applies the learned function to the input variables and generates the predicted output values. The accuracy of the predictions can be evaluated using performance metrics such as mean squared error (MSE), providing insights into the model's performance and generalization capabilities.

Table 3.3 includes some examples of regression algorithms, their advantage, disadvantages, and hyperparameter for optimization.

Table 3.4: Some examples of regression-based analysis in microwave sensors

Ref.	Sensor Type	f_res [GHz]	Objective	Best ML Algorithm	Metric
[146]	Coaxial line	0.3-3	Complex permittivity of 6 Alcohols	ANN	MSE = 2.5%
[147]	Grounded CPW	1-10	Complex permittivity of FR-4 substrate	DNN	MSE = 15%
[148]	Spiral resonators	1.8	Complex permittivity Of Water/Ethanol	ML based	$R^2 \approx 99.5\%$
[14]	SRR	1.19	Temperature of MUT	Linear Regression	$R^2 \approx 97.6\%$

Table 3.4, enumerates some noteworthy examples of regression-based analysis implemented in microwave sensors. It illustrates various sensor types, such as Coaxial line operating at 0.3-3 GHz, grounded CPW at 1-10 GHz, spiral resonators at 1.8 GHz, and SRR at 1.19 GHz. The objectives vary, including determining the complex permittivity of six different alcohols with an MSE of 2.5%, the complex permittivity of FR-4 substrate with an MSE of 15%, the complex permittivity of Water/Ethanol with an R^2 of approximately 99.5%, and the temperature of the MUT with an R^2 of approximately 97.6%. Correspondingly, the best machine learning algorithms employed are varied, such as ANN, DNN, Machine Learning-based methods, and Linear Regression, demonstrating the adaptability of machine learning techniques in this field. The metrics for each application, either in terms of MSE or the determination coefficient (R^2), further quantify the success and accuracy of these approaches in the context of microwave sensors.

3.3.3 Clustering

Clustering is an unsupervised ML technique that involves grouping similar data points together based on their intrinsic characteristics. The principle of clustering encompasses several steps shown in Fig. 3.9. First, the data set is collected, and preprocessing tasks, such as handling missing values and feature normalization, are performed. Next, an appropriate clustering algorithm is selected, considering factors such as the shape of the data, the desired number of clusters, and the assumptions made about the data. Popular clustering algorithms include K-means, hierarchical clustering, DBSCAN (Density-Based Spatial Clustering of Applications with Noise) and Gaussian mixture models. The selected algorithm undergoes training on the data that has been preprocessed, systematically allocating data points to specific clusters. This assignment is guided by the inherent similarities or proximity among the points. To gauge the integrity of the derived clusters, various evaluation metrics are employed. Subsequently, an analysis of the clusters is conducted to discern the intrinsic patterns that may be underlying within the data. Clustering has applications in various domains, such as crack detection in pipelines, image recognition, and anomaly detection. It enables the discovery of meaningful groupings and patterns in the data without the need for predefined labels.

Identifying the ideal count of clusters within a dataset represents a foundational challenge in partitioning clustering methods such as K-means clustering. This technique mandates the determination of the quantity of clusters k to be formulated. The process of discerning the most suitable number of clusters k is not straightforward and can be undertaken employing a diverse array of strategies and methodologies. The Elbow Method involves running the clustering algorithm for a range of values of k and, for each value of k, calculating the sum of squared distances from each point to its assigned center (distortions).

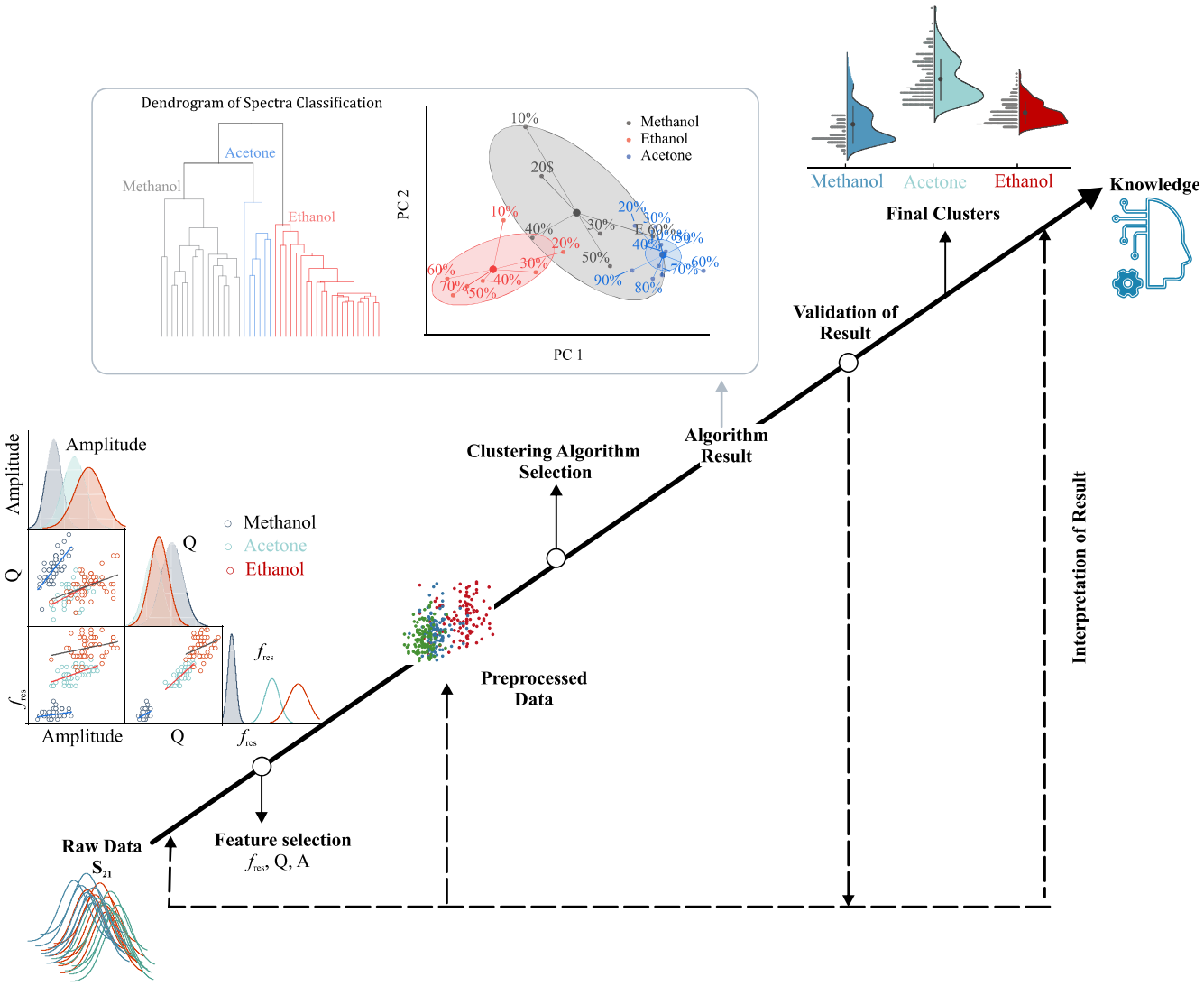


Figure 3.9: A roadmap to clustering algorithm in ML.

When these are plotted, the distortions will decrease with increasing k. The "elbow" point, where the rate of decrease sharply shifts, can be a good choice for k. The Silhouette Method measures how similar a point is to the points in its cluster compared to points in other clusters, with the optimal number of clusters maximizing the average silhouette over a range of possible values for k. The Gap Statistic compares the total intra-cluster variation for different values of k with their expected values under a null reference distribution of the data (i.e., a distribution with no obvious clustering), with the optimal number of clusters being the one that maximizes the gap statistic. In Gaussian Mixture Models and other probabilistic models, Bayesian Information Criterion (BIC) or Akaike Information Criterion (AIC) can be used to select the number of clusters, with the optimal model minimizing the criterion. Additionally, some methods, such as Gaussian Mixture Models, can be used with Cross-Validation to select the number of clusters by looking for the model with the highest cross-validation score. Each of these methods has its strengths and weaknesses (see Table 3.5), and often the results may not agree, especially with real-world

Table 3.5: Comparison of various clustering algorithms highlighting their advantages, disadvantages, and commonly used evaluation metrics.

Clustering Algorithm	**Advantages**	**Disadvantages**	**Common Evaluation Metrics**
K-Means	Simple, easy to implement and interpret results, efficient for large data	Assumes spherical clusters and equal cluster sizes, requires a pre-specified number of clusters	Inertia, Silhouette Score
Hierarchical Agglomerative Clustering (HAC)	No need to specify the number of clusters, yields a dendrogram that provides additional information about the data	Computationally expensive for large datasets, sensitive to noise and outliers	Cophenetic Correlation
DBSCAN (Density-Based Spatial Clustering of Applications with Noise)	No need to specify the number of clusters, identify noise points, can find arbitrarily shaped clusters	Difficult to choose parameters that work for all data types, sensitive to density estimation	Silhouette Score, DB Index
Gaussian Mixture Models (GMM)	Soft clustering (provides probabilities), allow different cluster shapes	assume data distribution is Gaussian, can converge to local optimum	BIC, AIC, Log-Likelihood
Spectral Clustering	Can identify complex cluster shapes, do not assume spherical clusters	Difficult to choose the number of clusters, computationally expensive	Silhouette Score, Rand Index

data that may not meet the assumptions of these methods. The best choice often depends on the specific characteristics of the data and the purpose of the clustering.

Assessment of a clustering algorithm's efficacy is a complex task, differing fundamentally from the mere enumeration of errors or the computation of *Precision* and *Recall* in supervised Machine Learning (ML) methods. Unlike supervised learning, where the outcomes are well-defined and measurable, clustering's unsupervised nature necessitates a distinct set of criteria. Within the field of ML, an array of specialized metrics have been developed and are routinely employed to quantify and interpret the performance of clustering algorithms, reflecting the algorithm's ability to correctly group data without explicit guidance.

Inertia, also known as within-cluster sum-of-squares, measures the compactness of the clustering, i.e., the sum of the squared distances from each sample to the nearest cluster center (lower values are better)

$$I = \sum_{i=1}^{n} \min_{\mu \in C} (||x_i - \mu||^2). \tag{3.12}$$

Silhouette Coefficient is a measure of how close each sample in one cluster is to the samples in the neighboring clusters. It is a measure of how much one sample is similar to its own cluster compared to other clusters. The values range from -1 for incorrect clustering to +1 for highly dense clustering. Values near zero indicate overlapping clusters. Davies-Bouldin index is defined as the average similarity measure of each cluster with its most similar cluster, where similarity is the ratio of within-cluster distances to between-cluster distances (lower values are better). . Normalized Mutual Information (NMI) quantifies the degree of interdependence between two variables by adjusting the Mutual Information (MI) score to a specific range between 0 and 1. A value of 0 indicates no mutual connection, and 1 signifies a complete correlation. On the other hand, the Fowlkes-Mallows Index is defined as the pairwise *Precision* and *Recall*'s geometric mean, which helps assess clustering algorithms. There is no one-size-fits-all "best" method and the results may vary for different metrics even with the same clustering process.

Table 3.6 outlines a selection of examples focusing on clustering-based analysis within the realm of microwave sensors. It showcases various sensor types, such as Horn antenna operating within 2-6 GHz,

CSRR array at 2.4-2.5 GHz, CSRR at 0.5-3.5 GHz, and FMCW Radar at 1 GHz. The materials under test (MUT) encompass various substances, including Alpha-cypermethrin, Glucose, Ethanol/Estrogen, and materials such as PVC, FR4, PE, and Acrylic. A variety of machine learning algorithms are employed for these tasks, including k-NN, SVM, GBM, RF, PCA, and K-means. Interestingly, the metric column mostly lacks specific values, except for the FMCW Radar analysis, where the Elbow method is utilized to determine the optimal number of clusters (best k). This table concisely illustrates the diversity and adaptability of clustering techniques in microwave sensor applications.

Table 3.6: Some examples of clustering-based analysis in microwave sensors.

Ref.	Sensor Type	f_{res} [GHz]	MUT	Best ML Algorithm	Metric
[149]	Horn antenna	2-6	Alpha-cypermethrin	KNN/SVM/GBM/RF	-
[150]	CSRR array	2.4-2.5	Glucose	PCA	-
[151]	CSRR	0.5-3.5	Ethanol/Estrogen	PCA/K-means	-
[152]	FMCW Radar	1	PVC/FR4/PE/Acrylic	K-means	Elbow to find best k

3.3.4 Deep Neural Networks

DNNs, a specialized category within the broader family of Neural Networks (NN), are characterized by their multilayered structure comprising interconnected neurons or nodes. They find extensive utility across diverse domains, including but not limited to, image interpretation, speech deciphering, natural language manipulation, and controlling autonomous vehicles. What sets DNNs apart is their ability to handle voluminous and intricate data, learning from and making predictions based on provided input-output pairs. The fundamental architecture of a DNN is crafted with an input layer, several intermediary hidden layers, and a concluding output layer. The hidden layers engage in nonlinear transformations of the input data, empowering the network to capture intricate features and patterns. The fine-tuning of the network occurs during the training phase, wherein the weights and biases are meticulously adjusted to curtail the disparity between the predicted and actual outputs. Various optimization techniques, such as backpropagation, stochastic gradient descent, and Adam optimization, facilitate this training. The evolution of deep learning has catapulted DNNs to a distinguished position within ML and AI paradigms. A visual portrayal of ML models accentuating the eminent role of DNNs among contemporary algorithms can be seen in Fig. 3.10.

Within the DNN domain, transfer learning often emerges as a pivotal concept. The essence of transfer learning lies in enhancing the predictive performance of a learning algorithm from a specific domain by leveraging knowledge from a correlated domain.

Refering to Fig. 3.10, conventional Supervised Learning models [153] encompass methods such as k-NN [154, 155], logistic regression [156], decision tree, SVM [157], and FFANN [158]. Conversely, the traditional Unsupervised Learning models [159, 160] incorporate methodologies such as k-means [161] and PCA [162]. Some of these classic models have been elaborated in previous Tables 3.1, 3.3, and 3.5.

In contrast, in the realm of modern ML models, techniques such as locally linear embedding (LLE) [163], t-distributed stochastic neighbor embedding (t-SNE) [164], isometric mapping (ISOMAP) [165], kernel principal component analysis (KPCA) [163], and multidimensional scaling (MDS) [166] have made

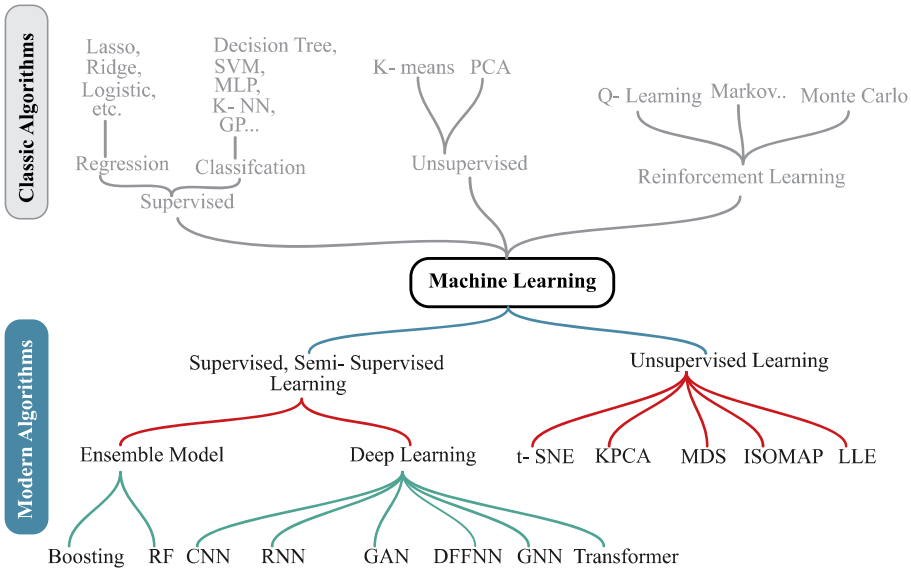


Figure 3.10: Different machine learning models including classic and modern models.

significant progress. Modern Supervised Learning (SL) [167], Semi-Supervised Learning (SSL) [168], RL [169], and DL [170] models embody the ensemble methods such as boosting [171], random forest [169], along with deep learning techniques such including feed-forward neural networks (DFFNN) [172], CNN [173], RNN [174], long short-term memory (LSTM) [175], GAN [176] and others. In the following sections, modern models such as GAN and LSTM are described as data-processing tools.

3.3.5 Convolutional Neural Networks

CNNs, depicted in Fig. 11, stand out as a specialized category of NN tailored for the interpretation of visual information such as images and videos, as shown in Fig. 3.11.. Their excellence in recognizing visual components such as edges, textures, and shapes sets them apart. The core mechanism of CNNs lies in the application of a convolutional operation, where a series of filters are successively employed on the input visual data to craft a feature map. By iterating this procedure through various layers of filters, the network constructs a progressively abstract representation of the input image.

In addition to convolutional layers, CNNs may also include pooling layers to reduce the dimensionality of feature maps and fully connected layers to perform classification (when the activation function of the last layer units is Softmax, Leaky ReLU, or ELU) or regression tasks (typically without any activation function). For a specific regression problem, the last layer can have a single neuron or multiple neurons, depending on the task. Without an activation function, the network can directly output the predicted numerical values without any constraints, allowing the model to learn and predict any real number as the target value. In case an activation function is required, it can be set to Linear. CNNs have been widely used in computer vision applications such as object detection, image classification, and facial recognition, among others. They can be employed to enhance the performance of planar microwave sensors in material characterization applications. One such application is for the characterization of dielectric materials.

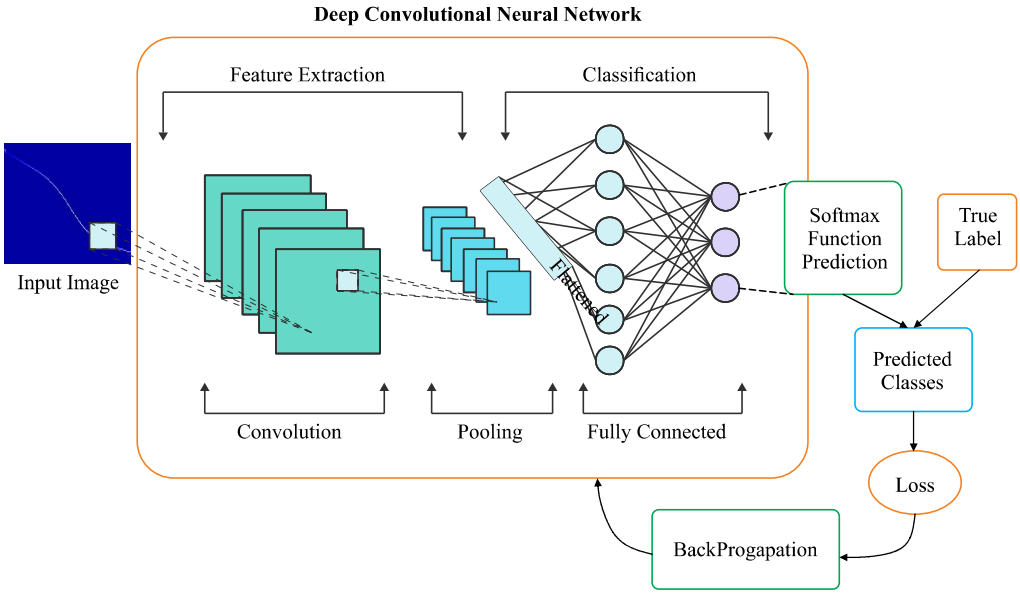


Figure 3.11: Structural Diagram of a Convolutional Neural Network (CNN). This diagram presents the various layers within a CNN, starting from the input layer, through the convolutional and pooling layers, and ending with the fully connected and output layers. The CNN architecture, designed to process grid-like data such as images, extracts features through successive layers to ultimately perform tasks like image classification or object detection.

CNNs can be trained to estimate the dielectric properties of materials based on the sensor data from a planar microwave sensor. The dielectric constant and the loss factor are crucial parameters in various applications such as material quality control and the detection of water content. The input data to the CNN can be either 1D (S_{21} profile) or 2D (heatmap, spectrogram).

CNNs can also be used for anomaly detection in the MUT. For instance, in non-destructive testing (NDT), a planar microwave sensor might be used to scan a structure. The sensor data can then be analyzed using a CNN to detect any defects or anomalies, such as those that might occur in oil and gas pipelines.

Another application of CNNs is in material classification. Similar to dielectric characterization, CNNs can be used to classify different types of materials. To classify the material using a CNN, one needs more input data than with a standard NN or DNN to tackle the overfitting issue.

Finally, CNNs can be used in the sensing of material changes. The planar microwave sensor can be used to monitor changes in material parameters under various conditions such as temperature, humidity, and mechanical stress. A trained CNN can classify and predict these changes. Some examples of CNNs in the microwave regime are included in Table. 3.7. This table illustrates various applications, focusing on different sensor types and MUT. Seven distinct studies are outlined, employing CNNs for tasks such as regression, classification, and clustering, across sensor types such as CCSRR, antenna arrays, CPW TL, CRLH, and FMCW radar. The MUTs vary from substances such as Water/Methanol/Ethanol to porous foam, tissues, and different chemical substances, represented in diverse input forms. The corresponding CNN tasks are associated with performance metrics including MSE values, SNR, and accuracies ranging from 90.7% to 93.3%. This overview shows the versatility and effectiveness of leveraging CNN-based approaches for diverse applications in MW sensor technology.

Table 3.7: Some examples of CNN-based applications in microwave sensors.

Ref.	Sensor Type	f_{res} GHz	MUT	Input	CNN Task	Metric
[82]	CCSRR	2.45	Water/Methanol/Ethanol	1D S_{21}	Regression	MSE ≈0.7%
[177]	Antenna array	3	Moisture in Porous Foam/Tomography	Image	Regression	SNR_{dB} = 28.7
[178]	CPW TL	1-6	ε of Tissue	2 Vectors: 1×1011 ε_r, tanδ	Regression	MSE ≈ 15%
[28]	CRLH	3-5.5	/Methanol/Ethanol/Acetone	Heatmap Image	Classification	Acc = 90.7
[179]	FMCW Radar	126-182	Acrylic Sheet/ Silicon Wafer/PVC	S_{11}	Clustering	MSE ≈ 0.18
[180]	Antenna	1-8	Tumour detection in Breast	2D Sinogram of S_{11}	Classification	ROC AUC = 90%
[29]	CCSRR	4.7	Glucose in ISF	Heatmap Image	Classification	Acc = 93.3%

3.3.6 AutoEncoder

Autoencoders and variational autoencoders (VAEs) are models underpinning generative AI systems. In an autoencoder (Fig. 3.12), the encoder learns how to reduce input down to its latent features using a convolution similar to CNNs. The input is often an image, but it could also be a 1D data such as S_{21} profile. Patterns in the input space are distilled down using convolution to obtain the latent features or latent vectors of that image/profile. The latent factors (Z) of the input data are shown in the gray boxes. On the other side of the network, the decoder acts in an opposite way.

VAE is a type of neural network that is used for unsupervised learning of hidden representations in high-dimensional data [181]. It is a type of generative model that can learn the probability distribution of a data set and generate new samples from that distribution. Unlike traditional AEs, VAEs learn a probabilistic latent space where each point represents a different possible input to the generator. This latent space allows for smoother interpolation between samples and the ability to generate new data points that do not necessarily correspond to any of the original data points. VAEs are commonly used in image generation, speech synthesis, and natural language processing. They are particularly effective when interpolation is important. If the goal is to interpolate between different generated instances in a smooth and controlled manner, VAEs are a good choice. This is because they enforce a structured and continuous latent space, allowing for more controlled transitions between data points.

VAEs are also beneficial when the goal is to model the overall data distribution. If the aim is to capture the full distribution of the data rather than focusing on generating a few high-quality samples, VAEs are typically a better fit compared to tranditional AEs. In addition, VAEs have been successfully used for dimensionality reduction and feature extraction tasks. They provide a principled framework for learning compact representations, making them a suitable choice for these tasks. Finally, for conditional generation, both VAEs and GANs can be extended to conditional models. However, VAEs offer a more straightforward and stable way to include conditioning information, making them a preferred choice in such scenarios.

Variational Auto Encoder

VAE is a type of neural network that is used for unsupervised learning of hidden representations in high-dimensional data [181]. It is a type of generative model that can learn the probability distribution of a dataset and generate new samples from that distribution. Unlike traditional autoencoders, VAEs

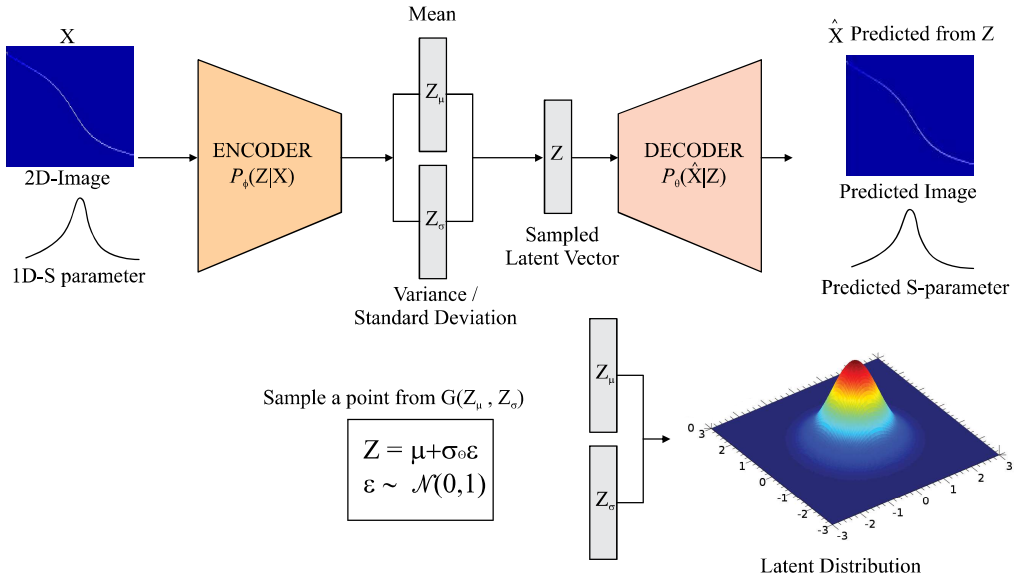


Figure 3.12: This diagram showcases the VAE's two main components: the encoder, which compresses input data into a latent representation, and the decoder, which reconstructs data from this latent representation. The central aspect of the VAE's operation is the probabilistic encoding and sampling process, enabling the model to generate new data similar to the inputs.

learn a probabilistic latent space where each point represents a different possible input to the generator. This latent space allows for smoother interpolation between samples and the ability to generate new data points that do not necessarily correspond to any of the original data points. VAEs are commonly used in image generation, speech synthesis, and natural language processing. VAEs are best used when:

Interpolation is Important: If you want to interpolate between different generated instances in a smooth and controlled manner, VAEs are a better choice. This is because they enforce a structured and continuous latent space, allowing for more controlled transitions between data points.

Modeling Overall Data Distribution: If the goal is to get a model that captures the full distribution of the data rather than focusing on generating a few high-quality samples, VAEs are typically a better fit.

Dimensionality Reduction: VAEs have been successfully used for dimensionality reduction and feature extraction tasks, as they provide a principled framework for learning compact representations.

Conditional Generation: Although both VAEs and GANs can be extended to conditional models, VAEs offer a more straightforward and stable way to include conditioning information.

3.3.7 Generative Adversarial Networks

Generative Adversarial Networks (GANs) represent a unique subset of deep learning frameworks comprised of two interdependent neural networks: the generator and the discriminator [176]. As illustrated in Fig. 3.13, the generator strives to create data that mimics the original training data, while the discriminator's role is to differentiate between these synthesized instances and genuine data. This interplay takes place in a competitive environment similar to a game. The generator aims to deceive the discriminator by generating increasingly authentic-looking data, while the discriminator enhances its proficiency

in recognizing counterfeit data. Through this iterative and adversarial training process, both networks continually refine their skills, resulting in the generator's ability to fabricate highly realistic data and the discriminator's heightened accuracy in spotting false data. GANs have been successfully applied to a variety of tasks, such as image generation, video synthesis, and text generation. They are also used in combination with other deep learning models to improve their performance, such as in semi-supervised learning and domain adaptation. GANs are typically chosen under certain conditions. If the priority is generating very high-quality individual samples, GANs typically outperform VAEs. They are known for their ability to generate sharp and realistic images, making them a preferred choice when the quality of generated samples is important.

GANs have also been used as powerful tools for unsupervised learning tasks. In some cases, such as with conditional GANs, labeled data is needed to add extra information as input to the generator. They can learn complex data distributions without explicit annotations, making them a suitable choice for unsupervised learning. In addition, GANs can be used to augment existing data sets by generating new, synthetic data samples. These generated samples can be used to improve the performance of other models, making GANs a valuable tool for data augmentation. Finally, while VAEs are often used for feature extraction, GANs can also serve for this purpose. In fact, features learned by the discriminator of a GAN can sometimes provide useful information about input data, making them a viable option for feature extraction tasks.

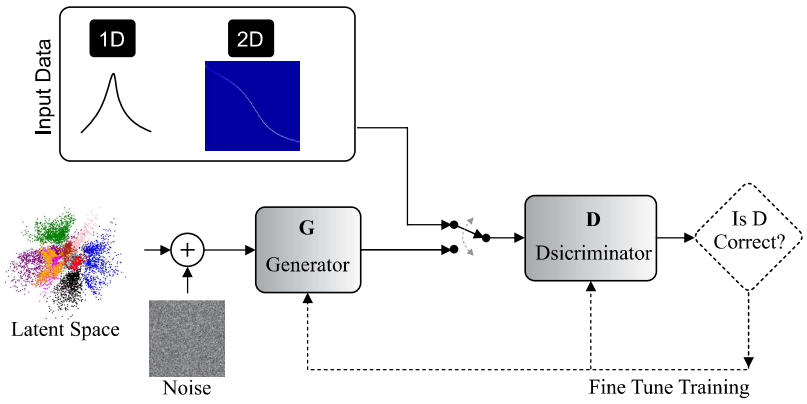


Figure 3.13: Architectural Diagram of a GAN. The diagram displays the two main components of the GAN: the Generator network, tasked with creating data that resemble the true data distribution, and the Discriminator network, which is responsible for distinguishing between real and generated data. The Generator and Discriminator networks engage in a continuous adversarial game, striving to outdo each other, ultimately leading to the generation of high-quality synthetic data.

In the realm of GANs, there are numerous variants of the basic model [182], each designed to address specific challenges or improve performance in unique ways. For instance, Deep Convolutional GANs (DCGANs) [183] use convolutional layers in the generator and discriminator, making them more suitable for image data. On the other hand, Conditional GANs (cGANs) [184] condition the generation process on external information, providing better control over data generation. Moreover, Wasserstein GANs (WGANs) [185] introduce a new method to measure the distance between real and generated distributions, improving training stability and sample quality.

Some GAN variants are designed for specific tasks, such as Cycle-Consistent Adversarial Networks

(CycleGANs) [186], which are used for tasks such as image-to-image translation. Similarly, the Progressive Growing of GANs (ProGANs) [187] gradually increases the complexity of the generated images by starting from a low-resolution image and progressively adding layers to increase the resolution. Notably, Style-Based GANs (StyleGANs) [188], introduced by Nvidia, provide control over the style of the generated images at different levels of detail.

Other variants of GANs include the Self-Attention GANs (SAGANs) [189], which incorporate a self-attention mechanism into the GAN architecture to focus on more global, long-range dependencies in the image data. BigGANs are larger-scale GANs that generate high-quality images, demonstrating the power of scaling up GANs in terms of model size and training data volume. Finally, Bidirectional GANs (BiGANs) [190] and Adversarial Autoencoders (AAEs) [191] learn to encode data into latent representations as well as generate data from these representations, effectively learning an inverse mapping to the generator.

CycleGAN

CycleGAN is a type of GAN that is used for unsupervised image-to-image translation [186]. Unlike traditional GANs, CycleGAN does not require paired input-output training data. Instead, it learns to translate images from one domain to another by minimizing a cycle consistency loss function that ensures that the translated image can be translated back to the original domain without loss of information. This results in a model that can generate realistic-looking images from an input domain, even when there is no direct correspondence between the input and output domains. CycleGAN has been used in a variety of applications, such as style transfer, object transfiguration, and image colorization.

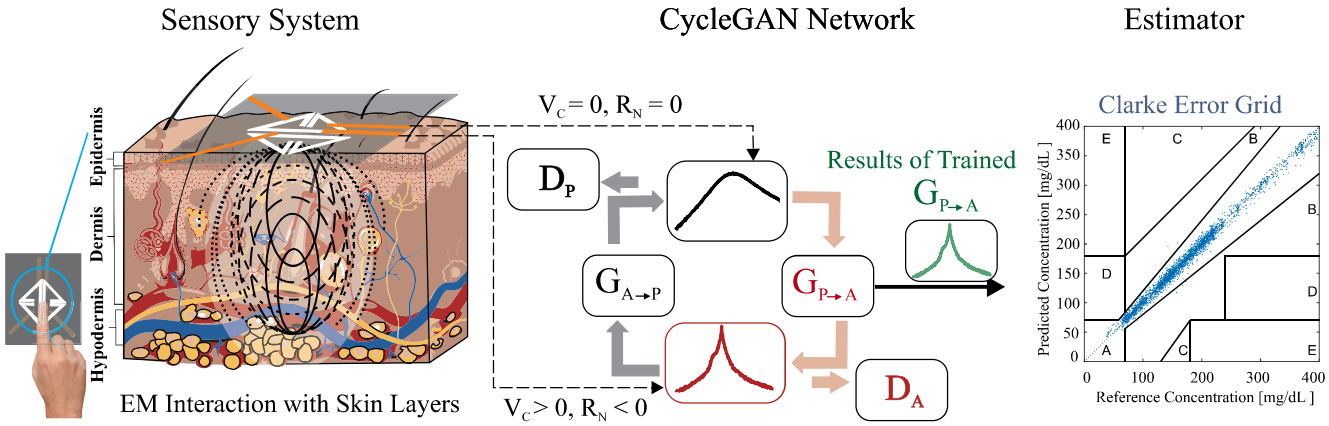


Figure 3.14: Transformation of Input Image to Target Domain using CycleGAN: Demonstrating the Power of Unsupervised Image-to-Image Translation. Microwave sensing scheme combined with CycleGAN network towards improved Clarke error grid.

There is an example of using CycleGAN in microwave sensor enhancement [1]. This article presents an application of CycleGAN to improve the performance of microwave planar sensors. The authors employ CycleGAN to convert passive sensor responses into their active counterparts, thereby enhancing the quality factor of sensor response [192–194]. This technique is also applied to a noninvasive glucose sensing scenario, where the passive sensor response is amplified using CycleGAN. This results in a high-

performance sensor that operates in passive mode, without power. The implementation of this approach involves a generator based on the U-net architecture and a discriminator using PatchGAN. The training processes have been optimized for computational efficiency and adaptive style normalization. The article reports on a human trial of the CycleGAN-boosted glucose sensor. The participant, who did not have diabetes, underwent continuous monitoring of their glucose levels. Sensing was non-invasive, implemented by placing the subject's index finger on the sensitive region of the sensor. The sensor response in both passive and active regimes was analyzed to extract the resonance frequency. The accuracy of the proposed sensor in detecting glucose levels was evaluated using Clarke error grid analysis. For the passive sensor response, 96.2% of paired sensor reference values fell into region A, which holds 20% of the reference values. The resolution of the CycleGAN-boosted response was close to that of the active sensor (≈20 mg/dL), showing a considerable improvement compared to the resolution of the passive sensor (≈70 mg/dL)

In another example [195], a novel application of Cycle-GAN is used to calibrate microwave imaging systems, eliminating the need to measure a known target. The authors demonstrate the use of Cycle-GAN to calibrate data from a 2-D scalar Microwave Imaging (MWI) system, successfully reconstructing targets. This approach addresses the challenges of calibration in MWI systems, which often stem from the discrepancies between the physical system and the computational model, referred to as modelling error. Traditional calibration methods often require the use of a known target or an empty chamber for measurements, which is not always feasible, particularly in large-scale or remote applications, such as grain bin imaging. The Cycle-GAN approach circumvents this limitation by using two sets of data: experimental and synthetic, with the experimental data not requiring labeling. The authors demonstrate that with an experimental calibration set of roughly 150 targets, the Cycle-GAN approach provides results comparable to known-target calibration for the class of targets used in the 2D near-field MWI systems. Although the Cycle-GAN calibration results were slightly inferior to those of traditional calibration, the authors suggested that the results could potentially be improved by collecting a larger set of unlabeled experimental data. The paper concludes by advocating for further investigation of the Cycle-GAN approach for other electromagnetic imaging systems where calibration targets cannot be used.

SRGAN

SRGAN, standing for Super-Resolution Generative Adversarial Network, is a specialized deep learning architecture developed for the purpose of image super-resolution, a technique aimed at enhancing the resolution of an image [196]. First unveiled in the 2017 paper "Photo-Realistic Single Image Super-Resolution Using a GAN" by Christian Ledig and colleagues, this model employs the GAN structure to master the transformation of low-resolution imagery into their high-resolution counterparts. Trained with sets of corresponding low and high-resolution images, SRGAN is engineered to create high-fidelity, lifelike images, emphasizing intricate details and textures. Capable of magnifying images by a multiple of 4, and 8, its performance has been shown to rival or even surpass traditional interpolation techniques. Its utility extends across various sectors, encompassing areas like medical imaging, satellite image analysis, and video processing technologies.

There is an example of using SRGAN in microwave sensor enhancement [29]. The paper uses SRGAN

to improve the resolution of a microwave planar sensor response using SRGAN. The sensor, which is highly sensitive, is used to measure minute variations of glucose in interstitial fluid. The proposed approach involves converting sensor responses into equivalent 2D images and using SRGAN to enhance resolution. This process significantly improves the classification accuracy from 62.1% to 93.3%. The paper also discusses the challenges associated with the response resolution when the deep-zero region of the sensor is exploited. In practical applications, the VNA is typically replaced by integrated readout circuitry that has a limited number of measurement points. To address the low-resolution issue caused by the limited number of points, the paper proposes using SRGAN to generate high-resolution responses with a considerably higher number of points from the measured data. This method is based on processing a 2D plot of phase response vs. frequency, with the magnitude of transmission represented by color. This paper demonstrates how a low-cost, low-profile measuring system can be deployed for commercial sensing while retaining the ability to achieve a high-resolution response of the sensor.

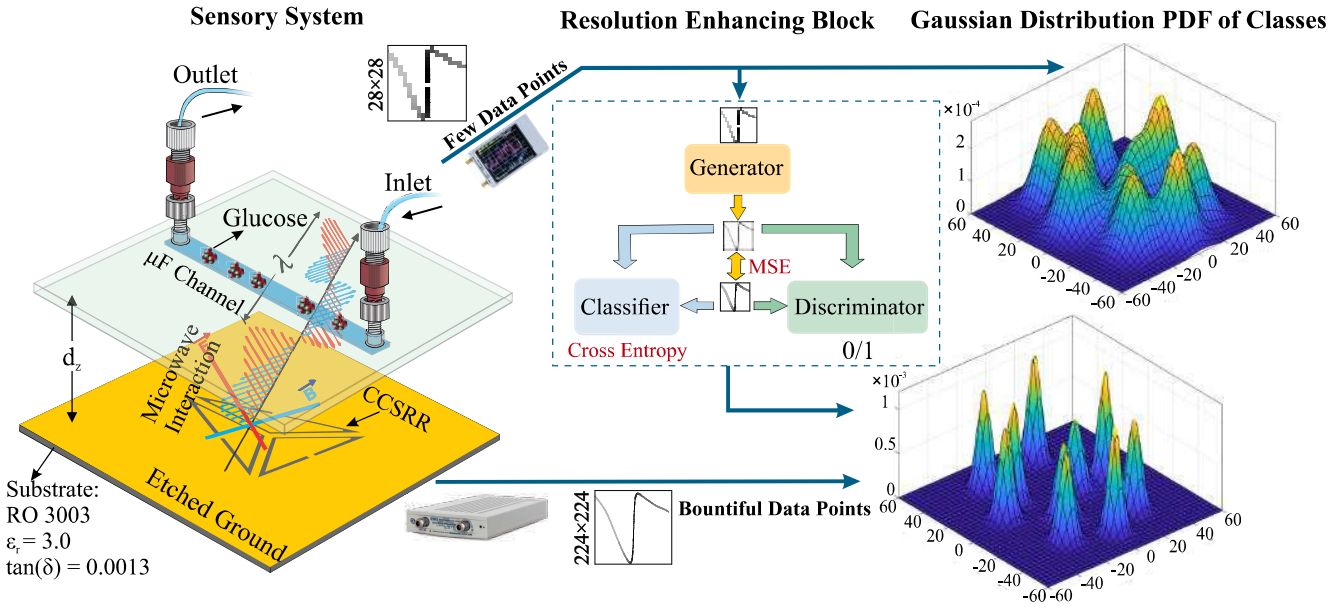


Figure 3.15: Conceptual Framework of a Super-Resolution Generative Adversarial Network (SRGAN). The diagram outlines the SRGAN's two core components: the Generator, which upscales low-resolution images to high-resolution, and the Discriminator, tasked with differentiating between the super-resolved images and original high-resolution images. The adversarial process between these networks ultimately leads to the generation of high-quality super-resolved images.

In summary, Generative models such as VAEs and GANs share the common goal of data generation but exhibit considerable differences in methodology, sample quality, training stability, and latent space properties. Employing variational inference and Bayesian statistics, VAEs offer an easier and more stable training process, and their objective function results in a continuous and structured latent space. However, the images generated might lack sharpness. Contrastingly, GANs create a competitive scenario between a generator and a discriminator, producing samples of higher quality. Training GANs can present difficulties due to issues such as mode collapse and unstable dynamics, and their latent spaces are not necessarily designed to be structured or continuous.

3.3.8 Recurrent Neural Networks

RNNs are a subclass of NNs crafted specifically to analyze data sequences, which might include applications such as time series analysis, recognizing speech patterns, and processing natural languages. Distinct from FFANNs, RNNs incorporate a unique feedback mechanism, facilitating the retention and transfer of information across sequential steps. This structure empowers the network to examine the current input in conjunction with prior inputs, establishing an understanding of their interconnections and influence on the current input. By considering past and present inputs, RNNs can discern intricate relationships within the data. They become an invaluable asset in scenarios where understanding context plays a critical role, for instance, in language construction, voice recognition, and translating machines. However, traditional RNNs encounter a challenge known as the vanishing gradient problem, restricting their ability to grasp long-term dependencies within data. To remedy this, variations such as LSTM and Gated Recurrent Unit (GRU) networks have been developed.

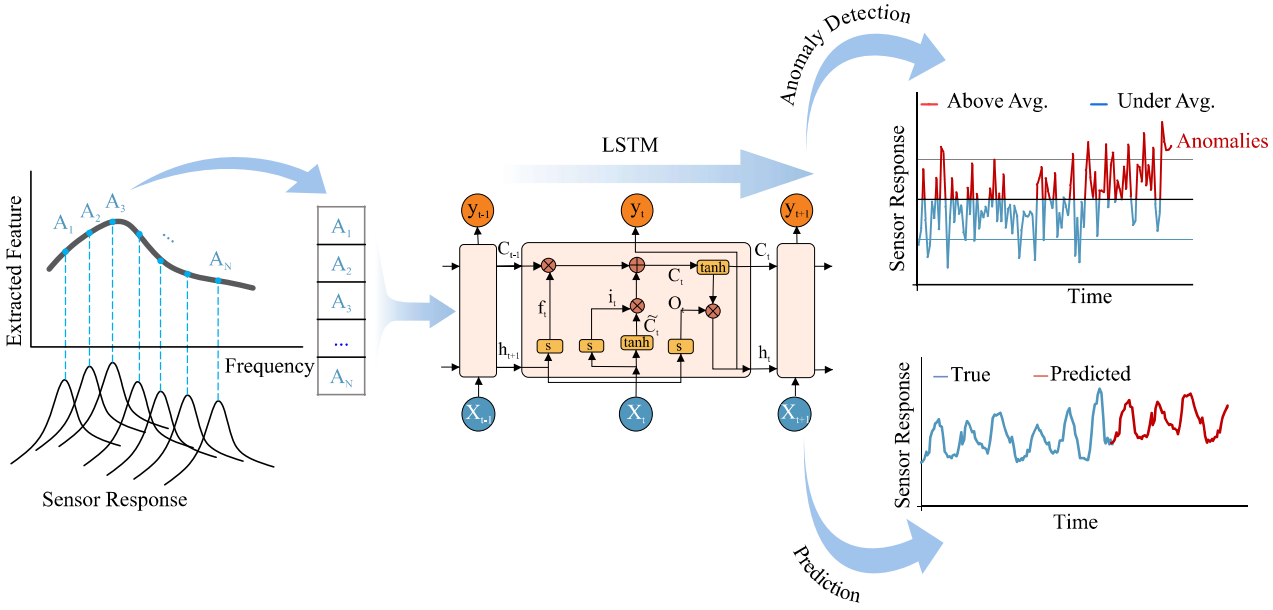


Figure 3.16: Architectural Diagram of a LSTM Cell. These mechanisms work together to regulate the cell's ability to remember or forget information over long sequences, making LSTM a powerful model for sequence-based tasks such as time series prediction or anomaly detection.

Specifically, Long Short-Term Memory (LSTM) networks were engineered to tackle the vanishing gradient problem that conventional RNNs faced, especially when training over elongated sequences. LSTMs possess the distinctive ability to regulate memory within the network's internal cell state selectively. Through a series of specialized gates, such as the forget and input gates, information flow is controlled, allowing the network to decide what to retain or erase. This singular feature has propelled LSTMs into widespread use across diverse fields involving sequential data interpretation, from understanding human language and recognizing speech to predicting trends in time series data.

Building on the capabilities of LSTM, a recent study Balasubramanian2023 demonstrates its application in the field of microwave sensing. The paper presents a real-time, non-destructive inspection system

for detecting erosive wear on coatings. The core of this system is the integration of LSTM with microwave resonators. The AI-enabled microwave resonators are used to detect erosion and compensate for external noises, making the system highly sensitive and accurate.

The ML part of the system is vital in distinguishing between layers in multi-layer coatings and estimating wear depth and rate. By employing ML algorithms, the system can adapt to various conditions and provide precise measurements, making it a practical solution for real-world coating erosion applications.

This approach showcases the potential of ML, specifically LSTM, in enhancing microwave sensor technology, particularly in the field of erosive wear detection. It represents a significant advancement in the application of AI in sensor technology, offering both safety and financial benefits in the aerospace industry.

In another example, a recent study [2] demonstrates its application in the field of microwave sensing, specifically for non-invasive continuous glucose monitoring. The authors use an LSTM neural network to analyze time-series data from a microwave glucose sensor. The LSTM model was trained to predict future glucose levels, providing an early indication of upcoming behavior. This predictive capability is instrumental in triggering preemptive measures in cases of unusual glucose level trends, showcasing the potential of LSTM in healthcare applications.

The LSTM model has also been employed for anomaly detection, an aspect critical to ensure the reliability of sensor data. Prediction errors, which are the differences between the predicted and actual sensor readings, were used as anomaly indicators. These errors were modeled using a Gaussian function, and anomalies were identified based on the likelihood of the observation being an anomaly, rather than accepting it as a truthful reading. This approach allowed the system to recognize and remove common artifacts from sensor data, such as excessive pressure on top of the sensor, sudden changes in temperature, data inconsistency, and hand movement [2]. However, the authors noted a trade-off between the accuracy of long-term prediction and that of anomaly detection. Therefore, the model parameters optimized for prediction may need further manual tuning for effective anomaly detection. The authors also suggested that collecting a larger set of data could potentially improve the performance of the LSTM model.

In conclusion, the integration of LSTM into the operation of the microwave glucose sensor significantly improved its accuracy, reliability, and predictive capabilities. This example underscores the versatility of LSTM and its potential in enhancing the performance of microwave sensors, making it a promising tool for various applications involving sequential data.

Transformers: are a deep learning architecture that revolutionized natural language processing tasks by effectively capturing long-range dependencies [197]. They were introduced in 2017 through the paper "Attention is All You Need". they employ self-attention mechanisms to compute contextual representations of words in a sequence, allowing for parallel processing and eliminating the need for recurrent connections. This means that Transformers can process the end of a sentence before the beginning, for example. With its attention-based approach, the Transformer has become the backbone of many state-of-the-art models in machine translation, text summarization, and other language-related tasks. Popular Transformer models include Bidirectional Encoder Representations from Transformers (BERT) and various Generative Pre-trained Transformer (GPT) models developed by OpenAI.

3.3.9 Deep Reinforcement learning

RL falls within the broader category of ML and involves an agent's ability to navigate within an environment to enhance a particular reward signal. The learning curve for an agent in RL is essentially dictated by a 'trial and error' method. The agent's performance in tasks is gauged through a reward signal, steering the learning towards actions that incrementally result in better rewards [198]. Numerous successful implementations of RL exist, such as in gaming scenarios [199], robotics engineering [200], the field of autonomous vehicles [201], and managing energy systems [202]. One distinct edge RL holds over other methodologies is its proficiency in adapting to complicated, ever-changing scenarios without the prerequisite of direct oversight or guidance.

There is an example of using RL in microwave sensor design [130]. The design process is complex and involves many variables, including the physical characteristics of the sensor, the environment in which it operates, and the specific tasks it needs to perform. Traditional design methods often struggle to find the optimal balance between these variables. To address this, the paper presents a novel approach to the automatic optimal design of resonant structures in high-sensitivity microwave microfluidic sensors using a deep deterministic policy gradient algorithm based on joint simulation. The agent's action strategy is decomposed into multiple movement actions against the pixelated structure and finally outputs the optimized configuration. This strategy dramatically improves the sensitivity of the designed sensor.

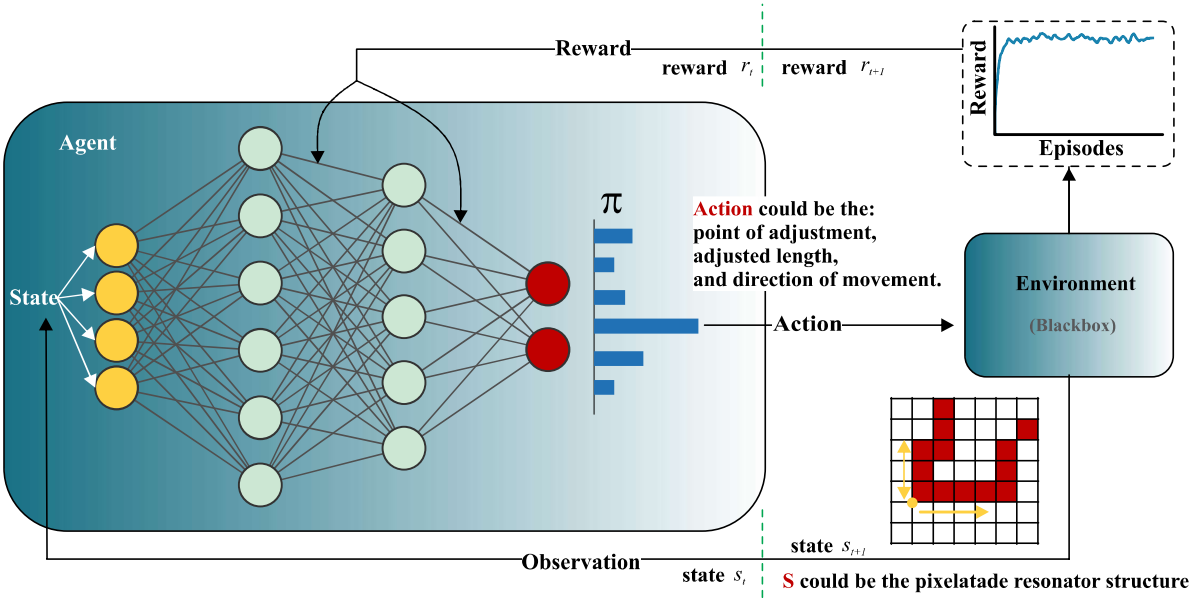


Figure 3.17: Illustration of the Reinforcement Learning Process. The diagram depicts the interaction between an agent and its environment. The agent takes actions based on its current state, influencing the environment, and subsequently receives rewards or penalties. This feedback guides the agent in learning the optimal policy over time to maximize the cumulative reward.

RL is introduced as a solution to this problem. In the context of microwave sensor design, the RL agent iteratively adjusts the sensor's design parameters to maximize its performance. The agent receives feedback (rewards or penalties) based on the sensor's performance, which guides its future actions. The document further explains that RL can handle the high-dimensionality and nonlinearity of the microwave sensor design problem. It can explore a large design space and find solutions that traditional methods

might miss. The RL agent can learn from its past experiences and improve its decisions over time, leading to a more efficient and effective sensor design. Two optimized prototypes have been prepared and tested, demonstrating the effectiveness of the approach. This study represents a meaningful attempt to develop an automatic design procedure for planar microwave sensors. The document concludes that RL has the potential to revolutionize the design of microwave sensors. It can automate the design process, reduce the time and effort required, and improve the performance of the sensor. However, the document also notes that more research is needed to fully realize this potential.

3.4 Potential Future Directions

Exploring the possible future paths for machine learning (ML) in the realm of microwave sensing leads us to the brink of various innovative technologies and methods that might redefine this domain. In the ever-changing world of ML, we encounter novel frameworks such as Capsule Networks, Neuroevolution, NeuroSymbolic AI, and Quantum Computing. These pioneering methods, each possessing its distinctive advantages and functionalities, could dramatically amplify the efficacy and potential of microwave sensors. Moreover, the dawn of the Internet of Things (IoT), accompanied by an unprecedented increase in sensor usage, will result in an overwhelming quantity of data. In this context, quantum optimization emerges as a pivotal technology. Its potential for accelerated processing could be instrumental in handling and interpreting this immense data pool, allowing for immediate analysis and more precise findings. This perspective, built on the analysis of existing and emergent trends, illustrates a promising direction for the integration of advanced technologies within the field of microwave sensing.

3.4.1 Capsule Networks

Capsule Networks represent an intriguing development in the field of deep learning, introduced by Geoffrey Hinton to tackle the shortcomings of conventional neural networks [203]. These limitations often arise in the context of recognizing spatial hierarchies and dealing with varying scales and viewpoints.

Unlike traditional neural networks, which are based on individual neurons as basic computational units, Capsule Networks adopt a more sophisticated approach. They introduce a structure called a "capsule" that groups neurons together to represent specific entities or features. This approach helps to understand the complexities of spatial and hierarchical relationships within the data.

Capsules encode information on the presence, pose, and instantiation parameters of specific entities or objects. Unlike single neurons, a capsule can encapsulate multidimensional information, offering a richer representation. One of the distinctive features of capsules is their ability to establish a hierarchy in the data. By understanding the part-to-whole relationships, they are more proficient at modeling complex structures, which is particularly important in fields such as computer vision.

Traditional neural networks often employ pooling layers to reduce dimensionality, potentially losing valuable information. Capsules, on the other hand, retain this information, enhancing the ability of the network to recognize spatial configurations and subtler details. Capsule Networks excel at handling variations in scale, orientation, and viewpoint. While traditional neural networks may struggle with these

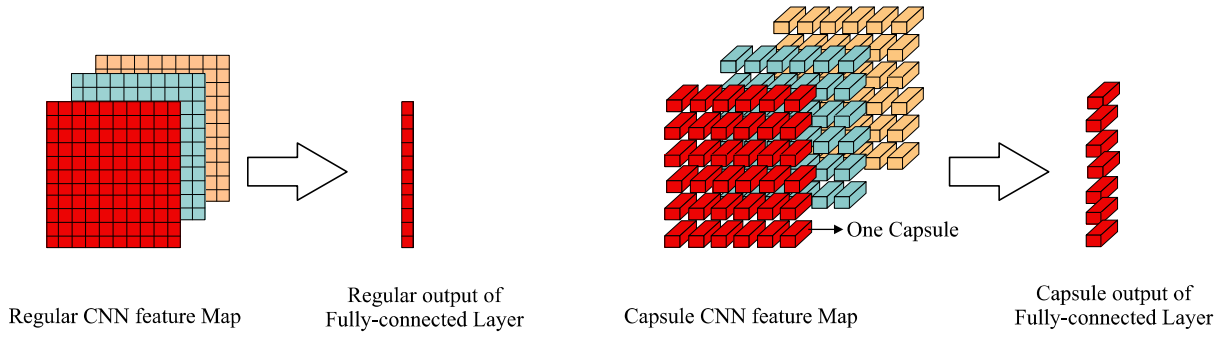


Figure 3.18: Comparative Analysis of Capsule Networks and Convolutional Neural Networks: This diagram illustrates the difference in feature map generation and output between Capsule Networks (CapsNet) and CNN. While CNNs focus on the presence of features with less regard for spatial hierarchies, CapsNet captures the spatial relationships between features, providing a richer, more holistic understanding of the input data.

changes, capsules can be more adaptive and flexible, providing a more robust representation.

A significant innovation in Capsule Networks is the dynamic routing algorithm. It ensures that the lower-level capsules agree with the higher-level capsules, thereby creating coherent predictions. This is in contrast to traditional networks, where fixed transformations might not adapt well to varying inputs. The capabilities of Capsule Networks could be particularly useful in fields requiring detailed recognition of spatial relationships, such as medical imaging, robotics, and 3D modeling. Their ability to maintain critical information while adapting to variations makes them a promising tool for these areas.

Capsule Networks have a lot of promise, but they also come with some challenges. They often need a lot of computer power and careful adjustments. More research and work are needed to make them work better and use them in more areas. To sum up, Capsule Networks are a big step forward in deep learning. They fix some important problems found in regular neural networks. By putting neurons into groups called capsules, they offer a more flexible way to deal with complicated spatial relationships and changes. People are still looking into what they can do, and they might become a key part of the future of deep learning.

3.4.2 Neuroevolution

Neuroevolution is a concept that combines evolutionary computation and neural networks. It has been studied since the early 1990s, with a focus on evolving neural networks using evolutionary algorithms [204]. Kenneth O. Stanley's research on the NeuroEvolution of Augmenting Topologies (NEAT) in the late 1990s contributed to the development of this area. It aims to evolve neural networks through evolutionary processes, offering advantages such as reduced chances of being trapped in local minima and potential integration with deep learning, deep RL, and ML, making it a critical tool in the pursuit of Artificial General Intelligence. Below is a general outline of how to apply neuroevolution to a classification problem:

1. **Task Definition**: Outline the specific classification problem, which might encompass anything from categorizing images to more intricate challenges. Assemble the dataset, comprising both the feature inputs and corresponding target labels.

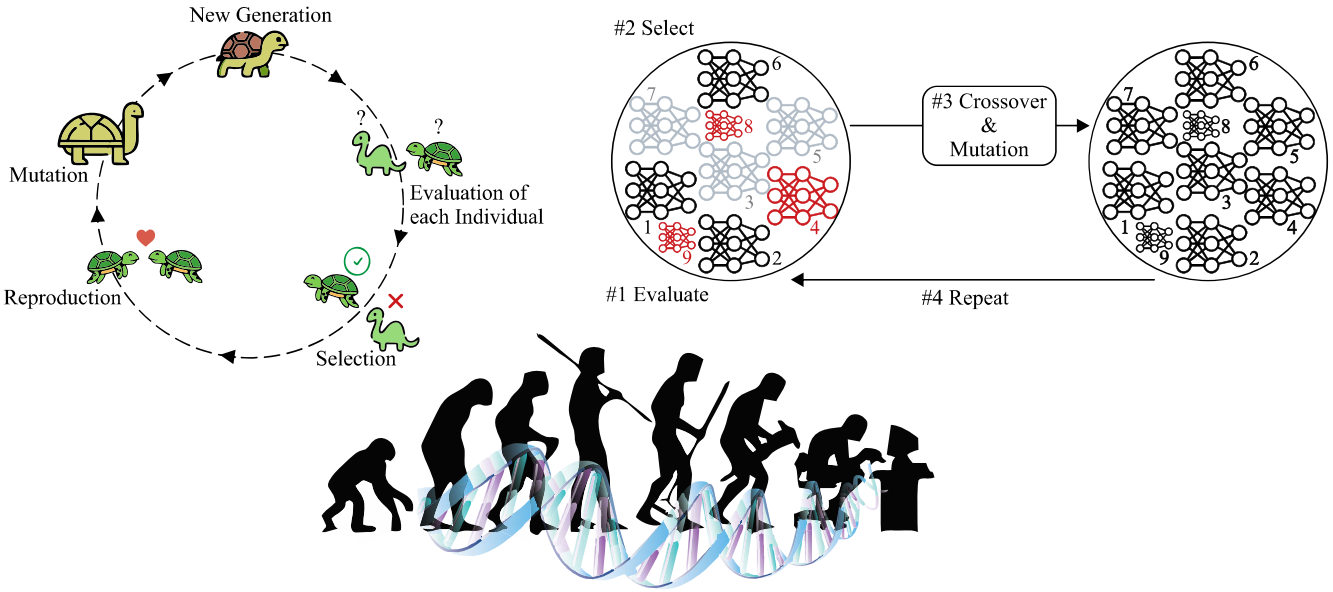


Figure 3.19: Flow Diagram of a Neuroevolution Process. This diagram illustrates the cyclical nature of the neuroevolution process, starting with the initialization of a population of neural networks. Each network is then evaluated based on its performance on a task. Selection, reproduction, and mutation operations are subsequently applied to generate a new generation of networks. This process repeats until a stopping condition is met. The end result is an evolved network that has adapted to perform the task more effectively.

2. **Population Initialization**: Generate an initial group of neural networks, furnishing them with random weights and possibly varied structures. The population's size will be contingent on the distinct problem and available computational means.
3. **Individual Evaluation**: Utilize each neural network within the population to forecast on the training data. Assess each network's efficacy using an appropriate criterion such as *Accuracy* or the cross-entropy loss for classification tasks.
4. **Selection Phase**: Identify the superior-performing networks for reproduction, relying on their results. The selection might be deterministic (such as consistently selecting the top achievers) or probabilistic (affording higher-performing networks an increased probability of selection).
5. **Reproduction Stage**: Forge a subsequent generation of networks through genetic procedures such as crossover (merging weights or structures of two parent networks) and mutation (sporadically modifying weights or structures).
6. **Replacement of Underperformers**: Substitute the least effective networks in the group with the freshly created ones.
7. **Iterative Process**: Persist with the cycle of assessment, selection, and reproduction for a predetermined number of generations or until no further enhancements in performance are noticed.
8. **Model Testing**: Upon completion of the neuroevolution procedure, the top-performing network from the final generation can be employed to forecast on the test data, and its performance can be appraised.

Neuroevolution can be computationally expensive, but it is an effective way to optimize neural networks when gradient-based methods are not applicable or do not perform well. It is also particularly

useful when thre is little prior knowledge about the appropriate network architecture for the task.

3.4.3 NeuroSymbolic AI

NeuroSymbolic AI is an interdisciplinary approach that combines symbolic AI and NNs to overcome their limitations [205]. It aims to integrate the abstract approach and explicit knowledge representation of symbolic reasoning with the learning and pattern recognition capabilities of NNs. By creating hybrid models that reason symbolically, manipulate structured knowledge, and simultaneously learn from data, NeuroSymbolic AI enables interpretable and explainable AI systems with improved generalization and robustness. MIT-IBM Watson AI Lab describes NeuroSymbolic AI as a fusion of neural networks and symbolic representations, allowing powerful AI systems that require less training data, can make inferences and draw conclusions, and transfer knowledge across domains, leading to AI systems that learn more like humans.

3.4.4 Quantum computing

Quantum computing represents an innovative area of study within the field of computational science, drawing upon the fundamental principles of quantum mechanics to carry out specific tasks [206]. In contrast to conventional computers, which utilize binary bits, quantum computing relies on quantum bits (qubits). Uniquely, qubits can exist in numerous states at once, a phenomenon known as superposition, affording them the ability to conduct parallel calculations. This parallelism can lead to a significant increase in computational efficiency, particularly in areas such as large number factorization, optimization problems, and quantum system simulations. As a burgeoning field, quantum computing holds promise for transformational changes across several sectors, including but not limited to cryptography, the development of pharmaceuticals, and the enhancement of optimization algorithms.

Additionally, the implications of quantum computing extend to AI. While quantum AI is still in its early stage, the fusion of quantum mechanics with AI is a focus of current research. Scientists are keen to explore how the distinctive attributes of quantum systems can be harnessed to augment AI methodologies and facilitate the resolution of multifaceted problems with greater efficacy.

One area of interest is quantum ML, where quantum algorithms and quantum computing techniques are used to improve learning and inference processes. Quantum ML algorithms aim to exploit quantum properties to enhance computational power, speed up training, and improve the performance of AI models. Quantum ML has the potential to enhance pattern recognition, clustering, and optimization tasks.

Additionally, quantum computing could play a role in tackling combinatorial optimization problems that are prevalent in AI. These problems involve searching for the best combination or arrangement of variables, which can be time consuming for classical computers. Quantum algorithms, such as quantum annealing or quantum approximate optimization algorithms, offer the potential for more efficient solutions to such optimization problems.

Furthermore, quantum computing may also impact areas such as data analysis and quantum-inspired algorithms. They leverage classical computers to mimic the behavior of quantum systems, enabling faster processing and more efficient data analysis.

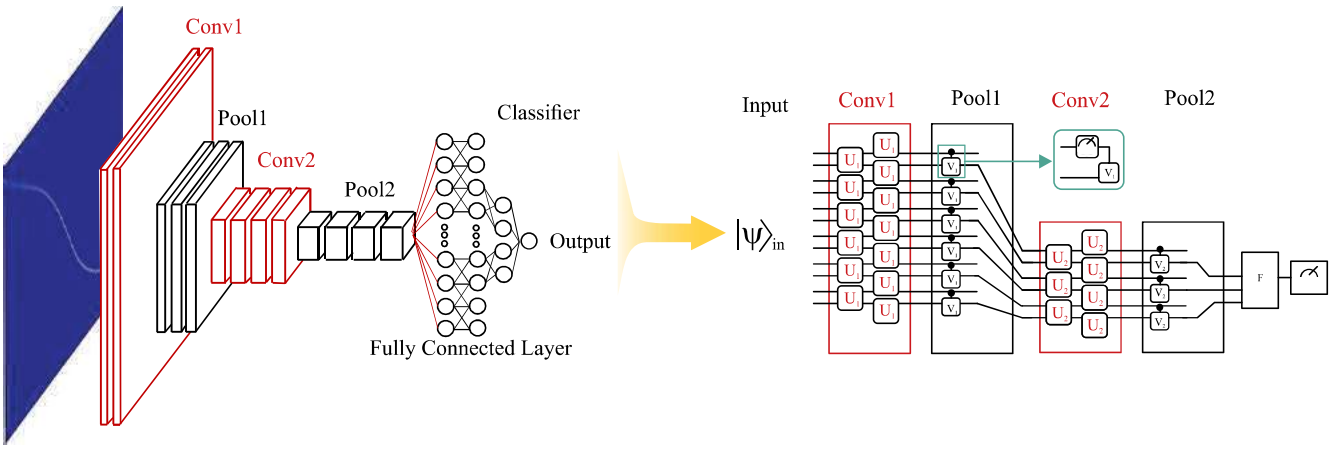


Figure 3.20: Visualizing CNN and QCNN Basics: From Conventional Convolution and Pooling Layers to their Quantum Counterparts, with the Added Complexity of QEC through Measurement-Based Pooling

In summary, while quantum computing is still in its early stages and faces several technical challenges, it holds promise for enhancing various aspects of AI, including ML, optimization, and data analysis, offering the potential for significant advancements in the field.

3.5 Conclusion

In the rapidly evolving field of microwave sensing, the integration of ML and deep learning models with microwave sensors has emerged as a transformative approach, offering innovative solutions and applications. This comprehensive review traces the historical development of artificial intelligence, highlighting key milestones from the birth of AI to the deep learning revolution, and delineating the relationship between artificial intelligence, machine learning, neural networks, and deep learning. This article also explores the intricate connection between these technologies and microwave sensing, focusing on material characterization.

The review specifically delves into the applications of ML in microwave sensing, showcasing how various techniques, from classic methods such as Support Vector Machines to advanced deep learning methods such as CNNs, RNNs, and adversarial autoencoders, are being employed in material classification, evaluation of structures, and detection of defects. These advancements have led to improved accuracy and efficiency in complex tasks such as material characterization, gas detection, blood glucose monitoring, and more. However, challenges persist in areas such as sensitivity, selectivity, and interference, with limitations in current hardware and software that require further research, innovation, and development. Ethical and regulatory considerations must also be considered, possibly with the establishment of a regulatory framework that keeps pace with technological development and ensures that applications are safe, ethical, and accessible.

The continuous improvement of existing technologies, coupled with the emergence of new methodologies and interdisciplinary approaches such as NeuroSymbolic AI, indicates that genuinely noninvasive sensing is on the horizon. This holds promise for more robust, adaptable, and precise systems that can

cater to diverse needs, such as differentiating between critical diabetic patients that require continuous monitoring and those that need only periodic checks. Furthermore, the application of ML in microwave sensors extends to broader contexts, such as tackling complex nonlinear dimensionality reduction and improving the carbon efficiency of ML models. Collaborative efforts between academia, industry, and regulatory bodies will be vital to accelerate the transition from research to real-world applications, and thoughtful consideration of ethical and societal implications will be key to realizing the full potential of this exciting frontier.

The integration of deep learning models with microwave sensors represents a dynamic and rapidly evolving field, where continued interdisciplinary collaboration, investment in research and development, and exploration of personalized and adaptive sensing solutions will be essential. The convergence of these technologies has already led to remarkable innovations in material characterization and other applications, yet the journey is far from complete. The ongoing exploration of these technologies is poised to unlock new capabilities and increase our understanding of complex systems.

Chapter 4

Resolution Enhancement by CycleGAN

4.1 Motivation (Overview)

This chapter introduces a novel system for noninvasive glucose monitoring, leveraging machine learning for high-resolution sensing without active sensors. The system utilizes the CycleGAN technique for mapping between passive and active sensor responses using unpaired datasets. It generates high-resolution profiles in the active domain from low-to-moderate quality factor profiles in the passive domain. The system employs cycle consistency loss to ensure consistency between the mapping $G : P \mapsto A$ and a reverse mapping $F : A \mapsto P$, combined with the adversarial loss for unpaired passive-to-active mapping (see Figure 4.1). The sensor response is represented as an image, enabling the use of image processing techniques for improved accuracy. This approach significantly enhances the performance of passive sensors, paving the way for the development of intelligent sensory systems for applications such as diabetic care.

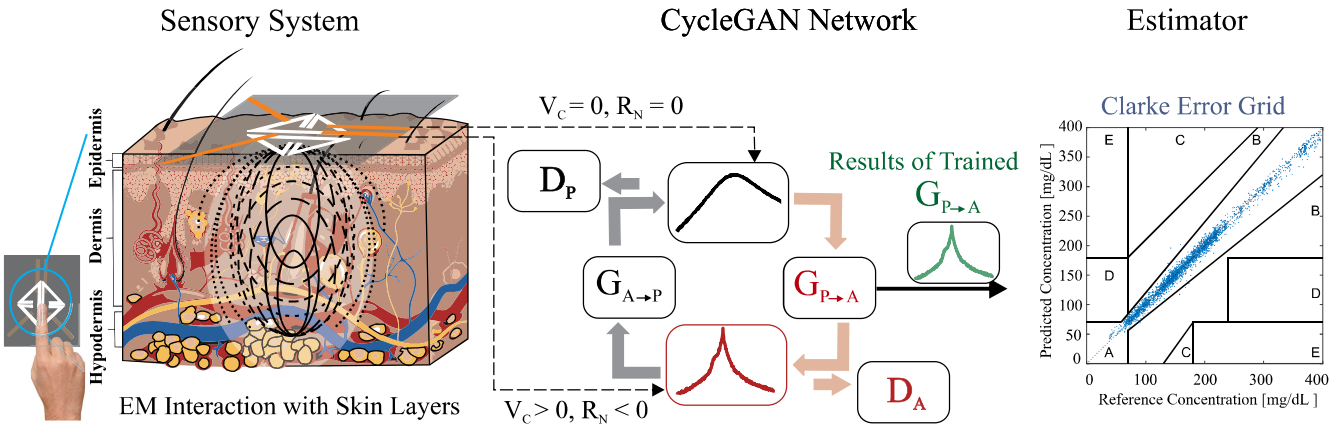


Figure 4.1: Microwave sensing scheme combined with CycleGAN network towards improved confusion matrix.

4.2 Sensor Design

4.2.1 Coupled CSRRs

In this work, a planar microwave sensor is designed using metamaterial-inspired inclusions, namely, complementary split ring resonators (CSRRs). These structures are normally etched out of the solid ground to leave a defected ground. Slots are sized to be 0.5 mm thick, which is an achievable accuracy through in-house etching process using ammonium persulfate. The CSRRs have been designed with triangular shape to achieve a high coupling with another similar resonator through capacitive coupling. A pair of CSRRs are placed in 0.5 mm distance from each other as shown in Fig. A.3(a) connected to the active circuit as given in Fig. A.3 (b). Hereafter, the two resonators lined this way are called coupled CSRRs (CCSRRs). This arrangement is designed to increase the sensitivity of the sensor when the coupled region between two parallel slots is used to interrogate the surrounding material. This feature is shown in Fig. A.3 (c) with two identical parallel resonators (R, C, and L). These resonators are coupled to each other using capacitive slots represented with equivalent J-inverter configuration given in Fig. A.3 (c) with coupling capacitor C_m. This coupling between two resonators results in separation of the original resonances to new values as follows:

$$\omega_1 = 1/\sqrt{L(C + C_m)}, \qquad \omega_2 = 1/\sqrt{L(C - C_m)} \tag{4.1}$$

As a results, the two identically coupled resonators are considered as a new system containing two non-coupled resonators with two different frequencies (ω_1, ω_2), where the modified resonances already contain the coupling information, this allows them to be cascade in the following analysis. The whole design is also coupled to the input/output transmission lines through coupling capacitors C_{TL}.

Another representation of the present network combines two individual back-to-back resonators that are in series. The transfer function representation of each resonator is given as follows:

$$H_1(s) = \frac{K_1 \frac{s\omega_1}{Q_1}}{s^2 + \frac{s\omega_1}{Q_1} + \omega_1^2}, \quad H_2(s) = \frac{K_2 \frac{s\omega_2}{Q_2}}{s^2 + \frac{s\omega_2}{Q_2} + \omega_2^2}, \tag{4.2}$$

where K_1 and K_2 represent maximum individual transmission amplitude, ω_1 and ω_2 are the resonance frequencies ($s = j\omega$), and Q_1 and Q_2 are the loaded quality factor of each resonator. Upon activating the amplifier of the circuit, the main contributing factors for resonator loss, also known as ohmic loss is compensated. The amplifier in the feedback loop injects extra power in the system in a constructive way, thereby increasing the loaded quality factor. This system can be summarized as follows:

$$H_{total} = \frac{Y(s)}{X(s)} = \frac{H_1(s)H_2(s)}{1 - A_v H_1(s)H_2(s)}, \tag{4.3}$$

$$H_1(s)H_2(s) = \frac{K_1 \frac{s\omega_1}{Q_1}}{s^2 + \frac{s\omega_1}{Q_1} + \omega_1^2} \times \frac{K_2 \frac{s\omega_2}{Q_2}}{s^2 + \frac{s\omega_2}{Q_2} + \omega_2^2}. \tag{4.4}$$

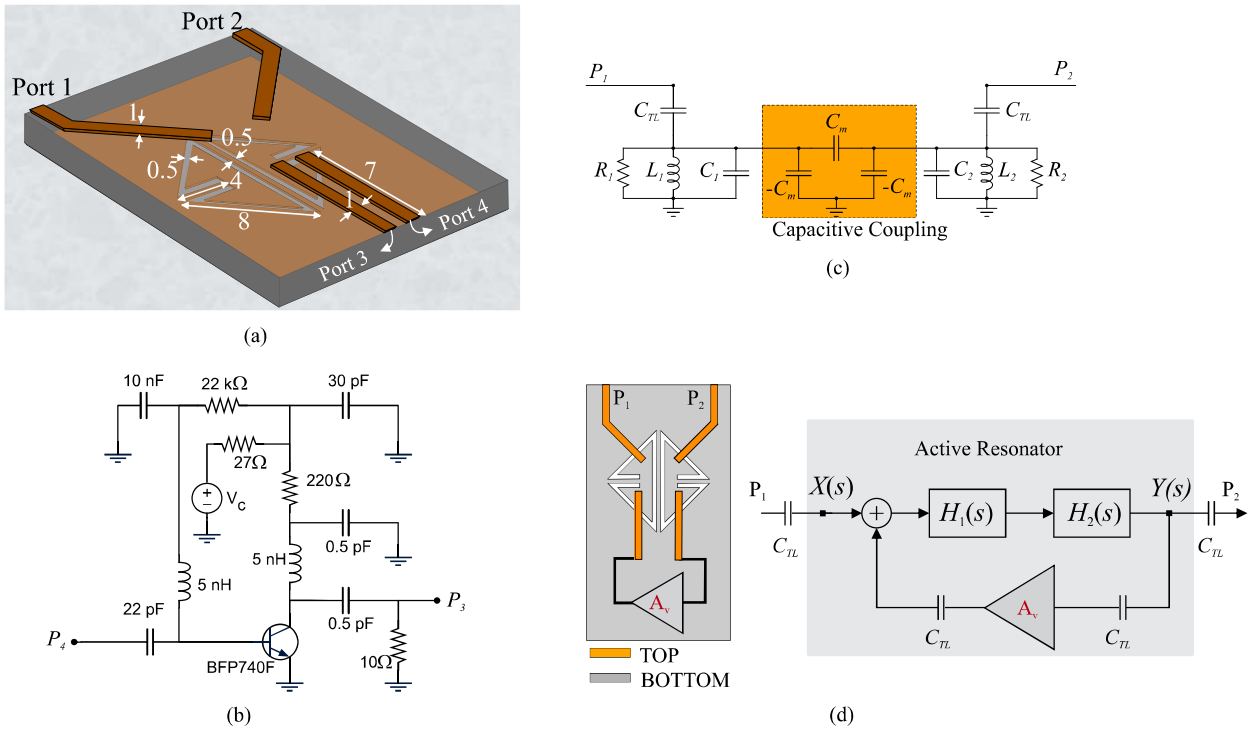


Figure 4.2: (a) Schematic of the active CCSRR with corresponding circuit. All dimensions are in [mm], (b) Amplifier circuit diagram.(c) Equivalent circuit model of the passive. (d) Equivalent block diagram of the proposed sensor

where the two coupling capacitors C_{TL} before and after the amplifier A_v are assumed to be included into the amplifier transfer function as elaborated in 4.3 as these frequency-independent elements act minimally as attenuators preceding and following the amplifier.

It can be shown that if the two original frequencies are close to each other ($\omega_1 \approx \omega_2$):

$$H_1(s)H_2(s) = H_1(s)\underbrace{\frac{\frac{\omega_1}{Q_1}}{\frac{Q_1}{\omega_1} - \frac{Q_2}{\omega_2}}}_{B_1} - H_2(s)\underbrace{\frac{\frac{\omega_2}{Q_2}}{\frac{Q_1}{\omega_1} - \frac{Q_2}{\omega_2}}}_{B_2}, \tag{4.5}$$

then (4.3) can be rearranged as:

$$H_{total} = \frac{B_1H_1(s) - B_2H_2(s)}{1 - A_vB_1H_1(s) - A_vB_2H_2(s)}. \tag{4.6}$$

Assuming that the resonators are identical, then (4.6) can be simplified considering $\omega_1 \approx \omega_2$, $Q_1 \approx Q_2$, and $K_1 \approx K_2$:

$$H_{total} \approx \frac{B_1H_1(s) - B_2H_2(s)}{1 - 2A_vB_1H_1(s)} \approx \frac{B_1H_1(s) - B_2H_2(s)}{1 - 2A_vB_2H_2(s)} = \frac{B_1H_1(s)}{1 - 2A_vB_1H_1(s)} - \frac{B_2H_2(s)}{1 - 2A_vB_1H_1(s)} \tag{4.7}$$

Each term in (4.7) can be simplified as follows:

$$\frac{B_1 H_1(s)}{1 - 2A_v B_1 H_1(s)} = \frac{\frac{B_1 K_1}{1-2A_v B_1 K_1} s \frac{\omega_1}{Q_1/(1-2A_v B_1)K_1}}{s^2 + s\frac{\omega_1}{Q_{1/(1-2A_v B_1 K_1)}} + {\omega_1}^2}. \tag{4.8}$$

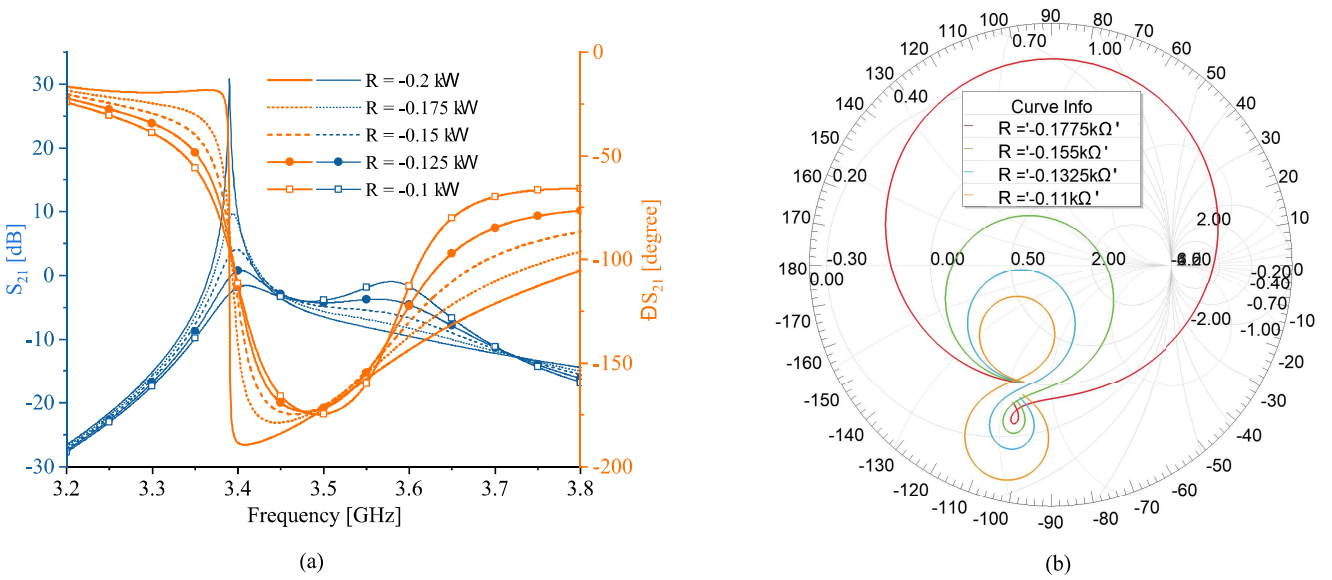
This shows that the original resonators' parameters, including amplitude (K_1 and K_2) and loaded quality factor, become amplified by the factor of $((1 - 2\mathrm{A_v B_1 K_1})^{-1})$ or $((1 - 2\mathrm{A_v B_2 K_2})^{-1})$ as follows:

$$K_1 \to \frac{B_1 K_1}{1 - 2A_v B_1 K_1}, \quad Q_1 \to \frac{Q_1}{1 - 2A_v B_1 K_1} \tag{4.9}$$

This amplification coefficient is controlled by the active circuit, which is triggered using the bias voltage V_C. The principle of operation of the regenerative amplifier relies on two main factors. First, the gain of the system denoted as A_v, must be high enough to compensate the transmission loss of the two series blocks H_1 and H_2. Second, the phase of the signal needs to incur a change of even multiples of 2π so that a positive construction occurs when the present signal in the loop joins the input signal. The signal's phase undergoes a negative sign within the amplifier. For the rest of the phase modification, extra transmission lines are used that connect the core passive resonator to the amplifier. The total length of 35 mm per CSRR, with the dimensions presented in Fig. A.3 (a), results in dual resonances around 3.5 GHz. The frequency of operation affects the depth of penetration into the free space and sensitivity of the sensor differently. A relatively high frequency is selected to gain more sensitivity [6], [7]. In addition, the size of the resonator is inversely linked to the frequency of operation. Finally, the CSRR edges are folded inwards, similar to the detailed development described in [8], to reduce the overall sensor size and contribute to its miniaturization.

4.2.2 Loss Compensation in CCSRRs

Microwave sensors need to have loaded quality factor sufficent to overcome high power loss when exposed to external materials. This inherent challenge has been addressed using various loss-compensating mechanisms, but only for single resonators. However, the proposed system uses a different type of resonator with intrinsic coupling mechanism. Loss-compensation treatment of this CCSRR is implemented using a regenerative amplifier design, as shown in Fig. A.3 (c). The input power from port 1 is coupled to the CCSRR and then flows on to the active circuitry through a transmission line. The input power is amplified and phase modified before returning back to the CCSRR for the second time. Once the returned power circulates in the CCSRR, it is sampled through the output transmission line. The phase of amplifier response is controlled to assure constructive feedback in the resonating system. This process also affects the transmission (S_{21}) and reflection (S_{11}) parameters of the sensor. Although the coupled design results in two adjacent resonances, only one resonance is designed to meet the criteria for correct phase compensation. The entire sensor is fabricated on Rogers RO3003 substrate with dielectric properties of $\varepsilon_r = 3.0$ and $\tan(\delta) = 0.0013$. A 0.8 mm substrate holds both the passive resonator design on one side



(a) (b)

Figure 4.3: (a) Simulation results of the CSRR with loss compensation representation in both amplitude and phase of resonance. (b) Smith chart characterization with various loss compensation states.

and the active circuitry on the other side. The active circuit is composed of commercially available surface mount electronic devices.

Sensor performance with an external loss compensation mechanism is presented in Fig. 4.3 (a), where both amplitude and phase responses are shown. The simulation model includes a shunt negative resistance in series with a variable phase element from both transmission lines (port 3 and port 4) exiting the CCSRRs to the ground. A fixed phase value of 126° with a variable negative resistance results in the sharpened transmission profiles shown in Fig. 4.3 (a). The resonance frequency of the sensor changes slightly; however, the loaded quality factor increases significantly. In an initial form of operation in active mode, where the negative resistance is close to $-100\ \Omega$, the two resonances yield a very low-resolution profile. The more compensation is applied to the coupled resonators, the sharper the resonance becomes. For this particular example, a typical sharp resonance can be realized with a negative resistance of $-200\ \Omega$. The nature of the coupling between the two resonances causes that only one of them is dominant. In this design, the resonance with lower frequency is set to be dominant. Once the sensor is exposed to a change in the environment, the entire resonance profile shifts downwards. In addition, the resonance frequencies of adjacent resonators are separated from each other as a result of the sensor being exposed to an external material. [207,208]. Consequently, the net downshift for the lower resonance is higher than that of the upper resonance. The reason for selecting the lower resonance is the ability to exploit its larger resonance dynamics in sensing.

The process of loaded quality factor enhancement can also be verified by monitoring the phase information in Fig. 4.3 (a). It can be seen that the phase diagram for the lower resonance becomes sharper as the negative resistance increases. The well-determined unloaded quality factor Q_0 for the sensor can be

computed using the following expression [209]:

$$Q_0 = \frac{f_0}{(f_2 - f_1)\sqrt{1 + \left(\frac{G_0(f_2 - f_1)}{4}\right)^2}}, \tag{4.10}$$

where f_0 is the center frequency, characteristic frequencies of f_1 and f_2 are frequencies where phase response has zero slope (excluding frequency-dependent phase contribution from the input/output transmission lines), and the gradient G_0 is the slope of phase at the center frequency. When the negative resistance increases, the quantity $(f_2 - f_1)$ reduces, and that results in an increased loaded quality factor. In this simulated example, the loaded quality factor rises from a nearly passive structure ($R = -100\ \Omega$) of $Q_0 \approx 15$ up to highly compensated version ($R = -200\ \Omega$) of $Q_0 \approx 683$. In cases where de-embedding transmission lines' effect becomes ambiguous an alternative solution is to consider the maximum transmission at resonance frequency f_0 ($\max|S_{21}| = S_{21\ 0}$) and the peak width (Δf) at a given level below $S_{21\ 0}$, that is evaluated at $S_{21\ U}$ as follows [210]:

$$|S_{21\ U}|^2 = \frac{S_{21\ 0}^2}{S_{21\ 0}^2 - 2S_{21\ 0} + 2} \tag{4.11}$$

that result in $Q_0 = \frac{f_0}{\Delta f}$. Another interesting observation of the sensor performance can be made by examining the Smith chart shown in Fig. 4.3 (b). The characteristics are obtained from port 1 of the simulator for various values of negative resistance. It is clear that the Smith chart represents better matching for the curves located close to the central point (1, 0). This is achieved with partial loss compensation for the case $R = -132.5\ \Omega$. However, even though the transmission loss is fully compensated, the total loaded quality factor is still not satisfactory. It needs further improvement, which entails higher negative resistance values resulting in higher values of Q. However, the reflection coefficient is also high, resulting in the Smith chart circle to move away from the central point. This high reflection, in some cases, may result in damage to the measurement devices. Therefore an attenuator is needed to quench the reflections in both pathways to the vector network analyzer. In a commercial implementation, this could be avoided with the inclusion of line couplers before both ports 1 and 2 of Fig. A.3 (d). It should also be noted that the resonance profiles in Fig. 4.3 (b) are shown with a different circulation of the Smith circle that corresponds to either a pole (left resonance) or a zero (right resonance).

The non-contact feature of the sensor enables it to interrogate the adjacent materials. The level of this interaction affects the dynamic range of the material location with respect to the sensor. This range is limited in the passive resonator design, mainly due to the lost power as a path loss when the input power is transmitted from the resonator to the material. In the proposed design, this problem is ultimately addressed using a custom active circuitry. It enables a wider range of operation, especially when the material under test is distant due to poor accessibility. Use of loss-compensation techniques provides higher sensitivity to materials when they are located at a distance with respect to the sensor. Illustration of the electric field magnitude in Fig. 4.4 shows different ranges of propagation on the surface (Fig. 4.4

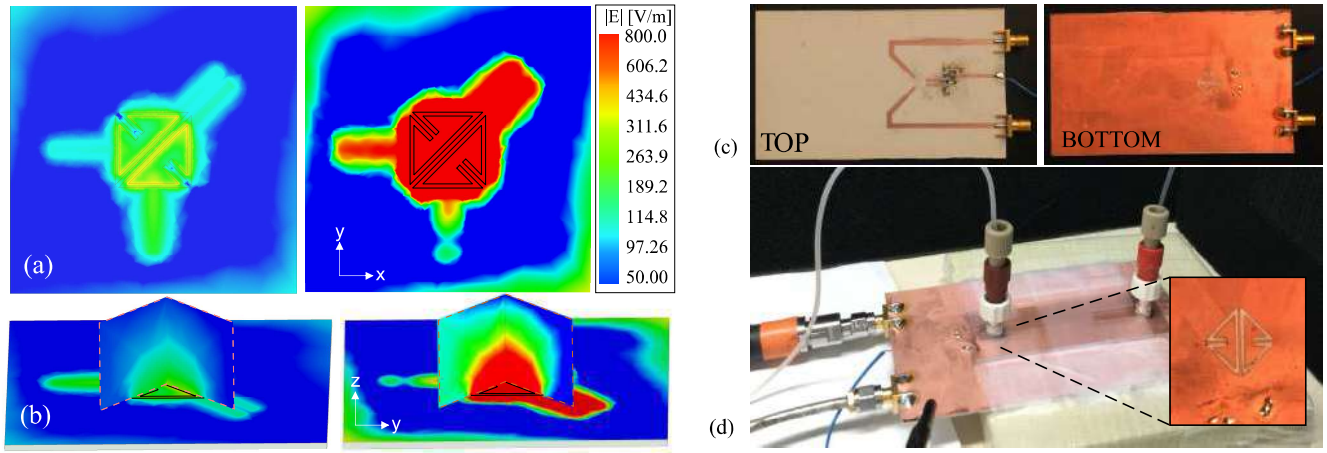


Figure 4.4: Electric Field concentration (a) on the surface and (b) surrounding the sensor in both passive (left) and active (right) states. (c) Fabricated sensor from top(left) and bottom(right) view, (d) Sensing deployment.

(a)), and in the surrounding space (Fig. 4.4 (b)), through a vertical cross section. It is evident that the slots are highly heated in the diagram, indicating a good potential for interaction with external materials. Moreover, the deeper extension of the electric field into the free space proves the sensor's capability in dealing with lossy environment and materials.

4.3 Measurement and Discussion

The proposed sensor is fabricated on Rogers RO3003 substrate as illustrated in Fig. 4.4 (c) that shows the top and bottom sections separately. The fabrication process includes printing patterns on both the top and bottom sides of the substrate using a laser printer and etching unmasked sections with ammonium per-sulphate in water solution. In this specific design, special care must be taken to consider the alignment of the two designs on both sides of the substrate to establish a correct coupling between the resonators and transmission lines of the active circuit. This in-house etching method is followed by soldering the surface mount devices on the substrate. In the following step, the performance of the fabricated sensor is verified in terms of working with external liquids. A μ-fluidic channel, purchased from ChipShop [9], is taped on the sensor with an embedded channel which is 150μm thick and 2.5 mm wide as shown in Fig. 4.4 (d). A small 175 μm lid separates the channel from the surface of the sensor. PTFE tubing and fitting are used to carry the injected fluid into and out of the μ-fluidic channel. Input port of the sensor is connected to an attenuator to remove the possible high reflections of the loss-compensated sensor back to the network analyzer. All connections are through phase-stabilized cables to preserve the sensitive phase information. The measurement results are controlled with various bias points through V_c, which essentially powers the amplifier on at different states.

Table 4.1: Dielectric Properties of Common Chemicals.

Chemical	**IPA**	**Ethanol**	**Methanol**	**Water**
Permitivity ε_r	4.13	6.8	20	76
Loss tangent tan(δ)	0.57	0.8	0.68	0.14

4.3.1 Sensor Calibration with Chemicals and Glucose

In this section, the sensor response is calibrated with respect to the simulation results obtained by standard permittivity values of the materials with the sensor in active mode. The loss compensation of the sensor is achieved with a bias voltage of $V_C = 1.8$ V, wherein the left resonance (see Fig. 4.3 (a)) is dominantly compensated. Various common chemicals including IPA, ethanol, methanol, and water are injected into the channel made of polymethyl methacrylate (PMMA) with 150 μm thickness that is separated from the sensor surface with a 175 μm lid. The measured resonance frequency shift from air is shown in Fig. 4.5 (a) for all samples. The sensitivity of the design $S = [f_0 - f_{res}]/[f_0(\varepsilon_r - 1)]$, where f_{res} is the the resonance frequency because of the material with ε_r and f_0 is that of the bare resonance. In parallel, simulation of the proposed design in Ansys HFSS is conducted on material permittivity values at 3.5 GHz as given in Table 4.1. The simulation results, shown in Fig. 4.5 (a), incorporate active circuit symbolized with the equivalent negative resistance of $R = -180\ \Omega$ and the phase of 126°. The resonance frequency shifts increases as a result of higher permittivity values, which demonstrates high agreement between simulation and measurement relative results. Inaccuracy of the sensor response can be evaluated from the percentage difference between the simulated and measured responses, which is capped at 0.02 % on average. The benefit of using a coupling region between two resonators is that the corresponding sensitivity drop at higher permittivity values is not drastic. This feature enables the proposed sensor to present a decent sensitivity up to 0.03 % as shown in Fig. 4.5 (a) considering a tiny sample volume of 4.5 μL inside a μ-fluidic channel. Also, the dynamic range of the sensor encompasses common permittivity range from 1 up to 80 that corresponds to $\Delta f = 24\ MHz$ frequency shift. In addition, the dynamic range of the sensor can be extended by optimizing this coupling, compared with a microwave sensor utilizing only one resonator. In the next section, the interaction of the proposed sensor with a lossy medium is discussed in greater detail.

In addition to common chemicals, the proposed sensor response is evaluated for various glucose concentrations in the range of 50 up to 400 mg/dL with 50 mg/dL increments. These samples are individually prepared with dissolving glucose powder in DI water as a baseline for sensor calibration. The corresponding dielectric properties of glucose solutions are imported into simulation model and the transmission amplitude $|S_{21}|$ is analyzed as shown in Fig. 4.5 (b). Measurements are replicated for 3 times to obtain repeatable results with inaccuracies of $< 10\ mdB$. The accuracy of AccuCheck glucometer is also verified on the prepared samples with the results shown in Fig. 4.5 (b) that represent < 10 % inaccuracy in the glucose level. This measurement depicts the high accuracy of the proposed sensor along with decent

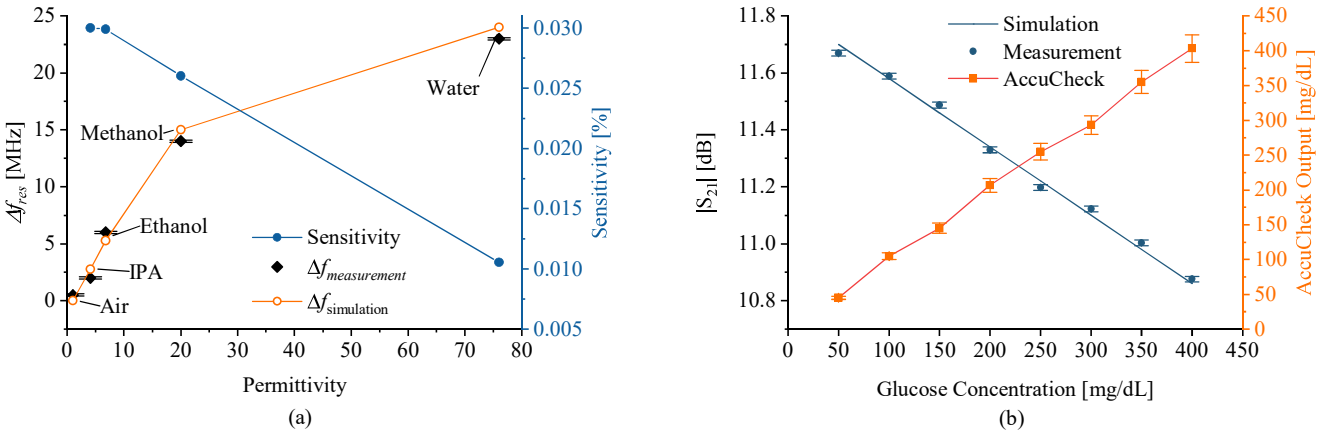


Figure 4.5: (a) Resonance frequency difference compared to air in simulation and measurement and corresponding sensitivity for various common chemicals inside μ-fluidic channel.(b) Amplitude shift in simulation/measurement of the proposed sensor with corresponding AccuCheck response with respect to different concentrations of glucose in DI water [50: 50: 400] mg/dL.

performance from the commercial gauge.

4.3.2 Water in Ethanol Sensing

To evaluate the ability of the proposed sensor to deal with minute changes in the environment (closer to the limit of detection), various concentrations of water in ethanol were prepared and tested with the sensor. A μ-fluidic channel was mounted on the sensor to carry the solution, as shown in Fig. 4.6. The binary solution is created using two syringe pumps containing water and ethanol. The pumps are set to flow at different speeds so that the water in the ethanol solution's concentration varies within the range of 1% up to 4%. This variation in the amount of water is smaller than the resolvable range in the passive mode of operation as shown in Fig. 4.6. A high-fidelity sensing response in bandpass filters, such as the current work, considers a variation that causes the shifted resonance profiles not to overlap within a 3-dB bandwidth region.

The passive mode of operation leads to a confusing response with the majority of the transmission profiles overlapping one another. This makes it very difficult to discern the differences between the individual profiles. To ensure a more robust performance of the sensor, its loss-compensated response is also measured with the same solutions of water in ethanol. As shown in Fig. 4.6, the high loaded quality factor of the active sensor results in a higher fidelity of its response. This remarkably enhances the applicability of sensor for detecting smaller quantities of materials.

The high fidelity of the sensor response is due to the active circuitry and the delicate phase compensation. However, this approach also requires high power consumption at higher frequencies and it hits high frequency-related technology bottlenecks for circuit design. In addition, scalability of the active sensor for use in other platforms including microwave, ultrasound, and optical, is burdensome. To cope with these

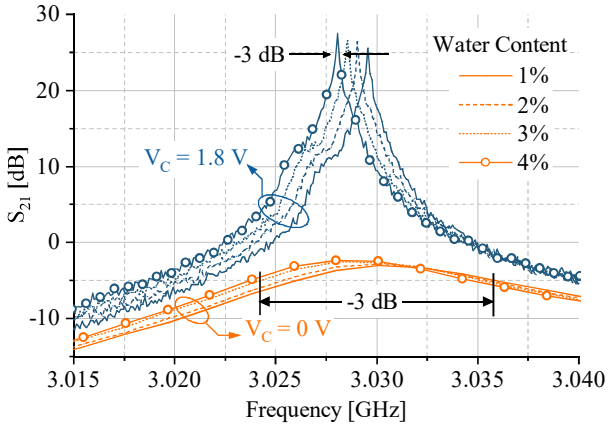


Figure 4.6: Measured sensor transmission response in passive and active mode of operation.

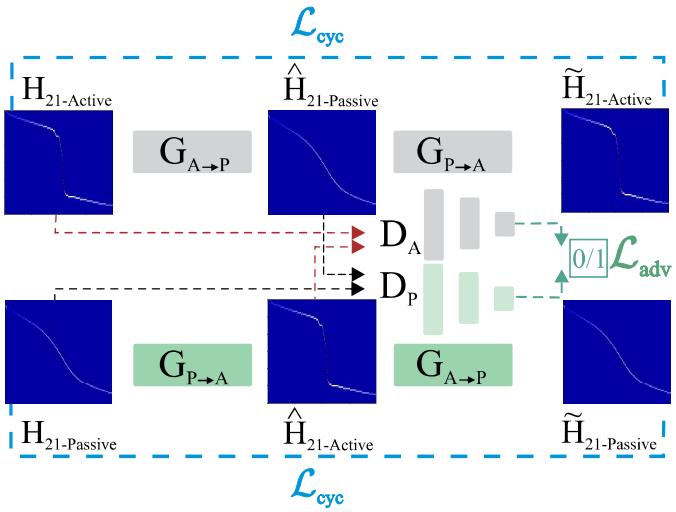


Figure 4.7: Structure of CycleGAN. $G_{A\to P}$ and $G_{P\to A}$ denote active-to-passive and passive-to-active generators.Discriminators of passive and active domain are shown as D_A and D_P, respectively. Cycle consistency loss is represented as $\mathcal{L}_{cyc}$.

limitations associated with the use of active circuit, in this work, we introduce a novel technique that expands the applicability of the sensor in passive mode using machine learning. The use of CycleGAN transformation enhances the response of a regular sensor, enabling its use for sensitive measurements and high-end applications. The principle of operation of CycleGAN and its application to the measured sensing data are described in the next section.

4.4 Architecture of CycleGAN Model

Mapping between an input image and an output image is a subset of image processing problems conventionally performed using a collection of paired images that share some common features. Recently, transfer of images from one domain into another without much feature overlap has become an interesting

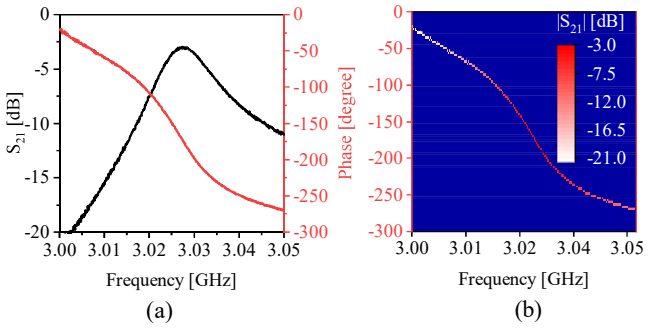


Figure 4.8: Developing heat map diagrams from S-parameter recordings. (a) Magnitude and phase of S_{21}, (b) Generated heat map diagram.

topic in machine learning [211–213]. However, in many situations, there are not enough paired images to uniquely define the mapping. In such cases, the learning process can be supervised on the level of the whole training by the CycleGAN algorithm.

In this work, CycleGAN is used to overcome the limitations of passive sensors outlined in the previous section. Drawing on the definitions common in image processing, the input/output images refer to the representations of the sensor response in the passive/active modes. CycleGAN algorithm is used to learn the relationship between the responses in the two modes. It captures the characteristics of active responses and applies them to passive responses, all without the need of a paired training set.

Conventionally, the low-Q factor of the passive resonance profile is the main attribute preventing high-resolution sensing. The proposed method transfers the low-Q profiles into the equivalent high-Q profiles to facilitate sensing with high-resolution. In this algorithm (see Fig. 5.1), a mapping $G_{P\mapsto A}$ is trained to transform an image p from the passive domain P into another image $\hat{a}$ in the active domain A. The new image $\hat{a} = G_{P\mapsto A}(p)$ is expected to be indistinguishable from other images in the active domain $a \in A$. The model includes two discriminators D_A and D_P such that D_P distinguishes between the input image p and the translated image $\hat{p} = G_{A\mapsto P}(a)$, while D_A discriminates between a and the translated images $\hat{a} = G_{P\mapsto A}(p)$. In order to use this scheme on the sensor dataset, the recorded transmission profiles are converted into images as follows. The main contributing factor in the measurement of a two-port system is the transmission parameter (S_{21}), which is a complex-valued vector. A simple plot of only amplitude of S_{21} or its phase over frequency creates line graphs. However, one can combine these two graphs as shown in Fig. 4.8 into a single plot by embedding the values of magnitude as RGB colors in the plot of phase vs. frequency. The generated heat map, called H_{21} hereafter, is then considered an input image. This conversion results in 512 bins for each frequency and phase. The pixels can accept a combination of three colors of red (R), green (G), and blue (B). Therefore, a single measurement is reshaped into an image with a data volume of size $512 \times 512 \times 3$.

The proposed method uses CycleGAN to translate the passive domain P into the active domain A, where the distribution of both are forced to be identical using adversarial loss [176]. Since this

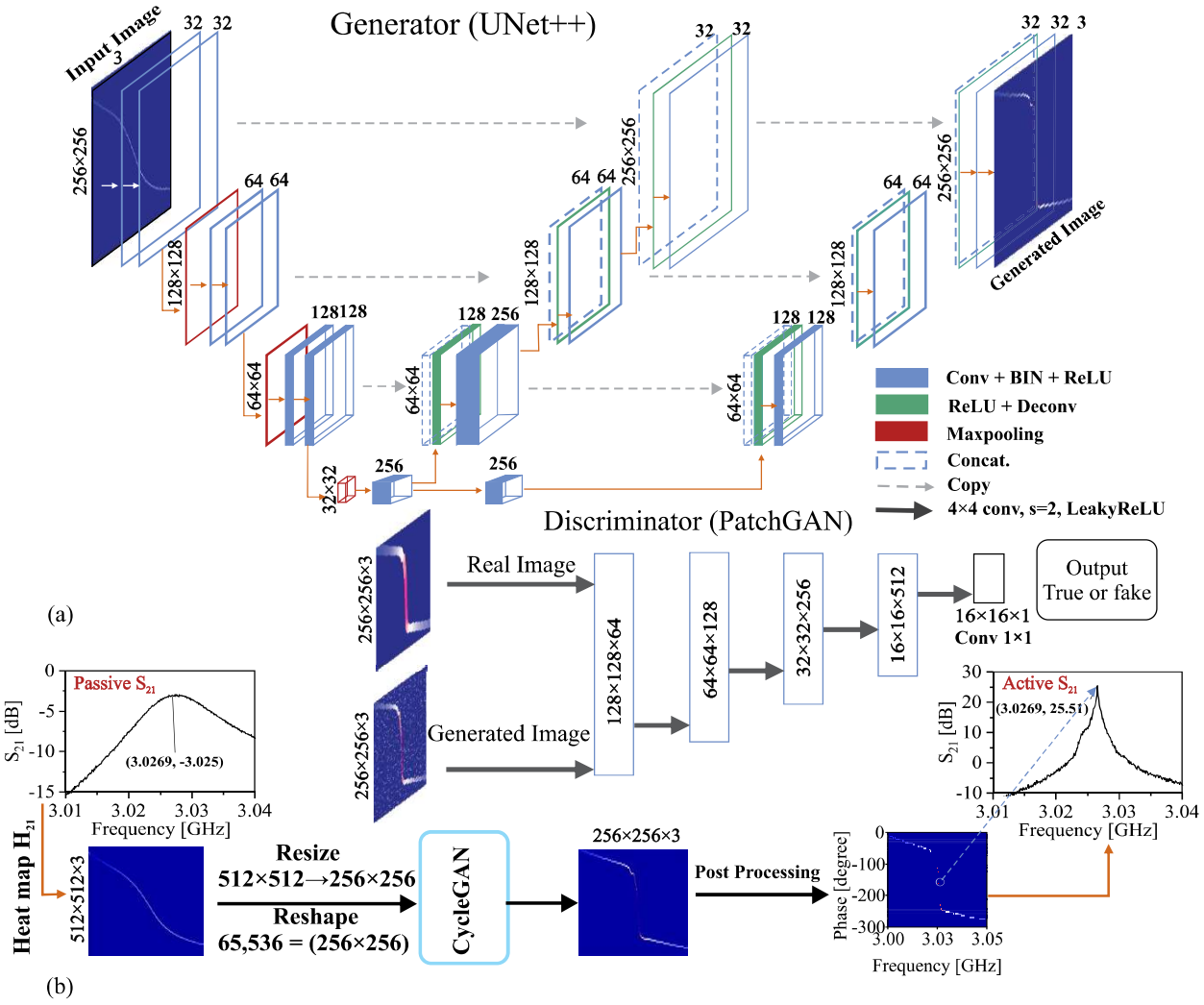


Figure 4.9: (a) The Architecture of generator and discriminator, (b) A workflow of data preparation to generate the active S_{21} profile from passive.

transformation model might result in many-to-one mappings, consistency of the transformation cycle must be preserved to avoid this ambiguity. To achieve this, the translator $G_{P\mapsto A}$ is accompanied with its inverse $G_{A\mapsto P}$ that converts the active domain into the passive one. This is implemented by using the output of one generator as the input of the other generator and combining the two mappings with a cycle consistency loss [186] that reinforces $G_{A\mapsto P}(G_{P\mapsto A}(p)) \approx p$ and $G_{P\mapsto A}(G_{A\mapsto P}(a)) \approx a$. Therefore, the objective function of CycleGAN is composed of adversarial loss as well as cycle consistency loss. The adversarial loss represents the difference between the real and estimated images, such that in generator G and discriminator D, the cost function is defined as follows:

$$\mathcal{L}(G_{A\mapsto P}, D_P) = \min_{\theta_g} \max_{\theta_d} \{\mathbb{E}_{\mathbb{P}}[\log D_P(h_{21P})] + \mathbb{E}_{\mathbb{A}}[\log(1 - D_P(G_{A\mapsto P}(h_{21A})))]\}, \tag{4.12}$$

where p and a are the unpaired training datasets, and $\mathbb{E}_P$ and $\mathbb{E}_A$ are the expectations in the output and

input images. The discriminator D_A and generator $G_{(P\mapsto A)}$ are trained synchronously within a min max game style with weights θ_g and θ_d.

Input and output images of P and A are trained simultaneously, even though they are different representations. Thus, cycle loss is used to encourage the consistency between forward and backward training. Then, a secondary consistency loss is defined as follows: Input and output images of P and A are trained simultaneously, even though they are different representations. Thus, cycle loss is used to encourage the consistency between forward and backward training. Then, a secondary consistency loss is defined as follows:

Algorithm 1 Minibatch stochastic gradient descent training of CycleGAN

Input: Minibatch H_{21} Heat Maps $\rightarrow H_{21Active} \in$ A (source domain) and $H_{21Passive} \in$ P (target domain)
Output: Update network weights $\theta_{g(A\rightarrow P)},\ \theta_{g(P\rightarrow A)},\ \theta_{d(A)},\ \theta_{d(P)}$
Initializing discriminator and generator weights.

1: **for** number of training iterations **do**
2: $(\hat{H}_{21Passive},\ \tilde{H}_{21Active}) \leftarrow (G_{(A\rightarrow P)}H_{21Active}),\ (G_{(P\rightarrow A)}\hat{H}_{21Passive})$
3: $(\hat{H}_{21Active},\ \tilde{H}_{21Passive}) \leftarrow (G_{(P\rightarrow A)}H_{21Passive}),\ (G_{(A\rightarrow P)}\hat{H}_{21Active})$
4: $\mathcal{L}_{D_{(A),(P)}} = \frac{1}{m}\sum_{i=1}^{m}[logD(H^i_{21-(A),(P)}) + log(1 - D(H^i_{21-(A),(P)}))]$ {Calculate D_P, D_A Loss}
5: $\theta_{d_{(A),(P)}} \leftarrow \theta_{d_{(A),(P)}} - \eta\nabla_{d_{(A),(P)}}\mathcal{L}_{cyc}$ {Update discriminator parameters}
6: $\mathcal{L}_{cyc} = \lambda_1 \left\|\tilde{H}_{21Active} - H_{21passive}\right\|_1 + \lambda_2 \left\|\tilde{H}_{21Passive} - H_{21Active}\right\|_1$
7: $\mathcal{L}_{G_{(A\rightarrow P)}} \leftarrow \frac{1}{m}\sum_{i=1}^{m}(1 - logD_p(\hat{H}^i_{21passive})) + \mathcal{L}_{cyc}$ {Calculate $G_{(A\rightarrow P)}$ Loss}
8: $\theta_{g(A\rightarrow P)} \leftarrow \theta_{g(A\rightarrow P)} - \eta\nabla_{\theta_{g(A\rightarrow P)}}\mathcal{L}_{G_{(A\rightarrow P)}}$ {Update $G_{(A\rightarrow P)}$ Parameters}
9: $\mathcal{L}_{G_{(P\rightarrow A)}} \leftarrow \frac{1}{m}\sum_{i=1}^{m}(1 - logD_A(\hat{H}^i_{21active})) + \mathcal{L}_{cyc}$ {Calculate $G_{(P\rightarrow A)}$ Loss}
10: $\theta_{g(P\rightarrow A)} \leftarrow \theta_{g(P\rightarrow A)} - \eta\nabla_{\theta_{g(P\rightarrow A)}}\mathcal{L}_{G_{(P\rightarrow A)}}$ {Update $G_{(P\rightarrow A)}$ Parameters}
11: **end for**=0

$$\mathcal{L}_{\text{cyc}}(G_{A\rightarrow P}, G_{P\rightarrow A}) = \|G_{P\rightarrow A}(G_{A\rightarrow P})(h_{21A}) - h_{21A}\|_1 + \|G_{A\rightarrow P}(G_{P\rightarrow A})(h_{21P}) - h_{21P}\|_1\,, \tag{4.13}$$

where $\|\cdot\|$ is L_1 loss and h_{21} denotes the heat map of the measured transmission response. The overall loss function of CycleGAN combines the two aforementioned losses as follows:

$$\mathcal{L}(G_{A\rightarrow P}, G_{P\rightarrow A}, D_A, D_P) = \mathcal{L}(G_{A\rightarrow P}, D_P) + \mathcal{L}(G_{P\rightarrow A}, D_A) + \lambda\mathcal{L}_{\text{cyc}}(G_{A\rightarrow P}, G_{P\rightarrow A}) \tag{4.14}$$

where λ is a hyperparameter to balance the loss terms, meaning that the consistency loss is λ times more important than the adversarial loss. In this case, the generators are trained to produce fake images as close to real ones as possible; and the discriminators are employed to distinguish more appropriately between real and fake images. The training stops when the discriminators reach close to 0.5 probability estimate for fake image discrimination.

Table 4.2: MSE, PSNR and SSIM Comparison of Different Algorithms. Reference: Active Sensor Response.

Methods	MSE	PSNR	SSIM
pix2pix	0.02	17.7194	0.4871
Conventional CycleGAN	0.0049	23.0437	0.7865
This work	0.00146	28.34	0.9167

4.4.1 Implementation

The proposed CycleGAN network consists of a generator and a discriminator in the source domain and another identical pair in the target domain. The passive-to-active generator network shares its characteristics with the active-to-passive network; similarly, the active discriminator D_A resembles to the passive one D_P. In the next sections, the generator and discriminator of the design are elaborated.

Generator Structure

The generator (see Fig. 4.9) is based on U-net [214] architecture; further inspired by U-net++ [215], the encoder and decoder are linked through a middle layer. This helps smoother integration of shallow and deep features. Normally-distributed weights in each layer allow deeper model connections to be used for complex scenarios. The structure of the generator can be summarized as follows. Let us consider CK as a 3×3 Convolution-BatchInstanceNormReLU layer (C) with K filters and stride 1. M denotes Maxpooling layer. DK denotes a 2×2 TransposedConvolutionalLayer-BatchInstanceNormReLU layer with K filters and stride 2, and 6 Concatenate blocks denoted as Concat.
C32, C32, M, C64, C64, M, C128, C128, M, C256, C256, D128, Concat., C128, D128, Concat., C128, D64, Concat., C64, D64, Concat., C64, D32, C32, D32, Concat., C32, C3.

Discriminator Structure

In the discriminator, real- or fake patches of the images are discriminated using Markovian discriminator (PatchGAN) [216]. Compared with regular full-sized image discriminators, this method uses less computational resources and is compatible with various image sizes. In addition, the use of least squares loss increases the robustness of its operation. Layers of the discriminator include filters with sizes of 64, 128, 256, and 512. The first four layers have a kernel size of 4×4 with LeakyRelu activation function (to introduce a small positive gradient when a neuron is not active), followed by the last layer with a sigmoid function. The discriminator architecture is:
C64, C128, C256, C512.

Training

The heat map images (512×512) are resized to (256×256) to be computationally efficient in the training process. Reshaped images are fed to the conventional UNet++ except for the batch normalization layer that is simply replaced with batch-instance normalization (BIN) [217] to normalize the styles adaptively to

the task and selectively to individual feature maps. Using BIN, the amount of style information that needs to be propagated through each channel of features is controllable by a learnable gating parameter. To make the model more general and benefit from both batch normalization (BN) in multiclass classification and instance normalization (IN) in style transfer networks, IN and BN are substituted by BIN. The networks are trained using stochastic gradient descent (SGD) with a batch size of 32, momentum of 0.9, weight decay of 0.0001, and initial learning rate of 0.1. The network is trained for 150 epochs, where the learning rate is divided by 10 at epochs 50 and 100. The entire procedure is formally presented in Algorithm 1.

4.4.2 High-resolution Sensor Response Using CycleGAN

In this section, the measured sensor response to water in ethanol solution is reevaluated with the CycleGAN network applied to the measured heat maps. Four solutions with 1%, 2%, 3%, and 4% water in ethanol are measured using the proposed sensor in both passive and active operating regime. The prepared solutions are passed through the μ-fluidic channel and the sensor responses are recorded using LabView with 5-second time resolution. The solution is passed through the channel to obtain a total of 200 sample measurements per recording before the syringe is emptied out. This process is repeated five times to acquire enough data for training purposes. Each recorded sample is then converted into its equivalent heat map (H_{21}) which combines the information on the amplitude and phase of resonance. Plots in the top and middle rows of Fig. 4.10 (a) show the heat maps obtained using the passive and active mode, respectively. Corresponding profiles obtained using CycleGAN trained for 150 epochs are shown in the bottom row of the figure. The output of this transformation is a new heat map trained to be similar to the heat map obtained from the active sensor data.

A magnitude-and-phase representation of the four measured solutions is shown in Fig. 4.10 (b), including the passive -domain measurements as well as the active-domain representation used as the target for CycleGAN training. The outputs obtained by mapping the passive-domain representations through the trained network are also included. The general shape of the high-resolution profiles clearly visible in the output response of CycleGAN, although there are some nuances in the absolute values of the S_{21} magnitude at the resonance frequency. The high-resolution response of the sensor using CycleGAN has a loaded quality factor of $Q_{\text{CycleGAN}} \approx 5040$, which exhibits resolution more than 20 times higher than that of passive domain $Q_{\text{Passive}} \approx 230$. The frequency dependent resolution of the sensor in active mode reduces down to $\approx 100\ kHz$ from $\approx\ 1\ MHz$ of passive case. The similarity between CycleGAN-boosted response and the active sensor response is evaluated in Table 4.2 in terms of several common metrics [218, 219]. Please refer to chapter (A) for more information on the structural similarity index (SSIM), mean-squared error (MSE), and peak signal-to-noise ratio (PSNR). The approach introduced in this contribution is compared to pix2pix cGAN [220] and to conventional CycleGAN [186]. The proposed network outperforms both algorithms with high SSIM of 0.9167 and low MSE of 0.00146. This is mainly due to replacing the ResNet (9 residual block for 256×256 generator) of conventional CycleGAN [186]

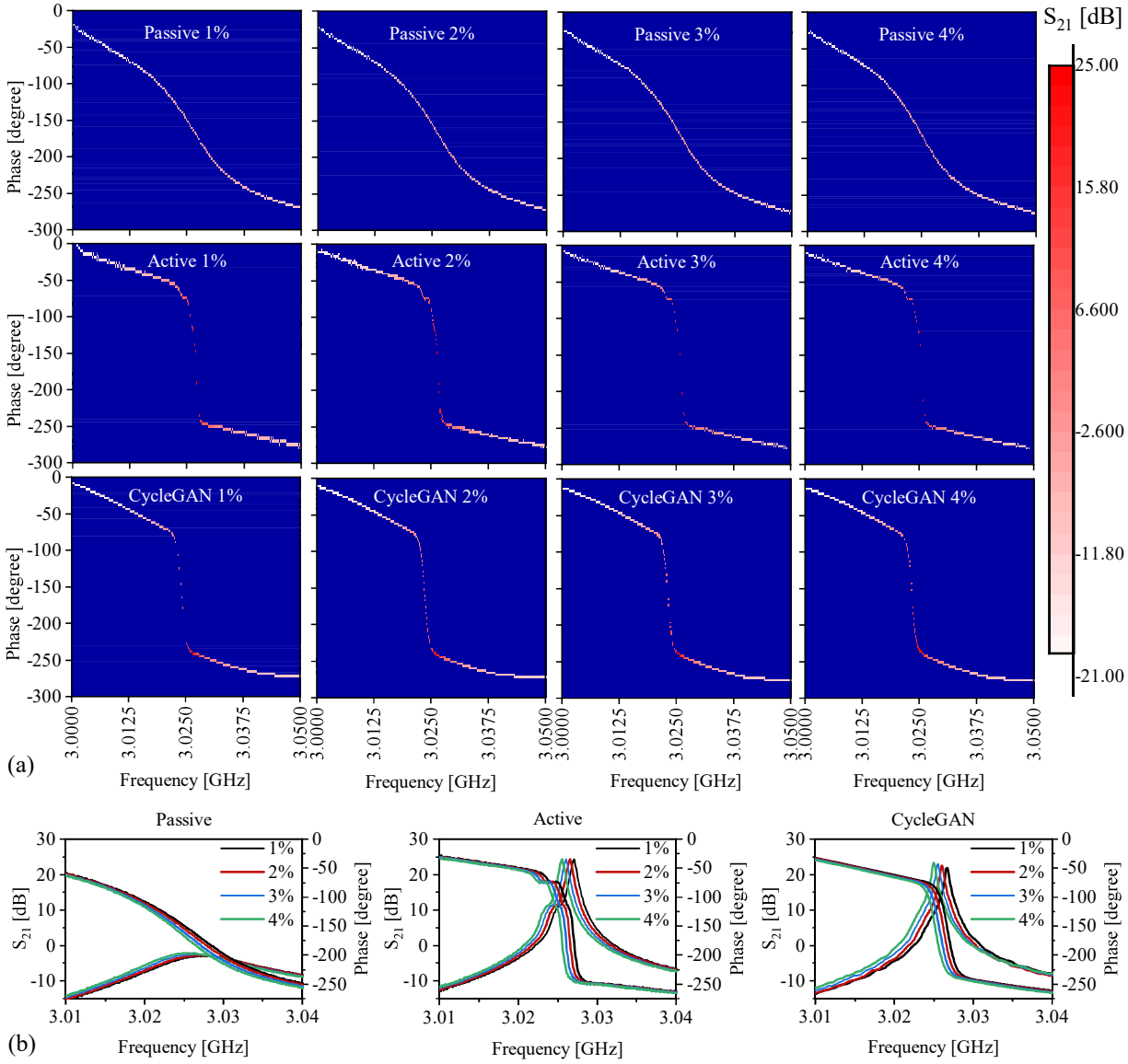


Figure 4.10: (a) Equivalent heat maps for one sample of 1%, 2%, 3%, and 4% water in ethanol using the sensor in passive mode (top row), active mode (middle row), and passive mode aided with CycleGAN (bottom row), (b) Measured sensor response for one sample of 1%, 2%, 3%, and 4% water in ethanol passive mode (left), active mode(middle), and CycleGAN-boosted passive mode (right)

with UNet++ [215] and also substituting instance normalization with batch-instance normalization [217]. BIN allows replacing Adam optimizer with SGD, which results in better performance of the network.

4.4.3 Glucose Sensing

In another experiment to evaluate the proposed approach, the CycleGAN-enhanced sensor is employed in an noninvasive glucose sensing application [4]. The glucose content increases in the blood as a result of food intake. It gradually drops over time while being absorbed by the cells. This fluctuation in the blood glucose

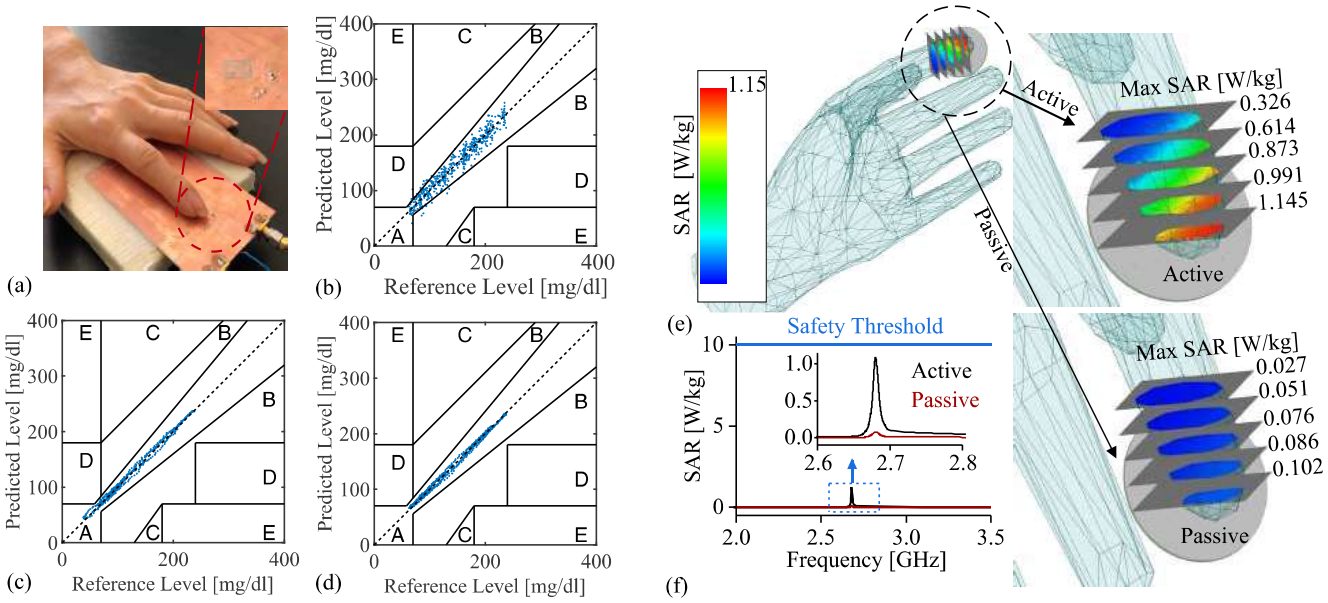


Figure 4.11: (a) Fingertip measurement Setup, Clarke error grid analysis for (b) Passive, (c) Active , (d) CycleGAN-boosted passive states of the sensor response.

level becomes apparent in the interstitial fluid under the skin with minutes of delay. This leaves a reliable trace of the actual glucose level deep inside veins. Non-contact sensing has already shown robustness in glucose sensing using active regime with high sensitivity and reasonable fidelity. This has been achieved using sophisticated active circuitry, which requires constant powering for continuous operation [221,222]. An alternative solution can be developed using the approach proposed in this article: the passive sensor response can be boosted using CycleGAN. This alternative preserves the high-Q performance normally offered by the active circuit design, while using zero-power sensors.

The performance of CycleGAN-boosted glucose sensor was evaluated using the following experiment. A participant, not suffering from diabetes, underwent continuous monitoring of their glucose level. The sensing was noninvasive, implemented by placing the subject's index finger on the sensitive region of the sensor as shown in Fig. 4.11 (a). Since the glucose content of the interstitial fluid changes the dielectric constant of the skin, the participant's right finger was placed in the close proximity of the sensor surface. While the participant was asked to minimize their movement during the experiment, the transmission profile of the sensor was measured continuously. During this process, 500 data points were recorded with 20-second resolution and the participant was fed at different times to control the glucose level.

An Accu-Check commercial glucometer was used to measure the blood glucose on the left hand finger of the participant every 5-minutes. The collected measurements serve as a reference for calibrating the sensor response. Using this technique, the resonance frequency and amplitude are correlated to the change in blood glucose level. Passive sensor response, due to its low-to-moderate loaded quality factor, is more prone to environmental amplitude noise. Considering this concept, measurements of the sensor

response in both passive and active regime (performed separately) are analyzed to extract the resonance frequency. For the passive resonance profile, its heat map is generated and fed into CycleGAN to obtain high-loaded quality factor equivalent heat maps. Subsequently, the resonance profile S_{21} is extracted from the corresponding heat maps. The resonance frequency is the most importance feature of the CycleGAN-generated heat map from CycleGAN as it must exactly correspond to the original frequency of operation. To ensure this correspondence, the cycle consistency loss is used to preserve the resonance frequency.

The accuracy of the proposed sensor in detecting glucose levels is evaluated using Clarke error grid analysis (EGA) [223]. The results of this analysis, performed on 500 recordings of the interstitial glucose level, are illustrated in Fig. 4.11. The horizontal axis represents the reference glucose level obtained from the commercial glucometer, while the vertical axis shows the level inferred using the proposed approach.

For the passive sensor response (see Fig. 4.11 (b)), 96.2% of paired sensor reference values fall into region A, which holds 20% of the reference values. Region B, which has next 20% of reference values in a form acceptable for treatment, contains 3.19% of pairs. In this evaluation, regions C and E do not contain any paired data. However, 0.59% of the data pairs fall in region D, which indicates potentially dangerous failure to detect low or high levels of glucose. The apparently scattered data points of the passive sensor Clarke EGA confirm that the low-resolution output can lead to unreliable detection of resonance frequency. This observation is further supported by relatively large values of root of mean squared error (RMSE) of 11.31, mean relative absolute deviation (mRAD) of 7.08, and mean absolute error (MAE) of 9.09.

For comparison, the same sensor was used in active mode of operation in the following day. The experimental setting including the measurement time, duration, fasting mode, food intake pattern, and the participant remained identical to the passive mode measurement conditions to avoid endogenous/exogenous masking effects. Since this measurement were conducted in a different session, the results do not correspond to the passive sensor measurements directly. However, they can be used to compare the general precision between the two sensor state operating modes. With loss compensation in the circuit, the prediction of glucose level became more precise with much less noise infused into the sensor response. The resonance frequency of each profile was extracted in correspondence to each reference glucose level as shown in Fig. 4.11 (b) and Clarke EGA was also applied to this data. It is shown that the spread of the predicted values is confined within a limited area, closer to the diagonal line. The compositional analysis of these measurement shows that all 100% datapoints are located in the zone A. The errors measures also attain much lower values of RMSE = 3.53, mRAD = 2.22, and MAE = 2.85.

Considering the improvements obtained using the active sensor data, the result of the passive sensor were also analyzed after CycleGAN transformation. As shown in Fig. 4.11 (c), the spread of datapoints of Clarke EGA is significantly reduced compared to that of datapoints obtained using passive sensor directly and all datapoints fall into zone A. The values of error metrics are also in line with active sensor response, i.e. RMSE = 4.013, mRAD = 3.14, and MAE = 3.17. The considerable improvement between the passive sensor response and the CycleGAN-boosted response confirms the high potential of

the proposed technique as a general method for improving sensing performance. This approach is not limited to microwave sensors but could be extended to various other applications.

The electromagnetic wave absorption of the proposed sensor when being applied to human skin is also simulated in full-wave simulator HFSS. While the skin is being monitored continuously, the specific absorption ratio (SAR) is computed at the frequency of operation and averaged over 10g of tissue that is considered with the average human hand model as shown in Fig. 4.11 (d). Maximum SAR values at different cross-sections of the finger tip is shown in both passive (R=-0.1 kΩ) and active states (R=-0.2 kΩ) of the sensor according to Fig. 4.11. The maximum SAR values across all hand volume over a wide range of frequencies is also shown in Fig. 4.11 (e), which depicts a fair margin for SAR values (max(SAR)<1 W/kg) compared with the safety threshold of 10 W/kg. This reassures safe range of applications for clinical trials and wearable applications.

4.5 Conclusion

In this research, we have demonstrated the potential of a microwave planar sensor for liquid material characterization and glucose sensing. The sensor, despite its moderate quality factor, has shown promising results, especially when enhanced with a regenerative amplifier. This enhancement significantly improves the quality factor of the sensor responses, thereby reducing the limit of water detection and improving the resolution of glucose sensing.

The application of CycleGAN, a machine learning algorithm, in processing the sensor responses has proven to be a game-changer. It has not only improved the quality factor but also allowed for the extraction of more accurate and reliable data from the sensor responses. This innovative approach of combining microwave sensing with machine learning algorithms opens up new avenues for research and development in the field of material characterization.

Furthermore, the advantages of microwave sensors over their optical and electrochemical counterparts have been highlighted in this study. The microwave sensors' ability to work in opaque environments, their non-invasive nature, and their potential for remote sensing make them a more versatile and practical solution for various applications.

However, it is important to note that while the results of this study are promising, further research is needed to optimize the sensor design and the machine learning algorithms used. Future work should focus on improving the sensitivity and selectivity of the sensor, as well as exploring other potential applications of this technology.

In conclusion, this study has made significant strides in the field of material characterization. It has shown that with the right enhancements and processing techniques, microwave planar sensors can be a viable and effective tool for these applications.

Chapter 5

Prediction-Anomaly Detection by LSTM

5.1 Motivation (Overview)

The research proposes a microwave sensor for continuous, instant glucose measurement in both microfluidic channel configurations and human trials, addressing a crucial requirement in diabetes management. Despite initial efforts to establish a closed-loop insulin delivery system to regulate hypo- and hyperglycemia, several problems (including sensor delays and malfunctions) persist, causing discrepancies between the measured and actual glucose levels. To circumvent these issues, the study suggests the integration of continuous glucose prediction, facilitated by machine learning algorithms, to mitigate natural latency and interpolate missing measurements. Notably, a long-short-term memory (LSTM) network is employed to forecast glucose levels 30 and 60 minutes into the future, aiding in anomaly detection arising from omitted readings or unintentional changes in the measurement setup. The effectiveness of these techniques is extensively scrutinized through in-vitro proof-of-concept experiments and human trials.

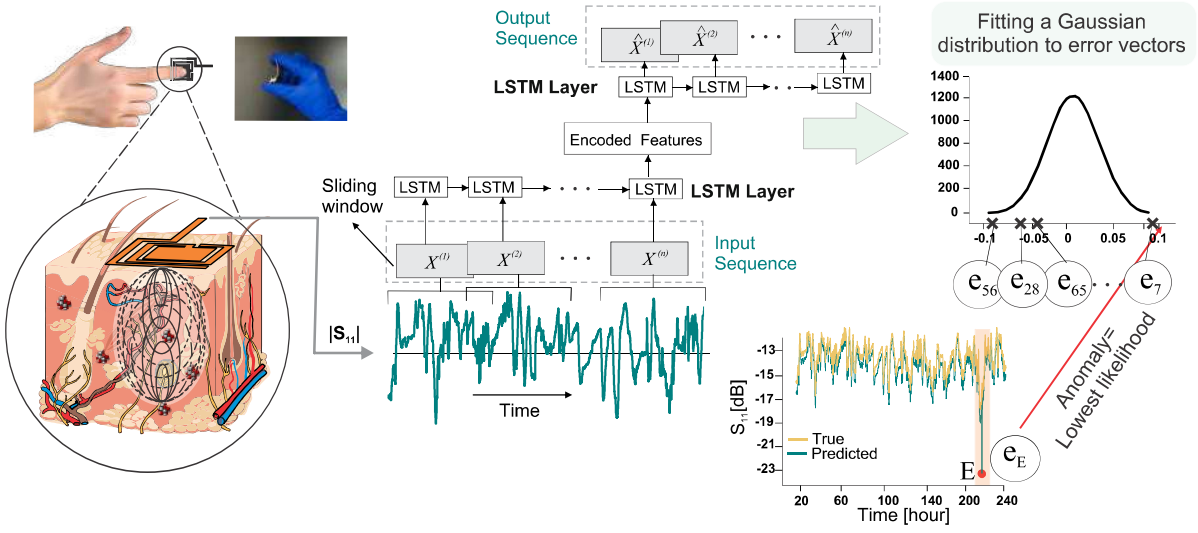


Figure 5.1: Schematic of the Multistep LSTM system: An advanced tool that uses deep learning to predict blood sugar levels and identifies potential recording anomalies caused by factors such as unusual pressure on the sensor or sudden temperature changes.

5.2 Methods

5.2.1 Sensor simulation and fabrication

The sensor is designed at low ultrahigh frequencies (UHF), 3 GHz, enabling deep penetration of electromagnetic (EM) waves into tissues. The in-depth interaction of wave and matter in this frequency range helps to achieve high sensitivity for small amounts of analytes (more details are provided in the Appendix. A). The initial conceptual design is based on a flexible substrate, which in this case is designed on a Rogers RO3003 substrate with the dielectric constant of $\varepsilon_r = 3.3$, and loss tangent of $tan(\delta) = 0.0013$, with 0.8 mm thickness. The sensor is printed on one side of the substrate, while the other side is fully covered with copper as an EM shield. This design increases the stability of the sensor response and reduces its susceptibility to environmental variations, such as the effects of proximity to surrounding objects.

The design criteria for the proposed sensor are the frequency of operation and the sensor size, which ensures its compatibility with flexible substrates. The loss factor from the glucose content affects the optimal frequency of operation as discussed in Fig. A.2. Geometrical form factor, on the other hand, is considered to be as compact as possible so that it can offer sensitive operation in a compact form and can be deployed at different locations on the body. The main concept behind the proposed sensor is the coupling between the underlying SRR and the loading patch that the SRR fully surrounds (as shown in Fig. 5.2 (a)) which eventually results in a miniaturized structure. This is a one-of-a-kind form of coupled resonators wherein SRR and patch are not completely identical resonators (as opposed to [224]). In this combination, SRR is responsible for a highly sensitive region nearby the SRR gap while the patch considers deeper levels of body tissue due to its high gain characteristics. This very feature imposes t interaction of different resonance modes that results in a higher sensitivity due to glucose content at both high and low depths from the skin. Also, as evident from Fig. 5.2 (b), individual constituent parts (SRR and patch) are resonating at high frequencies of 5.65 and 4.34 GHz, respectively. However, with respect to the coupling between individual resonators, this design gives rise to a highly sensitive resonance at a significantly reduced frequency of 3.6 GHz (only sensor), which denotes a size reduction of 36% and 17% with respect to SRR and patch, respectively.

An open-ended transmission line is used to excite the resonator, surrounding it for better energy transfer. This single-port sensor measures the reflection from a resonator loaded by an external MUT. Compared to other transmission-based sensor designs, including the peak of resonance, the proposed sensor can be highly matched with respect to the MUT to achieve a low reflection. Furthermore, the transmission-based counterpart sensors suffer from degradation in the sharpness of the transmission response because the input power is either absorbed into the MUT-loaded resonator or propagates into the free space, which leaves low power received by the output. It has been found that when the loading material becomes lossy, the distance between the resonator and the exciting transmission line can be modified to change the external loading of the transmission line and achieve a high input impedance matching, and

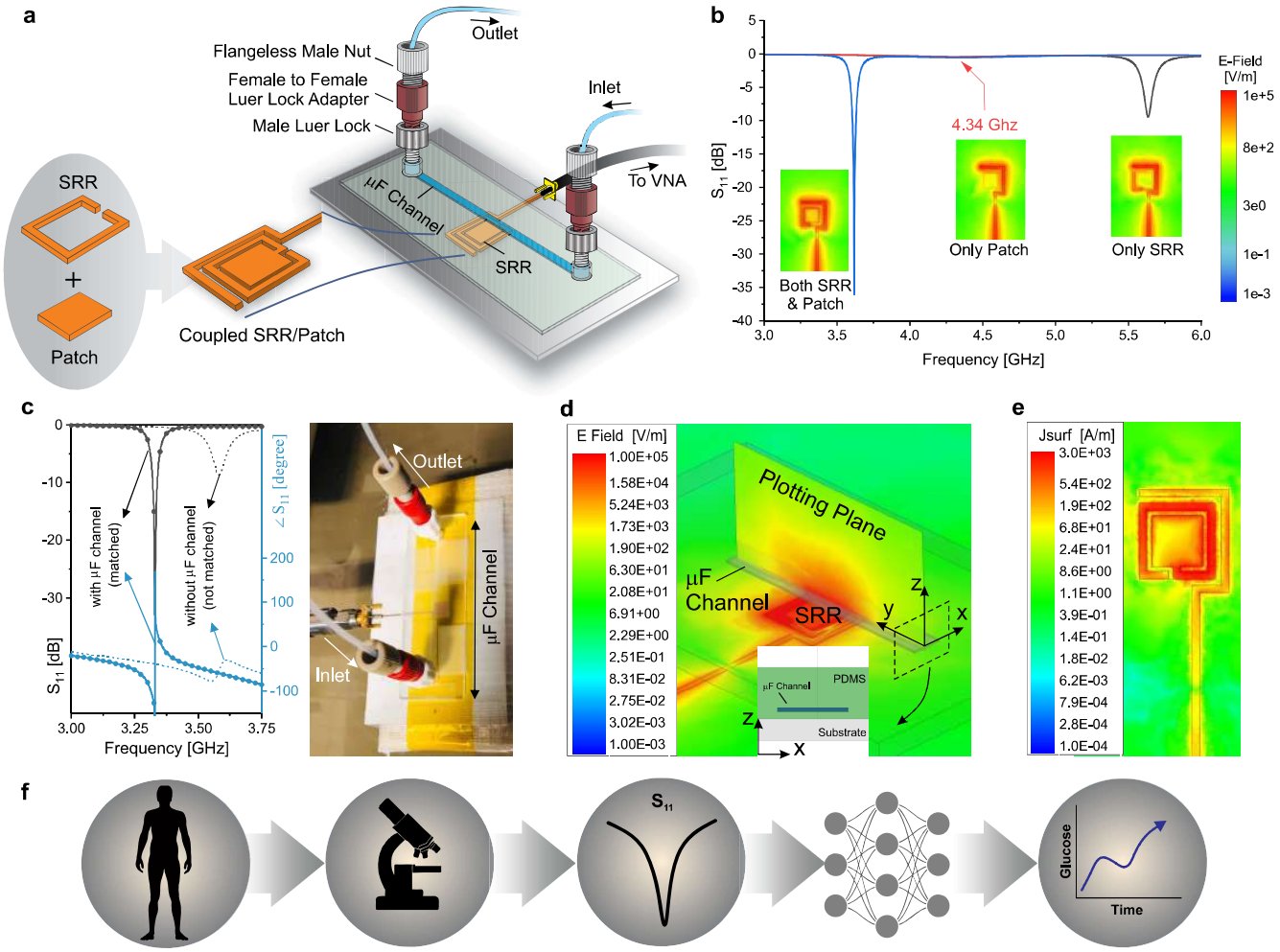


Figure 5.2: (a) Schematic of the proposes sensor with μF channel mounted. (b) The evolution of the senor design with respect to its constituents (SRR/Patch). (c) Magnitude and phase simulation of S_{11} for both the sensor without μF channel (not matched) and with μF channel (matched) on top of it using Ansys Electronics (HFSS) (left) as well as a fabricated sensor with μF channel on the right. (d) $|E|$-field distribution is plotted at the frequency of resonance over the surface of the sensor with its cross-section shown in the inset. The color bar shows the intensity of the electric field. (e) Surface current density of the sensor with the μF channel at the frequency of resonance. (f) Generic paradigm of glucose level prediction by deep learning with respect to data processing through sensor interaction with the human body.

thus a deeper resonance profile. Another variation of the proposed structure can be a loss-compensated resonator discussed in [15,82]; however, since in that case the depth of resonance becomes extremely low, the sensor response stability degrades substantially and becomes highly vulnerable to environmental noise. Therefore, the focus of optimized sensor design is to utilize a reflection-based passive sensor, wherein the MUT is used as a part of sensor response matching in receiving maximum input power. More specifically, the sensor is designed to handle MUT loading. A well-matched sensor depends on its input impedance which varies with respect to its loading. Normally, when the sensor is designed with a desirable matching with a bare μF channel, the performance drops significantly when the MUT is introduced.

A μF channel with dielectric properties of $\varepsilon_r = 2.9$ and $tan(\delta) = 0.013$ at low UHF frequencies is used as the MUT carrier for *in vitro* tests (see Fig. 5.2 (a)). The deployed μF channel is a hydrophilized

polymethyl methacrylate (PMMA), purchased from Microfluidic ChipShop as a straight channel. The fluidic channel is connected to a flat lid with adhesive tape. The thickness of the lid is 175 μm, and the channel inlet and outlet are 58.5 mm apart. The thickness and width of the channel are 150 μm and 2.5 mm. The channel is secured on top of the sensor with proper taping. The inlet and outlet of the channel, aided with polytetrafluoroethylene (PTFE) fittings, including a male Luer lock, a female-to-female Luer lock adapter, and a flangeless male nut, are connected to two programmable syringe pumps. The effect of μF channel on the bare sensor matching can be avoided by modifying the physical resonator parameters toward better input impedance matching specified. The same strategy is applied for clinical trials with a separate sensor designed to consider the dielectric properties of human skin. This essentially adds a degree of freedom to the final product that can be customized according to the personal characteristics of a specific subject.

High-frequency simulations are performed using the Ansys Electronics Desktop simulator (HFSS). Figure 5.2 (c) represents the magnitude and phase response of the sensor when optimized for low reflection with the channel above it. Deep resonances down to -34 dB are achievable for an empty channel at 3.326 GHz (matched), while the very resonator parameters yield a higher reflection when the channel is removed from the sensor (unmatched), shown with the dashed line. A fabricated prototype with the μF channel mounted on its top is shown in Fig. 5.2 (c). The sensor is fabricated using a high-precision milling machine from LPKF Laser and Electronics. The electric (E) field distribution at the resonant frequency is shown on the sensor surface in Fig. 5.2 (d). All gaps in the sensor are colored red, showing a high concentration of the E-field. This denotes a high coupling between the inner patch and the exterior SRR. For the purpose of this experiment, due to the limited width of the μF channel, only one gap is engaged. Sensor excitation is carried out through the open-ended transmission line that surrounds the sensor. This configuration allows for higher engagement of different resonator locations in receiving power from the input transmission line, leading to a better (deeper) matching of the sensor. The propagation of EM waves from the resonator to the surrounding medium is captured using E-fields in the plane that crosses the middle of the μF channel along the y-axis as shown in the inset of Fig. 5.2 (d). This shows how the sensor interacts with the surrounding materials non-invasively. Figure 5.2 (e) represents the surface current distribution at the resonance frequency, where highly concentrated currents flow around the edges of the gaps, indicating a high potential for the currents to be disturbed by an external medium. Due to the special configuration of the proposed sensor, one can use all sides of the SRR as a sensitive region with capacitive coupling, mainly due to the coupling between the SRR and the internal patch.

5.2.2 Data processing

In the proposed sensing system, depicted in Fig. 5.2 (f), the change in the total measured reflection profile (S_{11}) is expected to vary linearly since the variations in the GL are small (whether in water/ISF or in the body). Thus, using extra features in the analysis would only load the post-processing unit and make the proposed method computationally costly and unsuitable for real-time application. Therefore, only a single

parameter is used as a representation of the GL changes. Since the variation in the frequency of operation is negligible, specifically for finer variations, it is replaced by the amplitude of the S_{11} depth, which varies linearly with GL in both *ex vivo* and *in vivo* experiments. Each measurement includes 3001 frequency points (amplitude and phase of S_{11}), and the depth of the profile is processed as the single feature of the measurement. Measurements are performed using the S5065 vector network analyzer (VNA) with an IF bandwidth of 1 kHz; an average of 3 measurements is used for data processing. The initial postprocessing of the data is performed using a local LabView to find the minimum of S_{11}. The extracted feature is compared with the commercial sensor response and fitted with a line according to its linear behavior.

The sensor is used in successive patterns of increasing ISF GL concentrations (from 0 to 250, 275, and 500 mg/dL) for repeatability measurements. This pattern is called a 'cycle' with 90 data points at 5-s measurement intervals. The entire data set for this test is made up of repeating this cycle 164 times, resulting in a total of 14,760 data points. The prediction model, here LSTM, is trained only on normal data to learn normal patterns, and it is optimized for prediction accuracy. The dataset is divided into four subsets: 114 cycles for the training set (only normal values), eight cycles for the VN set (validation set without anomalies), 18 cycles for the VA (validation set that includes anomalies, 7 abnormal cycles augmented with 11 normal cycles), and 24 cycles for the test set (6 abnormal cycles augmented with 18 normal cycles). Several cycles are programmed differently to create anomalies in cycles: 65, 71, 73, 76, 79, 82, 147 in the VA set (with anomalies), 151, 153, 155, 158, 161, 164 in the test set.

For human trials, sensor data are recorded every 5 seconds for four subjects, one non-diabetic (female – F – age 35), one non-diabetic (male – M – age 30), and two diabetic patients (M – ages 34 and 28). Volunteers are asked to stay indoors and sit still in a chair with restricted and limited movement of their hands. Their fingers are securely attached to the sensor with proper tapping to ensure that volunteers have a comfortable time during measurement. Each subject undergoes several intakes of food/chocolate during measurement at random times. These recordings are recorded for eight days, three hours a day. Since removing the finger and placing it back over days affects the sensor response's starting point, they are calibrated each day according to the commercial sensor so that all S_{11} recordings correspond to a certain commercially measured GL match across all days. Thus, the total recorded data is 16000 data points for a single person. Because the very dataset is used for GL prediction and anomaly detection, it is divided into known sections for training/validation/testing, since training needs to be conducted on normal time series. In this case, for each person, 10000 data points are chosen for the training set, 1000 for VN, 2,000 for VA, and 3000 for the testing subset.

For both experiments, the training and VN data sets are transformed to a Gaussian distribution with zero mean and unit variance to standardize the data. A common oversight in this step could be to normalize the entire dataset, whereas the test set must remain intact during modeling. The statistics of the training set (mean and standard deviation) are then used to normalize the VA and testing sets. While transforming the data sets into the format required by the algorithm, each input sample is a roll of look-back number of timesteps to predict the desired future steps. This becomes a tensor (here a 3D

array) with the dimension of (samples, look-back, number of features = 1).

5.2.3 LSTM prediction methods

In this work, an encoder/decoder LSTM is used for the prediction model that takes the most recent p values, also known as lookback, and outputs q future values, also known as lookahead. The network consists of hidden recurrent layers followed by an output later. The number of hidden recurrent layers and the number of neurons in each layer are tuned using gradient descent optimization. Dropout is used to avoid overfitting between two consecutive layers, and the output layer is a fully connected dense layer with the number of neurons equal to lookahead, i.e. one neuron for each predicted future value. The activation function in the output layer is chosen to be linear for regression purposes, and the MSE is set as the loss function. With a look-ahead of q at time t, the model predicts the next q samples at $t+1, t+2, \ldots, t+q$. The algorithm contains the following steps: The training set is used to train the prediction model. The best values for the model parameters, such as lookback, dropout, and learning rate, are extracted from Bayesian optimization. To avoid overfitting, VN is used for early stopping. Next, anomaly detection is enabled by exploiting the prediction errors, the difference between the prediction made at time $t-1$ and the input at time t, as anomaly indicators. These errors are then modeled using a Gaussian function with mean and variance computed from MLE. Then, anomaly scores are defined as the log probability density (PD) of the error vector, such that lower values indicate a higher likelihood of the observation being an anomaly. To set a threshold for the log PD values, the validation set VA is used which includes both normal and abnormal data, to reduce the likelihood of false positives. The two main assumptions used here are as follows:

- The prediction model is trained on the normal data set with no anomalies involved.
- The prediction errors follow the Gaussian distribution.

The lookahead of 150 is used to predict the next 30 minutes. In case of no interest in the future time steps, one can gain high accuracy with a lookahead of 1. It should be noted that the model parameters are optimized to minimize the prediction MSE and are expected to exhibit higher errors when dealing with anomalies present in the VA and test sets. However, there is a trade-off between the accuracy of long-term prediction and that of anomaly detection. Therefore, the predictions-optimized parameters can be used as starting points for anomaly detection, with further manual tuning. A low MSE is required for the training set and a high MSE for the VA and the testing set since the latter contain anomalies. The cycles can be inspected from two perspectives. One with a longer time scale, and the other comprising three consecutive bumps with smaller time steps. If the LSTM states are wiped out at each batch or do not include information about the whole pattern period, then high prediction accuracy pushes the model to treat the anomalous data as normal ones; hence, this results in poor anomaly detection. Training the model is unsupervised with respect to the anomalies and there is no feedback to avoid fitting them. One solution to reduce false positives is to manually tune model parameters, e.g. lookback, number of

recurrent layers, and number of neurons in each layer. Since the period of peaks is known to be 90 timesteps, with a slight change in the length of each period, the window size (lookback) is set to 90 to hold the states. Another method of keeping the states throughout an epoch is to use a stateful LSTM mode with equivalent performance [225].

5.2.4 Selectivity analysis

To evaluate the selectivity of the sensor with respect to materials other than glucose (Product #G7021, Sigma-Aldrich, USA), several interfering items are added to the solution individually. The additional ingredients used in this analysis are fructose (100 mg/dL) (Product #PHR1002, Sigma-Aldrich, USA), caffeine (20 mg/dL)(Product #C0750, Sigma-Aldrich, USA), acetaminophen (15 mcg/mL = 1.51 mg/dL) (Product #A7085, Sigma-Aldrich, USA), galactose (5 mg/dL) (Product #G5388, Sigma-Aldrich, USA). ISF solutions with these contents are prepared and then glucose is added to the solution at a concentration of up to 100 mg/dL.

5.3 Results and Discussion

Initial sensor measurement includes bulk-medium sensing to evaluate the sensitivity at a fairly wide range of permittivity values. Figure 5.3 (a) depicts the reflection parameter for the empty channel (air) as well as isopropyl alcohol (IPA), ethanol, methanol, acetone, and water in magnitude (left) and phase (right). Water is the main ingredient of interstitial fluid (ISF) that is a medium to mimic the subcutaneous fluid in the human body, which is used in our verification measurements. The sensor matching is engineered for the μF channel so that a highly lossy medium such as water can also exhibit deep resonances down to -23 dB. The dielectric constant of water at low UHF frequencies is 76, which drops down to 61 due to the added glucose content (see Fig. 5.3 (b)). In fact, a portion of high permittivity deionized water is replaced with glucose; the chemical reaction of these two substances also leads to a reduction in the permittivity. Moreover, water molecules interact with the hydroxyl groups on glucose to allow the formation of an aqueous solution, which increases the effective loss factor of the solution. These changes are captured in each concentration's equivalent Debye model (more information is discussed in the chapter. A). Various glucose concentrations in ISF (detailed ingredients are given in the method section), including 100 mg/dL up to 400 mg/dL, are tested as shown in Fig. 5.3 (c). The resonant frequency shifts up slightly due to the permittivity reduction as the MUT loading reduces, while the amplitude of resonance change depends on the loading level from the MUT compared with the optimum load (matched case). In the case of lower loading impact from MUT, frequency, and amplitude variation trends are similar, whereas for higher MUT loading this trend becomes reversed. In this experiment, the amplitude increases as the solution becomes more lossy at higher concentrations (with loading impact less than matched load). Phase information also shows a reduction in the slope of the measured S_{11} values, indicating a higher loss content in the solution.

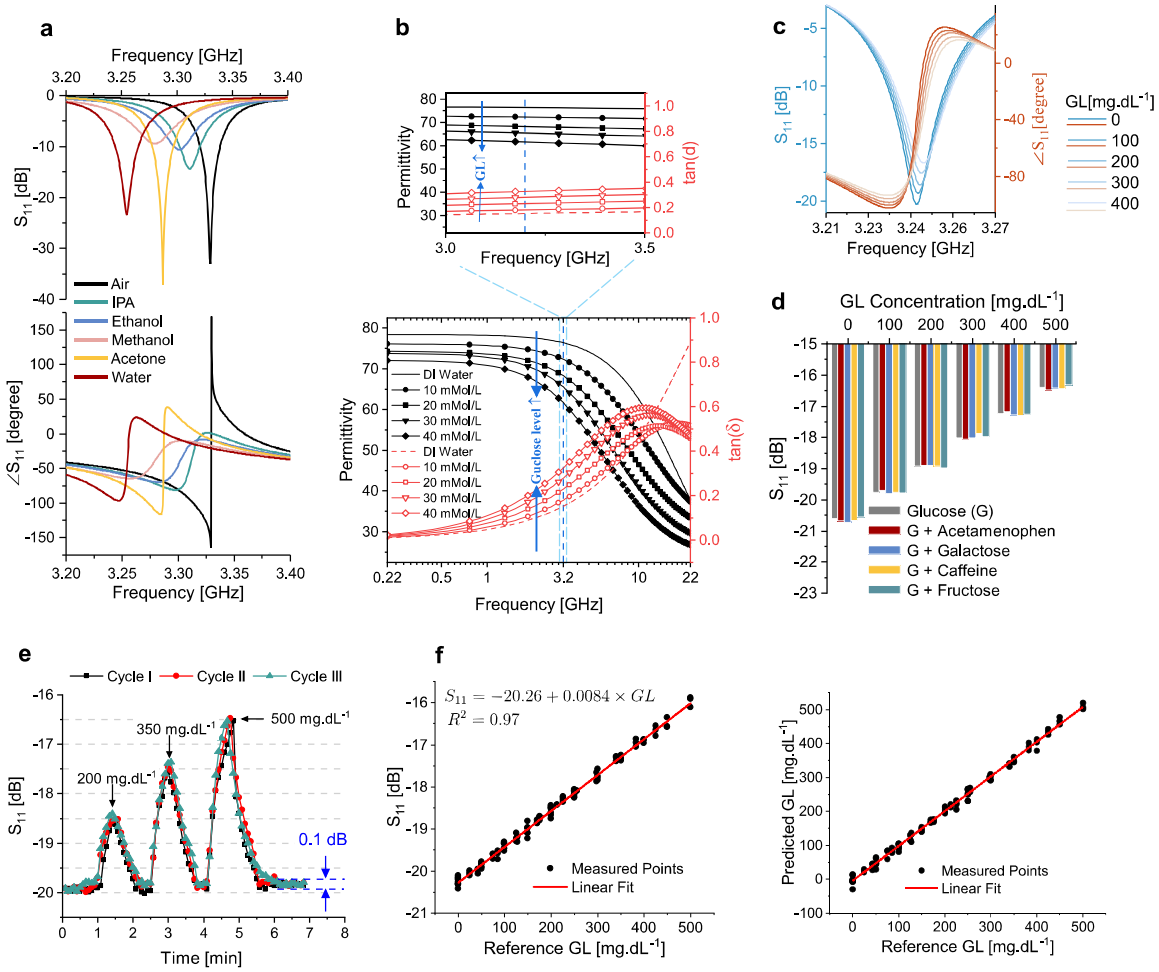


Figure 5.3: (a) S_{11} simulation result when the μF channel on the sensor is filled with common chemicals (IPA, ethanol, methanol, acetone, water) with magnitude (top) and phase (bottom). (b) Dielectric properties of glucose in water solution for various concentrations of 0, 10, 20, 30, 40 mMol/L obtained from the Debye model. The inset represents detailed information in the frequency span of interest (3-3.5 GHz). (c) Measured results for various concentrations of glucose in ISF solution with magnitude (left) and phase (right) of S_{11}. Individual samples are prepared and shaken well for homogeneity and injected separately. The experiment is conducted when the materials and sensor are in equilibrium with the room temperature. (d) Selectivity analysis of the sensor with respect to interfering ingredients. ISF solutions with different glucose content (0, 100, 200, 300, 400, 500 mg/dL) are mixed with interfering analytes including Acetaminophen, Galactose, Caffeine, and Fructose. (e) Sensor reliability test. Two syringe pumps, one with highly concentrated glucose-ISF solution and another with only ISF solution, are employed to create different concentrations continuously. A pattern is configured to transit smoothly between the states of 0 $\rightarrow$200$\rightarrow$0$\rightarrow$350$\rightarrow$0$\rightarrow$500$\rightarrow$0 mg/dL. This cycle is repeated two more times and the depth of S_{11} is measured at the frequency of resonance. (f) The sensor response is shown on the left, which is linearly fit with the reference GL. Estimation of GL with respect to measured S_{11} depth is shown on the right according to the linear fit for three cycles (right).

The selectivity of the proposed sensor is examined with respect to other similar ingredients, such as fructose and galactose. Also, with respect to the common lifestyle, acetaminophen, and caffeine are also considered to be a possible part of human consumption. As a result, the relevant concentrations of each ingredient are individually combined with glucose levels of 0 to 500 mg/dL (covering the range from hypoglycemia to hyperglycemia) with increments of 100 mg/dL. Figure 5.3 (d) shows the results of the μF channel experiment for this selectivity analysis. A slight variation is observed with respect to the level of extra interfering substances, while the main contribution is from the glucose itself. As a result, the proposed sensing technique is expected to respond selectively to glucose variations rather than to other

interferents in the complex matrix inside the body.

5.3.1 ISF tests with repeatability analysis

The ISF solution is used for the next experiments (Fig. A.6 (a)). The pH of the sample is modified to 7.2, close to the normal human level.

Table 5.1: Ingredients of ISF (pH = 7.2) expressed as concentrations in [mM/L].

Material	Conc.	Material	Conc.
HEPES	5	$NaHCO_3$	25
$MgCl_2$	1.0	$CaCl_2$	1.5
KCl	5.0	NaCl	95
Na_2HPO_4	1.2	NaOH	17.6
Na^+	140	Cl^-	105
HCO_3^-	25	K^+	5
Ca^{2+}	1.5	Mg^{2+}	1.0
HPO_4^{2+}	1.2		

Two pumps are engaged to produce the desired glucose concentrations in the ISF solution to feed the μF channel (Fig. A.6 (f)). To evaluate the sensitivity of the proposed sensor to glucose content, the whole range of the human body GL is covered. Fasting GL values are typically lower than 100 mg/dL for non-diabetic persons and greater than 126 mg/dL for diabetic patients. Concentrations of 0 to 500 mg/dL in 3 successive peaks (peak 1: 0-200-0, peak 2: 0-350-0, peak 3: 0-500-0) are created, and this pattern is repeated three times to ensure the stability and repeatability of the proposed test setup. It is clear from Fig. 5.3 (e) that there is a high coherence between similar peaks. The combination ratio between two pump injections creates the expected reference glucose content and is verified with a commercial glucometer. The recorded S_{11} values are compared with the reference values as shown in Fig. 5.3 (f) (left). Since the variation in glucose content is small, the senor operates linearly throughout the entire range of 0-500 mg/dL, with $R^2 > 0.97$ according to the following expression:

$$S_{11} \quad [dB] = -20.26 + 0.0084GL \quad [mg.dL^{-1}] \tag{5.1}$$

Therefore, mapping between the sensor response (in this case only the amplitude of resonance) and GL is possible using a linear expression. As a result, less computation is required. Data gathering for such a linear sensor response is faster, and thus better suited for commercial purposes. This hypothetical linear model between the sensor response and the expected GL is used to map all measured $|S_{11}|$ values to GL, resulting in predicted values as shown in Fig. 5.3 (b) (right). In a more complex environment, other sensor parameters such as resonance frequency, quality factor, and resonance bandwidth can be used. This would be relevant, especially when external parameters such as temperature, pressure, etc., adversely affected the sensor response. A concerning factor for microwave sensors is the safety limit of their wave propagation in the proximity of human tissues. This is typically measured with the specific

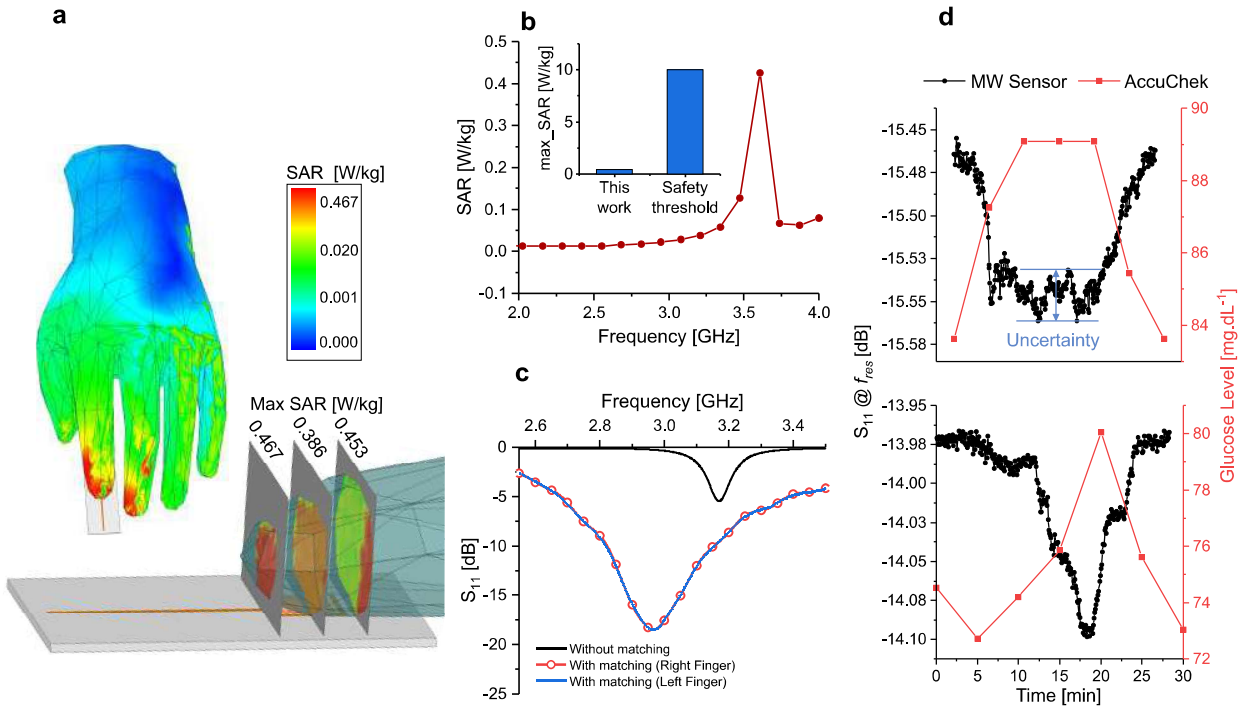


Figure 5.4: (a) SAR simulation results using HFSS on a human hand model that is positioned on the sensor. The inset represents the cross-sectional SAR plot on the index finger. (b) Maximum SAR values of the hand model are computed over a wide frequency range. The safety threshold of 10 W/kg is shown in the inset for comparison. (c) Initial verification of the sensor with a finger. The black curve shows measured S_{11} when the finger is placed on the sensor that was originally designed for μF channel. This experiment is repeated with a redesigned sensor while matching it with human skin dielectric properties. Measurement results for both right- and left-hand index fingers are shown in red/blue curves, respectively. (d) Calibration of the sensor with a human finger for two different cases with respect to commercial glucometer AccuChek. A non-diabetic volunteer is asked to expose their right-hand finger to the sensor for 30 minutes while ingesting chocolate (top). The sensor response is recorded every 5 seconds using LabView. A commercial glucometer is also used to sample the blood every 5 minutes for GL evaluation. The level of uncertainty in the sensor response for minimally changing GL is shown to be less than $\pm$ 1 mg/dL. This experiment is repeated in another instance (bottom) with a different chocolate intake profile.

absorption ratio (SAR), which indicates the level of absorbed power averaged over 10 grams of tissue. Full-wave simulations reveal that SAR values are high near the fingertip compared to other parts of the hand when the index finger is placed on top of the sensor (see Fig. 5.4 (a)). As shown in Fig. 5.4 (b), a fair margin is obtained for SAR values (with a maximum of 0.467 W/kg) compared to the safety threshold of 10 W/kg over a wide range of frequencies (2-4 GHz). This reassures the safety of the proposed sensor when used as a wearable device.

5.3.2 Sensor calibration for sequence-to-sequence modeling

In the next step, the sensor is exposed to the human body to evaluate its effectiveness in interrogating the blood GL and to confirm its real-time functionality. In this regard, four different persons including two non-diabetics, one type 1 diabetic, and one type 2 diabetic have been chosen for noninvasive monitoring. The sensor is confined within a small size such that it can be used on various locations of the body including, but not limited to, finger, arm, ear lobe, abdomen, etc. In this work, the sensor size is chosen small enough to be fully covered with a single index finger. It should be noted that the sensor has

been redesigned with a slightly modified size optimized for human body loading conditions (Fig. A.5), different from the μF channel loading. This is verified in an initial measurement, as shown in Fig. 5.4 (c). When the sensor for μF channel is used for finger measurement, the response dies out due to the excessive loss of the skin. However, the redesigned sensor exhibits a deeper resonance with a finger placed on it in comparison with a bare sensor. Similar results are obtained when the index fingers of the right and left hand of a person are exposed to the sensor. This confirms that the sensor provides consistent results as long as the positioning of the finger on the sensor surface is controlled and secured. For this purpose, the finger is placed inside a hollow partition inside a holder made of acrylic known as Poly Methyl Methacrylate (Fig. A.6 (d-e)). Minor discrepancies are expected due to the nature of planar design, but this can be avoided with firm wearability and secured configuration. Measured S_{11} values are shown when the right index finger is being monitored for 30 minutes and the sensor response is recorded every 5 seconds using LabView for calibration purposes. The sensor measures the effective change of the skin dielectric constant according to variations in the blood glucose content. The reflection depth of the sensor is compared to the commercial glucometer, Accu-Chek, every 5 minutes. Each person starts in fasting condition and is fed a specific amount of chocolate bar at known times. It is clear from Fig. 5.4 (d) that the trends are inversely proportional, mainly because of the way the sensor is loaded with human skin. Indeed, there is an optimum loading for each resonator that results in the lowest possible reflection (minimum of S_{11}) at matched condition resonance frequency f_0. While with a different loading from the environment, here the skin, the resonance frequency f_{res} might be slightly off to the left or the right side of f_0. In the case of $f_{res} < f_0$, the upshifts in the f_0 due to the reduction of effective permittivity at higher glucose concentrations lead to a lower $|S_{11}|$, which is expressed with inverse proportionality between S_{11} and the GL. In contrast, for the case of $f_{res} > f_0$, $|S_{11}|$ starts to increase in parallel with GL. This is not abnormal in a sensor designed to operate close to its optimum input matching point. To increase the trustworthiness of the sensor, it is also calibrated for all four persons involved (only two are shown for brevity). The GL measurements obtained using the commercially available glucometer are interpolated using cubic B-splines to generate intermediate results to match the number of S_{11} measurements. Once these two sets are compared, it is confirmed that the inverse proportionality follows a linear trend between the $|S_{11}|$ and GL with a high Pearson correlation r = -0.93 for Fig. 5.4 (d) (top) and r = -0.95 for Fig. 5.4 (d) (bottom).

In diabetic patients who have GL higher than 126 mg/dL while fasting (no food consumption for at least 8 hours), it is crucial to keep control of GL to avoid possible complications. These include vomiting, rapid heartbeat, vision concerns, and other symptoms. Patients at increased risk of hyperglycemia should keep their GL low, either through physical activities or by a limited food intake plan. In addition, continuous monitoring of GL in the blood can help prevent most hyperglycemia-related problems well in advance. In this work, we introduce a monitoring system composed of a machine learning algorithm coupled with a flexible microwave sensor. The fingertip is flushed on the planar sensor, which enables deeper interrogation of the skin layers through the wave propagation. The system tracks and predicts

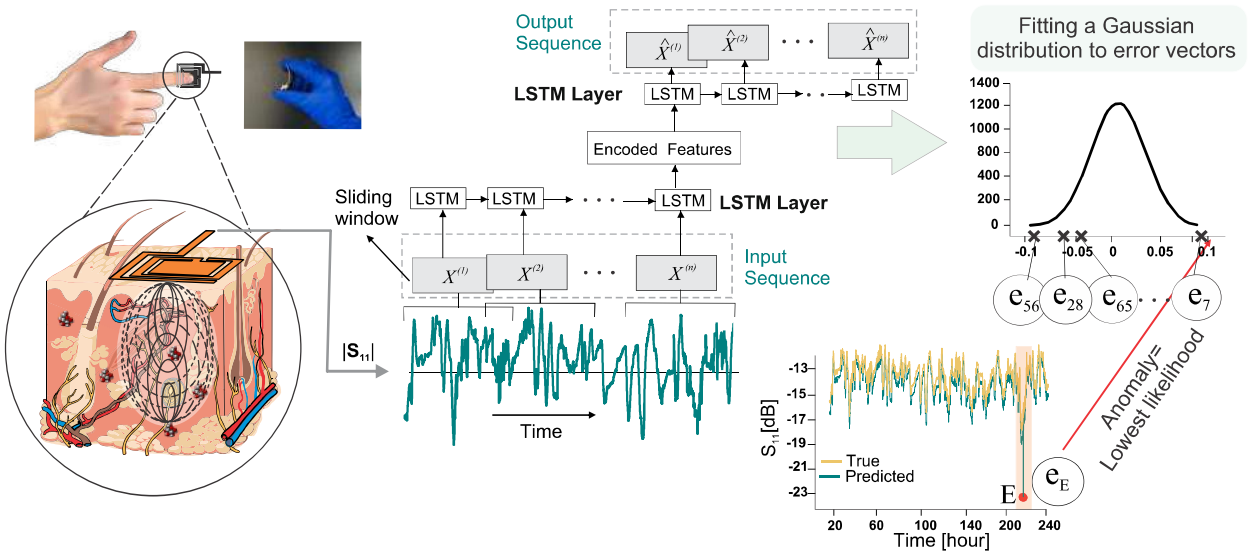


Figure 5.5: An overview of Multistep LSTM system, our predictive learning framework for glucose monitoring. The measured S_{11} amplitudes are combined and used as the input of a deep learning model that predicts GL. Next, the error vectors are modeled by a Gaussian distribution, where the lowest likelihood at the tail of this distribution represents an anomaly that occurred during data recording. Anomalies include excessive pressure on top of the sensor, sudden change in the temperature, data inconsistency, and hand movement; LSTM: long short-term memory network.

GL values and signals when they pass the undesired threshold or are likely to do so within the next 10-20 minutes as shown in Fig. 5.5. To demonstrate the predictive functionality of the proposed system, various concentrations of glucose are generated with the combination of two syringe pumps in a repetitive pattern representing low (max GL = 250 mg/dL), medium (max GL = 375 mg/dL), and high (max GL = 500 mg/dL) peaks of glucose content in the ISF solution as shown in Fig. 5.6 (a). The 3-peak cycles containing 90 points are repeated and measurements are recorded every 5 seconds with LabView. The corresponding $|S_{11}|$ values show normal behavior until certain periods when deliberate anomalies are introduced into the system. This is achieved by modifying the pattern program of the injecting pumps so that some of the regularly expected peaks are removed. In total, 14757 data points are recorded as time series. This data set is constructed with two major goals for the sensor system development: to predict future values and to detect possible anomalies. Predicting future steps provides an early indication of the upcoming behavior, and this can be used as a sign of an anomaly. An anomaly can be an outlier data point, which in this case can be due to a short-time environmental disturbance such as a sudden change in the temperature, humidity, or proximity of an adjacent object. Another common type of anomaly that can be present in time series is a behavior that introduces an unusual phenomenon disrupting the periodic pattern. Long short term memory (LSTM) is a neural network that can be used for time-series analysis. A multistep LSTM is built using the time-series data and deployed to reconstruct a new sample and predict the next step. In this model, a window of a limited number of samples, also known as lookback or timestep, is chosen as one segment of the input sequence to the LSTM network. This window slides over the entire dataset. The encoded features are then fed to the decoder to calculate the predicted output sequence, as shown in Fig. 5.5. The prediction error is calculated as the difference between the

prediction made at t-1 and the input value at time t of the training data. A loss function, formulated as the mean squared error (MSE), is used to evaluate the success rate of the model. The accuracy of the prediction versus lookback number is examined with the MSE loss for the train set and is shown in Fig. 5.6 (b) that is minimized at lookback of 90. The values of MSE loss for training, validation (including anomalies), and test (including anomalies) datasets are, respectively, 0.022, 0.0849, and 0.0514, after 500 epochs. In the case of detecting anomalies, the confusion matrix serves as a comprehensive evaluation tool encompassing four critical categories: 'True Positives' (TP), 'False Positives' (FP), 'False Negatives' (FN), and 'True Negatives' (TN). TP and TN signify correctly identified anomalies and normal data, respectively. In contrast, FP and FN represent inaccuracies where normal data is misinterpreted as an anomaly, and anomalies are overlooked as normal data, respectively. A detailed confusion matrix is shown in Fig. 5.6 (c). The error between predicted and measured $|S_{11}|$ values in the train set is obtained for the well-trained system as shown with histogram of Fig. 5.6 (d). The relationship between the measured and predicted $|S_{11}|$ values of this sequence is described as a linear fitting with $R^2 = 0.974$ (Fig. 5.6 (e)).

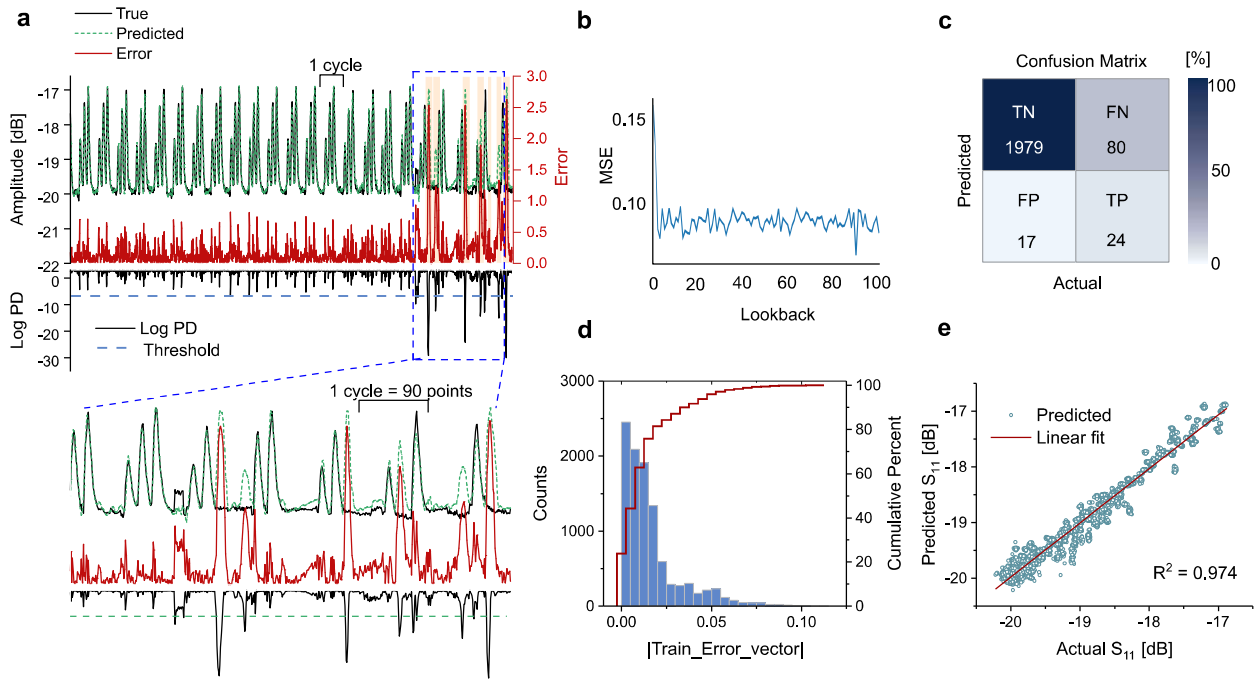


Figure 5.6: (a) The glucose content in the ISF is tested using a μF channel and a programmed pattern of glucose injections. At certain points, the regular pattern is intentionally broken to introduce anomalies. The error between the measured sensor response and the expected (predicted) response is also depicted. Log PD elaborates on the likelihood of a data point being an anomaly. A subsection of this pattern is shown at the bottom with more details on the anomalies. (b) Training error MSE on the repeatability dataset for different numbers of lookbacks. The optimum value for this parameter is achieved at 90 for a single timestep. Increasing the lookback number will not improve the prediction performance of the network. (c) Confusion matrix for anomaly detection in μF channel test. (d) The density distribution of the absolute error vector for the training set in the repeatability dataset is demonstrated as both histogram and cumulative percentage. (e) Predicted S_{11} correlation with the actual S_{11} values. The line shows a linear correlation (R^2) between experimentally measured S_{11} amplitudes and the predicted values using LSTM.

Microwave sensors with high sensitivity are prone to environmental interferents mainly due to their open structure and ease of interaction with the surrounding medium. This vulnerability limits the range

of applications and reduces its effectiveness in dealing with a complex medium. Anomalies introduced by the environment are likely to exacerbate the sensor output precision, hence an anomaly removal unit is crucial to ensure the sensor performance robustness. The anomaly detector uses the increased values of prediction errors as anomaly indicators. The prediction errors obtained using training data are modeled using a Gaussian distribution, with parameters (mean and variance) computed by maximum likelihood estimation (MLE). The errors at the end of this distribution are concluded to be anomalies (see the Appendix. A for more detailed information). On unseen data, the log probability densities (PDs) of errors are computed and used as anomaly scores. The detection of anomalies depends on the level of log PD values such that a lower log PD value (more negative) indicates a more unusual onset. A threshold of log PD is set using the validation dataset, which contains both normal data and anomalies. This threshold helps to separate anomalies from normal data and minimize the number of false positives. The predicted samples from the multistep LSTM are plotted in the top graph of Fig. 5.6 (a), along with the true test set values and several highlighted anomalies. In this case, peaks removed from the primary dataset using programmable syringe pumps are considered anomalies. A more detailed analysis of the μF channel using ISF is given in Fig. A.7.

5.3.3 *In vivo* test for non-invasive monitoring of glucose.

In this study, four volunteers were recruited after having human ethic approval, without discriminatory intent based on sex, age, occupation, race, or ethnicity. All used their right-hand index finger with the sensor while their hand was secured on the table with proper tapping and a fixed holder for the finger. They were asked to keep physical movement at the lowest possible level. GL of each participant was recorded every 5 seconds, while their finger-pricked blood GL was taken at random intervals due to the limited number of test strips available. At random points of measurement, the sensor response was compared against the Accu-Chek level; a high agreement was confirmed throughout the experiment. Furthermore, the participants ingested 45 g of chocolate bar with 30 g of total carbohydrates (22 g sugar content), and their corresponding blood GL was continuously monitored.

The blood GLs of four different individuals with various conditions are continuously monitored. The following subjects are included in the trials: a non-diabetic male, a non-diabetic female, and two diabetic males (type 1 and type 2). The sensor response measured for each person includes 16000 datapoints that are recorded at 5-second intervals for eight days. Each day, 2000 datapoints are recorded per 3-hour sitting. During the concatenation of daily measured data points, if there is a sharp contrast between the GL of the end and start points of adjacent days, it is considered as an anomaly. To verify that the measured data agree with the real GL, an AccuChek glucometer is used as a commercial gauge to measure the GL every 10 minutes. This frequency of measurements with AccuChek is high enough to capture all variations consider limited number of test strips. This verification is carried out for a few datapoints due to the limited number of test strips. This fact check confirms a high correlation between the commercial sensor and the proposed sensor response (Fig. A.13 - A.16). The data set shown in Fig. 5.7 (a) consists

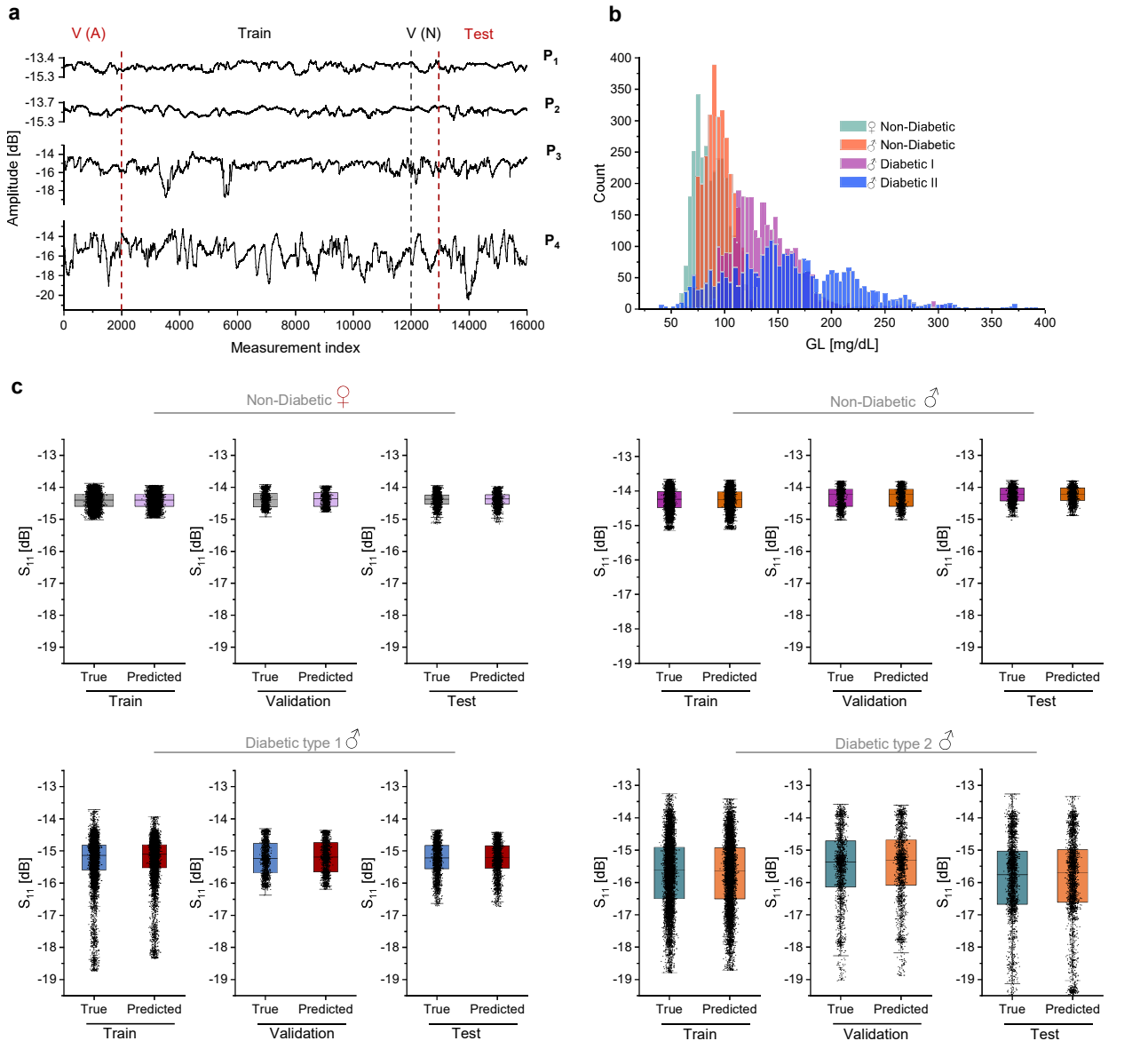


Figure 5.7: (a) The sensor is used to monitor the index fingers of 4 different persons with 5-second recording intervals and measure the depth of S_{11} amplitude values. Participants ($P_1 - P_4$) include one non-diabetic (P1: female – age 35), one non-diabetic (P2: male – age 30), and two diabetic participants (P3,4: male – ages 34 and 28). Each person receives chocolate at random times during the measurement. The whole time series data per person is obtained over eight days, three hours a day. (b) Measured S_{11} values are converted into corresponding GL values according to a linear fitting curve (GL [mg/dL] = -50 (S_{11} [dB] + 12.5)). Histograms of measured GL with the microwave sensor for all participants. (c) Distribution of actual and predicted S_{11} amplitude in three sets of train, validation (VN), and test for every person. In the box plots, the central line defines the median. The lower and upper hinges of the boxes correspond to the 25th and 75th percentiles, and the whiskers are the minimum and maximum values.

of a training set, a testing set, a validation set that includes anomalies VA, and a validation set without anomalies VN. The individual subsets are separated by vertical dashed lines. The measured $|S_{11}|$ values are converted into equivalent GLs according to the linear relationship obtained between measured $|S_{11}|$

and GL measured by the commercial AccuChek as follows:

$$GL\ [mg/dL] = -50\ (S_{11}\ [dB] + 12.5) \tag{5.2}$$

A histogram consisting of GLs computed from Eq. 5.2 for four participants is shown in Fig. 5.7 (b), which demonstrates a meaningful range for each person.

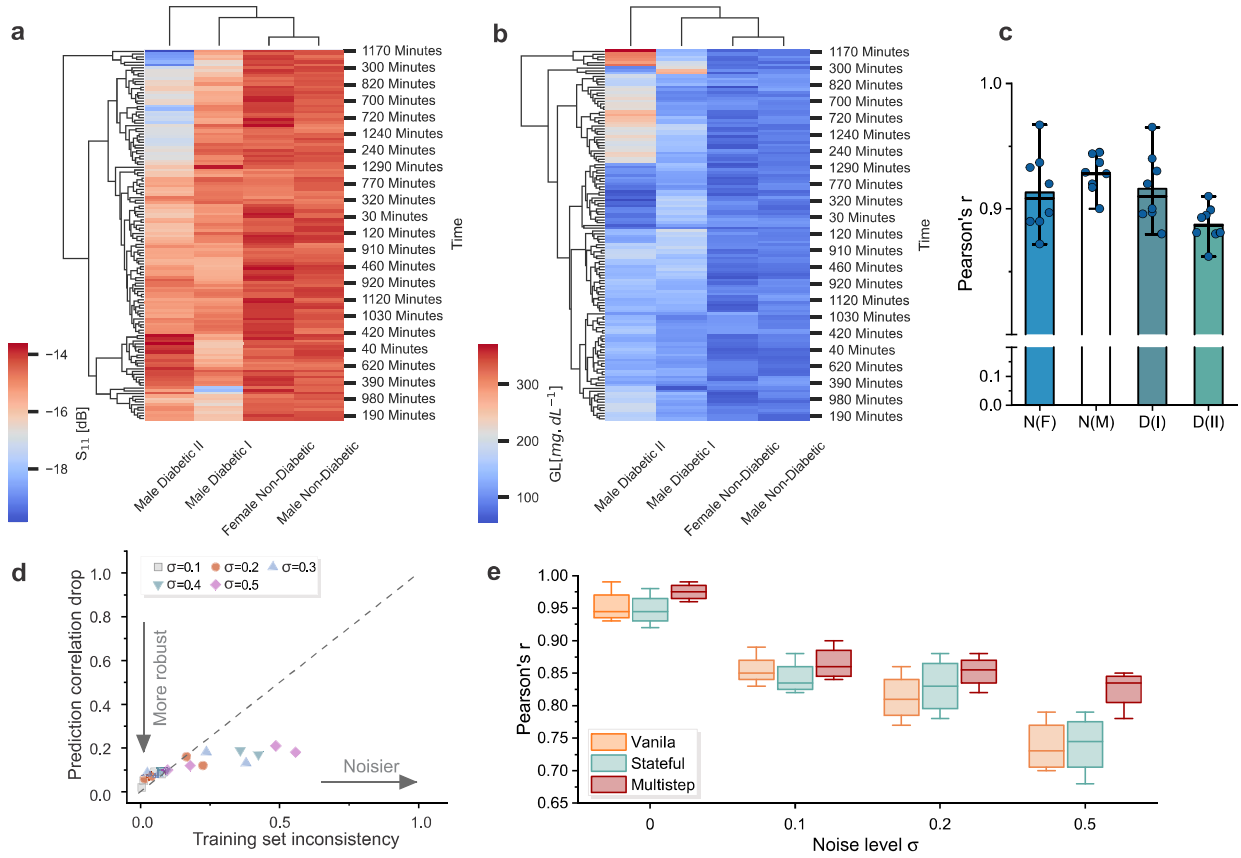


Figure 5.8: (a) Heatmap of all measurement time intervals are clustered by similarity. Both dendrograms are color-coded to represent S_{11} values on the left and corresponding measured GL using a commercial glucometer on the right. The color inversion between these illustrations is because of inverse proportionality between S_{11} value and GL. (b) Microwave sensor measurement results are split into three-hour intervals (Fig. A.8-A.11). Pearson correlation of each window between S_{11} (measured using MW sensor) and GL (measured using AccuChek) is obtained for all four participants (N(F): non-diabetic female, N(M): non-diabetic male, D(I): type 1 diabetic male, and D(II): type 2 diabetic male). (c) The training set related to each participant is perturbed by a random noise with normal distribution N(0, σ) to evaluate the model robustness. Five noisy training sets are simulated using σ = 0.1, 0.2, 0.3, 0.4, and 0.5, and a separate LSTM model is trained on each training set. Then the trained models are evaluated on the intact test sets. Pearson correlation between pre- and post-perturbed values of the training dataset is considered as r and the training set inconsistency is 1-r. A scatter plot shows the performance of the network for every participant at different noise levels. (d) Comparison between Vanilla, Stateful, and Multistep models on a noise-free training set and noisy test datasets at different noise levels with σ = 0.1, 0.2, and 0.5. The midline of the box plots represents the median. The lower and upper hinges of the boxes correspond to the 25th and 75th percentiles, and the whiskers are the minimum and maximum values.

The multistep LSTM is used to model this dataset, and the predicted values are obtained after 100 epochs for each person. The loss of the test set becomes as low as 0.0041, 0.0077, 0.015, and 0.091 for subjects 1-4, respectively. The predicted values of the sensor response are compared with the actual values in Fig. 5.7 (c). LSTM-based predictions for the train, test, and validation sets are summarized with a

box plot against their corresponding actual values. This comparison for all 4 participants reveals a similar range of datapoints that supports the high fidelity of the prediction system. To evaluate the correlation between various instances of a person's GL, the heat map of the measured $|S_{11}|$ and the GL measured by AccuChek is plotted in Fig. 5.8 (a). These heatmaps comprise each person's measured values sampled at 10-minute intervals. As a result, the similarity between the sensor response and the actual GLs is ensured according to the way in which the participants are clustered identically. Furthermore, the sensor-measured response ($|S_{11}|$) and the actual GLs determined by AccuChek are also shown in detail for each 3-hour sitting as given in the Fig. A.13 - A.16 for all four participants. In these figures, a separate mapping is used for each 3-hour window for high accuracy. The Pearson correlation value is then calculated for 8 windows per person and is plotted in Fig. 5.8 (b). The whiskers of the box plot illustrate the minimum and maximum correlations. A high overall correlation is achieved that suggests a high-fidelity sensor performance while being used by different participants provided that a separate calibration fit is used per person.

The proposed sensor is a generic prototype that is designed to be used by a broad range of customers. As a result of various operation environments, the measurement conditions might not be always ideal, which requires the consideration of noise in the training set. The LSTM model used in this work is examined with various noisy training sets. Next, a separate prediction is conducted over the total measured sequence of a participant. Then, the correlation between the actual sensor response and the predicted response according to a noisy training set is computed. The added noise follows a normal distribution with zero mean and different standard deviations of $\sigma = 0.1$, 0.2, 0.3, 0.4, and 0.5. In each case, the training set is correlated with the noise-free training set to calculate the Pearson correlation r, which is used to define the training set inconsistency as $1 - r$. It is shown in Fig. 5.8 (c) that the prediction model is robust against the added noise as the correlation drop is not linear with the data inconsistency. This result is also compared with other well-known LSTM models, namely, Vanila and Stateful. Figure 5.8 (d) represents the comparative Pearson correlation for all four participants using a box plot. It is found that the multistep LSTM outperforms the other methods, which suggests its potential as a highly stable and robust predictive system.

A brief analysis of this modelling is shown in Fig. 5.9 (a), which represents an exemplary test set for a non-diabetic male participant (more details for all participants' results can be found in the Fig. A.13 - A.16). The results show that an accurate prediction is obtained for each person in both upward and downward trends of the body GL. The real-time prediction of the output values from the sensor response indicates that there is no time lag between the measured GL of the body at the frequency of operation. For each subject, deliberate anomalies are introduced during each measurement. Anomalous periods are highlighted, and the plots confirm that the prediction errors also become large during the same intervals. The anomalies are created by several means, including sudden lifting off the finger and returning it to approximately the same position, strong tapping on the table holding the sensor, and temporarily changing the temperature in close proximity to the sensor. A suitable threshold is chosen for log PD

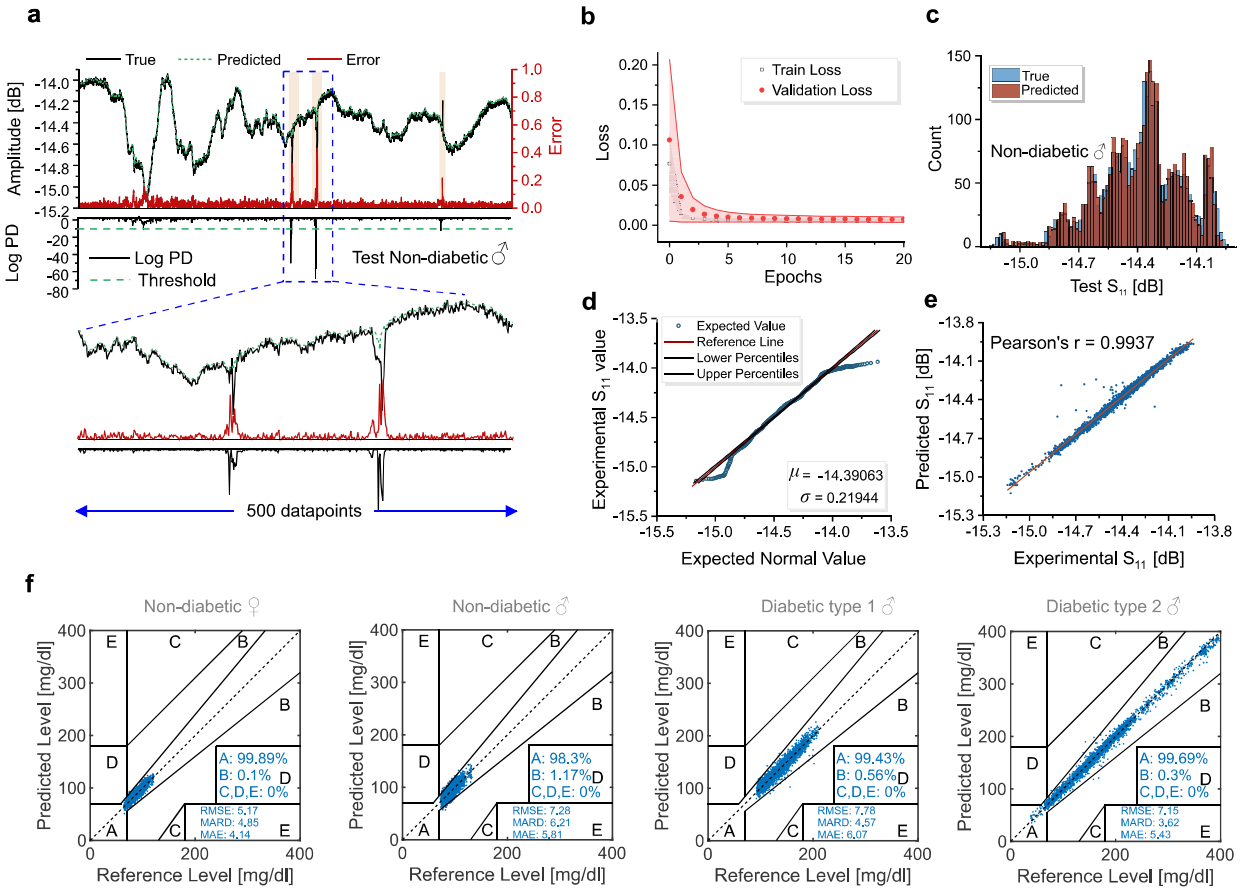


Figure 5.9: (a) Prediction and error vector of a male non-diabetic male participant is depicted with 3000 datapoints as an exemplary test set showing predicted S_{11} values from the LSTM model, error vector values, and Log PD. A more detailed view of the test set is extracted containing only 500 datapoint and is shown below. (b) The LSTM model convergence is depicted using training set- and validation set-loss. Out of total 100 epochs of this modelling, only 20 epochs are shown for presentation clarity. The convergence curves for each participant are used to create the band error plot containing mean values (symbols) and the error band (bounded by minimum and maximum values). (c) Comparison between the measured and predicted S_{11} values in the test set of the non-diabetic male participant is shown. This also showcases the measured values have a left-skewed distribution, which typically depends on the person's health factors. (d) A normal distribution is modelled based on the measured S_{11} values with μ=-14.39 and σ = 0.21944. A Quantile-Quantile plot compares the measured dataset and the modelled normal distribution, which shows deviations closer to the tails indicating the data to be flatter than a normal distribution (platykurtic). (e) Pearson factor between the measured S_{11} values and the predicted value by the LSTM model reveals a high correlation of r = 0.9937. (f) Clarke error grid analysis results are evaluated based on measured S_{11} values converted into GL using a linear fit for each participant. The measured datapoints reside mainly within zone A and marginally in zone B. Error analysis metrics including RMSE, MARD, and MAE are also given in the inset.

values of individual subjects to avoid false positive anomaly detection, which is showcased in detail in the Appendix. A, Fig. A.13 - A.16 The corresponding GL for each measured S_{11} depth is obtained according to the initial calibration of the sensor with respect to each subject, similar to the case shown in Fig. 5.9 (a). The train and validation loss for each person converge at 100 epochs; however, only 20 epochs are shown for brevity in Fig. 5.9 (b). The mean loss values from four persons are shown with symbols and the boundary is created with respect to the maximum and minimum loss values among all participants (see Fig. A.18 for train error analysis). The histogram of the measured sensor response ($|S_{11}|$) in the test set is compared with its predicted $|S_{11}|$ in Fig. 5.9 (c) for an exemplary case of the non-diabetic male participant that showcases the predicted values follow the same distribution as the actual ones. Detailed

histograms of the test/train/validation sets of all participants are given in Fig. A.19. The left-skewed distribution of this test set deviates slightly from a normal distribution. The measured $|S_{11}|$ is modelled with a normal distribution with respect to the test set's mean and standard deviation. Figure 5.9 (d) shows that the measured sensor response deviates from the normal distribution close to the tails, which also corroborates with the result of the histogram. The Pearson correlation r is also used to elaborate the similarity between the measured sensor response and the predicted values from the LSTM model in Fig. 5.9 (e), where a high r = 0.9937 reveals a good agreement between predicted values and the real ones. To elaborate on the repeatability of the proposed technique, the Clarke error grid (CEG) analysis for all four subjects is shown in Fig. 5.9 (f). It is clear from the inset information of the graphs that the estimated GLs fall 100 % within clinically acceptable zones. This can be broken down as A: 99.89 %, B: 0.1 % for non-diabetic person (F), A: 98.3 %, B: 1.17 % for non-diabetic person (M), A: 99.43 %, B: 0.56 % for type 1 Diabetic (M), and A: 99.69 %, B: 0.3 % for type 2 Diabetic (M). The parameters of interest in evaluating the performance of the proposed sensor are the root mean square error (RMSE), the mean absolute relative difference (MARD), and the mean absolute error (MAE). These measures are provided with detailed information in the inset of Fig. 5.9 (f). These percentages, obtained for a total of 3,000 test data points, show a good match between the measured sensor response and the reference GL.

Implementation of this sensor is based on single-frequency operation due to the high sensitivity of the depth of S_{11}. This reduces the complexity of using bulky VNAs into a simpler readout circuit similar to a low-frequency radar board [150] or frequency demodulation scheme [226]. In such a portable configuration, only single frequency generation is required with power as low as -10 dBm (100 μW) for reflection measurement purposes. Regarding deployment of LSTM and the anomaly detection algorithm, predictions typically occur within milliseconds, once the system has been trained. This is sufficiently fast for predicting future events in the next 30-60 minutes.

5.4 Conclusion

This research has demonstrated the potential of a non-invasive, continuous, low-energy microwave glucose sensor with advanced machine learning capabilities. The study has shown that the sensor, when combined with an LSTM model, can predict glucose levels in human subjects. The LSTM model was trained using a dataset transformed into a Gaussian distribution with zero mean and unit variance, which was crucial for standardizing the data and ensuring the accuracy of the predictions.

The study also highlighted the importance of anomaly detection in ensuring the robustness of the sensor's performance. Anomalies were detected by identifying increased values of prediction errors, which were then removed to improve the sensor's precision.

The results of the in-human testing were conducted in highly controlled conditions, with the estimated glucose levels falling 100% within clinically acceptable zones for all four subjects. The performance of the proposed sensor was evaluated using parameters such as the root mean square error (RMSE), the

mean absolute relative difference (MARD), and the mean absolute error (MAE), all of which showed satisfactory results.

However, it is important to note that while these results are encouraging, further research is needed to optimize the sensor design and the machine learning algorithms used. Future work should focus on improving the sensitivity and selectivity of the sensor, as well as exploring other potential applications of this technology.

In conclusion, this study has made a significant contribution to the field of non-invasive glucose monitoring. It has demonstrated that, with the appropriate enhancements and processing techniques in a highly controlled environment, a low-energy microwave sensor can be employed for applications requiring continuous monitoring.

Chapter 6

Conclusion and Future work

The first chapter of this thesis provides a comprehensive background on various methods of microwave sensors, with a particular focus on different types of planar microwave sensors. These sensors, including resonant and non-resonant types, have been thoroughly discussed in terms of their principles of operation, advantages, and disadvantages. The resonant sensors, for instance, offer high sensitivity and selectivity, making them suitable for applications that require precise material characterization. On the other hand, non-resonant sensors provide broader bandwidth and are less affected by environmental changes, making them more versatile for different applications. The chapter provides a detailed comparison of these sensor types, setting the stage for the subsequent exploration of machine learning applications in microwave sensors.

Following the exploration of planar microwave sensors, the second chapter delves into the application of various machine learning algorithms in engineering. It discusses both supervised and unsupervised learning methods, each with its unique advantages and applications. Supervised learning, with its ability to learn from labeled data and make predictions, has found extensive use in applications that require precise outcomes. Unsupervised learning, on the other hand, excels in discovering hidden patterns and structures from unlabeled data, making it ideal for exploratory and data-driven applications.

The chapter extends the discussion to deep learning, a subset of machine learning that mimics the neural networks of the human brain to interpret data patterns. Deep learning algorithms, with their ability to learn from large volumes of data, have revolutionized various fields, including image recognition, natural language processing, and even sensor technology.

The chapter also touches upon high-end algorithms in the literature, such as quantum optimization. These advanced algorithms have the potential to solve complex problems that are currently beyond the reach of traditional algorithms. They represent the cutting edge of machine learning and hold great promise for future applications.

In the context of microwave sensors, these machine learning algorithms, from basic supervised and unsupervised learning methods to advanced deep learning and quantum optimization algorithms, have been leveraged to boost sensor performance. State-of-the-art works that have successfully used these

algorithms to enhance microwave sensors have been provided, showcasing the potential of machine learning to revolutionize sensor technology. The chapter underscores the potential of machine learning to revolutionize sensor technology and sets the stage for the specific projects discussed in the subsequent chapters.

The third section of the thesis, the CycleGAN project, demonstrates the practical application of machine learning in enhancing sensor responses. The CycleGAN model was used to convert passive sensor responses into their active counterparts, thereby enhancing the quality factor of the sensor response. This technique was applied to a noninvasive glucose-sensing scenario, where the passive sensor response was amplified using CycleGAN. The resolution of the CycleGAN-boosted response approached that of the active sensor, showing a considerable enhancement when compared to the resolution of the passive sensor. The Clarke error grid (CEG) analysis for all four subjects showed that the estimated GLs fall 100 % within clinically acceptable zones. This project showcases the potential of machine learning to transform sensor technology, providing a practical demonstration of the concepts discussed in the previous chapters.

The fourth chapter presents an LSTM-based analysis of the sensor response. This algorithm has been instrumental in enabling continuous analysis of sensor responses, predicting future values, and importantly, identifying and removing undesired anomalies. This has significantly improved the reliability and accuracy of the sensor readings, paving the way for more precise material characterization. The LSTM model was trained and tested on various noisy training sets, demonstrating robustness against added noise. The multistep LSTM model outperformed other LSTM models, suggesting its potential as a highly stable and robust predictive system. The system was used to monitor the index fingers of different individuals, including non-diabetic and diabetic participants. The sensor response and the actual glucose levels were found to be highly correlated, indicating high-fidelity sensor performance. A Pearson correlation of $\rho = 0.9937$ reveals a good agreement between predicted values and the real ones.

In cases of low-resolution sensors, the thesis proposes a novel approach that combines the CycleGAN and LSTM methods. By first using CycleGAN to enhance the resolution and then applying LSTM to remove anomalies and forecast changes in material concentration, such as glucose (in highly controlled conditions), the performance of low-resolution sensors can be significantly improved.

In conclusion, the proposed system provides an approach for continuously monitoring and predicting glucose levels in highly controlled conditions, offering early indications of potential anomalies. The integration of LSTM and CycleGAN models significantly enhances the performance and resolution of the sensor, making it a viable tool for sensitive measurements and high-end applications. This could be particularly beneficial for patients at increased risk of hyperglycemia, helping to prevent most hyperglycemia-related problems well in advance. However, further research and testing may be needed to refine the system and validate its effectiveness in real-world settings.

The experiments conducted in vivo were carried out under highly controlled conditions. Participants were instructed to maintain minimal movement and stillness throughout the duration of the experiments. Furthermore, the ambient temperature/humidity were carefully maintained at a constant level to ensure

the stability of the measurements. To eliminate any unwanted interference, a radiofrequency absorber was strategically placed around our setup, effectively minimizing external disruptions. These meticulous measures were taken to guarantee the precision and reliability of the data collected during the experiments, allowing for a thorough and accurate analysis of the microwave sensor's performance in real-world scenarios.

In light of these controlled experiments, it becomes evident that, at their current stage of development, microwave sensors may not yet be suitable for wearable applications, especially in the context of continuous glucose sensing in dynamic human or animal environments. The challenges associated with maintaining minimal movement and consistent ambient conditions underscore the limitations of this technology for practical, real-world usage. While microwave sensors hold promise, further advancements and innovations are likely needed to make them viable for such demanding applications.

This thesis has made significant strides in the field of material characterization. It has demonstrated that with the right enhancements and processing techniques, microwave planar sensors can be a viable and effective tool for these applications. The findings of this thesis have the potential to make a substantial impact on various industries, particularly by leveraging machine learning to advance sensor technology.

6.1 Future Work

In the field of microwave sensors, there is still much to explore in terms of incorporating machine-learning techniques. Some potential areas of future work include:

Advanced signal processing: Machine learning algorithms can be applied to enhance the signal processing capabilities of microwave sensors. This can improve the accuracy and reliability of measurements, as well as the ability to distinguish between different materials.

Multimodal sensing: By combining data from multiple sensors, such as microwave and thermal sensors, machine learning algorithms can be used to provide a more comprehensive understanding of the properties of materials. This could lead to new applications in fields such as biomedical engineering and environmental monitoring.

Real-time control and optimization: Machine learning algorithms can be used to develop real-time control and optimization (Quantum Optimization) systems for microwave sensors. This can enable the sensors to adapt to changing conditions and optimize their performance for different materials and applications.

Deep learning for feature extraction: Deep learning techniques, such as convolutional neural networks, can be used to automatically extract relevant features from microwave sensor data. This can reduce the need for manual feature engineering and improve the accuracy and efficiency of the sensing system.

Robustness and reliability: Machine learning algorithms can be used to improve the robustness and reliability of microwave sensors. This includes developing models that can handle noisy or incomplete data, as well as identifying and mitigating sources of error and uncertainty.

Overall, the integration of machine learning techniques into microwave sensing has the potential to unlock new applications and improve the performance and accuracy of these sensors. As the field continues to develop, we can expect to see new breakthroughs and applications that transform how we measure and analyze materials.

Bibliography

[1] N. Kazemi and P. Musilek, "Enhancing microwave sensor performance with ultrahigh q features using cyclegan," *IEEE Transactions on Microwave Theory and Techniques*, vol. 70, no. 12, pp. 5369–5382, 2022.

[2] N. Kazemi, M. Abdolrazzaghi, P. Light, and P. Musilek, "In–human testing of a non-invasive continuous low–energy microwave glucose sensor with advanced machine learning capabilities," *Biosensors and Bioelectronics*, p. 115668, 2023.

[3] D. Kiela, M. Bartolo, Y. Nie, D. Kaushik, A. Geiger, Z. Wu, B. Vidgen, G. Prasad, A. Singh, P. Ringshia *et al.*, "Dynabench: Rethinking benchmarking in nlp," *arXiv preprint arXiv:2104.14337*, 2021.

[4] M. Abdolrazzaghi, N. Katchinskiy, A. Elezzabi, P. E. Light, and M. Daneshmand, "Noninvasive glucose sensing in aqueous solutions using an active split-ring resonator," *IEEE Sensors Journal*, 2021.

[5] M. Baghelani, Z. Abbasi, M. Daneshmand, and P. E. Light, "Non-invasive continuous-time glucose monitoring system using a chipless printable sensor based on split ring microwave resonators," *Scientific Reports*, vol. 10, no. 1, pp. 1–15, 2020.

[6] T. Athauda, P. C. Banerjee, and N. C. Karmakar, "Microwave characterization of chitosan hydrogel and its use as a wireless ph sensor in smart packaging applications," *IEEE Sensors Journal*, vol. 20, no. 16, pp. 8990–8996, 2020.

[7] M. Abdolrazzaghi, M. Daneshmand, and A. K. Iyer, "Strongly enhanced sensitivity in planar microwave sensors based on metamaterial coupling," *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 4, pp. 1843–1855, 2018.

[8] N. Kazemi, M. Abdolrazzaghi, P. Musilek, and M. Daneshmand, "A temperature-compensated high-resolution microwave sensor using artificial neural network," *IEEE Microwave and Wireless Components Letters*, vol. 30, no. 9, pp. 919–922, 2020.

[9] E. L. Chuma, Y. Iano, G. Fontgalland, L. L. B. Roger, and H. Loschi, "Pcb-integrated non-destructive microwave sensor for liquid dielectric spectroscopy based on planar metamaterial resonator," *Sensors and Actuators A: Physical*, vol. 312, p. 112112, 2020.

[10] A. Hedayatipour, S. Aslanzadeh, S. H. Hesari, M. A. Haque, and N. McFarlane, "A wearable cmos impedance to frequency sensing system for non-invasive impedance measurements," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 5, pp. 1108–1121, 2020.

[11] R. A. Alahnomi, Z. Zakaria, Z. M. Yussof, A. A. Althuwayb, A. Alhegazi, H. Alsariera, and N. A. Rahman, "Review of recent microwave planar resonator-based sensors: Techniques of complex permittivity extraction, applications, open challenges and future research directions," *Sensors*, vol. 21, no. 7, p. 2267, 2021.

[12] E. L. Chuma, Y. Iano, G. Fontgalland, and L. L. B. Roger, "Microwave sensor for liquid dielectric characterization based on metamaterial complementary split ring resonator," *IEEE Sensors Journal*, vol. 18, no. 24, pp. 9978–9983, 2018.

[13] M. Saadat-Safa, V. Nayyeri, M. Khanjarian, M. Soleimani, and O. M. Ramahi, "A csrr-based sensor for full characterization of magneto-dielectric materials," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 2, pp. 806–814, 2019.

[14] N. Kazemi, M. Abdolrazzaghi, and P. Musilek, "Comparative analysis of machine learning techniques for temperature compensation in microwave sensors," *IEEE Transactions on Microwave Theory and Techniques*, 2021.

[15] N. Kazemi, K. Schofield, and P. Musilek, "A high-resolution reflective microwave planar sensor for sensing of vanadium electrolyte," *Sensors*, vol. 21, no. 11, p. 3759, 2021.

[16] K. Schofield, N. Kazemi, and P. Musilek, "Vrf battery characterization using microwave planar complementary split ring resonators," in *2021 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*. IEEE, 2021, pp. 1–6.

[17] M. H. Zarifi, P. Shariaty, M. Abdolrazzaghi, Z. Hashisho, and M. Daneshmand, "Particle size characterization using a high resolution planar resonator sensor in a lossy medium," *Sensors and Actuators B: Chemical*, vol. 234, pp. 332–337, 2016.

[18] S. Mohammadi and M. H. Zarifi, "Differential microwave resonator sensor for real-time monitoring of volatile organic compounds," *IEEE Sensors Journal*, vol. 21, no. 5, pp. 6105–6114, 2020.

[19] K. K. Kazemi, S. Hu, O. Niksan, K. K. Adhikari, N. R. Tanguy, S. Li, M. Arjmand, and M. H. Zarifi, "Low-profile planar antenna sensor based on ti3c2tx mxene membrane for voc and humidity monitoring," *ADVANCED MATERIALS INTERFACES*, vol. 9, no. 13, 2022.

[20] O. Niksan, M. C. Jain, A. Shah, and M. H. Zarifi, "A nonintrusive flow rate sensor based on microwave split-ring resonators and thermal modulation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 70, no. 3, pp. 1954–1963, 2022.

[21] A. Mason, M. Soprani, O. Korostynska, A. Amirthalingam, J. Cullen, M. Muradov, E. N. Carmona, G. Sberveglieri, V. Sberveglieri, and A. Al-Shamma'a, "Real-time microwave, dielectric, and optical sensing of lincomycin and tylosin antibiotics in water: Sensor fusion for environmental safety," *Journal of Sensors*, vol. 2018, 2018.

[22] I. Frau, S. Wylie, P. Byrne, P. Onnis, J. Cullen, A. Mason, and O. Korostynska, "Microwave sensors for in situ monitoring of trace metals in polluted water," *Sensors*, vol. 21, no. 9, p. 3147, 2021.

[23] C. Jang, J.-K. Park, H.-J. Lee, G.-H. Yun, and J.-G. Yook, "Non-invasive fluidic glucose detection based on dual microwave complementary split ring resonators with a switching circuit for environmental effect elimination," *IEEE Sensors Journal*, vol. 20, no. 15, pp. 8520–8527, 2020.

[24] M. Abdolrazzaghi, N. Kazemi, and M. Daneshmand, "Machine learning to immune microwave sensors from temperature impact," in *2020 IEEE International Symposium on Antennas and Propagation and North American Radio Science Meeting*. IEEE, 2020, pp. 843–844.

[25] N. Kazemi, "Machine learning for error estimation and compensation for microwave sensors," 2021.

[26] M. Abdolrazzaghi and M. Daneshmand, "Multifunctional ultrahigh sensitive microwave planar sensor to monitor mechanical motion: Rotation, displacement, and stretch," *Sensors*, vol. 20, no. 4, p. 1184, 2020.

[27] M. Baghelani, Z. Abbasi, and M. Daneshmand, "High-dynamic-range chipless microwave resonator-based strain sensor," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–7, 2021.

[28] N. Kazemi, N. Gholizadeh, and P. Musilek, "Selective microwave zeroth-order resonator sensor aided by machine learning," *Sensors*, vol. 22, no. 14, p. 5362, 2022.

[29] N. Kazemi and P. Musilek, "Resolution enhancement of microwave sensors using super-resolution generative adversarial network," *Expert Systems with Applications*, vol. 213, p. 119252, 2023.

[30] D. M. Pozar, *Microwave engineering*. John wiley & sons, 2011.

[31] J. Baker-Jarvis, M. D. Janezic, J. H. Grosvenor Jr, and R. G. Geyer, "Transmission/reflection and short-circuit line methods for measuring permittivity and permeability," *Nasa Sti/recon Technical Report N*, vol. 93, p. 12084, 1992.

[32] F. Martín, P. Vélez, and M. Gil, "Microwave sensors based on resonant elements," p. 3375, 2020.

[33] A. L. de Paula, M. C. Rezende, and J. J. Barroso, "Modified nicolson-ross-weir (nrw) method to retrieve the constitutive parameters of low-loss materials," in *2011 SBMO/IEEE MTT-S international microwave and optoelectronics conference (IMOC 2011)*. IEEE, 2011, pp. 488–492.

[34] A. N. Vicente, G. M. Dip, and C. Junqueira, "The step by step development of nrw method," in *2011 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC 2011)*. IEEE, 2011, pp. 738–742.

[35] Y. Shi, T. Hao, L. Li, and C.-H. Liang, "An improved nrw method to extract electromagnetic parameters of metamaterials," *Microwave and Optical Technology Letters*, vol. 58, no. 3, pp. 647–652, 2016.

[36] L.-F. Chen, C. K. Ong, C. Neo, V. V. Varadan, and V. K. Varadan, *Microwave electronics: measurement and materials characterization*. John Wiley & Sons, 2004.

[37] E. J. Rothwell, J. L. Frasch, S. M. Ellison, P. Chahal, and R. O. Ouedraogo, "Analysis of the nicolson-ross-weir method for characterizing the electromagnetic properties of engineered materials," *Progress In Electromagnetics Research*, vol. 157, pp. 31–47, 2016.

[38] F. Costa, M. Borgese, M. Degiorgi, and A. Monorchio, "Electromagnetic characterisation of materials by using transmission/reflection (t/r) devices," *Electronics*, vol. 6, no. 4, p. 95, 2017.

[39] A. Nicolson and G. Ross, "Measurement of the intrinsic properties of materials by time-domain techniques," *IEEE Transactions on instrumentation and measurement*, vol. 19, no. 4, pp. 377–382, 1970.

[40] W. B. Weir, "Automatic measurement of complex dielectric constant and permeability at microwave frequencies," *Proceedings of the IEEE*, vol. 62, no. 1, pp. 33–36, 1974.

[41] J. Baker-Jarvis, E. J. Vanzura, and W. A. Kissick, "Improved technique for determining complex permittivity with the transmission/reflection method," *IEEE Transactions on microwave theory and techniques*, vol. 38, no. 8, pp. 1096–1103, 1990.

[42] V. Nguyen, M. Hoang, H. Phan, T. Hoang, and T. Vuong, "Measurement of complex permittivity by rectangular waveguide method with simple specimen preparation," in *2014 International Conference on Advanced Technologies for Communications (ATC 2014)*. IEEE, 2014, pp. 397–400.

[43] L. C. Maier Jr and J. Slater, "Field strength measurements in resonant cavities," *Journal of applied physics*, vol. 23, no. 1, pp. 68–77, 1952.

[44] A. Kraszewski and S. Nelson, "Resonant cavity perturbationsome new applications of an old measuring technique," *Journal of Microwave Power and Electromagnetic Energy*, vol. 31, no. 3, pp. 178–187, 1996.

[45] D. Kajfez and A. Gundavajhala, "Measurement of material properties with a tunable resonant cavity," *Electronics Letters*, vol. 29, no. 22, pp. 1936–1937, 1993.

[46] G. H. Bryant, *Principles of microwave measurements*. IET, 1993, vol. 5.

[47] T. W. Athey, M. A. Stuchly, and S. S. Stuchly, "Measurement of radio frequency permittivity of biological tissues with an open-ended coaxial line: Part i," *IEEE transactions on microwave theory and techniques*, vol. 30, no. 1, pp. 82–86, 1982.

[48] M. A. Stuchly, T. W. Athey, G. M. Samaras, and G. E. Taylor, "Measurement of radio frequency permittivity of biological tissues with an open-ended coaxial line: Part ii-experimental results," *IEEE Transactions on Microwave Theory and Techniques*, vol. 30, no. 1, pp. 87–92, 1982.

[49] D. Berube, F. Ghannouchi, and P. Savard, "A comparative study of four open-ended coaxial probe models for permittivity measurements of lossy dielectric/biological materials at microwave frequencies," *IEEE Transactions on microwave Theory and Techniques*, vol. 44, no. 10, pp. 1928–1934, 1996.

[50] J. Wang, E. G. Lim, M. P. Leach, Z. Wang, and K. L. Man, "Open-ended coaxial cable selection for measurement of liquid dielectric properties via the reflection method," *Mathematical Problems in Engineering*, vol. 2020, pp. 1–8, 2020.

[51] E. C. Burdette, F. L. Cain, and J. Seals, "In vivo probe measurement technique for determining dielectric properties at vhf through microwave frequencies," *IEEE transactions on microwave theory and techniques*, vol. 28, no. 4, pp. 414–427, 1980.

[52] I. S. Seo, W. S. Chin *et al.*, "Characterization of electromagnetic properties of polymeric composite materials with free space method," *Composite Structures*, vol. 66, no. 1-4, pp. 533–542, 2004.

[53] F. J. F. Gonçalves, A. G. Pinto, R. C. Mesquita, E. J. Silva, and A. Brancaccio, "Free-space materials characterization by reflection and transmission measurements using frequency-by-frequency and multi-frequency algorithms," *Electronics*, vol. 7, no. 10, p. 260, 2018.

[54] V. V. Varadan, R. D. Hollinger, D. K. Ghodgaonkar, and V. K. Varadan, "Free-space, broadband measurements of high-temperature, complex dielectric properties at microwave frequencies," *IEEE Transactions on Instrumentation and Measurement*, vol. 40, no. 5, pp. 842–846, 1991.

[55] D. K. Ghodgaonkar, V. V. Varadan, and V. K. Varadan, "A free-space method for measurement of dielectric constants and loss tangents at microwave frequencies," *IEEE Transactions on Instrumentation and measurement*, vol. 38, no. 3, pp. 789–793, 1989.

[56] J. Musil and F. Zacek, "Microwave measurements of complex permittivity by free space methods and their applications," *NASA STI/Recon Technical Report A*, vol. 87, p. 20354, 1986.

[57] D. Ghodgaonkar, V. Varadan, and V. K. Varadan, "Free-space measurement of complex permittivity and complex permeability of magnetic materials at microwave frequencies," *IEEE Transactions on instrumentation and measurement*, vol. 39, no. 2, pp. 387–394, 1990.

[58] J. G. Oliveira, J. G. D. Junior, E. N. Pinto, V. P. S. Neto, and A. G. D'Assunção, "A new planar microwave sensor for building materials complex permittivity characterization," *Sensors*, vol. 20, no. 21, p. 6328, 2020.

[59] G. Facer, D. A. Notterman, and L. Sohn, "Dielectric spectroscopy for bioanalysis: From 40 hz to 26.5 ghz in a microfabricated wave guide," *Applied Physics Letters*, vol. 78, no. 7, pp. 996–998, 2001.

[60] P. Queffelec, P. Gelin, J. Gieraltowski, and J. Loaec, "A microstrip device for the broad band simultaneous measurement of complex permeability and permittivity," *IEEE transactions on magnetics*, vol. 30, no. 2, pp. 224–231, 1994.

[61] A. Raj, W. Holmes, and S. Judah, "Wide bandwidth measurement of complex permittivity of liquids using coplanar lines," *IEEE Transactions on Instrumentation and Measurement*, vol. 50, no. 4, pp. 905–909, 2001.

[62] B. C. Wadell, "Transmission line design handbook," *(No Title)*, 1991.

[63] J. Hinojosa, "S-parameter broadband measurements on-coplanar and fast extraction of the substrate intrinsic properties," *IEEE microwave and wireless components letters*, vol. 11, no. 2, pp. 80–82, 2001.

[64] R. Waldron, "Perturbation theory of resonant cavities," *Proceedings of the IEE-Part C: Monographs*, vol. 107, no. 12, pp. 272–274, 1960.

[65] U. Raveendranath, S. Bijukumar, and K. Mathew, "Broadband coaxial cavity resonator for complex permittivity measurements of liquids," *IEEE Transactions on Instrumentation and Measurement*, vol. 49, no. 6, pp. 1305–1312, 2000.

[66] J. B. Thomas, "Cross-coupling in coaxial cavity filters-a tutorial overview," *IEEE Transactions on microwave theory and techniques*, vol. 51, no. 4, pp. 1368–1376, 2003.

[67] R. Johnson, J. Green, M. Robinson, A. Preece, and R. Clarke, "Resonant open ended coaxial line sensor for measuring complex permittivity," *IEE Proceedings A (Science, Measurement and Technology)*, vol. 139, no. 5, pp. 261–264, 1992.

[68] J. Hirokawa and M. Ando, "Single-layer feed waveguide consisting of posts for plane tem wave excitation in parallel plates," *IEEE Transactions on Antennas and Propagation*, vol. 46, no. 5, pp. 625–630, 1998.

[69] M. Abdolrazzaghi, N. Kazemi, and M. Daneshmand, "An siw oscillator for microfluidic lossy medium characterization," in *2020 IEEE/MTT-S International Microwave Symposium (IMS)*. IEEE, 2020, pp. 221–224.

[70] C. Yang, H. Yu, Y. Shang, and W. Fei, "Characterization of cmos metamaterial transmission line by compact fractional-order equivalent circuit model," *IEEE Transactions on Electron Devices*, vol. 62, no. 9, pp. 3012–3018, 2015.

[71] C. Caloz, "Dual composite right/left-handed (d-crlh) transmission line metamaterial," *IEEE microwave and wireless components letters*, vol. 16, no. 11, pp. 585–587, 2006.

[72] F. Martín, P. Vélez, J. Muñoz-Enano, and L. Su, *Planar Microwave Sensors*. John Wiley & Sons, 2022.

[73] M. Joler, A. N. J. Raj, and J. Bartolić, "A simplified measurement configuration for evaluation of relative permittivity using a microstrip ring resonator with a variational method-based algorithm," *Sensors*, vol. 22, no. 3, p. 928, 2022.

[74] Y.-H. Chou, M.-J. Jeng, Y.-H. Lee, and Y.-G. Jan, "Measurement of rf pcb dielectric properties and losses," *Progress In Electromagnetics Research Letters*, vol. 4, pp. 139–148, 2008.

[75] A. A. Al-Behadili, I. A. Mocanu, T. M. Petrescu, and T. A. Elwi, "Differential microstrip sensor for complex permittivity characterization of organic fluid mixtures," *Sensors*, vol. 21, no. 23, p. 7865, 2021.

[76] S. A. Andevari, J.-L. Olvera-Cervantes, and C. E. Saavedra, "Inclination sensor with a wide angle measurement range using half-wavelength microstrip resonator," *IEEE Access*, vol. 11, pp. 28 699–28 705, 2023.

[77] J. Muñoz-Enano, P. Vélez, M. G. Barba, J. Mata-Contreras, and F. Martín, "Differential-mode to common-mode conversion detector based on rat-race hybrid couplers: Analysis and application to differential sensors and comparators," *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 4, pp. 1312–1325, 2019.

[78] A. Ebrahimi, J. Scott, and K. Ghorbani, "Differential sensors using microstrip lines loaded with two split-ring resonators," *IEEE Sensors Journal*, vol. 18, no. 14, pp. 5786–5793, 2018.

[79] J. Munoz-Enano, P. Velez, M. G. Barba, and F. Martin, "An analytical method to implement high-sensitivity transmission line differential sensors for dielectric constant measurements," *IEEE Sensors Journal*, vol. 20, no. 1, pp. 178–184, 2019.

[80] A. A. Mohd Bahar, Z. Zakaria, A. Isa, Y. Dasril, R. A. Alahnomi *et al.*, "Real time microwave biochemical sensor based on circular siw approach for aqueous dielectric detection," *Scientific reports*, vol. 9, no. 1, pp. 1–12, 2019.

[81] M. Fayaz, M. H. Zarifi, M. Abdolrazzaghi, P. Shariaty, Z. Hashisho, and M. Daneshmand, "A novel technique for determining the adsorption capacity and breakthrough time of adsorbents using a noncontact high-resolution microwave resonator sensor," *Environmental Science & Technology*, vol. 51, no. 1, pp. 427–435, 2017.

[82] M. Abdolrazzaghi, N. Kazemi, V. Nayyeri, and F. Martin, "Ai-assisted ultra-high-sensitivity/resolution active-coupled csrr-based sensor with embedded selectivity," *Sensors*, vol. 23, no. 13, p. 6236, 2023.

[83] M. Abdolrazzaghi and M. Daneshmand, "Exploiting sensitivity enhancement in micro-wave planar sensors using intermodulation products with phase noise analysis," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 12, pp. 4382–4395, 2020.

[84] Z. Abbasi, H. Niazi, M. Abdolrazzaghi, W. Chen, and M. Daneshmand, "Monitoring ph level using high-resolution microwave sensor for mitigation of stress corrosion cracking in steel pipelines," *IEEE Sensors Journal*, vol. 20, no. 13, pp. 7033–7043, 2020.

[85] M. Fayaz, M. J. Lashaki, M. Abdolrazzaghi, M. H. Zarifi, Z. Hashisho, M. Daneshmand, J. E. Anderson, and M. Nichols, "Monitoring the residual capacity of activated carbon in an emission abatement system using a non-contact, high resolution microwave resonator sensor," *Sensors and Actuators B: Chemical*, vol. 282, pp. 218–224, 2019.

[86] M. Abdolrazzaghi, "Advances in active resonator based planar microwave sensors for material characterization," 2017.

[87] M. Abdolrazzaghi, M. H. Zarifi, C. F. Floquet, and M. Daneshmand, "Contactless asphaltene detection using an active planar microwave resonator sensor," *Energy & Fuels*, vol. 31, no. 8, pp. 8784–8791, 2017.

[88] M. Abdolrazzaghi and M. Daneshmand, "Dual active resonator for dispersion coefficient measurement of asphaltene nano-particles," *IEEE Sensors Journal*, vol. 17, no. 22, pp. 7248–7256, 2017.

[89] A. Litvinenko, R. Khymyn, V. Tyberkevych, V. Tikhonov, A. Slavin, and S. Nikitov, "Tunable magnetoacoustic oscillator with low phase noise," *Physical Review Applied*, vol. 15, no. 3, p. 034057, 2021.

[90] M. Abdolrazzaghi and M. Daneshmand, "A phase-noise reduced microwave oscillator sensor with enhanced limit of detection using active filter," *IEEE Microwave and Wireless Components Letters*, vol. 28, no. 9, pp. 837–839, 2018.

[91] M. Brandl and L.-M. Wagner, "Microwave oscillator design for a srr based biosensor platform," in *Proceedings*, vol. 2, no. 13. MDPI, 2018, p. 865.

[92] K. Finkenzeller, "Rfid handbook: radio-frequency identification fundamentals and applications," *(No Title)*, 1999.

[93] V. D. Hunt, A. Puglia, and M. Puglia, *RFID: a guide to radio frequency identification*. John Wiley & Sons, 2007.

[94] K. Finkenzeller, *RFID handbook: fundamentals and applications in contactless smart cards, radio frequency identification and near-field communication*. John wiley & sons, 2010.

[95] M. Chen and S. Chen, *RFID technologies for internet of things*. Springer, 2016.

[96] R. Want, "Enabling ubiquitous sensing with rfid," *Computer*, vol. 37, no. 4, pp. 84–86, 2004.

[97] A. Rida, L. Yang, and M. M. Tentzeris, *RFID-enabled sensor design and applications*. Artech House, 2010.

[98] S. Caizzone and E. DiGiampaolo, "Wireless passive rfid crack width sensor for structural health monitoring," *IEEE Sensors Journal*, vol. 15, no. 12, pp. 6767–6774, 2015.

[99] Y. Zhang, M. Shafiei, J. Z. Wen, Z. Abbasi, and C. L. Ren, "Simultaneous detection of pressure and bending using a microwave sensor with tag and reader structure," *IEEE Transactions on Instrumentation and Measurement*, 2023.

[100] Z. Abbasi, M. Baghelani, and M. Daneshmand, "High-resolution chipless tag rf sensor," *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 11, pp. 4855–4864, 2020.

[101] M. Baghelani, Z. Abbasi, and M. Daneshmand, "Noncontact high sensitivity chipless tag microwave resonator for bitumen concentration measurement at high temperatures," *Fuel*, vol. 265, p. 116916, 2020.

[102] Z. Abbasi, P. Shariaty, Z. Hashisho, and M. Daneshmand, "Silicagel-integrated chipless rf tag for humidity sensing," in *2018 18th International Symposium on Antenna Technology and Applied Electromagnetics (ANTEM)*. IEEE, 2018, pp. 1–2.

[103] S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, and F. Piccialli, "Scientific machine learning through physics–informed neural networks: Where we are and what's next," *Journal of Scientific Computing*, vol. 92, no. 3, p. 88, 2022.

[104] P. H. Winston, *Artificial intelligence*. Addison-Wesley Longman Publishing Co., Inc., 1984.

[105] T. Pham, T. Tran, D. Phung, and S. Venkatesh, "Predicting healthcare trajectories from medical records: A deep learning approach," *Journal of biomedical informatics*, vol. 69, pp. 218–229, 2017.

[106] K. Kourou, T. P. Exarchos, K. P. Exarchos, M. V. Karamouzis, and D. I. Fotiadis, "Machine learning applications in cancer prognosis and prediction," *Computational and structural biotechnology journal*, vol. 13, pp. 8–17, 2015.

[107] M. A. Shipp, K. N. Ross, P. Tamayo, A. P. Weng, J. L. Kutok, R. C. Aguiar, M. Gaasenbeek, M. Angelo, M. Reich, G. S. Pinkus *et al.*, "Diffuse large b-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning," *Nature medicine*, vol. 8, no. 1, pp. 68–74, 2002.

[108] M. W. Libbrecht and W. S. Noble, "Machine learning applications in genetics and genomics," *Nature Reviews Genetics*, vol. 16, no. 6, pp. 321–332, 2015.

[109] J. Vamathevan, D. Clark, P. Czodrowski, I. Dunham, E. Ferran, G. Lee, B. Li, A. Madabhushi, P. Shah, M. Spitzer *et al.*, "Applications of machine learning in drug discovery and development," *Nature reviews Drug discovery*, vol. 18, no. 6, pp. 463–477, 2019.

[110] A. D. Macpherson, "Using deep neural networks to classify astronomical images," 2023.

[111] A. L. Faisst, A. Prakash, P. L. Capak, and B. Lee, "How to find variable active galactic nuclei with machine learning," *The Astrophysical Journal Letters*, vol. 881, no. 1, p. L9, 2019.

[112] D. Rolnick, P. L. Donti, L. H. Kaack, K. Kochanski, A. Lacoste, K. Sankaran, A. S. Ross, N. Milojevic-Dupont, N. Jaques, A. Waldman-Brown *et al.*, "Tackling climate change with machine learning," *ACM Computing Surveys (CSUR)*, vol. 55, no. 2, pp. 1–96, 2022.

[113] A. Zarnani, S. Karimi, and P. Musilek, "Quantile regression and clustering models of prediction intervals for weather forecasts: A comparative study," *Forecasting*, vol. 1, no. 1, pp. 169–188, 2019.

[114] T. Barton and P. Musilek, "Day-ahead dynamic thermal line rating using numerical weather prediction," in *2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)*. IEEE, 2019, pp. 1–7.

[115] A. Zarnani and P. Musilek, "Learning uncertainty models from weather forecast performance databases using quantile regression," in *Proceedings of the 25th International Conference on Scientific and Statistical Database Management*, 2013, pp. 1–9.

[116] A. Zarnani and P. Musílek, "Non-parametric interval forecast models from fuzzy clustering of numerical weather predictions," in *2013 Joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS)*. IEEE, 2013, pp. 667–672.

[117] A. Lacoste, A. Luccioni, V. Schmidt, and T. Dandres, "Quantifying the carbon emissions of machine learning," *arXiv preprint arXiv:1910.09700*, 2019.

[118] S. Villon, C. Iovan, M. Mangeas, and L. Vigliola, "Confronting deep-learning and biodiversity challenges for automatic video-monitoring of marine ecosystems," *Sensors*, vol. 22, no. 2, p. 497, 2022.

[119] M. Roser, "The brief history of artificial intelligence: The world has changed fast – what might be next?" Dec 2022. [Online]. Available: https://ourworldindata.org/brief-history-of-ai

[120] S. Alaparthi and M. Mishra, "Bidirectional encoder representations from transformers (bert): A sentiment analysis odyssey," *arXiv preprint arXiv:2007.01127*, 2020.

[121] A. Radford, K. Narasimhan, T. Salimans, I. Sutskever *et al.*, "Improving language understanding by generative pre-training," 2018.

[122] A. Newell and H. Simon, "The logic theory machine–a complex information processing system," *IRE Transactions on information theory*, vol. 2, no. 3, pp. 61–79, 1956.

[123] A. Newell, J. C. Shaw, and H. A. Simon, "Report on a general problem solving program," in *IFIP congress*, vol. 256. Pittsburgh, PA, 1959, p. 64.

[124] E. Shortliffe, *Computer-based medical consultations: MYCIN*. Elsevier, 2012, vol. 2.

[125] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems and their applications*, vol. 13, no. 4, pp. 18–28, 1998.

[126] J. R. Quinlan, "Induction of decision trees," *Machine learning*, vol. 1, pp. 81–106, 1986.

[127] T. G. Dietterich, "Ensemble methods in machine learning," in *International workshop on multiple classifier systems*. Springer, 2000, pp. 1–15.

[128] A. Farooq, S. Anwar, M. Awais, and S. Rehman, "A deep cnn based multi-class classification of alzheimer's disease using mri," in *2017 IEEE International Conference on Imaging systems and techniques (IST)*. IEEE, 2017, pp. 1–6.

[129] M. F. Rabby, Y. Tu, M. I. Hossen, I. Lee, A. S. Maida, and X. Hei, "Stacked lstm based deep recurrent neural network with kalman smoothing for blood glucose prediction," *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, pp. 1–15, 2021.

[130] B.-X. Wang, W.-S. Zhao, D.-W. Wang, J. Wang, W. Li, and J. Liu, "Optimal design of planar microwave microfluidic sensors based on deep reinforcement learning," *IEEE Sensors Journal*, vol. 21, no. 24, pp. 27 441–27 449, 2021.

[131] Z. Azimi Dijvejin, M. C. Jain, R. Kozak, M. H. Zarifi, and K. Golovin, "Smart low interfacial toughness coatings for on-demand de-icing without melting," *Nature Communications*, vol. 13, no. 1, p. 5119, 2022.

[132] M. Alijani, B. D. Wiltshire, H. Sopha, Z. Sarpanah, J. Mistrik, L. Hromadko, M. H. Zarifi, and J. M. Macak, "Investigating the thickness-effect of free-standing high aspect-ratio tio2 nanotube layers on microwave-photoresponse using planar microwave resonators," *Applied Materials Today*, vol. 32, p. 101832, 2023.

[133] H. Mirzaei, J. McClelland, D. Sharma, M. Arjmand, and M. H. Zarifi, "A microwave voyage into swelling phenomenon: Investigation of polydimethylsiloxane and vocs interaction," *ACS Applied Materials & Interfaces*, 2023.

[134] T. Mosavirik, V. Nayyeri, M. Hashemi, M. Soleimani, and O. M. Ramahi, "Direct permittivity reconstruction from power measurements using a machine learning aided method," *IEEE Transactions on Microwave Theory and Techniques*, pp. 1–12, 2023.

[135] T. Mosavirik, M. Hashemi, M. Soleimani, V. Nayyeri, and O. M. Ramahi, "Material characterization using power measurements: Miracle of machine learning," in *2021 51st European Microwave Conference (EuMC)*, 2022, pp. 606–609.

[136] ——, "Microwave permittivity characterization using power measurements and machine learning," in *2021 IEEE Indian Conference on Antennas and Propagation (InCAP)*, 2021, pp. 618–620.

[137] L. Harrsion, M. Ravan, D. Tandel, K. Zhang, T. Patel, and R. K. Amineh, "Material identification using a microwave sensor array and machine learning," *Electronics*, vol. 9, no. 2, p. 288, 2020.

[138] E. L. Chuma and Y. Iano, "Novelty sensor using integrated fluorescence and dielectric spectroscopy to improve food quality identification," in *2022 IEEE Sensors*. IEEE, 2022, pp. 1–4.

[139] O. Niksan, K. Colegrave, and M. H. Zarifi, "Battery-free, artificial neural network-assisted microwave resonator array for ice detection," *IEEE Transactions on Microwave Theory and Techniques*, vol. 71, no. 2, pp. 698–709, 2022.

[140] E. L. Chuma and T. Rasmussen, "Metamaterial-based sensor integrating microwave dielectric and near-infrared spectroscopy techniques for substance evaluation," *IEEE Sensors Journal*, vol. 22, no. 20, pp. 19 308–19 314, 2022.

[141] K. Colegrave, M. C. Jain, O. Niksan, and M. Zarifi, "Artificial neural networks for antenna-based contactless liquid classification," in *2022 International Conference on Electrical, Computer and Energy Technologies (ICECET)*. IEEE, 2022, pp. 1–5.

[142] H. Sun, T. Tang, and G. Du, "Improved approach using symmetric microstrip sensor for accurate measurement of complex permittivity," *International Journal of RF and Microwave Computer-Aided Engineering*, vol. 28, no. 5, p. e21258, 2018.

[143] Y. Xie, X. Yang, P. Su, Y. He, and Y. Qiu, "A microwave time domain reflectometry technique combining the wavelet decomposition analysis and artificial neural network for detection of defects in dielectric structures," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–11, 2022.

[144] Z. Marinković, G. Gugliandolo, M. Latino, G. Campobello, G. Crupi, and N. Donato, "Characterization and neural modeling of a microwave gas sensor for oxygen detection aimed at healthcare applications," *Sensors*, vol. 20, no. 24, p. 7150, 2020.

[145] D. Covarrubias-Martínez, H. Lobato-Morales, J. M. Ramírez-Cortés, and G. A. Álvarez-Botero, "Classification of plastic materials using machine-learning algorithms and microwave resonant sensor," *Journal of Electromagnetic Waves and Applications*, vol. 36, no. 12, pp. 1760–1775, 2022.

[146] T. Mosavirik, M. Hashemi, M. Soleimani, V. Nayyeri, and O. M. Ramahi, "Accuracy-improved and low-cost material characterization using power measurement and artificial neural network," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–9, 2021.

[147] L. Nov, J.-Y. Chung, and J. Park, "Broadband permittivity characterization of a substrate material using deep neural network trained with full-wave simulations," *IEEE Access*, vol. 10, pp. 48 464–48 471, 2022.

[148] C. Gocen and M. Palandoken, "Machine learning assisted novel microwave sensor design for dielectric parameter characterization of water–ethanol mixture," *IEEE Sensors Journal*, vol. 22, no. 3, pp. 2119–2127, 2021.

[149] P. Kot, M. Muradov, S. Ryecroft, M. O. Pedrola, A. Shaw, J. Hemingway, R. Deb, and M. Coleman, "Identification of optimal frequencies to determine alpha-cypermethrin using machine learning fea-

ture selection techniques," in *2018 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2018, pp. 1–7.

[150] A. E. Omer, G. Shaker, S. Safavi-Naeini, H. Kokabi, G. Alquié, F. Deshours, and R. M. Shubair, "Low-cost portable microwave sensor for non-invasive monitoring of blood glucose level: Novel design utilizing a four-cell csrr hexagonal configuration," *Scientific Reports*, vol. 10, no. 1, pp. 1–20, 2020.

[151] S. Harnsoongnoen, P. Loutchanwoot, and P. Srivilai, "Sensing high 17β-estradiol concentrations using a planar microwave sensor integrated with a microfluidic channel," *Biosensors*, vol. 13, no. 5, p. 541, 2023.

[152] S. Abouzaid, T. Jaeschke, S. Kueppers, J. Barowski, and N. Pohl, "Deep learning-based material characterization using fmcw radar with open-set recognition technique," *IEEE Transactions on Microwave Theory and Techniques*, 2023.

[153] S. B. Kotsiantis, I. Zaharakis, P. Pintelas *et al.*, "Supervised machine learning: A review of classification techniques," *Emerging artificial intelligence applications in computer engineering*, vol. 160, no. 1, pp. 3–24, 2007.

[154] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992.

[155] S. A. Dudani, "The distance-weighted k-nearest-neighbor rule," *IEEE Transactions on Systems, Man, and Cybernetics*, no. 4, pp. 325–327, 1976.

[156] H.-F. Yu, F.-L. Huang, and C.-J. Lin, "Dual coordinate descent methods for logistic regression and maximum entropy models," *Machine Learning*, vol. 85, pp. 41–75, 2011.

[157] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, pp. 273–297, 1995.

[158] G. P. Zhang, "Neural networks for classification: a survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 30, no. 4, pp. 451–462, 2000.

[159] H. B. Barlow, "Unsupervised learning," *Neural computation*, vol. 1, no. 3, pp. 295–311, 1989.

[160] T. Hastie, R. Tibshirani, J. Friedman, T. Hastie, R. Tibshirani, and J. Friedman, "Unsupervised learning," *The elements of statistical learning: Data mining, inference, and prediction*, pp. 485–585, 2009.

[161] J. MacQueen *et al.*, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, no. 14. Oakland, CA, USA, 1967, pp. 281–297.

[162] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.

[163] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[164] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne." *Journal of machine learning research*, vol. 9, no. 11, 2008.

[165] J. B. Tenenbaum, V. d. Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *science*, vol. 290, no. 5500, pp. 2319–2323, 2000.

[166] I. Borg and P. J. Groenen, *Modern multidimensional scaling: Theory and applications*. Springer Science & Business Media, 2005.

[167] G. Synnaeve, Q. Xu, J. Kahn, T. Likhomanenko, E. Grave, V. Pratap, A. Sriram, V. Liptchinsky, and R. Collobert, "End-to-end asr: from supervised to semi-supervised learning with modern architectures," *arXiv preprint arXiv:1911.08460*, 2019.

[168] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.

[169] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.

[170] N. Q. K. Le and Q.-T. Ho, "Deep transformers and convolutional neural network in identifying dna n6-methyladenine sites in cross-species genomes," *Methods*, 2021.

[171] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics*, pp. 1189–1232, 2001.

[172] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural networks*, vol. 61, pp. 85–117, 2015.

[173] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision," in *Proceedings of 2010 IEEE international symposium on circuits and systems*. IEEE, 2010, pp. 253–256.

[174] L. Mou, P. Ghamisi, and X. X. Zhu, "Deep recurrent neural networks for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 7, pp. 3639–3655, 2017.

[175] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[176] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.

[177] T. Lähivaara, R. Yadav, G. Link, and M. Vauhkonen, "Estimation of moisture content distribution in porous foam using microwave tomography with neural networks," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 1351–1361, 2020.

[178] V. Mattsson, M. D. Perez, D. Dematties, and R. Augustine, "Neural network approach for dielectric characterization of tissues in microwave frequencies using coplanar waveguide transmission," in *2020 14th European Conference on Antennas and Propagation (EuCAP)*. IEEE, 2020, pp. 1–5.

[179] S. Abouzaid, T. Jaeschke, J. Barowski, and N. Pohl, "Fmcw radar-based material characterization using convolutional neural network and k-means clustering," in *2022 24th International Microwave and Radar Conference (MIKON)*. IEEE, 2022, pp. 1–4.

[180] T. Reimer and S. Pistorius, "The diagnostic performance of machine learning in breast microwave sensing on an experimental dataset," *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology*, vol. 6, no. 1, pp. 139–145, 2021.

[181] D. P. Kingma, M. Welling *et al.*, "An introduction to variational autoencoders," *Foundations and Trends® in Machine Learning*, vol. 12, no. 4, pp. 307–392, 2019.

[182] M. Sabuhi, M. Zhou, C.-P. Bezemer, and P. Musilek, "Applications of generative adversarial networks in anomaly detection: A systematic literature review," *IEEE Access*, 2021.

[183] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015.

[184] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional gans," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8798–8807.

[185] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of wasserstein gans," *Advances in neural information processing systems*, vol. 30, 2017.

[186] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.

[187] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of gans for improved quality, stability, and variation," *arXiv preprint arXiv:1710.10196*, 2017.

[188] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 4401–4410.

[189] H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena, "Self-attention generative adversarial networks," in *International conference on machine learning*. PMLR, 2019, pp. 7354–7363.

[190] A. Jaiswal, W. AbdAlmageed, Y. Wu, and P. Natarajan, "Bidirectional conditional generative adversarial networks," in *Computer Vision–ACCV 2018: 14th Asian Conference on Computer Vision, Perth, Australia, December 2–6, 2018, Revised Selected Papers, Part III 14*. Springer, 2019, pp. 216–232.

[191] A. Makhzani, J. Shlens, N. Jaitly, I. Goodfellow, and B. Frey, "Adversarial autoencoders," *arXiv preprint arXiv:1511.05644*, 2015.

[192] M. H. Zarifi, M. Fayaz, J. Goldthorp, M. Abdolrazzaghi, Z. Hashisho, and M. Daneshmand, "Microbead-assisted high resolution microwave planar ring resonator for organic-vapor sensing," *Appl. Phys. Lett.*, vol. 106, no. 6, p. 062903, Feb. 2015.

[193] M. Abdolrazzaghi and M. Daneshmand, "Exploiting sensitivity enhancement in micro-wave planar sensors using intermodulation products with phase noise analysis," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 67, no. 12, pp. 4382–4395, Dec. 2020.

[194] M. H. Zarifi, S. Farsinezhad, M. Abdolrazzaghi, M. Daneshmand, and K. Shankar, "Selective microwave sensors exploiting the interaction of analytes with trap states in TiO2 nanotube arrays," vol. 8, no. 14, pp. 7466–7473, Apr. 2016.

[195] B. Martin, K. Edwards, I. Jeffrey, and C. Gilmore, "Experimental microwave imaging system calibration via cycle-gan," *IEEE Transactions on Antennas and Propagation*, 2023.

[196] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4681–4690.

[197] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, and M. Shah, "Transformers in vision: A survey," *ACM computing surveys (CSUR)*, vol. 54, no. 10s, pp. 1–41, 2022.

[198] N. Gholizadeh, N. Kazemi, and P. Musilek, "A comparative study of reinforcement learning algorithms for distribution network reconfiguration with deep q-learning-based action sampling," *IEEE Access*, vol. 11, pp. 13 714–13 723, 2023.

[199] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.5602*, 2013.

[200] J. B. Tenenbaum, V. d. Silva, and J. C. Langford, "A survey of manifold-based learning methods," *International Journal of Robotics Research*, vol. 33, no. 9, pp. 1279–1300, 2014.

[201] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep cnn denoiser prior for image restoration," *arXiv preprint arXiv:1704.02532*, 2017.

[202] Z. Tchir, M. Z. Reformat, and P. Musilek, "Home energy management with v2x capability using reinforcement learning," in *2023 IEEE Conference on Artificial Intelligence (CAI)*. IEEE, 2023, pp. 89–91.

[203] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," *Advances in neural information processing systems*, vol. 30, 2017.

[204] D. Floreano, P. Dürr, and C. Mattiussi, "Neuroevolution: from architectures to learning," *Evolutionary intelligence*, vol. 1, pp. 47–62, 2008.

[205] M. K. Sarker, L. Zhou, A. Eberhart, and P. Hitzler, "Neuro-symbolic artificial intelligence," *AI Communications*, vol. 34, no. 3, pp. 197–209, 2021.

[206] A. Steane, "Quantum computing," *Reports on Progress in Physics*, vol. 61, no. 2, p. 117, 1998.

[207] L.-V. Herrera-Sepulveda, J.-L. Olvera-Cervantes, A. Corona-Chavez, and T. Kaur, "Sensor and methodology for determining dielectric constant using electrically coupled resonators," *IEEE Microwave and Wireless Components Letters*, vol. 29, no. 9, pp. 626–628, 2019.

[208] C. G. Juan, B. Potelon, C. Quendo, E. Bronchalo, and J. M. Sabater-Navarro, "Highly-sensitive glucose concentration sensor exploiting inter-resonators couplings," in *2019 49th European Microwave Conference (EuMC)*. IEEE, 2019, pp. 662–665.

[209] J. Aitken, "Swept-frequency microwave q-factor measurement," in *Proceedings of the Institution of Electrical Engineers*, vol. 123, no. 9. IET, 1976, pp. 855–862.

[210] J. Bray and L. Roy, "Measuring the unloaded, loaded, and external quality factors of one-and two-port resonators using scattering-parameter magnitudes at fractional power levels," *IEE Proceedings-Microwaves, Antennas and Propagation*, vol. 151, no. 4, pp. 345–350, 2004.

[211] P. Naseri and S. V. Hum, "A generative machine learning-based approach for inverse design of multilayer metasurfaces," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 9, pp. 5725–5739, 2021.

[212] V. Avkhimenia, T. Weis, and P. Musilek, "Generation of synthetic ampacity and electricity pool prices using generative adversarial networks," in *2021 IEEE Electrical Power and Energy Conference (EPEC)*, 2021, pp. 225–230.

[213] M. Sabuhi, M. Zhou, C.-P. Bezemer, and P. Musilek, "Applications of generative adversarial networks in anomaly detection: A systematic literature review," *IEEE Access*, vol. 9, pp. 161 003–161 029, 2021.

[214] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[215] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," in *Deep learning in medical image analysis and multimodal learning for clinical decision support*. Springer, 2018, pp. 3–11.

[216] C. Li and M. Wand, "Precomputed real-time texture synthesis with markovian generative adversarial networks," *ArXiv*, vol. abs/1604.04382, 2016.

[217] H. Nam and H.-E. Kim, "Batch-instance normalization for adaptively style-invariant neural networks," *arXiv preprint arXiv:1805.07925*, 2018.

[218] A. Z Wang, H. C Bovik, R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[219] D. R. I. M. Setiadi, "Psnr vs ssim: imperceptibility quality assessment for image steganography," *Multimedia Tools and Applications*, vol. 80, pp. 8423–8444, 2021.

[220] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," 2018.

[221] M. Abolrazzaghi, R. Genov, and G. V. Eleftheriades, "2d scanning by subwavelength superoscillatory focused beams for wireless power delivery," in *2023 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (USNC-URSI)*. IEEE, 2023, pp. 533–534.

[222] M. Abdolrazzaghi, R. Genov, and G. V. Eleftheriades, "Antenna array for wireless power transfer to deep-brain implants," in *2022 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (AP-S/URSI)*. IEEE, 2022, pp. 2018–2019.

[223] I. M. Wentholt, J. B. Hoekstra, and J. H. DeVries, "A critical appraisal of the continuous glucose–error grid analysis," *Diabetes Care*, vol. 29, no. 8, pp. 1805–1811, 2006.

[224] M. Abdolrazzaghi, R. Genov, and G. V. Eleftheriades, "Microwave planar sensor antenna for glucose sensing in aqueous solutions," in *2021 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (APS/URSI)*. IEEE, 2021, pp. 127–128.

[225] A. Singh, "Anomaly detection for temporal data using long short-term memory (lstm)," 2017.

[226] M. Hitzemann, K. J. Dehning, A. V. Gehl, E.-F. Sterr, and S. Zimmermann, "Fast readout of split-ring resonators made simple and low-cost for application in hplc," *Electronics*, vol. 11, no. 7, p. 1139, 2022.

[227] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[228] F. Reiterer, P. Polterauer, M. Schoemaker, G. Schmelzeisen-Redecker, G. Freckmann, L. Heinemann, and L. Del Re, "Significance and reliability of mard for the accuracy of cgm systems," *Journal of diabetes science and technology*, vol. 11, no. 1, pp. 59–67, 2017.

[229] C. Polk and E. Postow, *Handbook of Biological Effects of Electromagnetic Fields, -2 Volume Set*. CRC press, 1995.

[230] S. Wild, G. Roglic, A. Green, R. Sicree, and H. King, "Global prevalence of diabetes: estimates for the year 2000 and projections for 2030," *Diabetes care*, vol. 27, no. 5, pp. 1047–1053, 2004.

Appendix A

Appendix

Supplementary for CycleGAN

Evaluation metrics

In this section, the common metrics that are used in this work are explained as follows:

Structural similarity index (SSIM)

The similarity between two images is normally measured with SSIM. First, a Gaussian filter ($\sigma = 1.5$ pixels) is used to smooth both the target and generated images. Then SSIM is calculated for each pair of subimages with a side length of 11 pixels as follows:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_x^2 + c_2)} \tag{A.1}$$

where μ_x and μ_y are the average pixel intensity of the subimages x and y. σ_x, σ_y, and σ_{xy} are the standard deviation for the x and y subimages, and the covariance of the two subimages, respectively. Constant values $c_1 = (0.01 \times L)^2$ and $c_2 = (0.03 \times L)^2$ are used to avoid zero denominators [227], and it is recommended to consider L equal to 255 as default values. Then $SSIM(x, y)$ is averaged over all subimages to cover each pair of images.

Peak-signal-to-noise ratio (PSNR) and mean squared error (MSE)

PSNR has been in use before SSIM in image quality assessment. In RGB images, MSE is calculated based on three ($M \times N \times O$) dimensions, where M and N are the image resolution and O is the number of image channels. It should be noted that in the case of the gray scale representation, shown in Fig. A.1, one needs to stack the single channel to make it rank 3 that is acceptable for VGG. $I(x, y, z)$ is the pixel value of the original image in the x, y coordinates and channel z. Y and $\hat{Y}$ are the target and generated

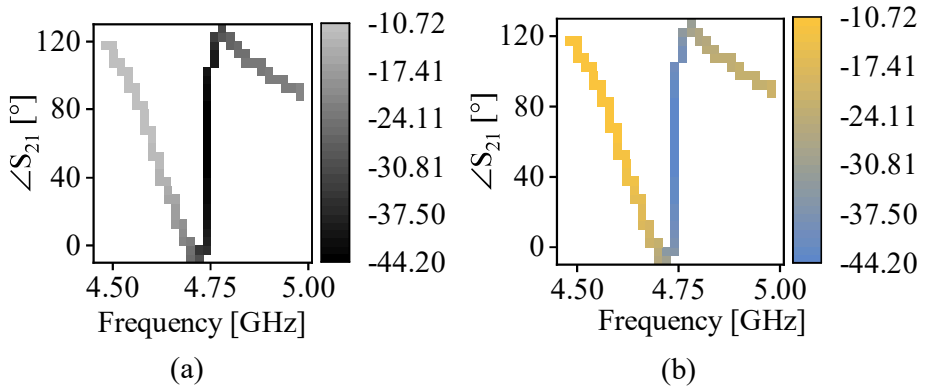


Figure A.1: Two possible representations of (a) grey scale and (b) color heatmaps.

images, respectively.

$$\text{MSE} = \frac{1}{MNO}\sum_{x=1}^{M}\sum_{y=1}^{N}\sum_{z=1}^{O}(Y_{(x,y,z)} - \hat{Y}_{(x,y,z)})^2 \tag{A.2}$$

Using MSE, the PSNR is calculated as follows:

$$PSNR = 10log_{10}(\frac{MAX_i^2}{MSE}) \tag{A.3}$$

where MAX is the highest scale value of the pixels. The error value is calculated considering the difference in pixel values at the same coordinate and channel [219].

Mean absolute relative difference (MARD)

MARD, which stands for the mean value of the absolute relative differences (ARD), is a metric to evaluate the accuracy of continuous glucose monitoring (CGM) of a system. MARD is based on the comparison between the reference (commercial) gluocometer data and a given continuous glucose sensor and is computed as follows [228]:

$$ARD = 100\% \cdot \frac{|y_{CGM} - y_{ref}|}{y_{ref}} \tag{A.4}$$

$$MARD = \frac{1}{N_{ref}}\sum_{k=1}^{N_{ref}} ARD_k \tag{A.5}$$

where y_{CGM} is the GL value of the commercial glucometer, and y_{ref} is the value measured using microwave sensor during the continuous measurement.

Supplementary for LSTM

Frequency of Choice

In EM-based sensors, the frequency of operation has always been a point of discussion. Various studies have reported good performance of EM-based sensors in both the MHz and GHz range. In this regard, the depth of penetration and the sensitivity of the sensor are considered the determining parameters. To evaluate the impact of each material at different frequencies, the human trial case is used to examine the sensor for generic and applicable cases. Assuming that the planar sensor sits on the surface of a finger (or other parts of the body) and that the glucose content is predicted according to the effective change of the dielectric constant in the skin, it is useful to have a dielectric constant of normal skin. The parametric model of the skin, elaborated by the double Cole-Cole model [229, 230], is given as:

$$\varepsilon_r(f) = \varepsilon_r{}'(f) - j\varepsilon_r{}''(f) = \varepsilon_\infty + \sum_{l=1}^{2} \frac{\Delta l}{1 + (\frac{jf}{f_{r,l}})^{1-\alpha_l}} + \frac{\sigma_s}{j2\pi f \varepsilon_0{}'} \tag{A.6}$$

where ε_∞ and σ_s are the permittivity at infinity and the DC conductivity of the samples, respectively, Δl, f_r, l, and αl are the amplitude, the relaxation frequency, and the distribution parameter for l-th dielectric dispersion, respectively. The Cole-Cole parameters for the dermis and epidermis are listed in Table 1.

Tissue	ε_∞	$\sigma_s[S/m]$	$f_{r,1}[GHz]$	Δ_1	α_1	$f_{r,2}[GHz]$	Δ_2	α_2
Dermis	4.17	0.553	23.6	34.7	0.0073	3.96	364	0.292
Epidermis	3.06	0.208	28.7	23.8	0.020	22.2	67	0.488

Table A.1: Cole-Cole parameters of skin. These parameters are suitable for a frequency range of 0.5 GHz – 110 GHz.

Supplementary Fig. A.2 shows the frequency dependence of the permittivity and loss tangent of the skin layers as shown in Fig. 1e. The permittivity values of the dermis and epidermis of the skin are calculated using the Cole-Cole model. There is a clear difference between permittivity and loss tangent in the single-digit GHz region.

It is clear that the permittivity of both dermis and epidermis drops with increasing frequency. However, the loss tangent has a local minimum around single-digit GHz frequencies. This region would be beneficial for microwave sensors to penetrate the skin and interrogate deeper layers. In this sense, frequencies around 3 GHz seem likely to be optimal for analyte sensing in the human body.

It is challenging to evaluate the dielectric properties of skin with various glucose content. However, the properties of glucose in water by itself illustrate all the required information about its impact in the skin layers. A single Debye model is used to model the frequency behavior of glucose in a water solution as follows:

$$\varepsilon_r{}' = \varepsilon_\infty + \frac{\varepsilon_0 - \varepsilon_\infty}{1 + \omega^2 \tau^2} \tag{A.7}$$

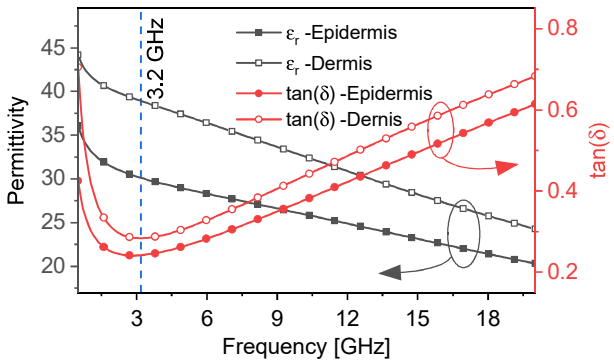


Figure A.2: **Dielectric properties of the dermis and epidermis**. The permittivity values of the dermis and epidermis of the skin are calculated using the Cole-Cole model. There is a clear difference between permittivity and loss tangent in the single-digit GHz region.

With the parameters given in Table 2, the results in Fig. 1e show that a change in glucose content leads to a change in the dielectric properties of the glucose-in-water solution, and that the variation increases at higher frequencies. This shows that working at relatively high frequencies increases the sensitivity of the sensory system.

Concentration	ε_0	ε_∞	$\tau[ps]$
0 (DI water)	78.4	9.8	8.7
10 mM/L	76.1	27.8	14.6
20 mM/L	74.3	27.7	19.3
30 mM/L	73.8	25.2	22.7
40 mM/L	72.1	23.2	26.1

Table A.2: Parameters of the Debye model for glucose in water solution. The single Debye model can accurately predict the dielectric properties of the glucose solution.

Combining the information of both glucose-in-water and skin layers suggests that higher frequencies would result in more distinct permittivity values for the solution. However, the skin loss tangent also grows, which precludes the use of higher frequencies. Therefore, a low-GHz frequency range is considered to be optimal because the passive sensor can interact with the glucose content. On the other hand, the penetration depth inside a medium is inversely proportional to the frequency of operation. This also confirms that the low-GHz range is the optimal region for this application.

Circuit Model Analysis:

In a simpler view of the circuit model, as shown in Fig. A.3 a, the material under test is represented as a parallel R and C, the patch resonator is considered as a parallel RLC resonator, and the coupling between the transmission line and the resonator is shown with C_M (see Fig.A.3 b-c). To derive the input impedance of the circuit, the coupling is modeled with a J-Inverter. According to Fig. A.3 d, Z_1 and Z_2

are calculated as follows:

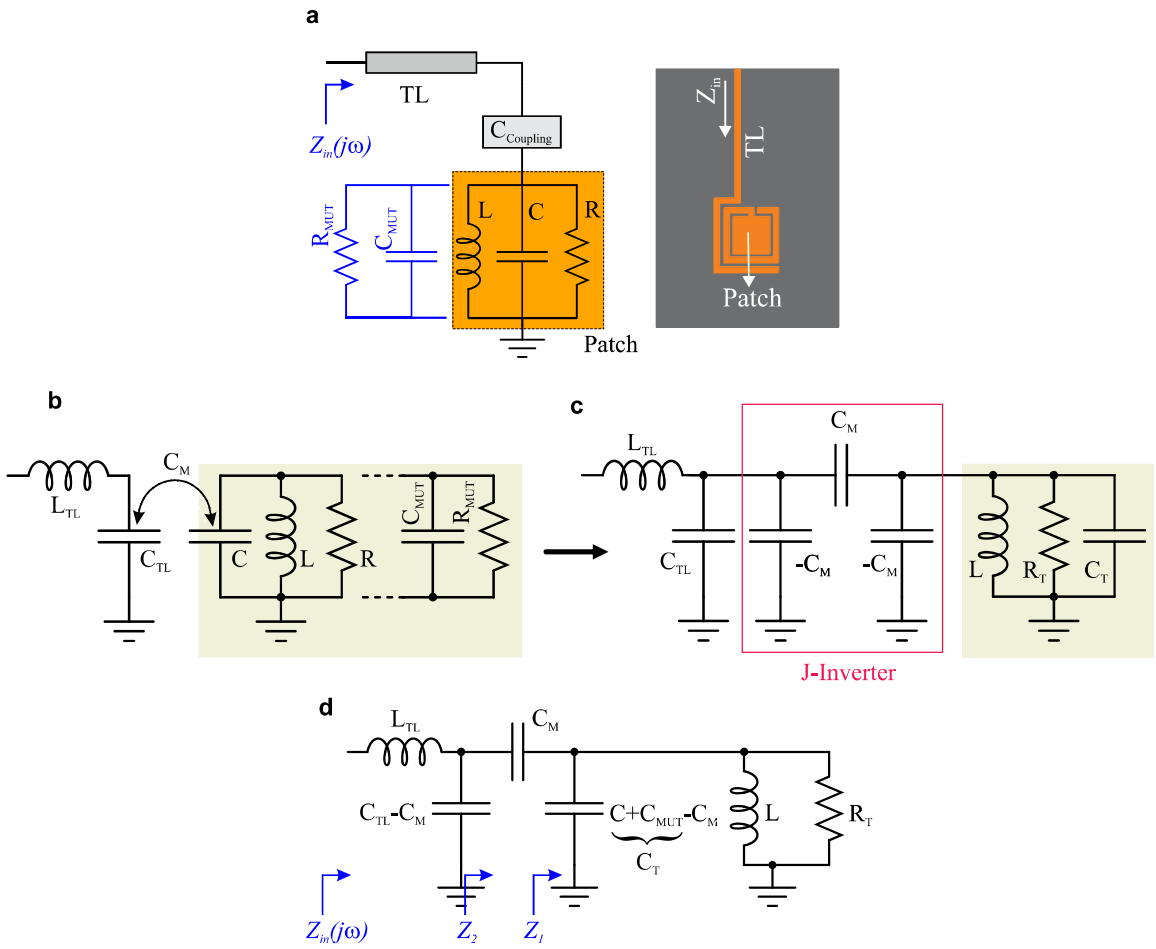


Figure A.3: Dielectric properties of the dermis and epidermis.

$$\frac{1}{Z_1} = \frac{1}{R_T} + \frac{1}{j\omega L} + j\omega(C_T - C_M), \tag{A.8}$$

$$Z_2 = Z_1 + \frac{1}{j\omega C_M} \tag{A.9}$$

Considering the transmission line's distributed model with L_{TL} and C_{TL}, the input impedance of the whole circuit is given as follows:

$$Z_{in} = j\omega L_{TL} + \frac{1}{1/Z_2 + j\omega(C_{TL} - C_M)} \tag{A.10}$$

The resonance frequency of the circuit is achieved when the imaginary part of the input impedance Z_{in} equals zero.

$$\text{Imag}(Z_{in\ @\omega_{res}}) = 0 \tag{A.11}$$

Data Analysis:

Data Collection:

In the experiments, the reflected parameter (S_{11}) of the passive sensor is measured with the vector network analyzer (VNA). This complex-valued parameter is measured at 3,001 equally spaced frequency points and then fed to LabView in the computer as an array. This allows identification of the resonance frequency and measurement of the depth of S_11 at this frequency. As a result, each measured profile 3,001 points can be reduced to a two-dimensional feature point characterizing the sensor response, and these feature points are eventually used by the prediction model.

Data Preprocessing:

The values measured by the VNA carry uncorrelated noise that can be suppressed by averaging the sensor response. In this case, the amplitude of the resonance can be measured more accurately. To this aim, the VNA measurements are repeated for 3 consecutive recordings. For human trials, since the measurement time was long and there were not enough test strips to cover the entire recording time, only selected data points were verified with the commercially available glucometer. This comparative measurement was repeated every 30 minutes. Great agreement is achieved between the measured values of the sensor response and the actual blood glucose level measured using the Accu-check glucometer.

Performance Metrics:

The accuracy of the predictions is evaluated with respect to the expected glucose values in the human body. The following is a list of the metrics used in our study:

1. Root of the mean of squared error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\frac{y_{ri} - \hat{y_i}}{y_{ri}})^2} \tag{A.12}$$

2. Mean absolute Relative difference (MARD):

$$\text{MARD} = \frac{1}{N}\sum_{i=1}^{N}|\frac{y_{ri} - \hat{y_i}}{y_{ri}}| \times 100 \tag{A.13}$$

3. Mean of absolute error (MAE):

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_{ri} - \hat{y}_i| \tag{A.14}$$

where y_r denotes the reference glucose level, $\hat{y}$ represents the predicted values, and N is the number of data points.

4. Pearson correlation (R) between predicted and reference sets:

$$R = \rho_{x,y} = \frac{cov(X,Y)}{\sigma_x \sigma_y} \tag{A.15}$$

where cov is the covariance, and σ_x is the standard deviation of$X$,R can take values between -1 and +1, with zero meaning no correlation, and ± 1 denotes extreme positive/negative correlation. The analysis of human trials for glucose prediction is performed using Clarke error grid analysis, which considers statistical performance of the predictions and their differences from real values under clinical mindset. It shows the accuracy of the predicted glucose values vs. reference values graphically. It has five different zones that capture scattered data points based on their level of accuracy as follows: Zone A shows data points with 20% accuracy. Zone B contains measurements outside the 20% accuracy region, but with an acceptable final prediction. Unlike zones A and B, zones D/E/F are not safe. They are clinically unacceptable and might lead to dangerous conclusions.

Prediction Methods:

Feed-forward vs. Recurrent Neural Networks

The main disadvantage of feedforward neural networks (NNs) is that they do not have an explicit memory of the past. Their modeling is restricted by the size of the input window, where the input data is considered as a multidimensional feature vector instead of a data sequence. This approach is only helpful for simple datasets and observation sequences that do not have long time-based dependencies, i.e. where a limited number of recent inputs suffice. The µF channel test not only has a long repeating sequence of up to 90 data points, but it also consists of nested patterns of three peaks followed by a rest time. This large number of past values needs to be remembered to achieve an accurate prediction. Therefore, feedforward NNs are not suitable for this application, and recurrent neural networks (RNNs) are considered instead.

Multistep LSTM Architecture

LSTM is the most popular variant of RNNs that avoids problems with long-term dependencies and vanishing gradient problem. The LSTM code is developed in Python programming language and Keras on top of TensorFlow for fast prototyping. Keras provides out-of-the-box solutions to problems using

deep learning building blocks. Training RNNs requires unfolding the network at each time step, which is computationally ineffective. Thus, in Keras, RNN unfolding is extended up to the size of the input window, also known as lookback. The input data are then a three-dimensional array (tensor) with the shape of: (batch-size, lookback, input-dimension). In this project, the LSTM is used in the default mode for its state maintenance. In this case, the samples of a batch are independent and the state of LSTM (memory) is maintained for a limited number of steps (=lookback). For instance, to ensure that the LSTM remembers the information of the nested pattern in the µF channel experiment, a lookback of 90 cells is used to reconstruct the next sample, since the slowest pattern takes 90 data points.

Optimization Algorithm

The best values for parameters such as lookback, dropout, learning rate, and network architecture are found using the Bayesian optimization process. This algorithm uses the multistep LSTM model as an objective function. The constraints of the problem are gathered in a dictionary as follows:

Parameter	type	domain
Dropout	continuous	(0.1, 0.9)
Learning rate	continuous	(0.0001, 0.1)
Batch size	discrete	(32, 64, 128, 256, 512, 1024)
Look back	discrete	(1, 2, 4, 8, 16, 32, 64)
layers	discrete	(0, 1, 2)

Where the layer types (0, 1, 2) are:

Layers
{'input': 1, 'hidden1': 80, 'output': 1},
{'input': 1, 'hidden1': 80, 'hidden2': 40, 'output': 1},
{'input': 1, 'hidden1': 80, 'hidden2': 40, 'hidden3': 20, 'output': 1}

Also, the acquisition type is set to 'expected improvement'. The performance of Bayesian optimization is shown in Fig. 3. It shows the distance between consecutive parameter sets (from multistep domain dictionary) as well as the validation loss value (Best y). It can be seen that the optimization is completed with only 15 iterations, and the final validation loss is 0.0175. The resulting optimal parameter values are as follows:

Optimal parameters					
dropout:	0.400000	Learning rate:	0.10	batch size: 32	Lookback: 6
Layers:	{'input': 1, 'hidden1': 80, 'hidden2': 40, 'hidden3': 20, 'output': 1}				

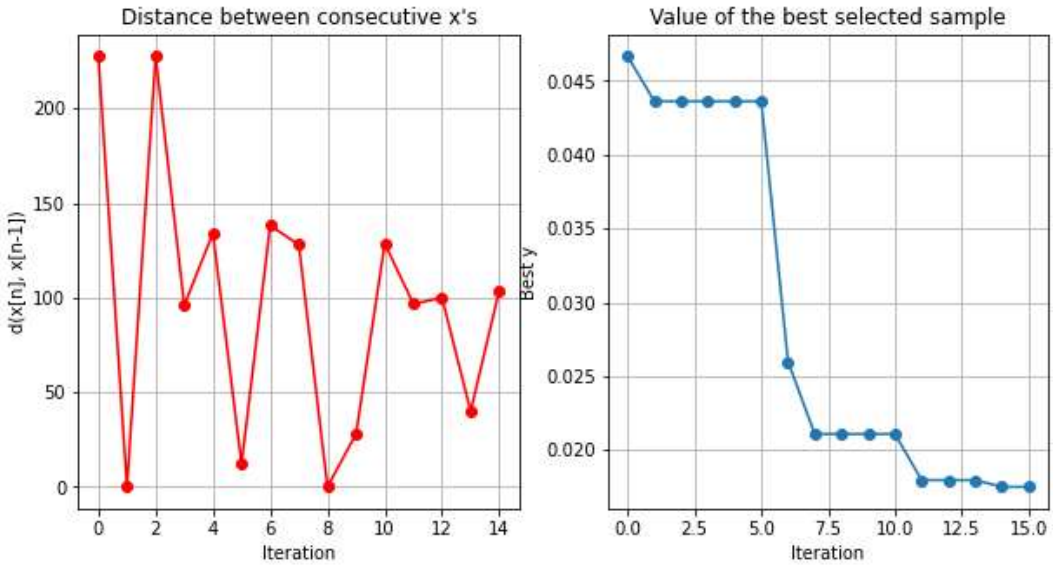


Figure A.4: **Performance metrics of the LSTM hyperparameter optimization**. The prediction model is tested with various parameters and the best performance is obtained with the hyperparameter optimization using the Bayesian method. The closeness of the distance between consecutive parameters (left) after $\sim$ 10 iterations as well as the lower validation loss (right) demonstrate that the fitness of the model converged with the optimal parameter values.

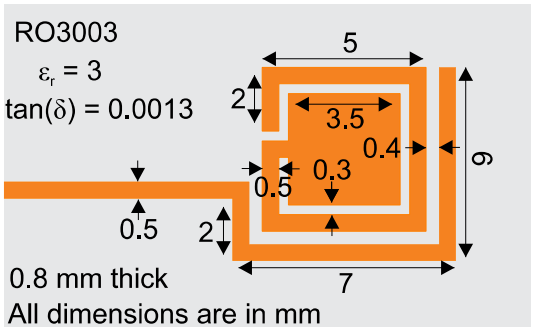


Figure A.5: **Sensor dimensions.** The sensor is fabricated on Rogers RO3003 substrate with the dimensions given in the figure, all in millimeter. The sensor is fabricated on Rogers RO3003 substrate with the dimensions given in the figure, all in millimeter.

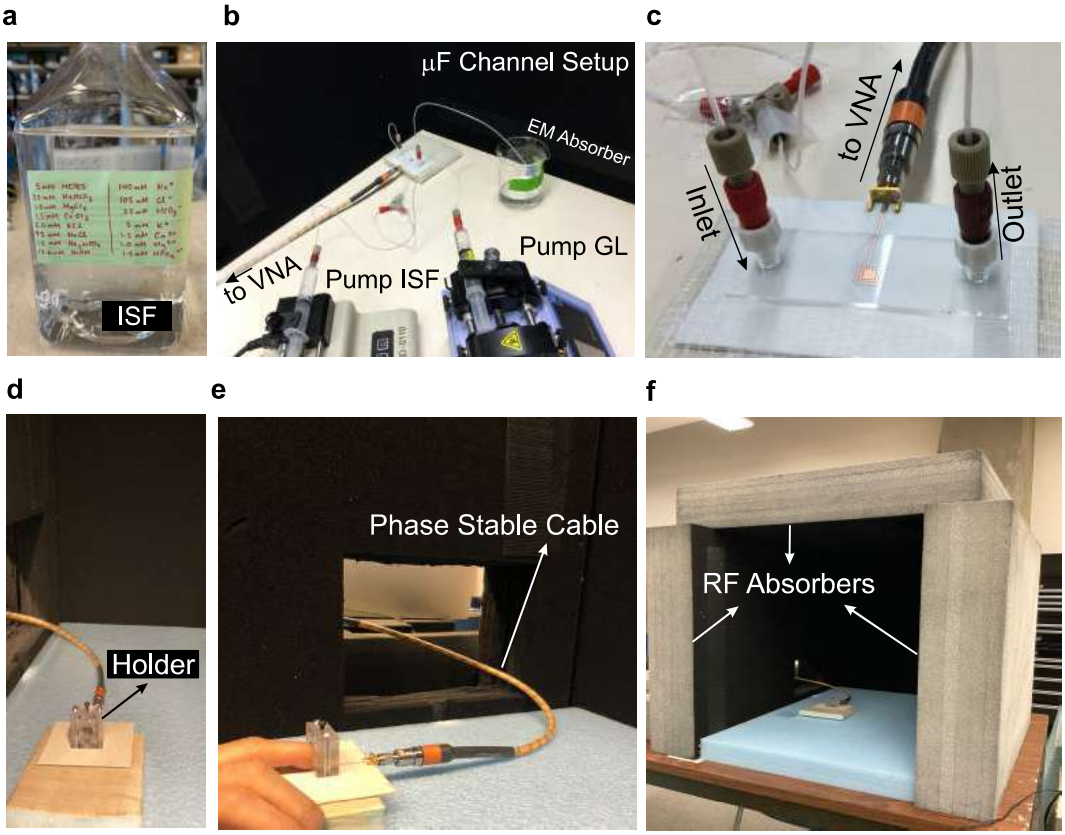


Figure A.6: **Experimental setup for *in-vivo* and *in-vitro* measurement.a** Prepared solution to mimic interstitial fluid,**b** Two programmable syringe pumps are used to make desired concentrations of glucose in ISF. **c** μF channel mounted on top of the sensor. **d** Finger holder to fix the position of the hand on the sensor, using which the location of finger remains constant through all measurements. **e** VNA is connected to the SMA connector using a phase-stable cable. **f** Four RF absorbers surround the sensor to block any undesired noise including changes in the air flow, movement of other persons, etc.

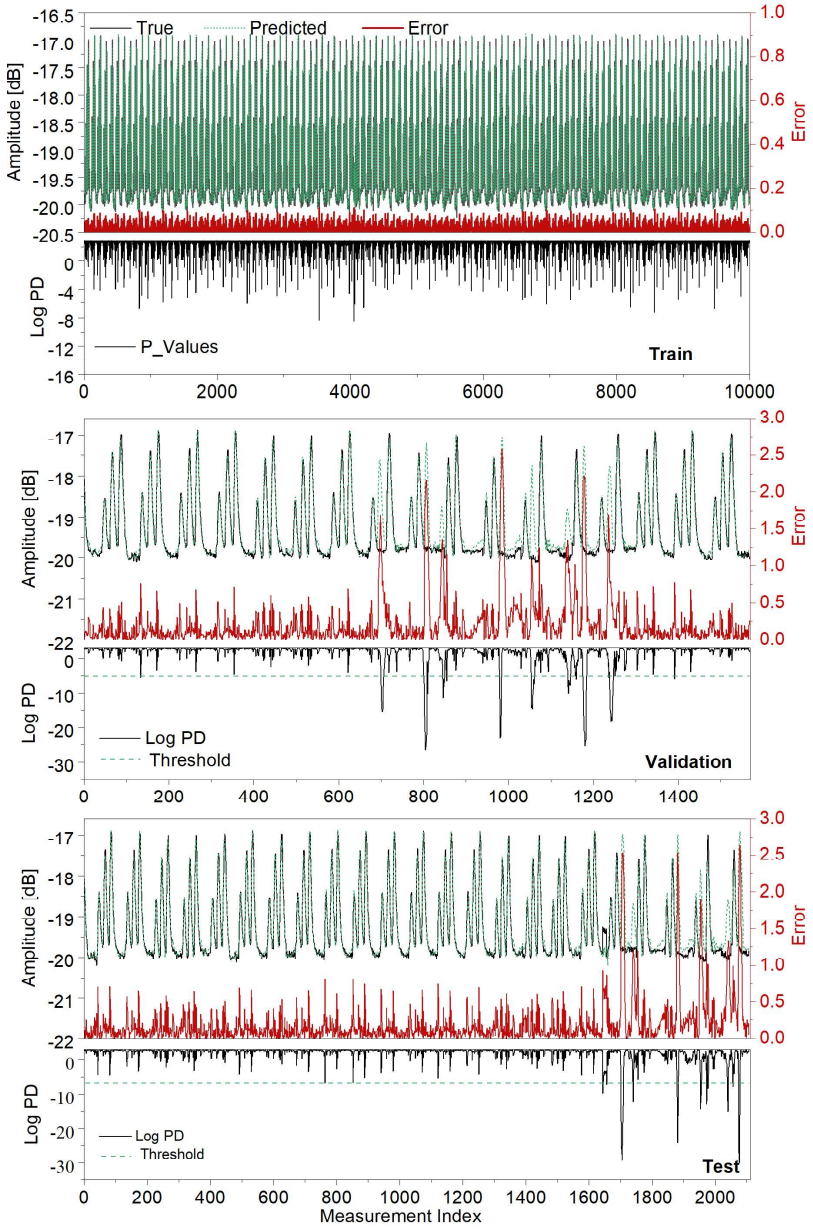


Figure A.7: **μF channel test results:** The glucose content in the ISF is tested using μF channel and a programmed pattern of glucose injections is used. At certain points, the regular pattern is modified to introduce anomalies, and the anomalies are detected properly using the LSTM prediction results. The error between the measured sensor response and the expected (predicted) response is also depicted. LogPD is the probability density function of the error values in log format. It elaborates on the likelihood of a data point being an anomaly according to the magnitude of its logPD. The results for the validation and test sets are shown separately.

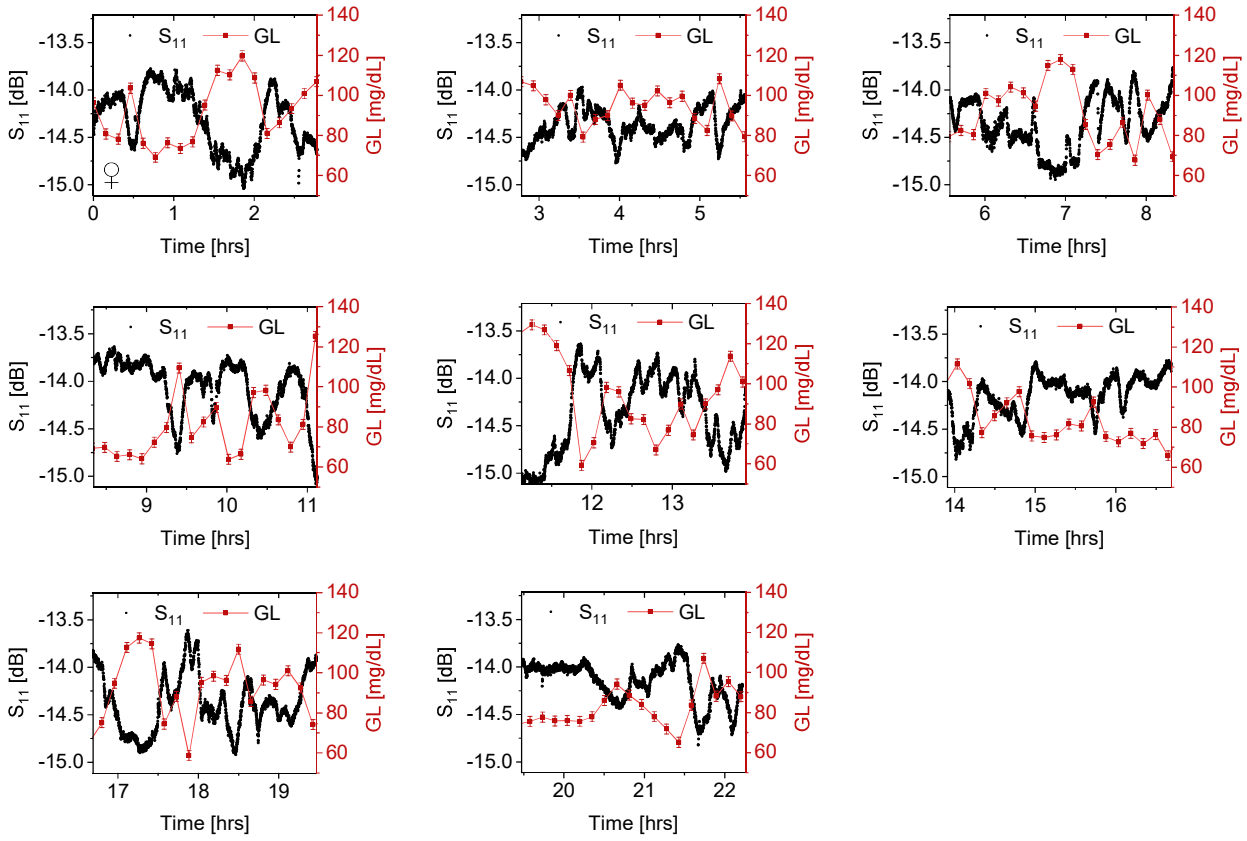


Figure A.8: **Comparison between S_{11} values of the patch sensor and commercial glucometer for a non-diabetic female participant.** Microwave sensor measurement response is recorded using LabView every 5 seconds, which continues per participant sitting. In addition, a commercial glucometer (AccuChek) is used to measure the actual blood GL by pricking the finger of the other hand every 10 minutes. This profile is repeated for eight days. Measured S11 values for each sitting are plotted along with the corresponding GL measured by AccuChek. Successive trials of the commercial glucometer in a short span ¡ 5 min revealed an uncertainty of ±2.5 mg/dL, which is represented by an error bar. To evaluate the relationship between the microwave sensor response and AccuChek, the Pearson correlation factor is computed for each separate sitting.

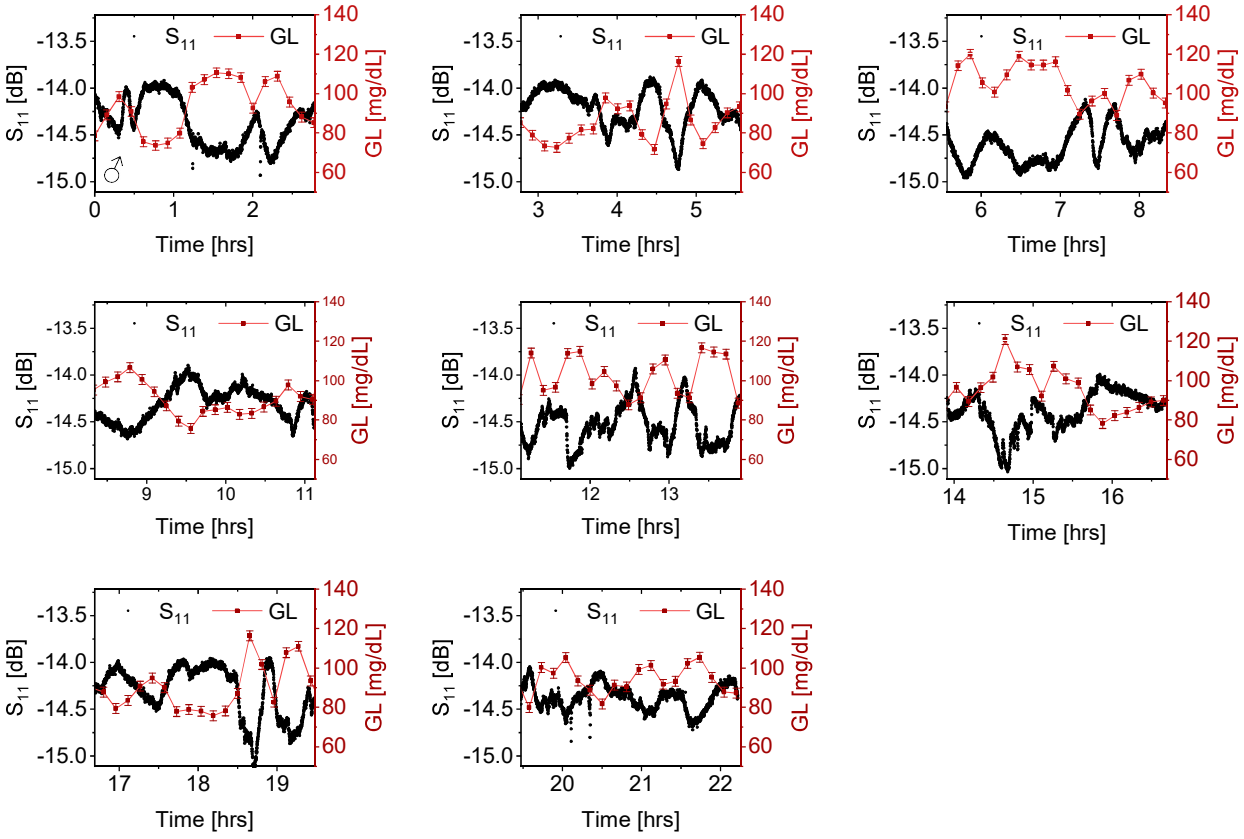


Figure A.9: Comparison between S_{11} values of the patch sensor and commercial glucometer for a non-diabetic male participant.

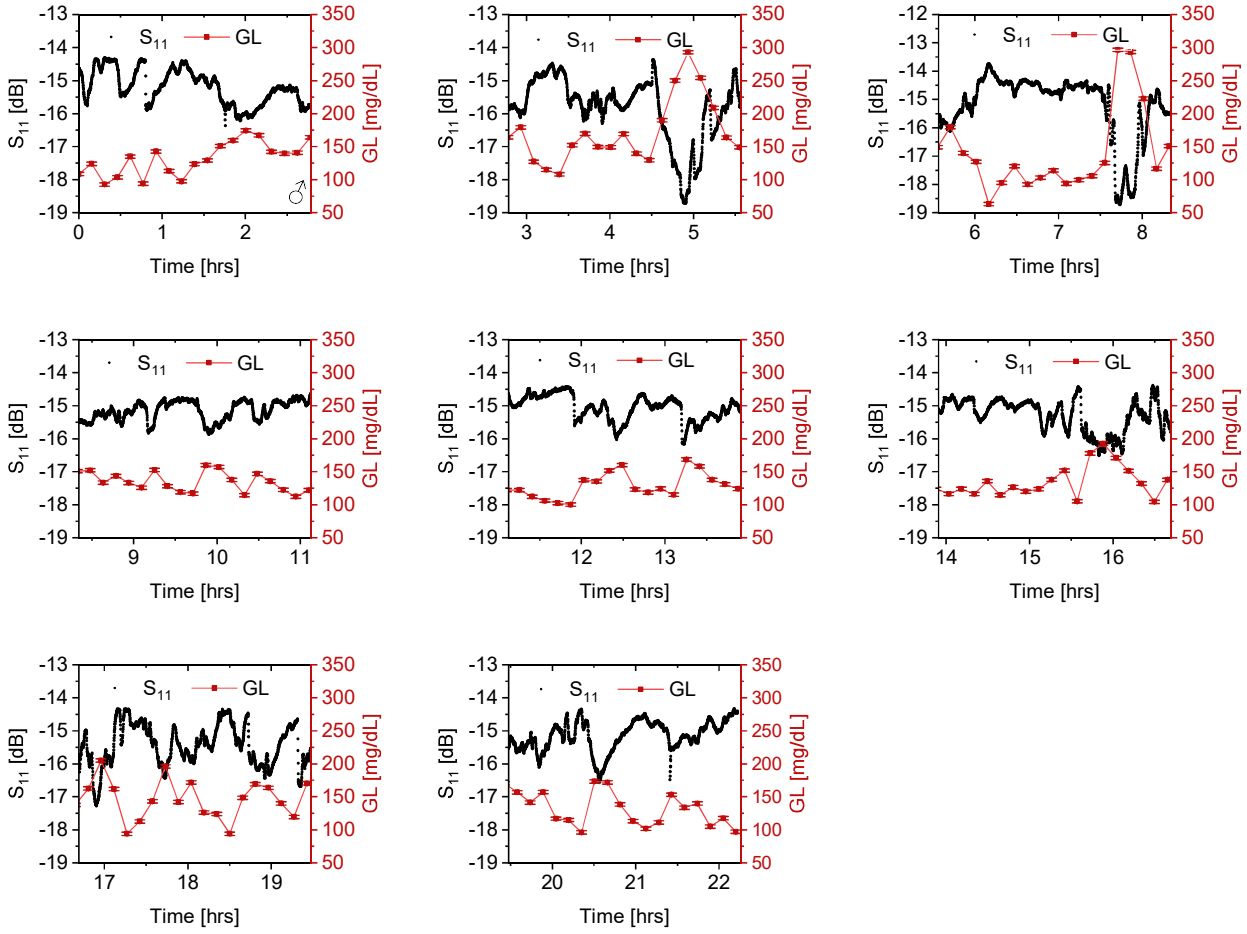


Figure A.10: Comparison between S_{11} values of the patch sensor and commercial glucometer for a male diabetic type 1 participant.

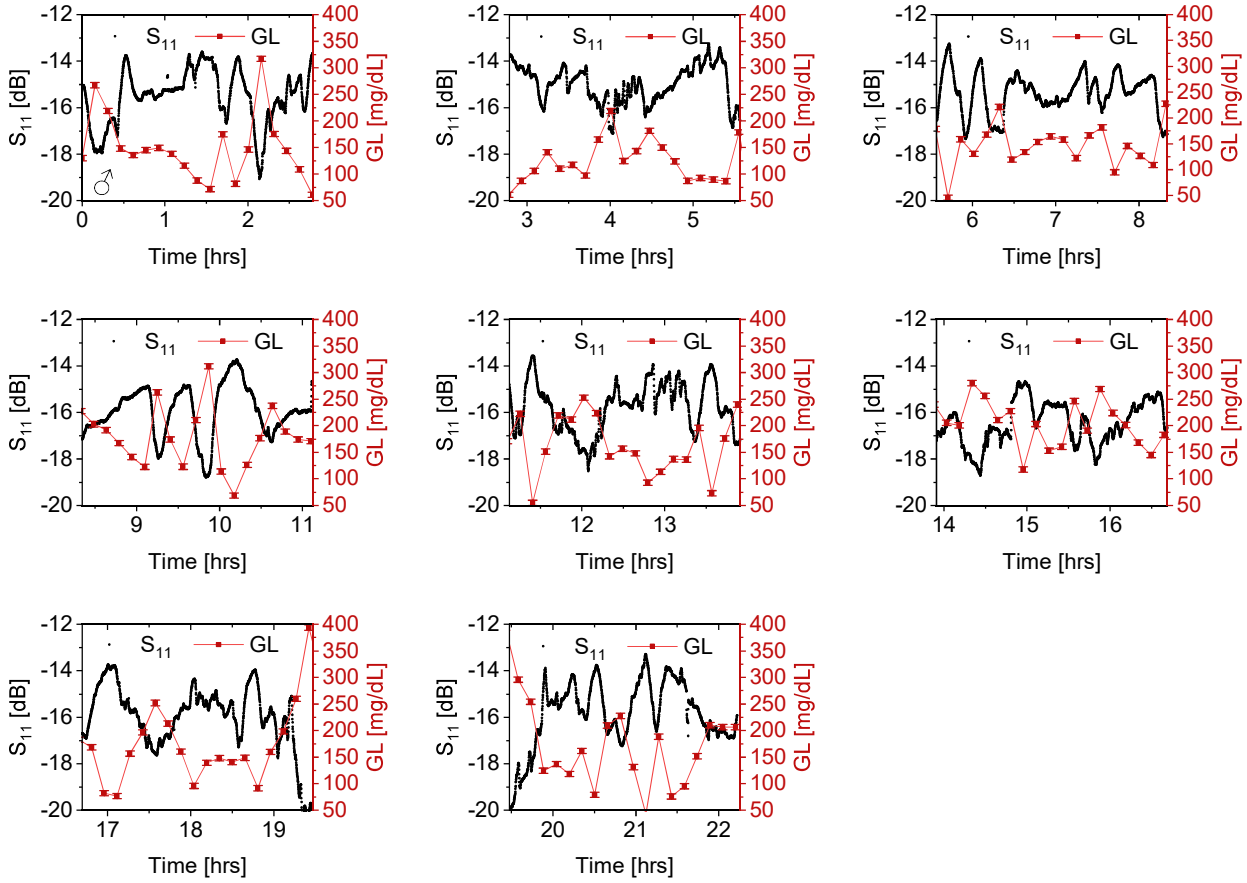


Figure A.11: Comparison between S_{11} values of the patch sensor and commercial glucometer for a male diabetic type 2 participant.

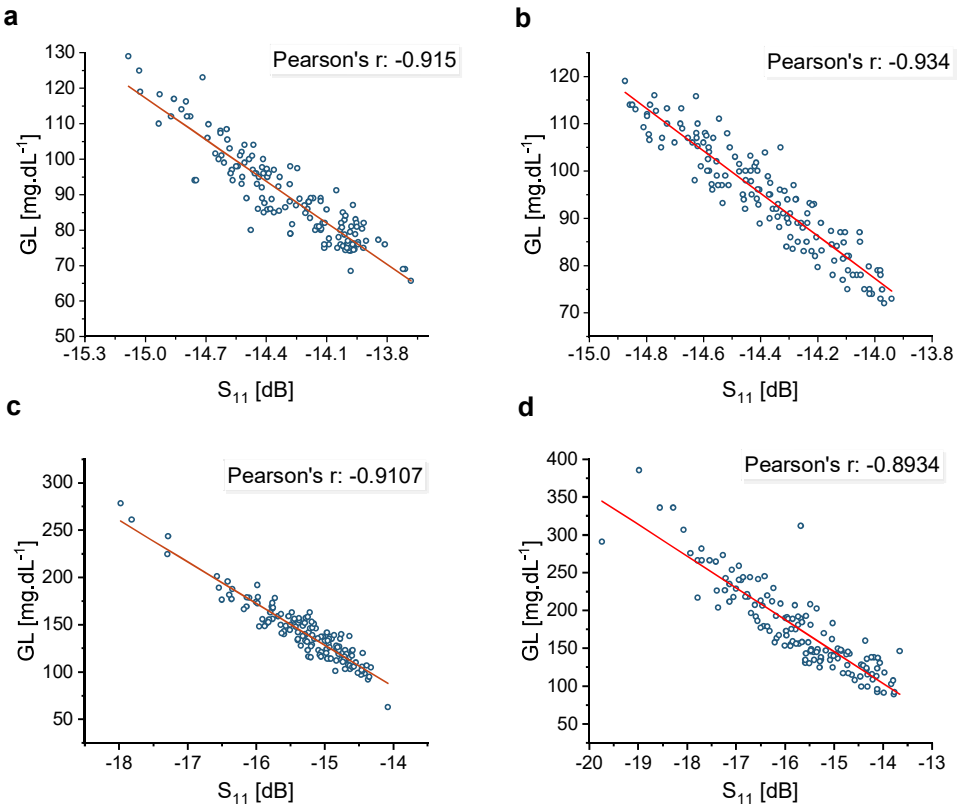


Figure A.12: **Pearson's correlation between S_{11} values and corresponding Accucheck glucometer.** The inverse relation between MW sensor response and commercial glucometer is modelled with a linear fit. Pearson's correlation factor is calculated for each person. It is shown that with a proper calibration per person **a** Non-diabetic (female) **b** Non-diabetic (male) **c** diabetic-typeI (male) **d** diabetic-typeI (male) a high correlation is achieved.

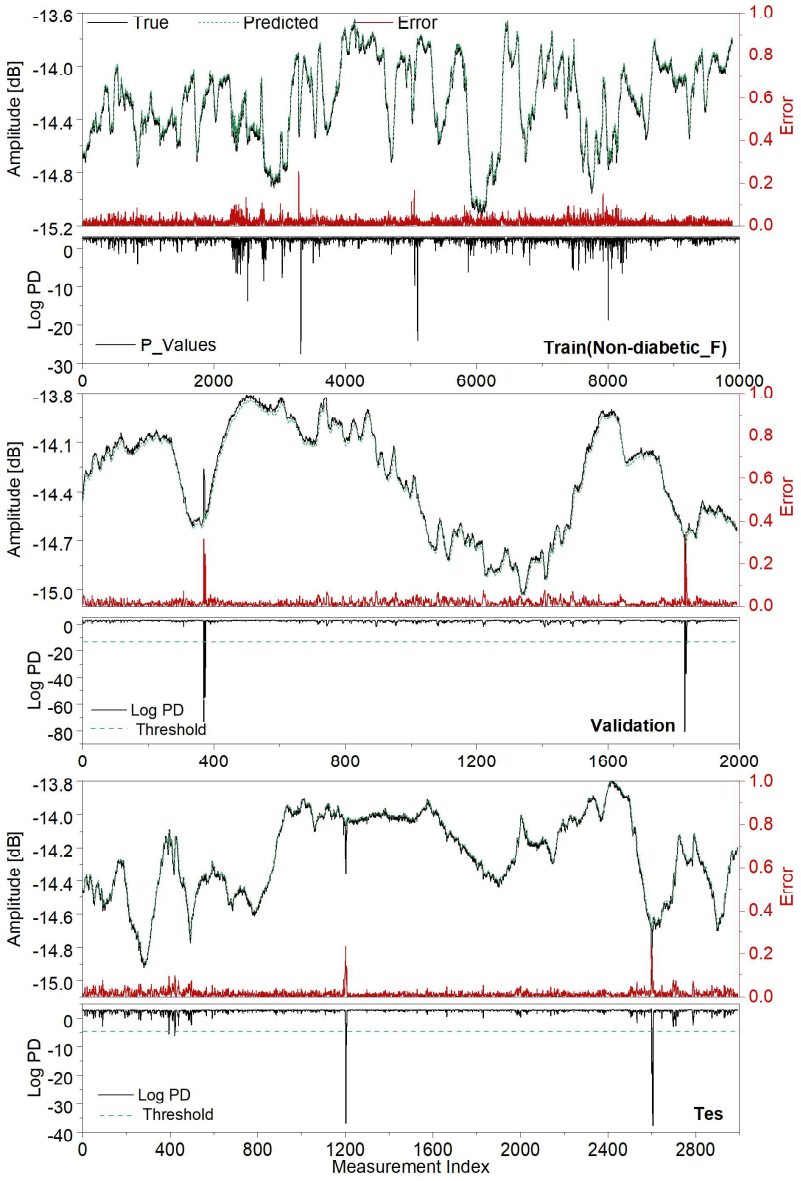


Figure A.13: **Measured sensor response of person 1.** The reflection parameter of the sensor is used to record the glucose level in human trials. The sensor response is then predicted using LSTM with error function and logPD values. The results for the validation and test sets are shown separately.

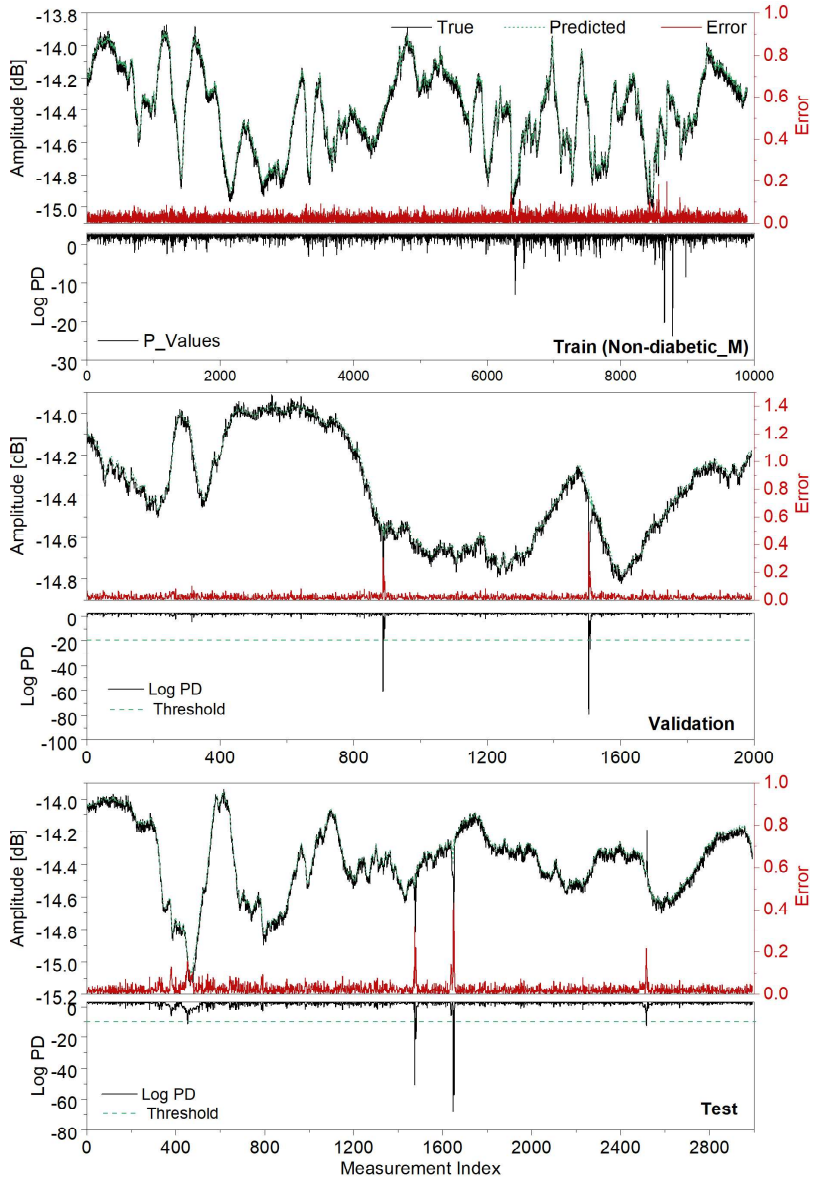


Figure A.14: **Measured sensor response of person 2.** The reflection parameter of the sensor is used to record the glucose level in human trials. The sensor response is then predicted using LSTM with error function and logPD values. The results for the validation and test sets are shown separately.

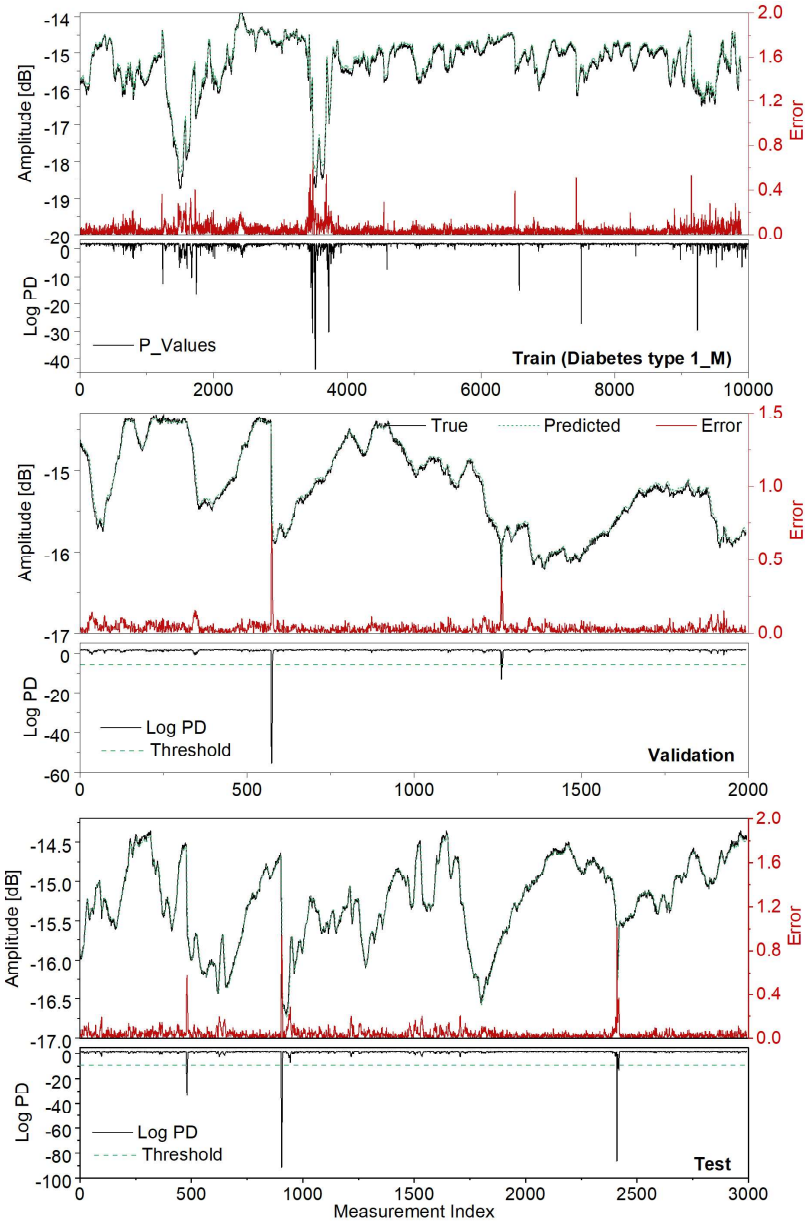


Figure A.15: **Measured sensor response of person 3.** The reflection parameter of the sensor is used to record the glucose level in human trials. The sensor response is then predicted using LSTM with error function and logPD values. The results for the validation and test sets are shown separately.

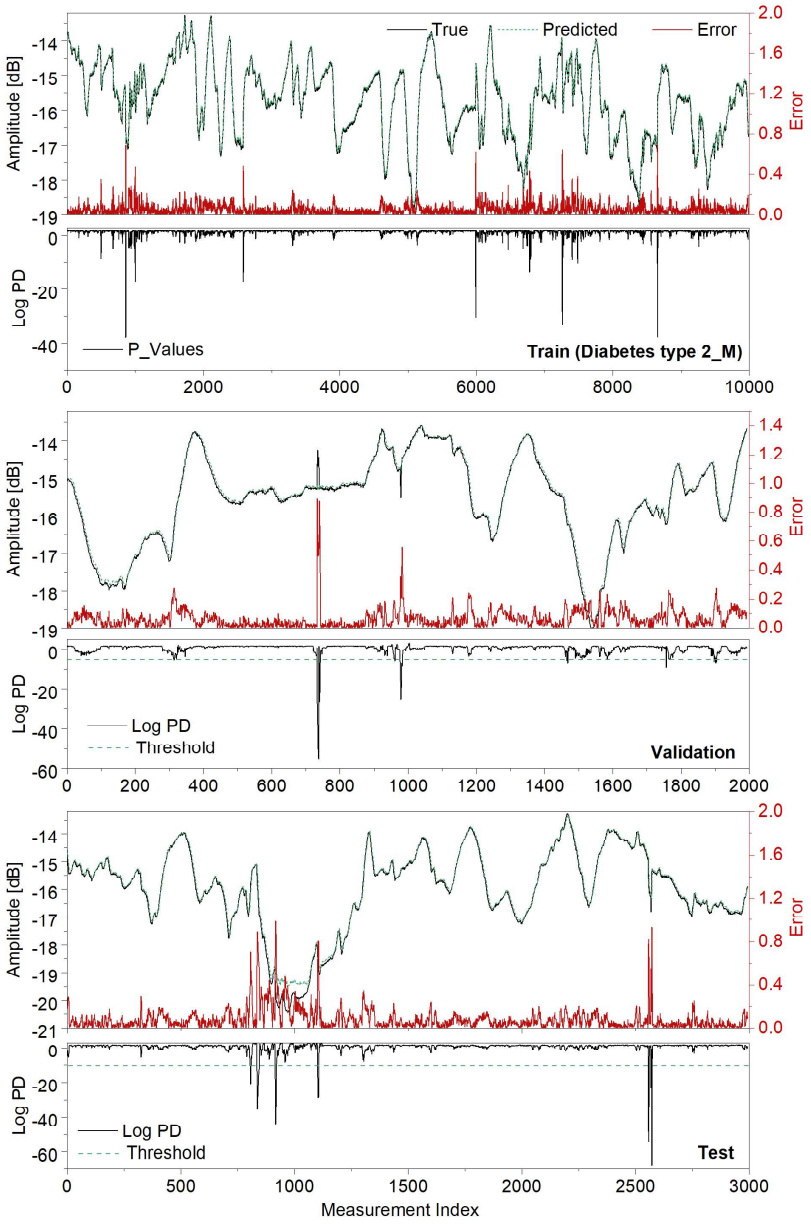


Figure A.16: **Measured sensor response of person 4.** The reflection parameter of the sensor is used to record the glucose level in human trials. The sensor response is then predicted using LSTM with error function and logPD values. The results for the validation and test sets are shown separately.

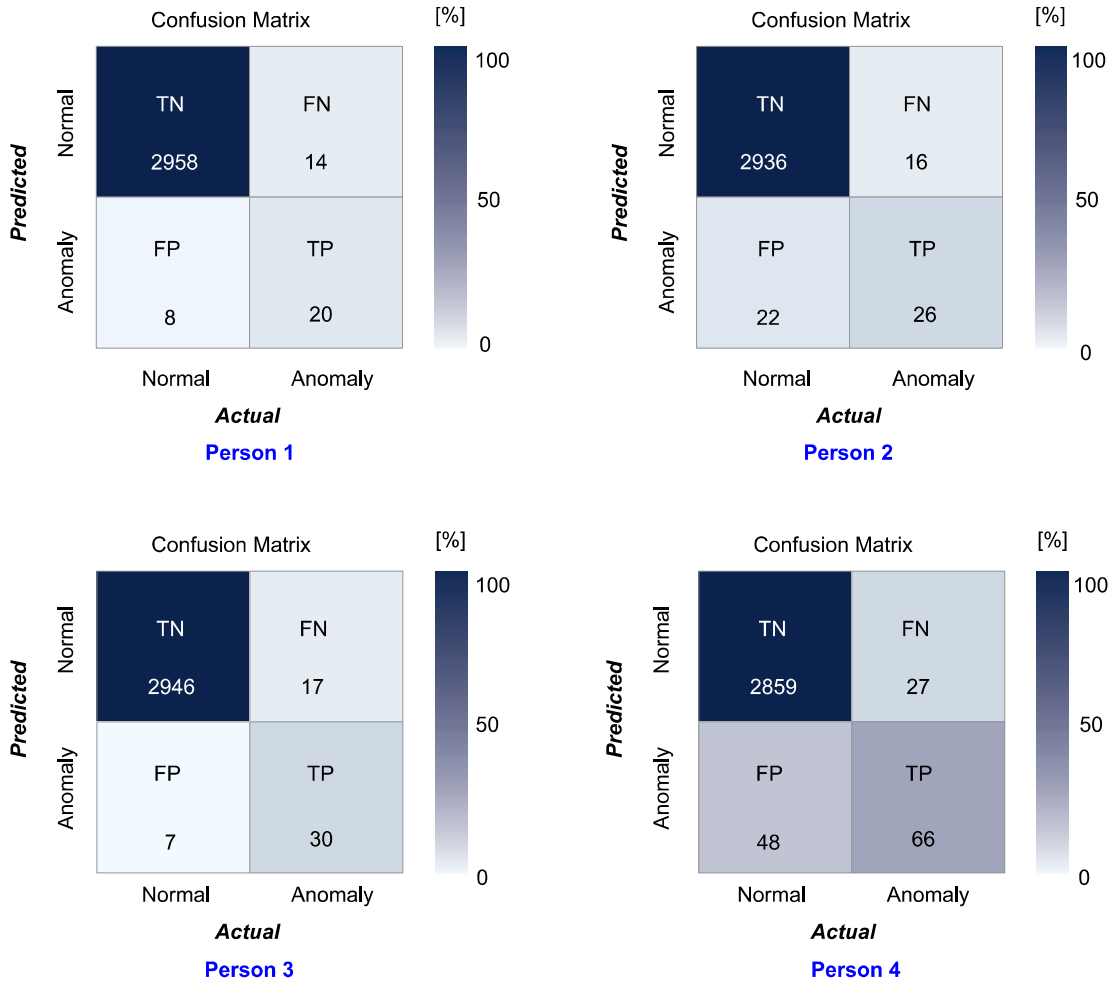


Figure A.17: **Confusion matrix for anomaly detection on test data set for all individuals $P_1 \rightarrow P_4$.** The confusion matrix functions as an extensive assessment instrument, covering four essential classifications: 'True Positives' (TP), 'False Positives' (FP), 'False Negatives' (FN), and 'True Negatives' (TN). In this context, TP and TN represent accurately distinguished anomalies and typical data, correspondingly. Conversely, FP and FN denote erroneous instances where conventional data is misconstrued as an anomaly and instances where anomalies are mistakenly recognized as normal data, correspondingly.

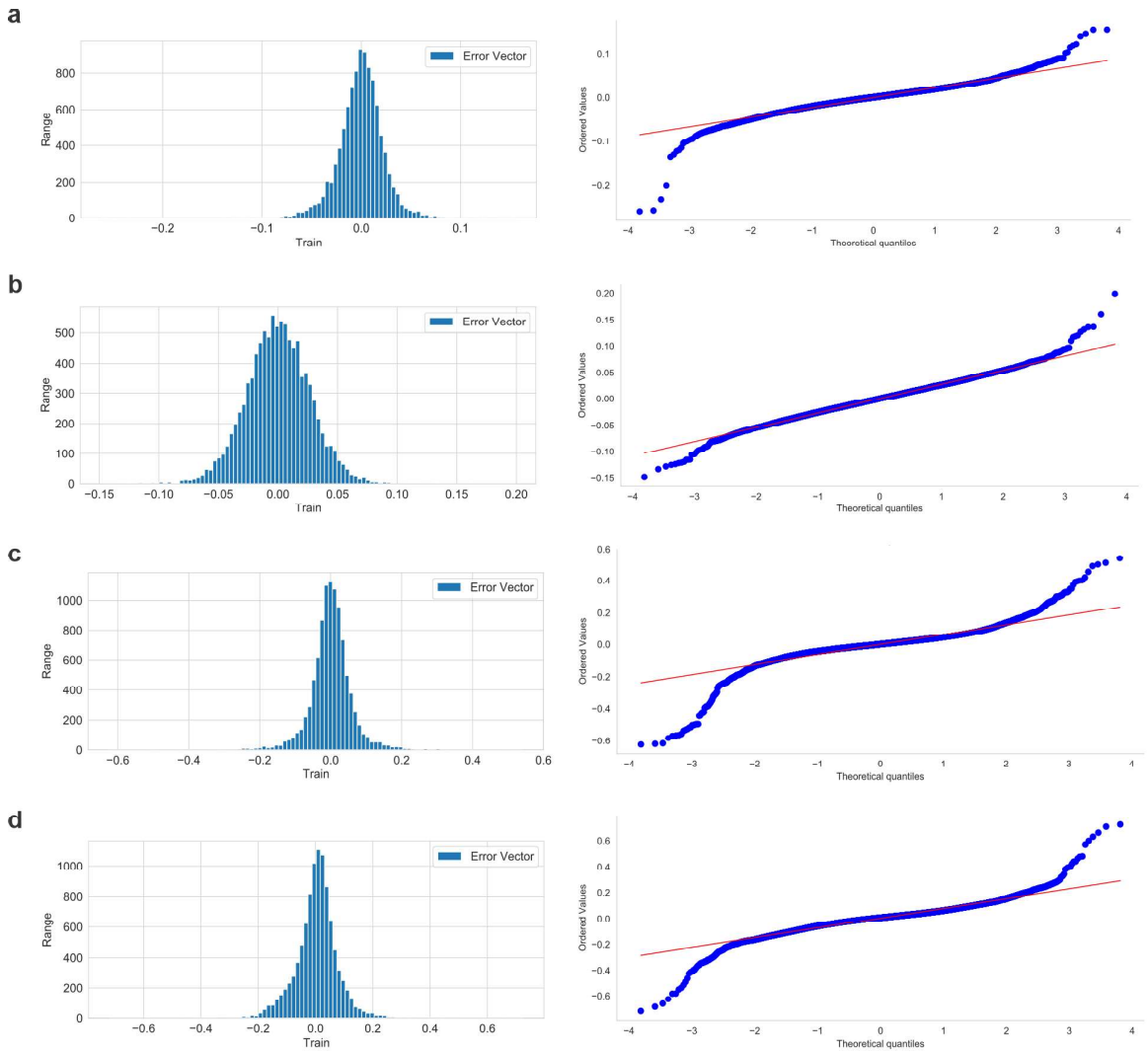


Figure A.18: **Human trial measurement and prediction error analysis.** The LSTM model is trained on each person's dataset for 100 epochs and the resultant histogram of train error vectors is plotted on the left. Next, a normal distribution is comprised based on the mean and standard deviation of the train error vector. The Quantile-Quantile (Q-Q) plot summarizes the relationship between the error vector values and the expected normal distribution, which showcases a high similarity for a wide range of error vectors.

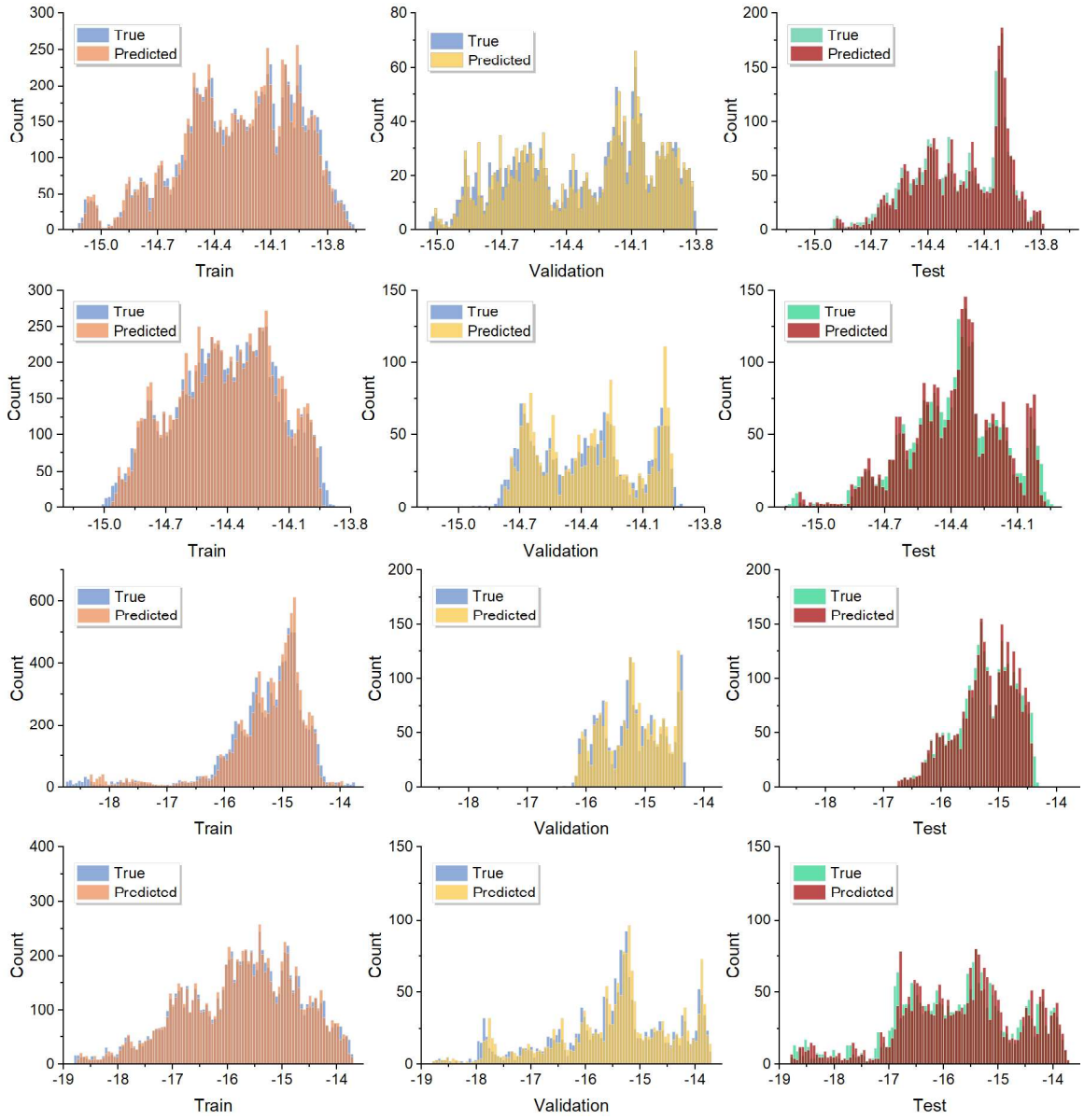


Figure A.19: **Human trial measurement and prediction histogram.** Measured S_{11} amplitude values, as well as the corresponding predicted values for each person divided into three sections of train/validation/test, sets with 10000/2000/3000 datapoints, respectively. It is clear that the predicted histograms are highly aligned with that of the measured values.