Application of a deep learning model to determine

midpalatal suture maturation stage on CBCT

BY

Mahshid Nik Ravesh

A thesis submitted in partial fulfillment of the

requirements for the degree of

Master of Science

Medical Sciences - Dentistry

University of Alberta

ABSTRACT

Transverse maxillary deficiency is a condition characterized by a reduced transverse dimension of the upper jaw, commonly associated with posterior cross-bite, dental crowding, pharyngeal airway narrowing, and mouth breathing. Accurate staging of the mid-palatal suture (MPS) fusion is crucial for determining the appropriate treatment approach, whether surgical or non-surgical maxillary expansion. Traditionally, MPS staging is performed using cone beam computed tomography (CBCT), a technique that heavily relies on the practitioner's experience and is inherently subjective, leading to potential variability in assessment and treatment decisions.

This study addresses these challenges by automating the classification and staging of MPS fusion through advanced deep learning (DL) techniques. We developed and trained both 2D and 3D convolutional neural network (CNN) models to enhance the accuracy, efficiency, and consistency of MPS evaluation. The 2D CNN model demonstrated remarkable performance with a high-test accuracy of 96.49% and excellent precision, recall, and F1-score values across all classification stages (AB, C, DE). This model highlights the effectiveness of traditional 2D approaches in handling MPS classification tasks.

In contrast, the 3D CNN model, designed to capture the volumetric information of the MPS, achieved a test accuracy of 78.26%. Although this accuracy is lower compared to the 2D

model, the 3D approach offers a more comprehensive evaluation by considering the full spatial context of the MPS, which could lead to more accurate staging in complex cases. The performance metrics for precision, recall, and F1-score in the 3D model were found to be acceptable, underscoring its potential for future refinement and optimization.

The findings from this study underscore the potential of DL methods to revolutionize MPS fusion assessment by providing a more reliable, objective, and repeatable classification system. Such advancements could significantly enhance orthodontic treatment planning by offering clinicians a powerful diagnostic tool, improving patient outcomes, and reducing variability in treatment approaches. Furthermore, the techniques developed in this research have broader implications for medical image classification tasks beyond orthodontics, paving the way for the integration of AI-driven solutions in various medical imaging applications.

PREFACE

This thesis is an original work by Mahshid Nik Ravesh. The research projects, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name "Application of deep learning models to determine mid-palatal suture maturation stage on CBCT", No. Pro00125920, on 12/20/2022

DEDICATION

To my husband, who has been a great source of motivation and comfort to me.

ACKNOWLEDGEMENTS

I would like to express gratitude to my supervisor, Dr. Hollis Lai, for introducing me to the world of artificial intelligence and guiding me on how to navigate it safely and intelligently. Your expertise and mentorship have been invaluable throughout this journey.

To Dr. Manuel Lagravere, my other supervisor, thank you for keeping my passion for orthodontics alive. Your unwavering support and patience in reviewing my countless CBCTs made me feel incredibly fortunate to be part of this research team. Your dedication has been an inspiration.

I am also profoundly grateful to my committee member, Dr. Ida Kornerup, for being the kindest and most gentle presence at school. Your compassion and support have meant more to me than words can express.

A special thank you to Dr. Silvia Cpekanas, who generously reviewed my CBCTs during her breaks, always providing valuable insights and never hesitating to answer my many questions.

To my classmate and brilliant friend, Dr. Victor Ladewig, thank you for spending endless hours classifying CBCTs with me and for always motivating me to keep going. Your camaraderie has been essential to this work.

Lastly, to my best friend and mentor, Dr. Nazila Ameli, thank you for being the coolest PhD student and orthodontist. Your help in reading and editing my thesis, along with the countless lessons you've taught me, has been indispensable. I owe much of what I know to you.

I would also like to acknowledge the financial support I received from the Mike Petryk School of Dentistry, which provided me with Fund for Dentistry 2023. This support has been crucial in allowing me to complete this research.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF COMMON ABBREVIATIONS

Abbreviation	Definition
MPS	Mid-Palatal Suture
MTD	Maxillary Transverse Deficiency
CNN	Convolutional Neural Network
DL	Deep Learning
ROI	Region Of Interest
CBCT	Cone Beam Computed Tomography

Chapter 1- Introduction and Background

1.1 introduction

Craniofacial growth and development are complex processes influenced by genetic, environmental, and biomechanical factors. (Roosenboom, Hens et al. 2016) Cranial sutures are fibrous joints connecting the bones of the skull and they are among the critical structures involved in the craniofacial growth and development. These sutures play a vital role in the growth and expansion of the craniofacial complex, allowing for the necessary flexibility and movement during early development and enabling the skull to accommodate the growing brain. (Moazen, Peskett et al. 2015) One of these cranial sutures is the midpalatal suture (MPS) which is crucial for maxillary growth, influencing the overall symmetry and function of the craniofacial skeleton. (Manjusha, Jyothindrakumar et al. 2017)

The midpalatal suture, located in the center of the palate, is a key site for orthodontic intervention of maxillary expansion. During adolescence, the midpalatal suture remains patent, meaning it has not yet fused, allowing for the possibility of non-surgical expansion through orthodontic appliances such as rapid palatal expanders.(Haghanifar, Mahmoudi et al. 2017) This approach is essential for correcting transverse maxillary deficiencies, which, if left untreated, can lead to various malocclusions, dental crowding, and even functional issues like obstructive sleep apnea.(Jimenez-Valdivia, Malpartida-Carrillo et al. 2019, Akbulut, Bayrak et al. 2020). Understanding the maturation stage of the midpalatal suture is, therefore, crucial for determining the appropriate timing and method of intervention, as the suture's fusion marks the transition from a malleable to a more rigid state, limiting the effectiveness of non-surgical treatments. (Sayar and Kılınç 2019)

Accurately assessing the maturation stage of the midpalatal suture is paramount for optimizing orthodontic outcomes in expansion treatments. (Silva-Montero, Faus-Matoses et al. 2022, Erten, O., & Yılmaz, B. 2018) The ability to predict the suture's stage of fusion more accurately could lead to more targeted and effective interventions, reducing the risk of relapse and improving long-term stability of the treatment outcomes. As orthodontics continues to move toward personalized care, there is a growing need for aiding tools and techniques that can enhance the precision of diagnostic and treatment planning processes.

Deep Learning (DL) has emerged as a powerful tool in medical diagnostics, offering the potential to revolutionize the way clinicians assess and predict clinical outcomes.(Glanz, Dudenkov et al. 2022, Mishra, Tiwari et al. 2023, Abou Ghayda, Cannarella et al. 2024) In orthodontics, DL can be employed to analyze complex imaging data, such as Cone Beam Computed Tomography (CBCT) scans, to identify patterns and features indicative of midpalatal suture maturation. (Zhu, Yang et al. 2024) By training algorithms on CBCT images, DL models can learn to recognize subtle differences in suture morphology that may not be apparent to the human eye (Ashkani-Esfahani, S., Mojahed-Yazdi, R., et al. 2022), thus providing a more accurate and reliable assessment. This advancement could significantly improve decision-making in orthodontics, leading to better patient outcomes and more efficient use of resources.

The aim of this thesis is to enhance orthodontic diagnosis through the integration of DL techniques. Specifically, it seeks to develop a classification tool that accurately determines the maturation stage of the MPS using CBCT images. By employing DL to analyze these images, the tool will assist orthodontists in providing individualized diagnoses and tailored treatment plans based on each patient's unique needs. This is done by:

Introducing a convolutional neural network (CNN) architecture for classifying two-dimensional MPS images.

Introducing a convolutional neural network (CNN) architecture for classifying three-dimensional MPS images.

To achieve these objectives, the following automated methods were developed and implemented:

1. Employing a CNN model to detect CBCT slices containing the MPS for subsequent classification.
2. Designing a 2D CNN model for individual MPS image classification.
3. Creating three-dimensional (3D) arrays of CBCT images to encompass the entire suture.
4. Designing a 3D CNN model for MPS classification.

1.2 Background

1.2.1 Malocclusion

1.2.1.1 Definition

Malocclusion, defined as any systematic abnormal relationship of upper and lower dental arches, has been extensively studied in orthodontics. Historically, malocclusion has been classified primarily by dental features. Angle's classification, which is the most widely used system, focuses on anterior-posterior discrepancy, categorizing occlusions into Class I, Class II, and Class III (Mageet 2016). This system is primarily dental in nature, concentrating on the alignment and relationship of the dental arches. In contrast, skeletal classifications, which consider a broader range of parameters, have been introduced to provide a more comprehensive understanding of malocclusions. For example, Ackerman and Proffit (1969) introduced a very comprehensive system of classification. This classification considered five characteristics and their inter-relationships: alignment, profile, transverse, class, and overbite. By incorporating multiple dimensions, including the previously overlooked transverse discrepancy of the jaws, this system provided a more holistic approach to understanding malocclusions (Ackerman and Proffit 1969) (Koo, Choi et al. 2017).

1.2.1.2 Maxillary Transverse Discrepancy

Regardless of the classification system used, transverse discrepancies between the upper and lower jaw constitute approximately 30% of all orthodontic cases (Sawchuk, Currie et al. 2016) (Ackerman and Proffit 1969) Maxillary transverse deficiency (MTD), characterized by a reduced width of the upper jaw, is a significant concern (McNamaraa 2000, Calvo-Henriquez, Megias-Barrera et al. 2021)

The normal transverse relationship between the maxilla and mandible involves a harmonious alignment where the maxillary arch width is proportionate to the mandibular arch, ensuring proper occlusion and esthetics. Ideal ratios for esthetic transverse relationships include an intercanine width/oral fissure breadth ratio of 0.638 and an oral fissure breadth/interparopia width ratio of 0.617. Reduced maxillary width is considered one of the primary causes of upper dental crowding and tooth protrusion, which are common complaints among patients seeking orthodontic treatment. (McNamaraa 2000) (Howes 1957, Nimkarn, Miles et al. 1995)

1.2.1.3 Prevalence of adversities associated with MTD

Posterior crossbite is defined as the irregular bucco-lingual or buccal-palatal relationship of one or more posterior teeth with opposing teeth in centric occlusion. It affects a significant portion of orthodontic patients and the general population (Iodice, Danzi et al. 2016). Its prevalence, particularly in primary and early mixed dentition, is reported to be between 8% to 22% among orthodontic patients and 5% to 15% in the general population The most frequent cause of posterior crossbite is the reduced width of the maxillary arch.(Iodice, Danzi et al. 2016) Furthermore, MTD can be associated with other orofacial adversities such as alterations in tongue posture, mouth breathing with a prevalence of 11 to 56%, and pharyngeal airway narrowing, which can contribute to conditions like sleep apnea(Akbulut, Bayrak et al. 2020) (Jimenez-Valdivia, Malpartida-Carrillo et al. 2019). These associations have a negative impact on oral health-related quality of life (OHRQL) and are significant reasons why patients seek orthodontic treatments (Kiyak 2008).

1.2.1.4 Etiology of Maxillary Transverse Deficiency

The etiology of MTD involves both environmental factors and genetic predispositions. Environmental influences such as thumb sucking, pacifier use, mouth breathing, and abnormal tongue functions can impact maxillary development, contributing to the narrowing of the upper jaw characteristic of MTD. (Darawsheh, Kolarovszki et al. 2023) Additionally, genetic predispositions are significant contributors to MTD, often appearing within broader craniofacial syndromes or conditions such as cleft lip and palate. These genetic factors further exacerbate the structural abnormalities associated with MTD, highlighting the multifactorial nature of its etiology. (Nervina, Kapila et al. 2014, Duque-Urióstegui, Reyes et al. 2023)

1.2.1.5 Diagnosis of MTD

Diagnosis of MTD is achieved through a series of clinical, model and radiographic evaluations. One of the examples of clinical evaluations is how maxillary arch form and symmetry are examined by clinicians to identify any deviations or irregularities. A well-developed arch with appropriate symmetry suggests normal transverse dimensions, whereas asymmetry or constricted arch form may indicate MTD. (Dakhil and Salamah 2021)

Assessment of the palatal vault shape provides valuable insights into transverse discrepancies and is another example of clinical evaluations for diagnosis of MTD. A narrow or V-shaped vault is indicative of transverse deficiency, whereas a broader U-shaped vault is associated with normal transverse dimensions. (Dakhil and Salamah 2021)

The occlusal relationship between the upper and lower arch is also evaluated to detect any signs of crossbite, unilateral or bilateral posterior crossbite being a common manifestation of MTD (Duque-Urióstegui, Reyes et al. 2023). Additionally, signs of dental compensations, such as buccal or lingual displacement of teeth, which may indicate underlying transverse deficiency is also helpful in diagnosing MTD. (Kim, Park et al. 2021, Duque-Urióstegui, Reyes et al. 2023)

Predominant breathing mode, whether oral or nasal, is also assessed, as mouth breathing is often associated with MTD. Mouth breathing can contribute to maxillary constriction and exacerbate transverse deficiencies. (Darawsheh, Kolarovszki et al. 2023) Other steps and evaluations used in diagnosing MTD is as follows:

Anteroposterior Cephalograms: These radiographs provide valuable information regarding skeletal relationships and transverse dimensions. Measurements such as the maxillomandibular width differential and maxillomandibular transverse differential index help quantify transverse deficiencies and guide treatment planning.

Occlusal Radiographs: These radiographs were previously used to visualize the MPS (Wehrbein and Yildizhan 2001) but lacked clarity and accuracy as they provided a 2D view of a potentially thin structure embedded deep in palatal bone. (Forst, Nijjar et al. 2014)

Cone-Beam Computed Tomography (CBCT) Scans: CBCT scans were eventually introduced into dentistry and offer detailed 3D imaging of the craniofacial region, allowing for precise assessment of skeletal structures and transverse dimensions and becoming crucial for diagnosis and treatment planning. (13) Cross-sectional evaluation of CBCT images facilitates the identification of asymmetries and the localization of discrepancies as well as providing a clearer visualization of the MPS. (Zahra and Samih 2017) (Mahdian, Shaykhian et al. 2019) (Katti, Shahbaz et al. 2020)

Model Analysis: Study models allow for direct measurement of dental arch dimensions and evaluation of tooth inclinations. They are however not effective and limited in diagnosing the skeletal component of the MTD (Sawchuk, Currie et al. 2016)

Understanding the implications of MTD and its associated conditions is crucial for effective orthodontic management and improving patient outcomes (Nowak and Zawiślak 2016).

1.2.1.6 Treatments of MTD

Treatments of MTD include surgical and non-surgical approaches that are implemented by orthodontists based on several clinical indications, including the degree of correction needed, whether skeletal or dentoalveolar correction is necessary, and the perceived effectiveness of expansion based on the maturation phase (closure) of the MPS(Hernández-Alfaro and Valls-Ontañón 2021) (Hernandez-Alfaro, Bueno et al. 2010) (Franchi, Baccetti et al. 2010) Failure to accurately stage the MPS maturation can lead to patients' with fully mature MPS receiving non-surgical palatal expansion treatments. These treatments are incapable of treating MTD while causing pain, mucosal ulceration or necrosis, accentuated buccal tipping and gingival recession of posterior teeth. Likewise, wrongfully prescribing surgical treatment to patients who are still in non-surgical MPS stages can lead to expensive and lengthy treatments. (Suri and Taneja 2008, Agarwal and Mathur 2010) Thus, the first critical step in treatment planning for MTD is correctly classifying the maturation stage of the MPS. (Isfeld, Flores-Mir et al. 2019) (Isfeld, Lagravere et al. 2017)

1.2.1.6.1 Non-surgical Treatments

Non-surgical treatments for maxillary transverse deficiency aim to address the reduced transverse dimension of the upper jaw without the need for invasive procedures. Among the most effective methods are Rapid Maxillary Expansion (RME), Slow Maxillary Expansion (SME), and Mini-Implant Assisted Rapid Palatal Expansion (MARPE).

1.2.1.6.1.1 Rapid Maxillary Expansion (RME)

One of the non-surgical treatments for maxillary transverse deficiency, primarily RME, relies on understanding the maturation stages of the MPS. (Haghanifar, Mahmoudi et al. 2017) RME corrects transverse discrepancies by applying force to stimulate the separation of the MPS, promoting collagenous fiber stretching and new bone formation, ultimately increasing the transverse width of the maxilla. Its effectiveness is influenced by the maturation stage of the MPS. RME is particularly effective in patients where the MPS has begun to ossify but is not yet fully fused. (Haghanifar, Mahmoudi et al. 2017, Shayani, Merino-Gerlach et al. 2023)

Moreover, the therapeutic effects of RME extend beyond dental alignment, potentially improving nasal breathing and facial aesthetics. (Kwak, Kim et al. 2016) Thus, a comprehensive understanding of MPS maturation stages and treatment considerations is essential in managing maxillary transverse deficiency with non-surgical approaches.

1.2.1.6.1.2 Slow Maxillary Expansion (SME)

SME is a non-surgical method used to correct transverse maxillary deficiencies by gradually widening the maxilla. Unlike RME, which can cause significant dental tipping and other side effects, SME aims to achieve more controlled and stable skeletal changes. SME is generally recommended for patients with a less mature MPS, where the suture is still relatively open and requires less force for separation. (Zong, Tang et al. 2019)

SME involves the use of a maxillary expansion appliance that applies gentle, continuous pressure to the maxillary sutures, promoting gradual bone remodeling. The expansion rate is typically slower, with follow-up intervals shortened to one week to monitor progress and manage any adverse events.

As with RME, in SME a comprehensive understanding of MPS maturation stages and treatment considerations is essential in managing MTD with this approach.

1.2.1.6.1.3 Micro-implant assisted Rapid Palatal Expansion (MARPE)

MARPE is a contemporary orthodontic treatment that addresses MTD, particularly in skeletally mature patients. Central to its efficacy is the understanding of the maturation stage of the MPS, which significantly influences both its approach and effectiveness. As the MPS matures, its degree of ossification and fusion increase, necessitating greater force for successful expansion. (Brunetto, Sant'Anna et al. 2017, Nguyen 2017)

The success of MARPE in patients with more mature MPS, may require adjustments in protocol and appliance design. The increased resistance posed by advanced suture maturation necessitates careful planning and customization to ensure optimal outcomes. (de Melo Quintela, Rossi et al. 2021) (Brunetto, Sant'Anna et al. 2017)

Clinical studies have reported a high rate of suture separation and stable long-term results with MARPE, indicating its efficacy in achieving skeletal expansion.(Jia, Zhuang et al. 2021, Chun, de Castro et al. 2022) (Zong, Tang et al. 2019) (Brunetto, Sant'Anna et al. 2017)

It has been suggested that future research endeavours should focus on optimizing MARPE protocols and appliance designs to enhance outcomes in patients with advanced suture maturation by determining the accurate suture maturation stage. By refining techniques and developing tailored approaches, orthodontic practitioners can continue to improve the effectiveness and applicability of MARPE in treating MTD in skeletally mature patients. (Brunetto, Sant'Anna et al. 2017, Zong, Tang et al. 2019, de Melo Quintela, Rossi et al. 2021)

1.2.1.6.2 Surgical

MTD poses a significant challenge in orthodontic treatment in skeletally mature patients where nonsurgical methods may prove insufficient. (Rachmiel, Turgeman et al. 2020, Shih, Ho et al. 2022) Surgical approaches, including Surgically Assisted Rapid Palatal Expansion (SARPE), segmental LeFort I osteotomy, and mandibular midline osteotomy with constriction, offer effective solutions tailored to the severity of the deficiency and the patient's skeletal maturity.

(Bloomquist and Joondeph 2019)

SARPE serves as a primary option for skeletally mature patients with substantial transverse deficiencies. This technique involves osteotomies to reduce resistance at the MPS, facilitating bone-borne expansion. Typically, SARPE is followed by orthodontic treatment to stabilize occlusion. While SARPE is reliable in expanding the maxillary arch in adults, it may present complications such as hemorrhage, pain, and relapse, and it may also involve higher costs due to the complexity of the procedure. (Brunetto, Sant'Anna et al. 2017, Rachmiel, Turgeman et al. 2020)

Segmental LeFort I osteotomy in another surgical option that addresses patients requiring less extensive transverse corrections while allowing for additional adjustments in the vertical or sagittal dimensions. This surgical approach enables simultaneous alterations in transverse, vertical, and sagittal dimensions, making it suitable for complex cases. (Nowak, Strzałkowska et al. 2015, Dakhil and Salamah 2021)

The choice between SARPE and LeFort I osteotomy is influenced by a combination of factors, including the severity of the deficiency, the maturity of the MPS, patient-specific anatomical considerations, and the risks and costs associated with each procedure. These considerations are critical in determining the most appropriate surgical intervention for managing maxillary transverse deficiency in skeletally mature patients. (Nowak, Strzałkowska et al. 2015, Dakhil and Salamah 2021)

1.2.1.7 The Consequences of Untreated MTD

Untreated MTD can have significant consequences, particularly concerning the pharyngeal airway and its relation to sleep apnea. Pharyngeal airway narrowing, which can be associated with MTD, along with altered tongue posture and mouth breathing, are factors that can contribute to the development of sleep apnea. (Jimenez-Valdivia, Malpartida-Carrillo et al. 2019) Sleep apnea not only impacts overall health but also has a notable effect on oral health-related quality of life (OHRQL). This condition often serves as a motivator for patients to seek orthodontic treatment, as it is closely linked to various oral health issues. (Kiyak 2008)

Additionally, untreated MTD can lead to orofacial and functional impairments. One of the common consequences of untreated posterior crossbite, which can result in significant discrepancies in the transverse dental arch relationship. This type of malocclusion negatively affects occlusion and dental alignment, leading to functional problems such as improper chewing and aesthetic concerns due to misaligned teeth. (Akbulut, Bayrak et al. 2020)

Furthermore, the long-term stability of the dental arch and occlusion may be compromised if MTD is left untreated. This increases the risk of relapse and may necessitate more complex orthodontic or surgical interventions later in life (Akbulut, Bayrak et al. 2020). These insights emphasize the critical importance of early diagnosis and treatment of MTD to prevent long-term adverse effects on dental health and overall well-being.

1.2.2 Midpalatal Suture

The MPS is a crucial suture in the roof of the mouth, connecting the palatine and maxillary bones. (Angelieri, Franchi et al. 2016, Shayani, Merino-Gerlach et al. 2023)Its visibility and morphology vary depending on its maturation or interdigitation stage, which can be assessed using Cone Beam Computed Tomography (CBCT). The maturation stages range from Stage A to Stage E, each characterized by distinct features.(Angelieri, Franchi et al. 2016) (Kwak, Kim et al. 2016)

During Stage A, the MPS appears as a straight high-density line with minimal interdigitation. As maturation progresses, it becomes irregular and scalloped in Stage B. Stage C the suture presents as two parallel, scalloped lines close to each other. Fusion initiates in the palatine bone during Stage D, with complete fusion occurring in both the palatine and maxillary portions by Stage E, rendering the suture invisible. (Angelieri, Franchi et al. 2016)

The treatment for stages A and B is clinically similar while they can differ from stage C, and the treatment for stage D and E greatly differ from stages A, B, and C. (Ladewig, Almeida-Pedrin et al. 2022) (de Miranda Ladewig, Capelozza-Filho et al. 2018)

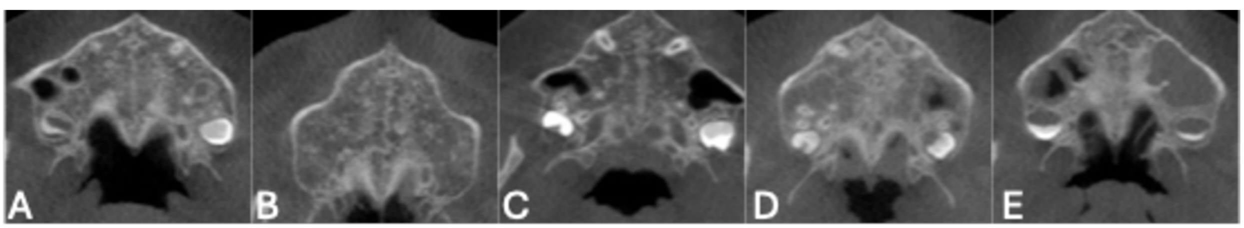

Figure 1 CBCT Axial images showing MPS maturation stages. A: stage A, B: stage B, C: stage C, D: Stage D, E: stage E

The treatment approaches for Stages A and B primarily involve slow or rapid maxillary expansion methods.(Silva-Montero, Faus-Matoses et al. 2022) These stages benefit from the pliability of the suture, allowing for effective skeletal changes.(Gao, Chen et al. 2022) However, as the suture progresses to Stage C, where it becomes more interdigitated, the treatment methods begin to differ, requiring more precise assessment to avoid unnecessary trauma.(Sayar and Kılınç 2019)

For Stages D and E, where the suture has largely or fully ossified, treatment differs significantly from the earlier stages. Surgically assisted rapid palatal expansion (SARPE) or miniscrew-assisted rapid palatal expansion (MARPE) may be necessary to achieve the desired expansion in these advanced stages. (Loddi, Pereira et al. 2008, Schwarz and Watzke 2009). These methods which are based on current clinical practices and understanding, are designed to overcome the increased resistance of the ossified suture, ensuring effective outcomes while minimizing complications. However, it is important to note that the treatment options used today are based on current clinical findings and may be subject to change in the future as new research and advancements in the field emerge.

Assessing MPS maturation aids in determining the appropriate timing and method for maxillary expansion treatments. CBCT is used for this assessment due to its detailed imaging capabilities without superimposition of surrounding structures. (Silva-Montero, Faus-Matoses et al. 2022) Additionally, the purpose of the maturation stages of the MPS can influence the choice between conventional RME and surgically assisted RME (SARME), potentially preventing unnecessary surgical interventions and complications.(Sayar and Kılınç 2019)

While chronological age and gender are not reliable indicators of MPS fusion status, CBCT-based assessments offer a more accurate evaluation. (Shayani, Merino-Gerlach et al. 2023) The morphology observed on CBCT images correlates well with other findings such as hand-wrist radiographs, providing a reliable method for assessing suture maturation. (Sayar and Kılınç 2019) Furthermore, orthopedic widening of the maxilla without surgical intervention may still be possible in advanced maturation stages, although the window for effective treatment may be limited as sutural fusion progresses. (Shayani, Merino-Gerlach et al. 2023)

1.2.3 Deep Learning

Artificial intelligence (AI) is a rapidly emerging field within computer science, involving the development of machines capable of performing tasks that traditionally require human intelligence. (Chiu, Chang et al. 2023) Within AI, deep learning (DL) is a powerful subset of machine learning (ML) that utilizes artificial neural network architectures to analyze complex data. DL is characterized by its multiple hidden layers, often referred to as deep neural networks (DNN), which excel in representing intricate data structures, such as images, by extracting and modeling complex features like edges, corners, and shapes (Schwendicke, Cejudo Grano de Oro et al. 2022) (Mehta and Patnaik 2021, Mehta, Shah et al. 2022)).

One of the key strengths of DL, particularly in the context of DNNs, is its ability to automate the feature extraction process. Feature extraction involves identifying and isolating the most important pieces of information from raw data. Traditionally, this process required significant manual effort, but DL models can automatically extract these features, reducing the need for human intervention (Monsarrat, Bernard et al. 2022). This capability allows DL to effectively model non-linear relationships among hidden variables in various data formats, including images. (Chollet 2019)

Among the various types of DL algorithms, convolutional neural networks (CNNs) are particularly well-suited for tasks involving image analysis, such as inference, recognition, segmentation, and classification of medical images. (Ker, Wang et al. 2017) (Singha, Thakur et al. 2021) CNNs are composed of artificial neurons arranged in layers through mathematical functions, forming the foundation of modern AI architectures. (Schwendicke, Cejudo Grano de Oro et al. 2022) (Wan, Shen et al. 2018)

In DL, data flows through the network in a feed-forward manner, where each layer's output serves as the input for the next layer until the final layer is reached. This iterative process allows for the transformation of data across layers, fine-tuning the model through various training approaches, including supervised, unsupervised, or semi-supervised methods (Schmidhuber 2015, Janiesch, Zschech et al. 2021) (Alloghani, Al-Jumeily et al. 2020)). In supervised learning, which relies on labeled datasets, the model learns to predict outcomes by minimizing the error between its predictions and the actual labels, making it particularly effective for tasks such as classification and regression.(Ayodele 2010)

Supervised learning relies on labeled datasets, where each input is paired with a corresponding output label. This method is particularly effective for tasks such as classification and regression, as the model learns to predict outcomes by minimizing the error between its predictions and the actual labels. (Ayodele 2010) (Goodfellow 2016)

Unsupervised learning, by contrast, involves training models on unlabeled data, focusing on identifying patterns or structures within the data. This approach is often employed for tasks such as clustering, dimensionality reduction, and anomaly detection. Semi-supervised learning combines elements of both supervised and unsupervised learning, utilizing a small amount of labeled data alongside a larger pool of unlabeled data. This method is especially advantageous when obtaining labeled data is costly or time-consuming, as it allows the model to leverage the unlabeled data to enhance its performance. Each of these learning paradigms offers unique benefits and is chosen based on the specific requirements and constraints of the task at hand. (Ayodele 2010, Goodfellow 2016)

1.2.4 Computer Vision

Computer vision is a field of study that uses computers to interpret and comprehend visual information from the world, akin to human visual perception. By employing deep learning techniques, computer vision systems can perform a variety of tasks such as object classification, detection, and scene understanding. (Esteva, Chou et al. 2021) Advancements in deep learning have significantly advanced the field of computer vision. (Rebecq, Ranftl et al. 2019, Szeliski 2022) One significant milestone in this progression was the ImageNet challenge, which pushed the boundaries of object classification accuracy to achieve human-level performance.(Esteva, Chou et al. 2021, Guo, Xu et al. 2022) (Rebecq, Ranftl et al. 2019)Key tasks in computer vision encompass object classification, localization, and detection, which involve identifying object types, their locations, and attributes simultaneously within images. (Li, Huang et al. 2021) These tasks are facilitated by deep learning models, particularly Convolutional Neural Networks (CNNs). (Voulodimos, Doulamis et al. 2018, O'Mahony, Campbell et al. 2020)

1.2.4.1 Image Classification

Image classification is a task in computer vision that involves the process of categorizing and labelling groups of pixels or vectors within an image based on specific rules. The primary goal is to identify and categorize all objects in an image according to their characteristics and assign them to specific labels.(O’Mahony, Campbell et al. 2020, Li, Huang et al. 2021) Features such as edges, corners, and textures are extracted from images to help distinguish between different object classes (O’Mahony, Campbell et al. 2020) During classification, these features are compared against new images to identify specific objects. Supervised learning models rely on accurate labelling, often based on human observation or specific labelling tools designed for computer vision tasks (42) which learn to recognize patterns and features from the training data. The quality and quantity of labelled data affect the model's performance. (Budach, Feuerpfeil et al. 2022)

Deep learning techniques, specifically CNNs, have significantly enhanced image classification performance, often achieving accuracy levels that surpass human capability across various domains, including remote sensing and medical imaging, due to their ability to handle large-scale datasets and complex image features. (O’Mahony, Campbell et al. 2020)

Image classification finds applications in various fields such as medical imaging, autonomous driving, and surveillance, enabling efficient and accurate image analysis. (Jaswal, Vishvanathan et al. 2014, Hadi, Ajel et al. 2021) Deep learning models are robust, which enables them to excel despite challenges like handling variability in object appearance due to changes in size, rotation, and lighting conditions. (O'Mahony, Campbell et al. 2020)

The application of machine learning in image classification not only automates the process but also enhances the ability to handle vast amounts of image data with varying features, which would be impractical and inefficient with manual techniques. (Afzaal, Bhattarai et al. 2021, Al-Badri, Ismail et al. 2022)

1.2.4.2 Object Detection

Object detection in computer vision is the task of identifying and localizing objects within an image or a video frame. (Cazzato, Cimarelli et al. 2020, Li, Huang et al. 2021) Unlike image classification, which assigns a single label to an entire image, object detection goes further by not only recognizing what objects are present but also precisely locating them by drawing bounding boxes around them. (Feng, Zhong et al. 2021)

The primary goal of object detection is to determine the presence of multiple objects of interest within an image or a video and accurately delineate their positions. This enables machines to understand and interact with visual data in a manner similar to humans. (Cazzato, Cimarelli et al. 2020)

Object detection typically involves two main components:

1. Localization: This involves identifying the spatial location of objects within an image by drawing bounding boxes around them. These bounding boxes specify the coordinates of the object's position, usually defined by a combination of a top-left corner point, width, and height.(Gidaris and Komodakis 2016, Jiang, Luo et al. 2018)

2. Classification: After localizing objects, the next step is to classify each detected object into predefined categories or classes. This classification process assigns a label to each object, indicating what type of object it is (e.g., car, person, dog). (Jiang, Luo et al. 2018, Cazzato, Cimarelli et al. 2020)

Object detection is a fundamental task in computer vision with a wide range of applications, (Tasnim and Qi 2023) including Autonomous vehicles, Surveillance systems, Robotics, and medical imaging. (Namdev, Agrawal et al. 2022, Tasnim and Qi 2023, Talwandi, Thakur et al. 2024)

1.2.4.3 Image Segmentation

Image segmentation in computer vision is the process of partitioning an image into multiple segments or regions to simplify its representation and facilitate further analysis. (Ghosh, Das et al. 2019) The goal is to divide the image into meaningful parts based on certain characteristics, such as color, texture, or intensity, in order to extract valuable information or identify objects within the image. (Minaee, Boykov et al. 2021)

This technique is fundamental in various computer vision tasks, including object recognition, image understanding, and scene analysis. (Isensee, Jäger et al. 2019) By segmenting an image into distinct regions, it becomes easier to identify and analyze specific objects or features within the image. (Ghosh, Das et al. 2019)This technique is crucial for a wide range of applications, including medical image analysis, where it helps identify and segment anatomical structures; autonomous driving, where it assists in detecting and understanding the surrounding environment; and content-based image retrieval, where it aids in searching for images based on their visual content.(Ghosh, Das et al. 2019, Isensee, Jäger et al. 2019, Minaee, Boykov et al. 2021)

Image segmentation algorithms typically assign labels or identifiers to pixels or groups of pixels that share similar attributes, effectively delineating boundaries between different regions or objects. (Kim, Kanezaki et al. 2020) Various methods can be used for image segmentation, including:

1. Thresholding: Dividing the image into regions based on pixel intensity values above or below a certain threshold.(Minaee, Boykov et al. 2021)

2. Clustering: Grouping pixels with similar attributes into clusters using techniques such as k-means clustering or mean-shift clustering.(Kim, Kanezaki et al. 2020)

3. Edge detection: Identifying boundaries between different regions based on changes in pixel intensity or gradients.(Minaee, Boykov et al. 2021)

4. Region growing: Starting with seed points and iteratively growing regions based on similarity criteria until boundaries are reached. (Minaee, Boykov et al. 2021)

5. Watershed segmentation: Treating the image as a topographic surface and flooding it from certain seed points to delineate regions. (Ghosh, Das et al. 2019)

1.2.4.4 Region of Interest

In deep learning and computer vision tasks, the concept of a Region of Interest (ROI) plays a pivotal role, especially when working with complex images that require precise analysis. As computer vision tasks, like image classification, object detection, and image segmentation, focus on understanding the entire image or locating specific objects within it, the ROI narrows down this focus to specific areas within an image that are deemed most significant for the task at hand. Identifying an ROI helps models improve accuracy, reduce computational load, and enhance interpretability by concentrating on the most relevant parts of the image (Vijay, Saini et al. 2021, Alqazzaz, Sun et al. 2022, Venkatachalam and Chandrabose 2023) ROI detection methods consist of Manual annotation and segmentation networks such as U-Net, R-CNN, and SegNet. (Vijay, Saini et al. 2021, Alqazzaz, Sun et al. 2022, Li, Liu et al. 2022)

Manual annotation involves human experts marking the areas of interest in images. This method is highly accurate and is considered the gold standard but labour-intensive and time-consuming. It is particularly used in medical imaging, where specialists such as radiologists annotate regions like tumors or other abnormalities on medical scans. (Baughan, Li et al. 2023, Li, Chou et al. 2023)

1.2.5 Convolutional Neural Network (CNN)

CNNs represent a specialized form of deep learning architecture, particularly effective in image processing tasks. (Figure 2) (Knaak, von Eßen et al. 2021) CNNs leverage convolution operations to recognize patterns within images or frames, employing learnable filters to extract essential features from input data(Naranjo-Torres, Mora et al. 2020) (Roy, Song et al. 2018, Knaak, von Eßen et al. 2021). Characterized by multi-layered structures comprising convolutional, pooling, activation, and fully connected layers, CNNs demonstrate versatility in handling diverse tasks, with more complex architectures enhancing performance on challenging objectives. (Naranjo-Torres, Mora et al. 2020, Aljuaid and Anwar 2022)

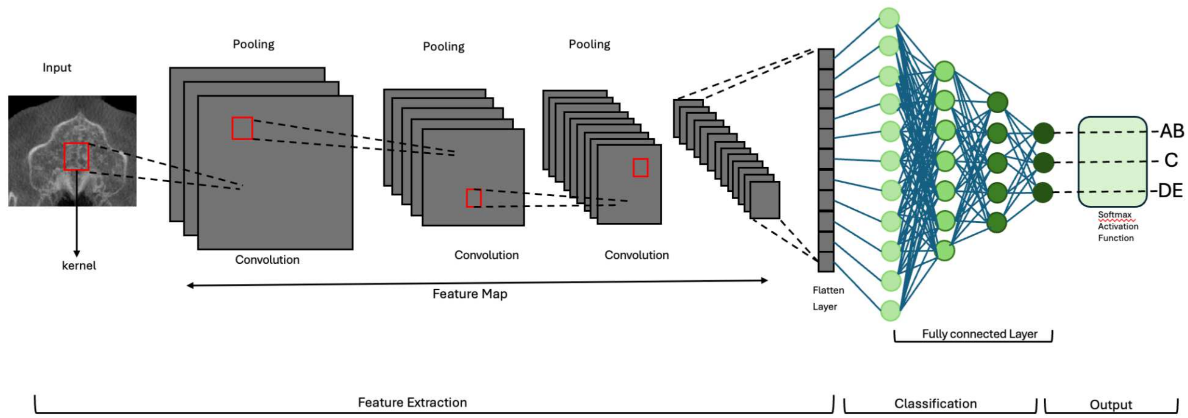

Figure 2 Convolutional Neural Network breakdown

Convolutional layers, fundamental components of CNNs, perform convolution operations on input data using learnable filters to generate feature maps capturing significant patterns of the input image (Knaak, von Eßen et al. 2021, Aljuaid and Anwar 2022). These layers exhibit properties like parameter sharing, sparse interactions, and translation invariance, which contribute to their effectiveness in image tasks (Wang and Shen 2017, Aljuaid and Anwar 2022) . Pooling layers, often integrated after convolutional layers, down-sample input data to reduce spatial dimensions while preserving essential features through operations like max pooling and average pooling (Naranjo-Torres, Mora et al. 2020, Aljuaid and Anwar 2022). Activation layers apply non-linear functions to layer outputs, facilitating the learning of complex relationships between input and output data. (Cha, Choi et al. 2017) Common activation functions include Rectified Linear Unit (ReLU) and sigmoid, with the softmax function favoured for multiclass classification tasks(Cha, Choi et al. 2017, Naranjo-Torres, Mora et al. 2020, El-Ghany, Azad et al. 2023). Fully connected layers establish connections between neurons across layers, facilitating classification or regression tasks. (Cha, Choi et al. 2017)

ResNet, or Residual Network, is a type of CNN deep learning model designed to tackle the challenges of training very deep neural networks, especially for image classification tasks. One major challenge in deep networks is the vanishing gradient problem, where the gradients used to update the network's weights become too small as they move through the layers, slowing down learning. ResNet addresses this by using residual learning, where the model learns the difference (or residual) between the input and the desired output, rather than trying to learn the output directly. This makes it easier to train deep networks. (He, Zhang et al. 2016, Gao, Chen et al. 2022)

Another highly used CNN architecture for imaging tasks is VGG, or Visual Geometry Group network. This model was developed to enhance the performance of image classification tasks by increasing the depth of the network. VGG networks are known for their simplicity in design, using small 3x3 convolutional filters consistently throughout the network, which allows them to capture intricate patterns in the data as the network deepens. (Simonyan 2014) A key characteristic of VGG models is the use of deep architectures with a large number of layers, which helps improve the accuracy of image classification. However, this comes with the trade-off of increased computational cost and memory usage. (Simonyan 2014, Luo, Wu et al. 2017)

1.2.6 Small Vs. Large CNN Models

The evolving landscape of neural network architectures, particularly Convolutional Neural Networks (CNNs), reflects a critical shift towards optimizing model efficiency amidst escalating computational demands. Researchers have increasingly focused on developing lightweight and efficient CNN models tailored for edge computing and low-power devices. Insights from recent studies underscore the practical implications of adopting simpler CNN architectures, emphasizing their effectiveness in resource-constrained environments.

1.2.6.1 Advantages of Small CNNs

Small or lightweight CNNs, characterized by fewer layers and parameters, are designed to be less computationally intensive compared to their larger counterparts. They offer several advantages:

Computational Efficiency: Small CNNs require less computational power and memory, making them suitable for deployment on devices with limited resources. This is particularly advantageous in edge computing scenarios, where processing power and energy consumption are critical factors (Howard, Zhu et al. 2017, Sandler, Howard et al. 2018).

Speed: Due to their reduced complexity, small CNNs can process data faster, enabling real-time applications. This is crucial in fields like autonomous driving and medical diagnostics, where timely decision-making is essential (Zhang, Zhou et al. 2018).

Generalization: Smaller models are less prone to overfitting, especially when trained on limited datasets. They can achieve competitive performance while maintaining a balance between bias and variance (Iandola, Han et al. 2016).

Deployment Flexibility: The lightweight nature of small CNNs allows them to be deployed on a variety of devices, including mobile phones, embedded systems, and other low-power devices (Howard, Zhu et al. 2017).

In one comparative study, researchers investigated various lightweight neural network architectures, including Long Short-Term Memory (LSTM), SimpleRNN, and multiple iterations of 1D CNNs, alongside more complex models like 1D ResNet and DenseNet. Their findings highlight that simpler CNN models not only offer computational savings but also maintain competitive performance across diverse problem domains. This research underscores the feasibility of deploying lightweight CNNs in scenarios where computational efficiency is paramount, such as edge computing applications. (Rong 2023) However, the effectiveness of lightweight CNNs can vary depending on the specific task and context. For instance, computational complexity and efficiency considerations may differ between image classification and object detection tasks, as well as between different deployment environments (e.g., edge vs. cloud computing) (Cheng, Wang et al. 2017, Yoo, Kim et al. 2019)

Moreover, the customization of CNN architectures based on specific environmental conditions further optimizes performance in specialized domains. For instance, in autonomous vehicle navigation (Cheng, Wang et al. 2017), researchers propose adaptive CNN models tailored to different terrains. They advocate for lighter CNN architectures for navigating simpler road conditions, optimizing computational resources without compromising driving efficacy. In contrast, more intricate terrain necessitates deeper CNNs capable of extracting nuanced features essential for safe and efficient navigation. (Pal and Khaiyum 2019)

This approach challenges the traditional one-size-fits-all paradigm by advocating for adaptable CNN architectures that dynamically adjust their complexity and computational footprint based on environmental demands. By aligning model complexity with task requirements, these studies advocate for a more efficient utilization of neural networks in real-world applications, balancing performance with resource constraints effectively.(Howard, Zhu et al. 2017, Elsken, Metzen et al. 2019)

In the field of medical diagnostics, small CNNs have been successfully used for tasks such as detecting diabetic retinopathy and classifying skin lesions, where quick and accurate analysis is critical.(Esteva, Kuprel et al. 2017) Similarly, in autonomous driving, lightweight models like SqueezeNet have been used for real-time object detection, demonstrating that small CNNs can achieve high performance without the computational burden of larger models.(Iandola, Han et al. 2016)

In summary, the pursuit of lightweight or small and adaptive CNN models not only addresses the escalating computational demands but also expands the applicability of neural networks in diverse settings. These advancements pave the way for broader integration of AI technologies in edge computing, autonomous systems, and other resource-constrained environments, fostering innovation and efficiency across various domains(Cheng, Wang et al. 2017, Yoo, Kim et al. 2019).

1.2.7 Examples of applications of image classification with CNN in Medicine and Dentistry

Image classification tasks with Deep learning, particularly convolutional neural networks (CNNs), is making significant strides in the realms of medicine and dentistry, revolutionizing diagnostic processes and patient care. These advancements are reflected in various applications, from image analysis to predictive modelling, demonstrating the transformative potential of deep learning across these fields.(Salunke, Joshi et al. 2022, Al-Khuzaie and Al-Jawher 2023)

In medicine, CNNs have shown remarkable efficacy in the classification and diagnosis of complex diseases. For instance, in diagnosing Interstitial Lung Disease (ILD) from high-resolution computed tomography (HRCT) images, CNNs have been adapted to handle the texture-like qualities of ILD patterns. By simplifying the architecture to a single convolutional layer and incorporating optimizations like random neural node dropout, CNNs have effectively managed the classification of ILD lung image patches(Huang, Lee et al. 2020). Additionally, the classification of pneumonia from chest X-rays using CNN-based algorithms highlights the efficiency of transfer learning and data augmentation. Techniques such as transfer learning with pre-trained models (e.g., VGG16) and training from scratch with capsule networks have shown that proper feature retraining and network complexity matching are crucial for optimal outcomes.(Yadav and Jadhav 2019)

Similarly In dentistry, CNNs are revolutionizing diagnostic and treatment approaches by enabling precise, personalized, predictive, and preventive care. One example is the classification of midpalatal suture maturation stages using cone-beam computed tomography (CBCT). Various CNN models, including ResNet18, ResNet50 and ResNet101, were trained to perform multi-class classification of suture maturation stages and detecting the CBCT plane containing the suture. Among these, ResNet50 has demonstrated superior performance, achieving an accuracy of 99.74% in identifying the targeted CBCT plane.(Zhu, Yang et al. 2024) For the five maturation stages (A/B/C/D/E) of the midpalatal suture, the models showed varying levels of accuracy. ResNet18 achieved an accuracy of 73.13%, while ResNet50 had a slightly higher accuracy at 73.51%. ResNet101, on the other hand, demonstrated an accuracy of 66.79%. Inception-v3 exhibited a lower accuracy of 54.85%.

For the simplified task of classifying three maturation stages (AB/C/DE), the CNN models demonstrated varying levels of accuracy. ResNet18 (42.73 MB and 11.69 million parameters) achieved an accuracy of 79.10%, while ResNet50 (90.04 MB and 25.56 million parameters) showed a higher accuracy at 82.84%. ResNet101 (162.82 MB and 44.55 million parameters) closely followed with an accuracy of 80.60%. Furthermore, the assessment of maxillary transverse development, which relies heavily on the maturation and ossification status of the midpalatal suture, has been enhanced through other deep learning techniques. Using CBCT image fusion and texture feature analysis algorithms, CNNs have provided comprehensive insights into suture maturation during rapid growth and development, particularly excelling in classifying age ranges of 4-10 and 17-23 years. (Gao, Chen et al. 2022)

In a similar vein, CNNs have been successfully applied to the classification of dental diseases using Radiovisiography images (Bitewing, periapical radiographs and OPG) . A study explored the use of ResNet, to classify three common dental conditions: dental caries, periapical infection, and periodontitis. By employing transfer learning, the VGG16 network significantly outperformed the custom CNN, achieving a classification accuracy of 97.43% compared to 92.15% for the custom model. The study's results not only improved diagnostic precision but also underscored the adaptability of CNNs in various dental imaging tasks, reinforcing their role in advancing modern dental care. (Prajapati, Nagaraj et al. 2017)

Another example is the use of VGG based models in dental imaging tasks. They have been applied to teeth detection and numbering on panoramic radiographs, achieving performance levels close to those of experts. This approach, which does not rely on hand-crafted features, is suitable for integration into clinical practice. (Tuzoff, Tuzova et al.

2019) Additionally, a modified VGG16 model, known as TVGG16, has been used to identify dental implant manufacturers from X-ray images. (Guo, Tsai et al. 2022) VGG16 has also been evaluated for classifying dental restorations using panoramic radiographs. Although it did not achieve the highest accuracy compared to models like ResNet, it still performed significantly well. (Top, Özdoğan et al. 2023)

In forensic odontology, VGG16 has been utilized for age estimation and individual identification tasks, demonstrating its versatility across different dental imaging applications. Furthermore, VGG-based models have potential applications in detecting various dental pathologies, such as caries, periodontitis, and dental cysts, although larger datasets are required to achieve expert-level accuracy. (Mohammad, Ahmad et al. 2022)

Beyond imaging, AI integration in dentistry encompasses a wide range of applications. AI techniques are improving patient management, diagnosis, prediction, and decision-making, enhancing the efficiency and efficacy of dental care delivery. This integration facilitates data-enriched clinical care, leading to comprehensive real-time assessments and precise interventions(Schwendicke, Cejudo Grano de Oro et al. 2022, Alzaid, Ghulam et al. 2023)For example, AI models exhibit proficiency in predicting orthodontic extractions, determining the necessity for orthodontic treatments, conducting cephalometric analysis, and discerning age and gender, showcasing AI's broad applicability in dental diagnostics and treatment planning.(Khanagar, Al-Ehaideb et al. 2021)

Overall, the integration of deep learning, particularly CNNs, in medical and dental imaging is poised to drive significant improvements in diagnostic precision, efficiency, and predictive capabilities. As research and technology continue to advance, these methodologies are expected to become integral to clinical practice, transforming patient care and outcomes across both fields. (Top, Özdoğan et al. 2023)

In this thesis, we explore the application of deep learning techniques, specifically CNNs, in the classification of MPS maturation stages. By harnessing the robust capabilities of AI, particularly in image analysis, we aim to enhance medical diagnostic processes, illustrating the potential of deep learning in advancing healthcare technologies.

1.3 Research hypothesis and objectives

1.3.1 Objectives

1. To develop and evaluate small CNN models for classifying midpalatal suture maturation stages on CBCT images.
2. To develop a method for determining the Region of Interest of MPS from CBCT slices of the palate based on the opinion of orthodontists.
3. To develop a method for determining the most appropriate slices of the palate for classification based on the opinion of orthodontists.
4. To develop a method for classifying the MPS as a 3D structure from CBCTs.

5. To validate the models and classify the MPS into specific stages according to the level of suture fusion.

1.3.2.1 Hypothesis 1

The utilization of a small 2D CNN model with the least number of layers will accurately classify the maturation stage of the MPS on slices of CBCT images.

1.3.2.2 Rationale for Hypothesis 1

The manual evaluation of MPS fusion using CBCT images is subjective and heavily reliant on practitioners' experience. To address this challenge, we propose the application of deep learning techniques, specifically a small CNN model to establish an objective and standardized approach for classifying the fusion stage of the mid-palatal suture. This automated classification process aims to reduce variability in diagnosis and enhance consistency across evaluations.

Accurate staging of mid-palatal suture fusion is imperative for selecting appropriate treatment strategies for transverse maxillary constriction. Leveraging a small CNN model enables efficient and precise classification of 2D slices of the mid-palatal suture on CBCT images. By automating this process, clinicians can streamline treatment planning, saving time while ensuring patients receive interventions tailored to their suture maturation stage.

Small CNN models are recognized for their efficiency, requiring fewer computational resources compared to larger models. Employing a small CNN model allows for the development of a classification system that runs efficiently on standard hardware without compromising accuracy. This approach maximizes computational efficiency while maintaining high levels of diagnostic precision.

1.3.2.3 Hypothesis 2

The implementation of a 3D CNN model with the least number of layers will effectively classify 3D arrays of the MPS, providing comprehensive classification of the entire suture volume without the need to exclude curved palates from the data set.

1.3.2.4 Rationale for Hypothesis 2

Effective treatment of transverse maxillary constriction relies on precise staging of mid-palatal suture fusion, necessitating a comprehensive assessment of the entire suture volume. Our proposed solution involves deploying a 3D CNN model, allowing for a holistic evaluation of the MPS by analyzing volumetric data. This model captures spatial relationships and contextual information across multiple slices, ensuring a thorough classification that covers the entire suture volume. This approach is particularly advantageous for curved palates where the suture may not be visible in a single slice alone.

By analyzing 3D arrays of the MPS, the CNN model can identify subtle patterns and features that may be missed in individual 2D slices. This heightened diagnostic precision enables more accurate classification of the suture's fusion stage, facilitating treatment decisions.

The use of a CNN model streamlines the classification process for MPS fusion, promoting standardization and objectivity across practitioners. This consistency enhances the reliability of treatment planning and outcomes by reducing variability in diagnosis.

Despite handling volumetric data, the 3D CNN model maintains computational efficiency, enabling timely processing of 3D arrays of CBCT images. This efficiency minimizes computational burden while maximizing diagnostic capabilities, making it a practical solution for clinical implementation.

1.4 Scope of the thesis

The two hypotheses and objectives are tested in this study as outlined in two chapters (chapters 2-3), with one additional chapter for the introduction and literature review (chapter 1) and a final chapter (chapter 4) for general discussion and conclusion.

Chapter 2 tested the first hypothesis and focused on the application of a small CNN model to classify the maturation stage of the MPS from 2D slices of CBCT images.

Chapter 3 tested the second hypothesis and focused on the application of a 3D CNN model for analyzing and classifying the MPS from 3D arrays of the palate containing the MPS.

Lastly, Chapter 4 discusses the outcome of these studies, their interrelationship, and their significance in the dental setting. Study limitations and future directions are highlighted as well.

2.1 Introduction and Background

2.1.1 Introduction:

Maxillary transverse deficiency (MTD) is a condition involving a narrowed upper jaw relative to the mandible, often leading to issues such as posterior crossbite and dental crowding, and it can contribute to sleep apnea due to associated airway narrowing and altered tongue posture (Calvo-Henriquez et al., 2021; Akbulut et al., 2020). Treating MTD requires accurate evaluation of the mid-palatal suture (MPS) maturation, as it guides the decision between surgical and non-surgical expansion methods (Franchi et al., 2010; Hernandez-Alfaro et al., 2010). Misclassification of MPS maturation can result in suboptimal treatment outcomes (Suri & Taneja, 2008).

The five-stage classification of MPS maturation, proposed by Angelieri et al. using CBCT, offers a framework for this assessment, though variability in orthodontists' evaluations remains a challenge (Barbosa et al., 2019). Accurate classification is critical, as non-surgical methods are appropriate in earlier stages (A and B), while surgical interventions are necessary for later stages (D and E) (de Miranda Ladewig et al., 2018; Gao et al., 2022).

Artificial Intelligence (AI), particularly deep learning (DL) algorithms, offers promising avenues for automating data extraction and analyzing complex datasets. (Monsarrat, Bernard et al. 2022) Deep learning models like Convolutional Neural Networks (CNNs) excel in image recognition, segmentation, and classification tasks, making them potentially valuable tools for MPS maturation assessment. In 2024, Zhu et al. published a study titled "Convolutional neural network-assisted diagnosis of MPS maturation stage in cone-beam computed tomography" in this study ResNet18 was one of the models evaluated. ResNet18, along with other CNN models like ResNet50 and ResNet101, was trained to classify MPS staged into three groups (AB/C/DE) based on image features.

The key findings from the study indicate that ResNet18 with 11.69 parameters and size of 42.73 MB, while being the smallest model and a part of the evaluation, had the lowest accuracy among the ResNet models used. Specifically, ResNet18's performance in the three-stage evaluation of midpalatal suture maturation stages showed an accuracy rate of 79.10%, which was lower compared to ResNet101 (162.82 MB and 44.55 million parameters), which achieved the highest accuracy of 80.60%. This suggests that while ResNet18 was effective in classifying the stages, it was outperformed by more complex models in the same family, such as ResNet101. (Zhu, Yang et al. 2024)

The growing computational demands of traditional Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) highlight the need for lightweight, efficient models, especially for edge computing and low-power devices. (Rong 2023) explored the performance and computational costs of various lightweight neural network architectures, including LSTM, SimpleRNN, 1D CNNs, 1D ResNet, and DenseNet. Their study found that simpler CNN models can be effective and computationally economical, which is beneficial for resource-constrained environments.

Khayuim et al. proposed customizing CNN architectures based on data conditions to balance computational efficiency. Simpler CNN models are sufficient for easy terrains, while complex terrains require more elaborate networks. This tailored approach reduces network size, trainable parameters, memory footprint, and refresh rate, aiming to improve efficiency by adapting CNN models to specific conditions rather than using a one-size-fits-all model. (Pal and Khaiyum 2019)

Building on these insights, our study introduces a deep learning method utilizing a lightweight CNN architecture to classify the maturation stages of the MPS. By employing supervised learning with a dataset annotated by expert orthodontists, we aim to achieve high accuracy in identifying MPS maturation stages from CBCT images while minimizing computational costs. This approach is particularly relevant for practical applications in clinical settings where computational resources may be limited.

Each of the CBCTs from patients normally consist of 450 to 500 axial slices, however only a few slices contain the maxilla and even fewer slices from the palatal region contain the MPS. In order to cultivate a dataset for classifying the maturation stage of MPS which consists exclusively of optimal slices that contain the MPS, our study introduces a small

CNN model for classifying axial slices containing the Region of Interest. The objective is to discern preferred slices exhibiting clear visibility of the MPS from non-preferred slices.

2.1.2 Objectives:

1. To develop and evaluate a 2D CNN model for classifying midpalatal suture maturation stages on CBCT images.
2. To develop a method for determining the Region of Interest (MPS) from CBCT slices of the palate based on the opinion of orthodontists.
3. To develop a method for determining the most appropriate slices of the palate for classification based on the opinion of orthodontists.

2.2 Methodology

This research was approved by the Health Research Ethics Board (HREB) of the University of Alberta (Approval number: Pro00125920) CBCT from 155 patients between the age of 7-21 who underwent CBCT (120 kVp, 5 mA and 4 sec, 0.3 voxel) for orthodontic treatments at the university clinic were initially collected for this study.

Exclusion criteria were as follows:

1. Prior orthodontic treatment
2. Impacted upper teeth in the mid-palatal region
3. Congenital malformations of the maxilla (i.e. cleft palate)

4. Palates where the curvature prevents the suture from being viewed in a single slice, requiring multiple consecutive slices to visualize the entire suture.
5. Blurry or low-quality CBCTs due to scatter

Out of the 155 patients, CBCTs from 111 patients did not have the exclusion criteria and were thus included in the study. The CBCT images were in DICOM format and anonymized prior to collection for this study, thus, they were first converted from DICOM format to PNG image format with the ITK-SNAP software (726 * 644 pixels). The axial view of the patient images' which consisted of an average number of 450 slices were classified into 5 groups of MPS maturation stage as first stated by Angeliere et al. by two orthodontists, achieving an inter-rater reliability of 68%, comparably higher than similar a study at 43%.(Barbosa, Castro et al. 2019) In the case of any conflicts, a third orthodontist evaluated the images to determine the class of MPS. Once the patients were classified based on the Angeliere method, the slices belonging to patients' palates were selected and saved in a separate folder.

The palate slices were then categorized into three groups AB (maturation stages A and B) C and DE (maturation stages D and E). This was done as a way to reduce the amount of variability for the DL model as the aim of this study was to help in reaching a diagnosis for an optimal treatment plan as: the treatments for stages A and B are clinically the same, but can differ from stage C, and stages A B and C have different treatments compared to stages D and E.(de Miranda Ladewig, Capelozza-Filho et al. 2018)

2.2.1 Data preprocessing:

In the data preprocessing phase, the images underwent several transformations to enhance the CNN model's ability to focus on the actual suture rather than surrounding structures. The initial step involved cropping the images so that the maxilla is the only structure in the image. This is done on the following coordinates which ensures the entirety of the maxilla is in the Region of Interest (ROI) [(300,100),(520, 100), (520, 290), (300,290)], resulting in a cropped area with a height of 190 and a width of 220. Subsequently to reduce the noise of the images and help determine the image contour a Gaussian blur was applied using the Cv2 and NumPy libraries. (Bradski and Kaehler 2000, Harris, Millman et al. 2020) Gaussian blur is a noise reduction technique commonly used in image preprocessing.

To correct for angular differences, the angle of the image was corrected using rotation. The contour with the maximum width, representing the transverse dimension of the maxilla, was first identified. A rectangle was then drawn around this contour, and the rotation angle of the rectangle was calculated for each image. The images were rotated to align with the horizontal line. Following rotation, Cv2 and NumPy were employed again to identify the Region of Interest (ROI) for the MPS. This was achieved by drawing a rectangle around the entire image and calculating the distance between the rectangle's edge and the image center. Subsequently, coordinates were coded to crop an area 47 pixels around the midline on the X-axis and 140 pixels around the midline on the Y-axis, which is the smallest area that effectively encompasses the entire suture length in all the CBCT images, which was manually verified on each CBCT. This final processing step resulted in an ROI with dimensions (140, 47). (Figure 2)

The choice of the dimensions for cropping the images to focus on the MPS was guided by both practical considerations and technical requirements to ensure consistent and accurate analysis across different CBCT images. The CBCT images from different patients were relatively consistent in terms of output size. To ensure that the ROI captured the entire MPS across all samples, it was essential to determine a dimension that would be universally applicable. To minimize the amount of irrelevant information and maximize the focus on the MPS, the smallest possible dimension that consistently encompassed the MPS across all samples was selected. This approach reduces computational load and enhances the model's ability to learn relevant features.

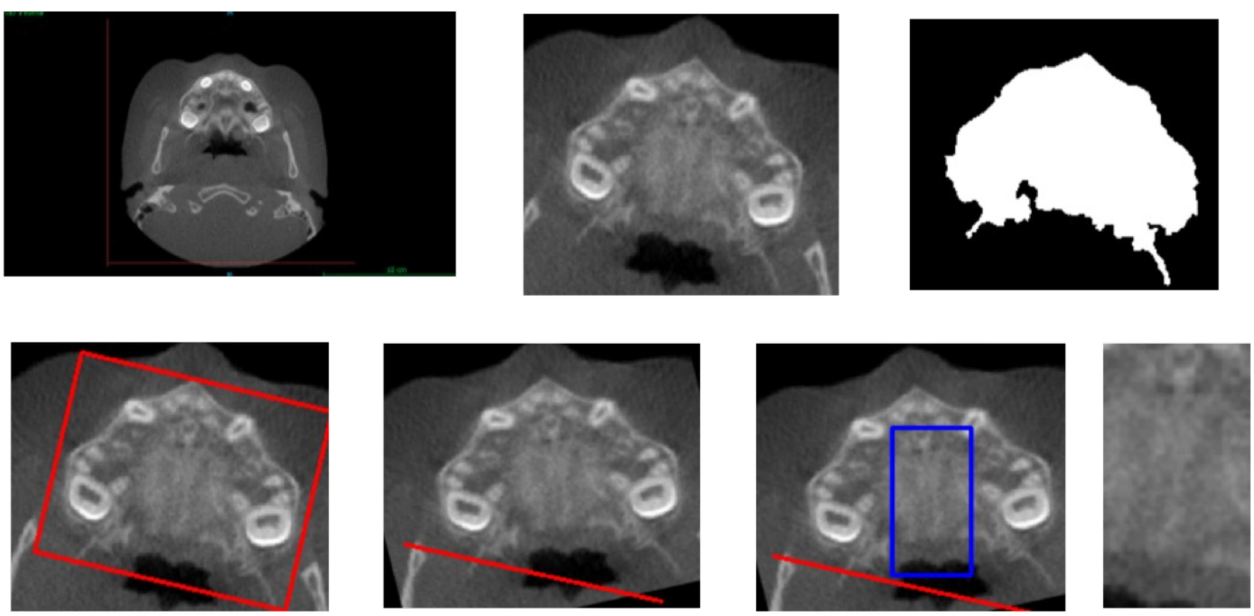

Figure 3 Determining ROI on axial slice of CBCT

2.2.2 Determining the ideal CBCT slice containing the MPS (ROI classifier):

To create a ROI classifier to identify the ideal axial slices of CBCT which contain the MPS, the following steps are required.

2.2.2.1 ROI Classifier: Data Collection and Preprocessing

For the development of the ROI (Region of Interest) classifier, a specialized dataset using CBCT scans from 111 patients was constructed. This dataset consists of two types of slices:

1. **Preferred Slices**: A total of 575 slices were selected, each containing a clearly visible mid-palatal suture (MPS). These slices were carefully chosen by orthodontists from the CBCT scans of the 111 patients.
2. **Non-Preferred Slices**: To complement the preferred slices, 421 random slices were also included from the same set of patients. These slices were selected from the axial view of the skull but did not exhibit the clear visibility of the MPS.

This ROI dataset serves as the foundation for building the classifier that will be used to create the primary dataset for the MPS classification model.

2.2.2.2 ROI Classifier: Data Splitting

The dataset for ROI classifier which consisted of 996 images was split into training, validation, and testing sets using a stratified approach to maintain representative class distributions. Specifically, 70% of the data was allocated for training, 20% for validation, and 10% for testing. This split percentage is one of the most simple and widely used splits in deep learning algorithms which ensures a balanced training.(Brea, De Jesus et al. 2020, Al-Sarem, Al-Asali et al. 2022)This resulted in 402 preferred images and 294 non-preferred images for training, 115 preferred images and 84 non-preferred images for validation, and 57 preferred images and 42 non-preferred images for testing.

2.2.2.3 ROI Classifier: Model Architecture

A Convolutional Neural Network (CNN) architecture was employed for binary classification of preferred and non-preferred slices. The CNN consisted of two convolutional layers followed by max-pooling layers for feature extraction. The first convolutional layer had 32 filters with a kernel size of (3, 3), followed by a max-pooling layer with a pool size of (2, 2). The second convolutional layer comprised 64 filters with the same kernel size, followed by another max-pooling layer. Subsequently, the feature maps were flattened and passed through two dense layers with ReLU activation functions. The final layer of the CNN used a softmax activation function, producing a probability distribution over the

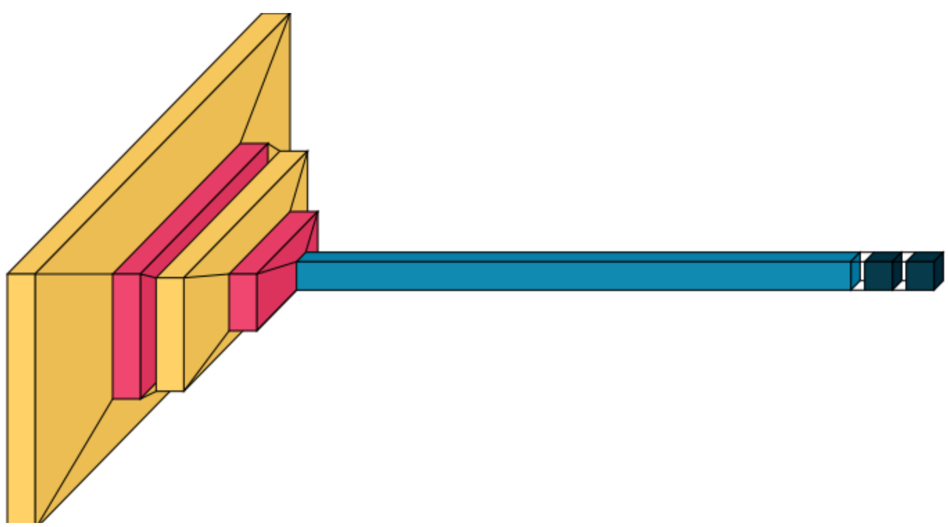

Figure 4 ROI Classifier CNN model architecture overview

two classes (preferred and non-preferred). The class with the higher probability was selected as the predicted class. (Figures 3 and 4)

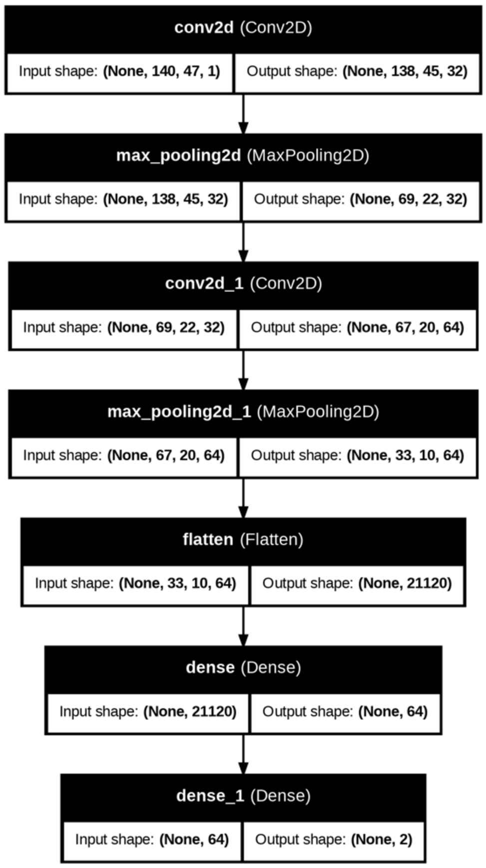

Figure 5 ROI Classifier CNN model architecture breakdown

2.2.2.4 ROI Classifier: Model Training

The model was trained using the Adam optimizer and binary cross-entropy loss function. Early stopping was implemented with a patience of 5 epochs to prevent overfitting, and model checkpoints were saved to retain the best-performing model based on validation loss. Training utilized a batch size of 32 and ran for 18 epochs. The training dataset was shuffled before each epoch to enhance model generalization.

2.2.3 Training, Validation and Testing split:

111 CBCTs from 111 patients were included in this study. The ROI classifier will be employed on the CBCTs to determine the ideal slices containing the MPS. Once the ROI classifier has been used on the patient CBCTs a dataset will be created. The dataset will be split for training, validation and test in patient-wise manner, where 70% of the data will be used for training, 20% for validation and 10% for testing the CNN architecture.

2.2.4 Deep Learning Frameworks

This study utilized TensorFlow and Keras for developing and training the convolutional neural network (CNN) models. TensorFlow is an open-source deep learning framework developed by Google Brain and released in November 2015. It facilitates the creation, training, and deployment of machine learning models across a range of hardware configurations, including CPUs and GPUs. (Hoeser and Kuenzer 2020)

Keras, introduced in May 2015, serves as a high-level API integrated into TensorFlow, providing an accessible and user-friendly interface for building deep learning models. Originally designed to work with frameworks like Theano and TensorFlow, Keras's modular and extensible design makes it particularly effective for complex tasks such as semantic segmentation. Following the end of Theano's official support in 2017, Keras has become one of the most popular frameworks for deep learning, particularly in fields such as computer vision, autonomous driving, and medical imaging. (Hoeser and Kuenzer 2020, Gupta 2023)

2.2.5 CNN architecture:

The convolutional neural network (CNN) architecture utilized in this study is a custom 9-layer CNN, implemented using TensorFlow and Keras. (Keras 2015, Abadi, Agarwal et al. 2016) (Figure 5 and 6) This model is designed with simplicity and computational efficiency in mind, incorporating key elements such as dropout and batch normalization to enhance performance.

The architecture consists of nine convolutional layers, each followed by Batch Normalization (BN) and rectified linear unit (ReLU) activation functions to introduce non-linearity. ReLU is a widely used activation function in CNNs. It outputs the input directly if it is positive; otherwise, it outputs zero, mathematically expressed as $f(x) = \max(0, x)$. ReLU helps prevent the vanishing gradient problem by providing a gradient of either 0 or 1, which facilitates efficient training of deep networks and induces sparsity in the activations, leading to more efficient computations. BN is a technique used in deep learning to enhance the training of neural networks by normalizing the inputs to each layer within a mini batch. This process involves standardizing the activations by adjusting their mean to zero and variance to one.

A Max-pooling layer is strategically placed in this CNN model to perform spatial down sampling, which aids in both computational efficiency and feature extraction. To reduce overfitting and improve generalization, dropout layers with a dropout rate of 0.5 are added after each block of layers.

In the initial convolutional layer, input tensors of shape (140, 47, 1) undergo convolution with 64 filters of size 7x7, followed by batch normalization and ReLU activation. A max-pooling operation with a pool size of 3x3 and a stride of 2x2 is then applied to reduce the spatial dimensions. The subsequent convolutional layers continue this pattern, progressively extracting features from the input data.

After the convolutional layers, global average pooling is applied to aggregate spatial information across feature maps, reducing the dimensionality of the data. Finally, a dense layer with softmax activation is employed to produce class probabilities. Softmax is an activation function used in the output layer for classification tasks. It converts raw output scores (logits) into probabilities by exponentiating each score and normalizing by the sum of all exponentiated scores. This results in a probability distribution over the classes, which is useful for multi-class classification problems.

This custom CNN architecture is specifically designed to meet the needs of the study, balancing complexity and computational efficiency. The model comprises a total of 4,733,187 parameters, of which 4,729,219 are trainable, and 3,968 are non-trainable parameters.

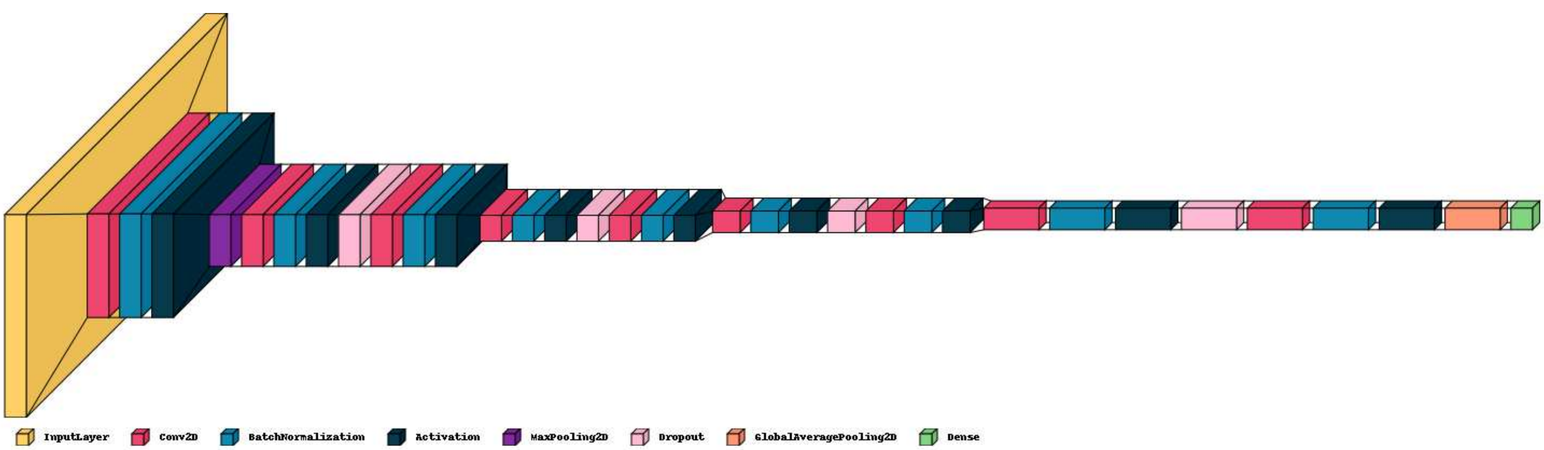

Figure 6 2D CNN architecture overview for MPS maturation stage classifier

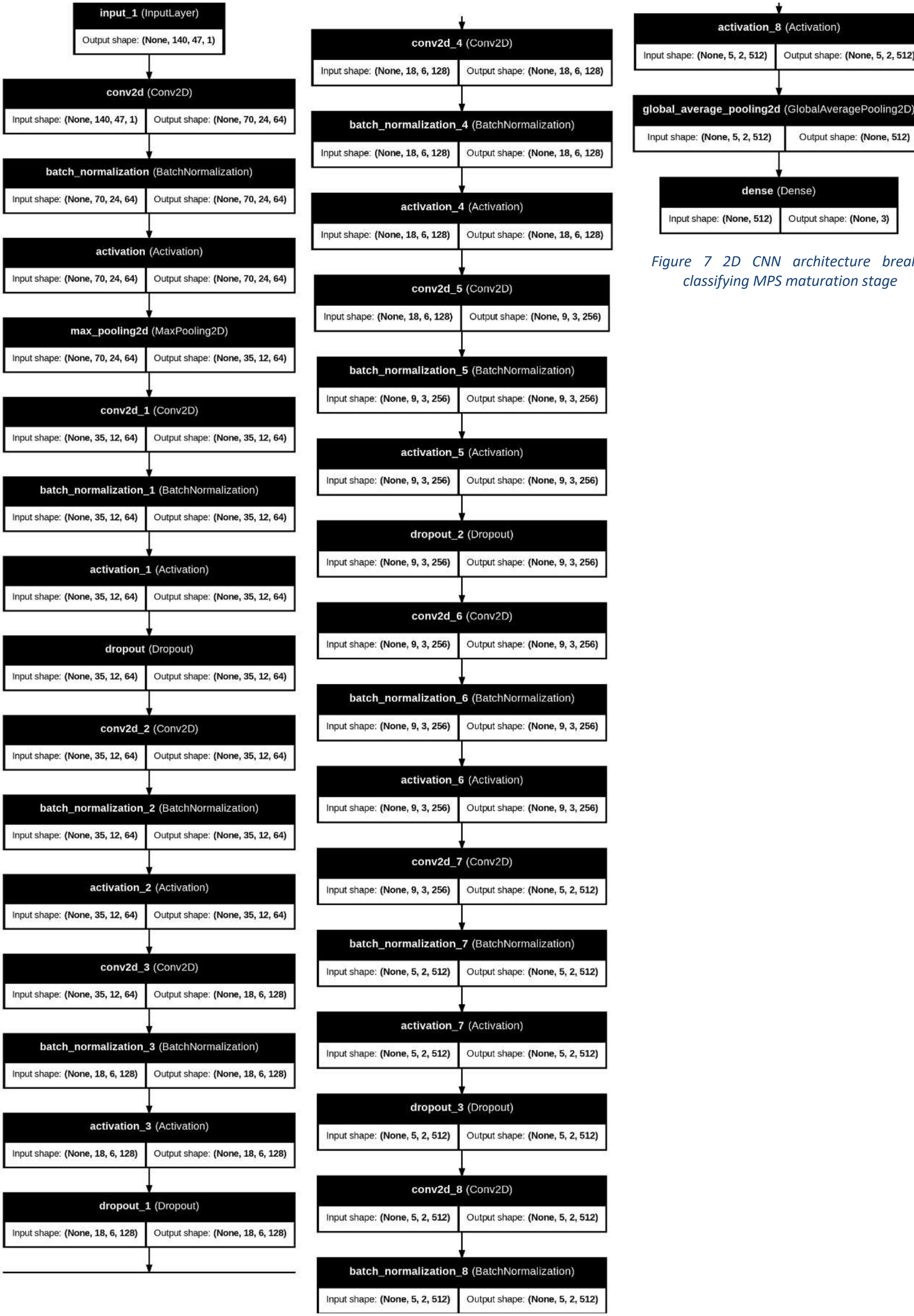

Figure 7 2D CNN architecture breakdown for classifying MPS maturation stage

2.2.6 Evaluation Metrics

To evaluate the performance of the model, precision, recall, and F1-score metrics will be used alongside accuracy and loss. These metrics are calculated for each class (AB, C, DE) to provide a comprehensive assessment of the model's classification abilities.

- **Precision** is the percentage of correct positive predictions out of all the positive predictions the model makes. High precision indicates that the model has a low false positive rate.
- **Recall** (also known as sensitivity) is the percentage of actual positive cases that the model correctly identifies. High recall indicates that the model has a low false negative rate.
- **F1-Score** is the combined measure of precision and recall. It gives a single score that reflects both how many mistakes the model avoids and how many positive cases it correctly finds.

2.3 Results

2.3.1 ROI classifier Model Evaluation

The ROI classifier was evaluated on the test set comprising 57 preferred images and 42 non-preferred images. The evaluation results demonstrated exceptional performance, with a test accuracy of 99%. Additionally, the model's accuracy and loss were monitored during training and validation, with both training and validation accuracies reaching 100% and negligible losses.

Once the ROI classifier was trained and validated, all the axial CBCT slices from the 111 patients were deployed through the classifier. This resulted in a dataset which consisted of 580 images. These images were compared to the images identified as ideal axial slices containing the MPS by orthodontists, which showed that the model was 99% accurate.

This dataset was then split patient-wise 70% (training) 20% (validation) and 10% (testing) to be used for the 2D CNN classification of MPS model. The training data was a total of 406 arrays (AB: 128 C: 158 DE: 120)

2.3.1 Training of Convolutional Neural Network (CNN)

The CNN model was trained for the classification of midpalatal maturation stages on CBCT images. The input shape of the images was set to 140x47x1, with a total of 405 training data samples distributed among three groups: AB (127 samples), C (158 samples), and DE (120 samples).

The training process utilized early stopping with a patience of 15 epochs and model checkpointing to save the best weights and model. The model was set for 100 epochs with a batch size of 16 but training stopped at 90 epochs where the validation loss failed to improve from epoch 75 onwards. Thus, the model was saved on epoch 75 which lasted 5 seconds with 218ms/step with a training accuracy of 97.81 %, validation loss of 0.0626 and validation accuracy of 98.08%. (Figure 7 and 8)

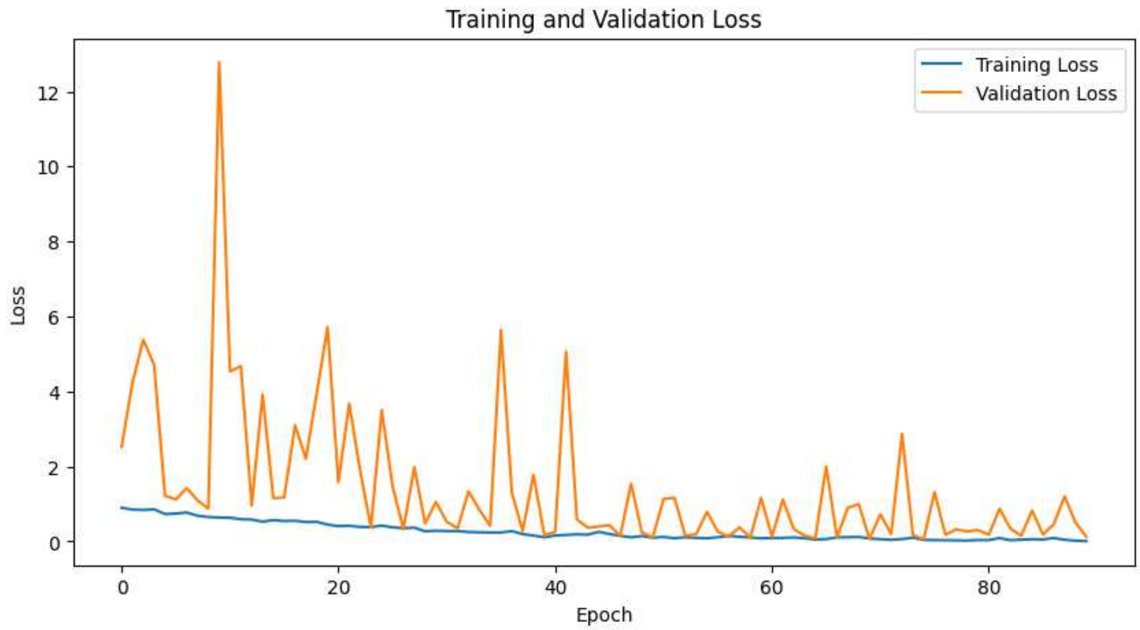

Figure 8 Training and Validation Loss 2D CNN model

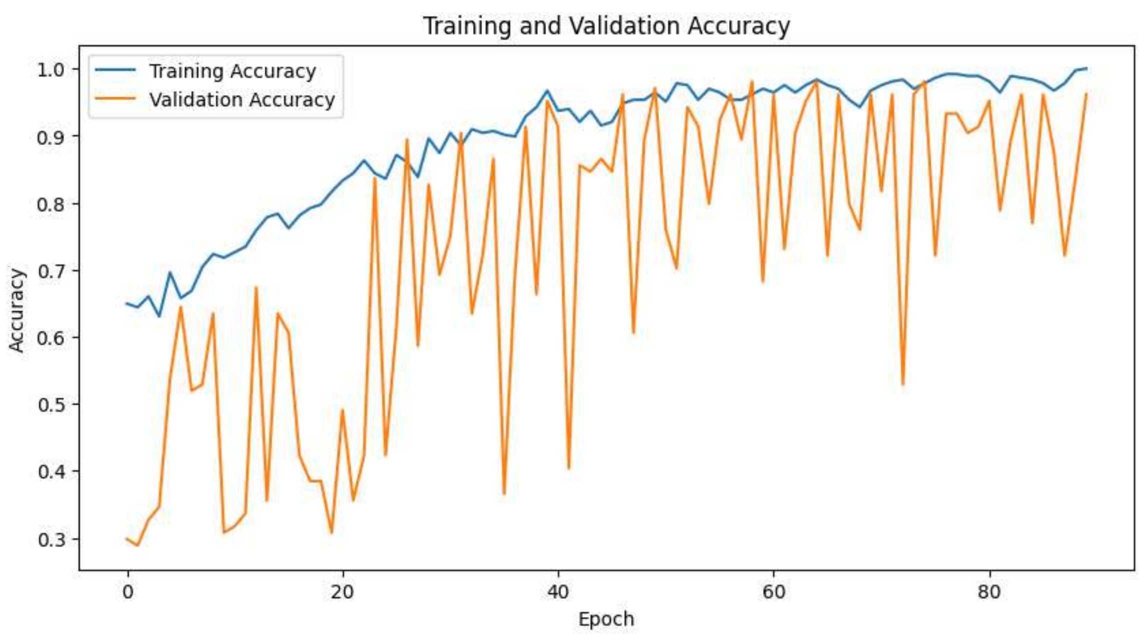

Figure 9 Training and Validation Accuracy 2D CNN model

2.3.2 Testing Results

The trained model was evaluated on a separate test dataset consisting of 57 images, distributed among the three groups: AB (18 samples), C (19 samples), and DE (20 samples).

The testing results indicated a test loss of 0.2286 and a test accuracy of 96.49%. Furthermore, the classification performance of the model was assessed using precision, recall, and F1-score metrics for each class (AB, C, DE). The results demonstrated high precision, recall, and F1-score values for all classes, indicating the effectiveness of the model in accurately classifying midpalatal maturation stages.

The detailed classification metrics for each class are as follows:

Table 1 Evaluation Metrics for testing 2D CNN model

Class	Precision	Recall	F1-Score
AB (18)	0.90	1.00	0.95
C (19)	1.00	1.00	1.00
DE (20)	1.00	0.90	0.95
Overall (57)	0.97	0.97	0.96

The overall accuracy and classification metrics demonstrate the efficacy of the proposed CNN model in accurately classifying midpalatal maturation stages on CBCT images.

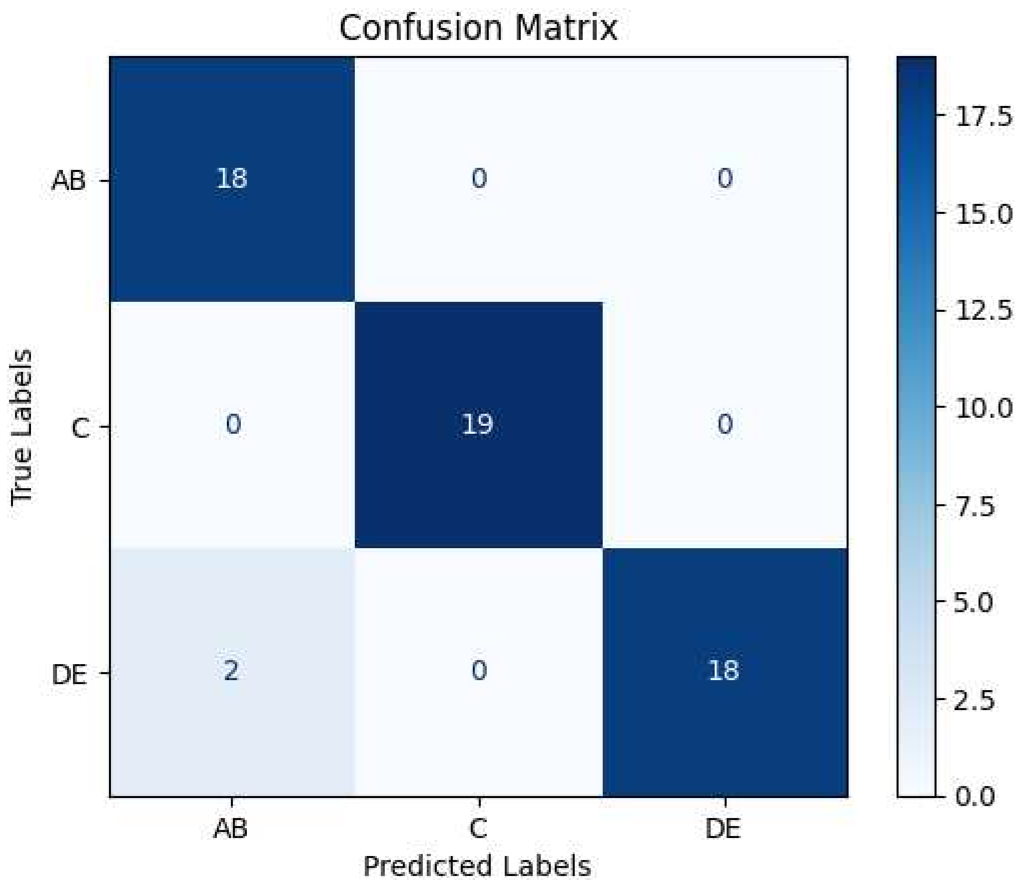

Figure 10 Confusion Matrix for 2D CNN model testing results

2.4 Discussion

Our study demonstrated the efficacy of this developed CNN architecture with 9 convolutional layers and 4.7 million parameters in classifying the midpalatal suture into three classes

(AB, C, DE) with a validation accuracy of 98.08%. Subsequent testing on a separate dataset consisting of 57 images yielded a testing accuracy of 96.49%.

The robust performance of our developed CNN model in accurately classifying midpalatal maturation stages on CBCT images underscores the potential of AI-based approaches, particularly deep learning algorithms like CNNs, in automating the assessment of MPS maturation. With a high training accuracy of 97.81% and validation accuracy of 98.08%, our model showcases its ability to distinguish between different stages of MPS maturation, thereby facilitating more precise treatment planning in orthodontics.

Comparing our results with those of Zhu et al., who also employed a lightweight CNN model (ResNet18) for midpalatal suture maturation stage classification from CBCT images, highlights the robustness of our model. Despite having fewer parameters (4,733,187 parameters vs. 11.69 million parameters), our model achieved a relatively higher training accuracy (97.81% vs. 79.10%). It is worth noting that our dataset consisted of 111 CBCTs, considerably smaller than the 785 CBCTs used in Zhu et al.'s study.

Moreover, our findings align with previous research demonstrating the effectiveness of deep learning techniques, particularly CNNs, in medical image analysis. The utilization of CNNs has revolutionized the interpretation of complex biomedical images, as evidenced by studies such as Mishra et al., and Schwendicke et al., which have shown CNNs' superior performance in tasks ranging from interstitial lung disease pattern classification to dental diagnosis and treatment planning. (Schwendicke, Cejudo Grano de Oro et al. 2022, Mishra, Tiwari et al. 2023)

Furthermore, the scalability and efficiency of our model offer practical advantages, especially in resource-constrained environments. The lightweight nature of our CNN architecture, coupled with its high accuracy, makes it well-suited for deployment on edge devices or low power computing platforms, expanding its potential applications in clinical settings where computational resources may be limited.

In conclusion, our study demonstrates the potential of deep learning, specifically CNNs, in automating the assessment of midpalatal suture maturation on CBCT images. By achieving high accuracy with a relatively lightweight model, we provide a valuable tool for orthodontic practitioners to enhance treatment planning accuracy and efficiency. Future research could focus on further refining the model and validating its performance across larger and more diverse datasets, ultimately advancing the integration of AI in orthodontic practice and improving patient outcomes.

However, it is important to acknowledge the limitations of our study. Firstly, the relatively small sample size of CBCT images may limit the generalizability of our findings to broader populations, which is a common issue in medical research. This underscores the need for future studies with larger datasets to validate the robustness of our CNN model. Notably, during testing, two samples determined to be stage AB were misclassified by the model as stage DE, indicating a type 2 error that may be attributed to the smaller sample size in the DE group.

Additionally, our proposed method solely focuses on analyzing individual slices of the midpalatal suture. While this approach is effective when the palate is not curved and the suture can be seen as a whole entity in a single slice, it may result in outliers if the palate curvature obstructs a comprehensive view of the suture in one slice. This limitation underscores the importance of considering alternative approaches, such as a three-dimensional (3D) analysis of the midpalatal suture, which may provide more comprehensive insights into its maturation stages.

2.5 Conclusion

Our study demonstrates the effectiveness of a 2 layer CNN model for binary classification of preferred CBCT slices with a training and testing accuracy of 100%. Also our study demonstrated the effectiveness of the lightweight CNN architecture in accurately classifying midpalatal suture maturation stages on CBCT images. With high training and validation accuracies of 97.81% and 98.08%, respectively, our CNN model showcases robust performance in distinguishing between different stages of midpalatal suture maturation, offering valuable insights for orthodontic treatment planning.

However, it is crucial to acknowledge that our study's findings are heavily dependent on Anglieri's staging for midpalatal suture maturation. This reliance presents a potential limitation, as the validity of our results could be compromised if future research were to challenge or disprove the accuracy of Anglieri's classification. This underscores the importance of continuous validation and the exploration of alternative classification systems to ensure the long-term relevance of AI-driven models in orthodontic diagnosis.

Addressing the limitations discussed in 2.4 Discussion, the next chapter will explore a 3D approach to midpalatal suture analysis, complementing the 2D approach presented in this study. By incorporating advanced imaging techniques and innovative methodologies, we aim to further enhance the accuracy and applicability of AI-driven solutions in orthodontic diagnosis and treatment planning.

Chapter 3- 3D CNN approach to classification of MPS on CBCTs

3.1 Introduction and Background

In the previous chapters, the clinical importance of correct classification of MPS was discussed. The study that was conducted in chapter 2 and similar studies in the literature such as Zhu et. Al looked at the suture in a two-dimensional (2D) view, whereas the suture is a three-dimensional (3D) structure with depth, which is evident in the response of the suture to expansion forces. (Bocklet, Ahmadi et al. 2024) Using a two-dimensional approach to a three-dimensional structure has the inevitable risk of loss of information. Finite element analysis (FEA) and other 3D modeling techniques used in studies show that 2D imaging cannot provide the detailed data required for accurate analysis of craniofacial structures. (Khanehmasjedi, Bagheri et al. 2020, Kaya, Seker et al. 2023)

Patients who have curved palates in which there is the inability to observe both the anterior and posterior segments of the midpalatal suture within a singular axial slice, mandating the utilization of two slices for sutural staging classification (Angelieri, Franchi et al. 2016), were excluded in studies that utilized 2D CNN approaches (Zhu, Yang et al. 2024).

In this study, I propose a 3D deep learning architecture and approach to the classification of midpalatal sutures to include curved palates. No similar study to classify MPS from 3D images has been done yet to the best of my knowledge.

3.2 Objectives

1. To develop and evaluate a 3D CNN model for classifying midpalatal suture maturation stages on 3D arrays from CBCT images.
2. To create 3D arrays from 2D axial slices of CBCT images

3.3 Methodology

This research was approved by the Health Research Ethics Board (HREB) of the University of Alberta (Approval number: Pro00125920). CBCT from 155 patients between the age of 7-21 who underwent CBCT (120 kVp, 5 mA and 4 sec, 0.3 voxel) for orthodontic treatments at the university clinic were initially collected for this study.

The exclusion criteria were as follows:

1. Prior orthodontic treatment
2. Impacted upper teeth in the mid palatal region
3. Congenital malformations of the maxilla (i.e cleft palate)
4. Blurry or low quality CBCTs due to scatter

Out of the 155 patients, 145 did not have the exclusion criteria and were thus included in the study. The CBCT images were in DICOM format and anonymized prior to collection for this study, thus, they were first converted from DICOM format to PNG image format with the ITK-SNAP software (726 * 644 pixels). The axial view of the patient images' which consisted of the average number of 450 slices was classified into 5 groups of MPS maturation stage as first stated by Angeliere et al. by two orthodontists. In the case of any conflicts, a third orthodontist evaluated the images for determining the class of MPS. Once the patients were classified based on the Angeliere method, the slices belonging to patients' palates were selected and saved in a separate folder.

The palate slices were then classified into three groups AB (maturation stages A and B) C and DE (maturation stages D and E), the same as chapter 2. This was done as a way to reduce the amount of variability for the DL model as the aim of this study was to help in reaching a diagnosis for an optimal treatment plan and the treatments for stages A and B are clinically the same, but can differ from stage C, and stages A B and C have different treatments compared to stages D and E. (Ladewig, M. et al., 2018)

3.3.1 Data preprocessing:

To prepare the CBCT axial slices containing the MPS to be stacked together into 3D arrays for this study, I first determined the ROI in the slices, as detailed in Chapter 2, Section 2.2.1. After identifying the ROI, I needed to classify and select the ideal slice containing the MPS from all the CBCT axial slices. This selection was done using the ROI classifier described in Chapter 2, Section 2.2.2.

3.3.2 Three dimensional arrays and Training, Validation and Testing

The original dataset consisted of 145 CBCT images. To create 3D arrays from the axial slices with the ROI cropped, the slices containing the midpalatal suture were stacked using the Numpy library. Since CNN models require uniform input shapes, all 3D arrays were reshaped to (12, 140, 47, 3), matching the size of the largest array, where 12 represents the depth of the arrays. (Figure 10)

145 CBCT images were used for training, validation, and testing. These were split into 50% for training, 20% for validation, and 30% for testing. To address the limited number of CBCTs that met the inclusion criteria, horizontal flip augmentation was applied to the training data alone, resulting in a total of 150 arrays (AB: 46, C: 62, DE: 42).

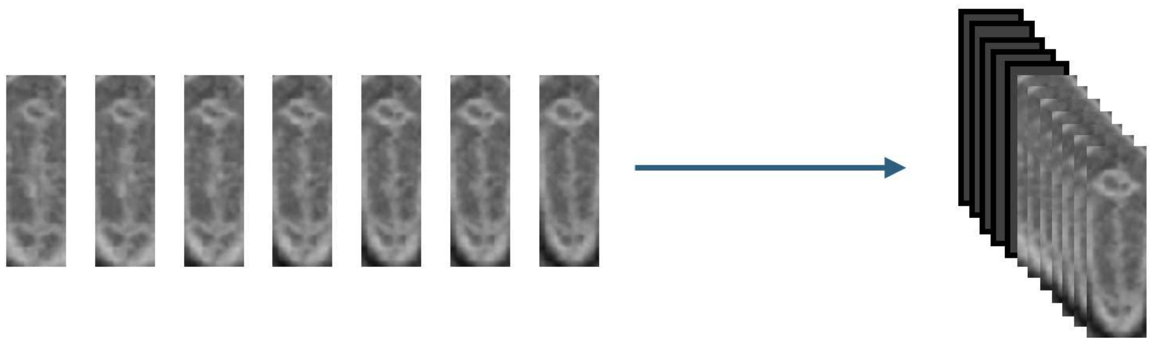

Figure 11 Creating 3D arrays from 2D axial slices of CBCT

3.3.3 CNN Architecture

The CNN architecture utilized in this study is designed to classify 3D data into one of three classes (AB, C, DE). This model consists of four convolutional blocks, each followed by batch normalization to stabilize training, rectified linear unit (ReLU) activation to introduce non-linearity, and max-pooling layers for spatial downsampling.(Figures 11 and 12)

In the first convolutional block, input tensors undergo convolution with 64 filters of size 3x3x3, followed by batch normalization, ReLU activation, and max-pooling with a stride of 2x2x2. Subsequently, in the following three convolutional blocks, the number of filters is increased progressively to 128, 256, and another 256 respectively, while the kernel size and pooling operations remain consistent.

After the convolutional layers, the tensor is flattened to be processed by fully connected layers. The flattened tensor is passed through two dense layers with 256 and 128 neurons respectively, both employing ReLU activation functions. Finally, the output layer utilizes softmax activation to produce class probabilities.

The model is compiled using the Adam optimizer with a categorical cross-entropy loss function, and accuracy is monitored during training.

Additionally, to facilitate model training, we implemented two callback mechanisms. The first is ModelCheckpoint, which periodically saves the model's weights during training. These checkpoints are stored in the directory, with the best performing model being saved based on validation accuracy. The second callback, EarlyStopping, terminates training if the monitored metric in our case validation accuracy and validation loss does not improve for a specified number of epochs (patience), thus preventing overfitting. Both callbacks contribute to efficient model training and ensure the preservation of the best performing model.

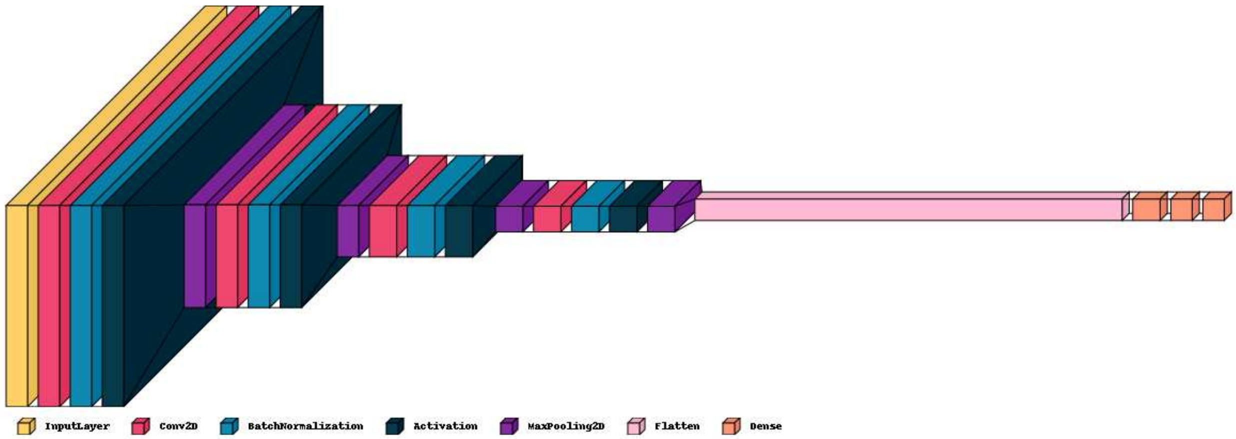

Figure 12 3D CNN architecture overview for classifying MPS maturation stage

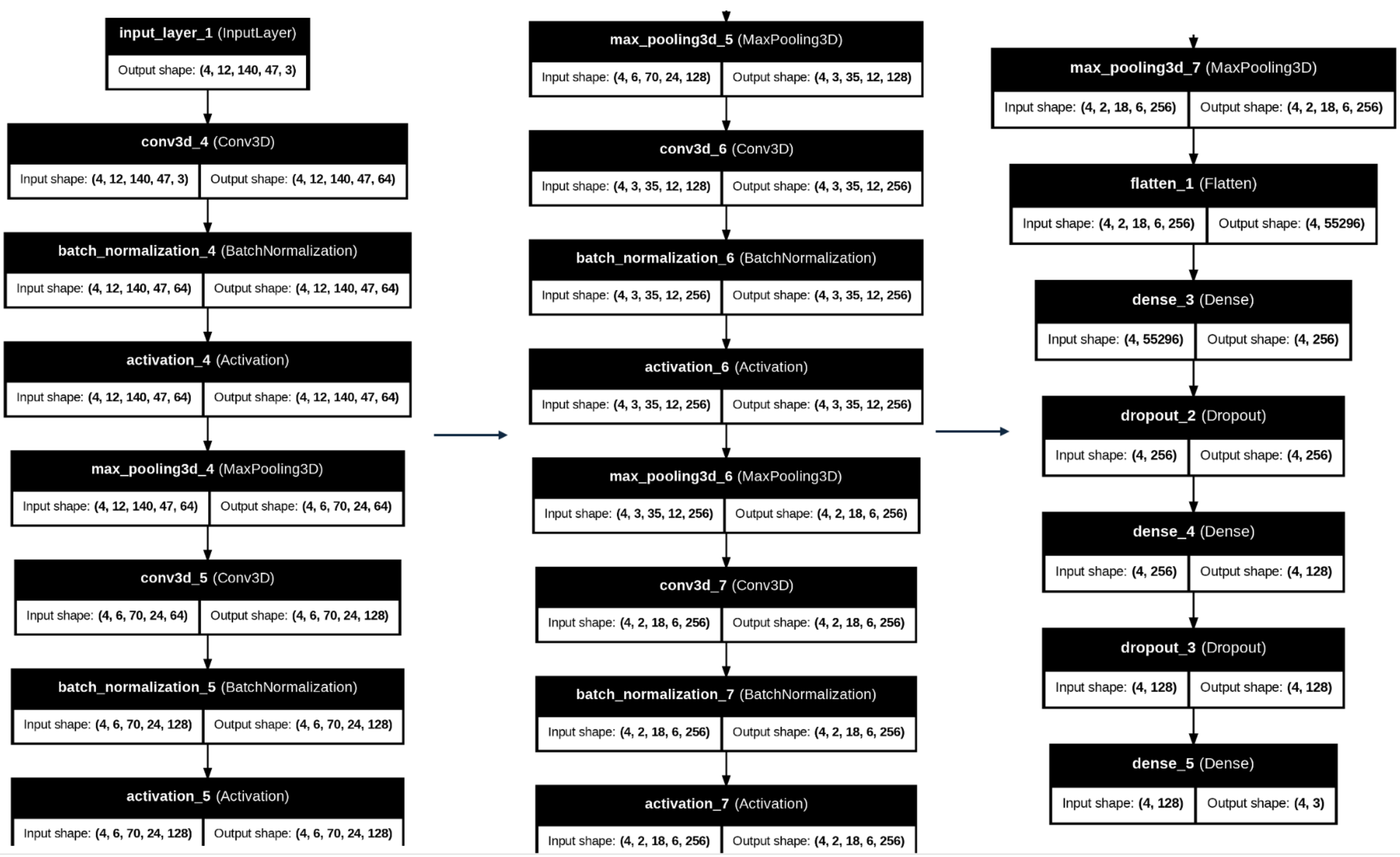

Figure 13 3D CNN architecture breakdown for classifying MPS maturation stage

3.4 Results

3.4.1 Training of Convolutional Neural Network (CNN)

Cone Beam Computed Tomography (CBCT) images from a cohort of 145 patients were analyzed, with each patient contributing between 3 to 12 slices on average. This is because each patient has a different thickness of palate, which means a patient might have a palate which shows the MPS in numerous slices while another patient with a thin palate only has the MPS visible in a few slices. Two orthodontists independently categorized these images into five maturation stages, achieving an inter-rater reliability of 68%. In instances of disagreement, a third orthodontist intervened to provide a final categorization. The images were then classified into three groups of AB, C and DE due to the clinical approach for the treatment of these stages. For the trained convolutional neural network (CNN) model, we observed a training accuracy of 64% and a validation accuracy of 71.43%. (Figure 13)

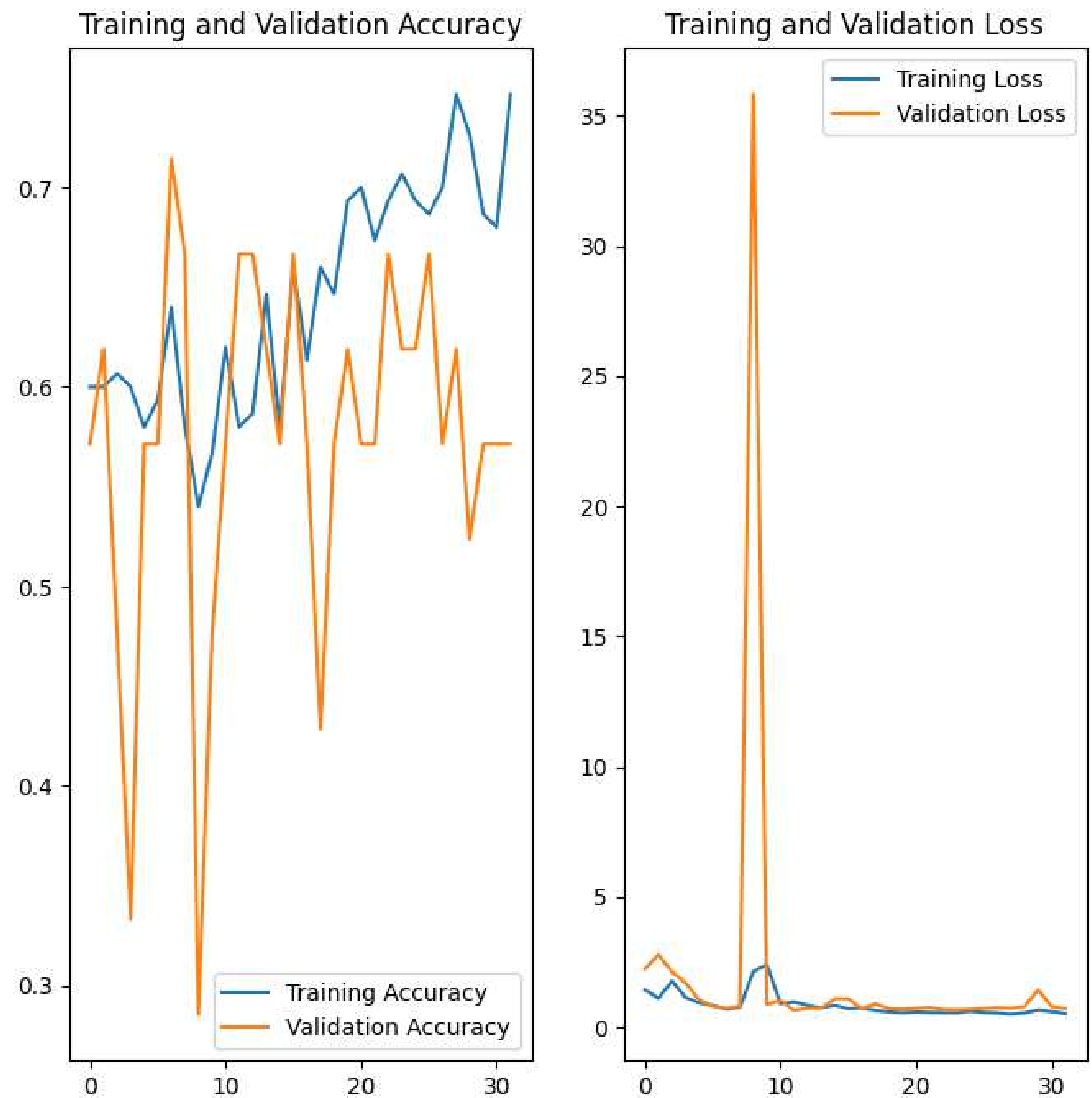

Figure 14 Training and Validation Accuracy and Loss for 3D CNN model

3.4.2 Testing Results

The trained model was evaluated on a test dataset. The set consists of 46 arrays, distributed among the three groups: AB (14 samples), C (23 samples), and DE (9 samples).

The testing results from the first test set indicated a test loss of 0.9 and a test accuracy of 78.26%. Furthermore, the classification performance of the model was assessed using precision, recall, and F1-score metrics for each class (AB, C, DE). The results demonstrated acceptable precision, recall, and F1-score values for all classes, indicating the effectiveness of the model in accurately classifying midpalatal maturation stages.

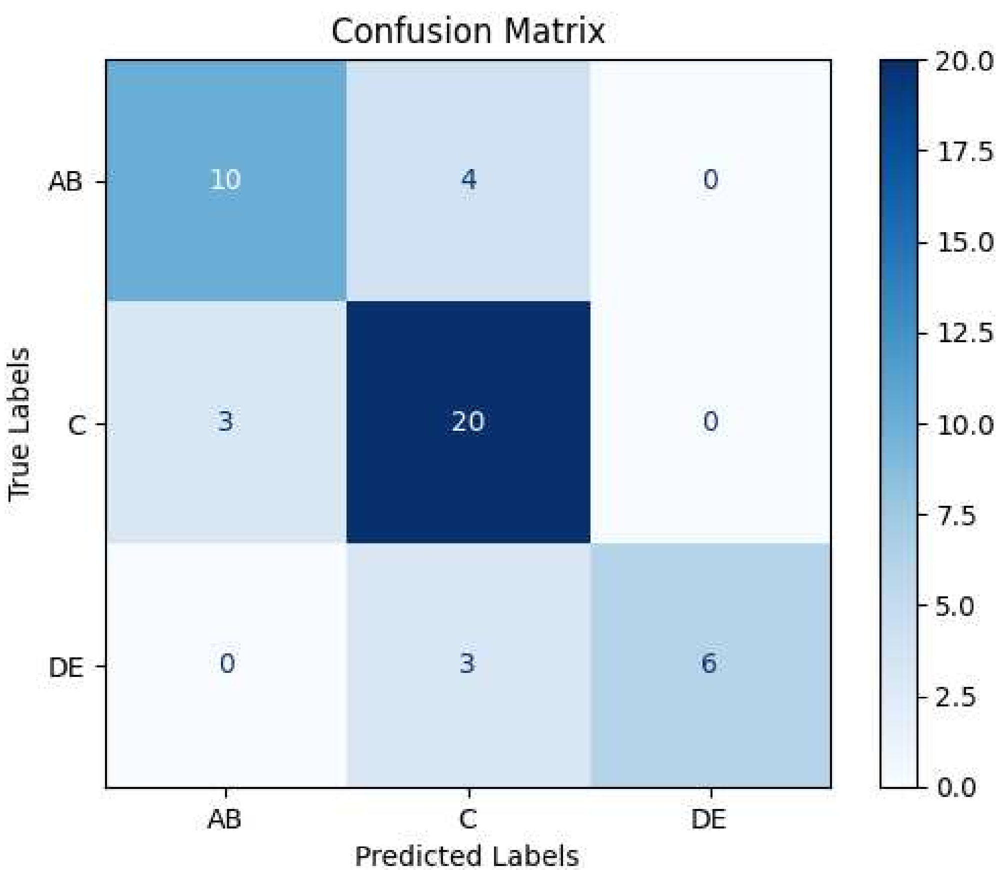

Figure 15 Confusion Matrix for testing 3D CNN model on test set

Table 2 Evaluation Metrics for testing 3D CNN model

Class	Precision	Recall	F1-score	Support
AB	0.77	0.71	0.74	14
C	0.74	0.87	0.80	23
DE	1.00	0.67	0.80	9
Overall	0.84	0.75	0.78	46

3.5 Discussion

Our proposed 3D CNN model, with a testing accuracy of 78.26%, shows promise in classifying midpalatal suture (MPS) maturation stages on CBCT images. While these results are promising, further analysis and comparison with existing methods are necessary to fully understand the model's performance and its implications in clinical settings.

One significant advancement of our approach is the transition from a 2D to a 3D analysis of the MPS. The 2D CNN model demonstrated high training and validation accuracies of 97.81% and 98.08%, respectively, with a test accuracy of 96.49%. However, the 2D approach inherently simplifies the complex anatomy of the MPS by analyzing only individual slices, which may overlook crucial spatial information that is only captured in

3D. This limitation is particularly relevant in cases where the palate is curved, or the suture does not present as a single entity in any given slice.

The 3D volumetric architecture, on the other hand, offers a more comprehensive evaluation by considering the entire volumetric structure of the MPS. This allows for a more accurate assessment of suture maturation stages, potentially leading to different staging results compared to 2D analysis. The 3D approach can capture subtle variations in the suture that may be missed in 2D slices, providing a more holistic view of the maturation process.

Despite these advantages, our 3D model's performance indicates the need for further refinement. The relatively lower testing accuracy compared to the 2D model suggests that the 3D CNN requires optimization and validation with larger datasets. Nonetheless, the shift towards 3D analysis is a crucial step forward, as it aligns more closely with the anatomical reality of the MPS, potentially improving the accuracy and reliability of orthodontic treatment planning.

In conclusion, while our study demonstrates the feasibility and potential of a 3D CNN approach for MPS classification, it also highlights the importance of continued research to optimize and validate these models. The comparison between 2D and 3D methods underscores the necessity of a volumetric approach in capturing the full complexity of craniofacial structures, paving the way for more accurate and effective AI-driven tools in orthodontics.

3.6 Conclusion

Our study underscores the potential of employing a 3D CNN model for the classification of 3D arrays from MPS maturation stages on CBCT images. With a testing accuracy of 78.57%, our proposed model demonstrates promising performance in this critical area of orthodontic diagnosis.

CHAPTER 4- DISCUSSION

This thesis aimed to improve orthodontic diagnosis by integrating machine learning techniques to tailor treatment plans to each patient's unique needs. Specifically, it focused on the application of deep learning models to classify MPS maturation stages using CBCT images. Two main approaches were introduced: a 2D CNN model and a 3D CNN model. The 2D CNN model demonstrated significant potential in classifying MPS maturation stages from individual CBCT slices, achieving a validation accuracy of 98.08% and a test accuracy of 96.49%. The 3D CNN model, developed to address the limitations of the 2D approach for curved palates, showed promise with a testing accuracy of 78.26% on a separate test set.

4.1 Detailed Discussion of Results

4.1.1 2D CNN Model

The 2D CNN model's high validation and test accuracies underscore its effectiveness in classifying MPS maturation stages. This model used a 9-layer CNN architecture which, despite being lightweight, yielded high accuracy. The use of a 2D approach was particularly effective for non-curved palates where the suture can be visualized as a whole in a single slice. This model's simplicity and computational efficiency make it suitable for deployment in resource-constrained environments, such as dental clinics with limited access to high-end computing resources.

This model included dropout layers and batch normalization. This simplification, combined with effective preprocessing steps such as data augmentation and ROI classification, allowed the model to focus on the suture rather than surrounding structures. The resulting model demonstrated a robust ability to distinguish between different MPS maturation stages, which is crucial for accurate treatment planning.

Comparing our 2D model's performance with existing studies, particularly Zhu et al.'s work, highlights its competitive edge. Zhu et al. employed a ResNet-18 model for MPS classification and achieved an accuracy of 79.10%. Despite having fewer parameters (4.73 million vs. 11.69 million), our model achieved higher accuracy. Zhu et al.'s study involved a larger dataset of 785 CBCTs, yet our model, with a smaller dataset of 111 CBCTs, demonstrated superior performance. This comparison underscores the effectiveness of optimized lightweight models for specific medical imaging tasks.

Additionally, Zhu et al. explored the use of deeper models like ResNet-50 and ResNet-101, which achieved higher accuracies but at the cost of significantly increased computational complexity. Our study, by contrast, demonstrates that with proper optimization, a smaller model can achieve comparable or even superior performance, thus highlighting the practicality of using small CNNs in resource-constrained clinical settings.

4.1.2 3D CNN Model

The 3D CNN model was developed to address the inherent limitations of the 2D approach, particularly for curved palates where the suture is not visible in a single slice. The 3D model aimed to capture the three-dimensional nature of the MPS, providing a more comprehensive analysis. This model used a volumetric approach, stacking CBCT slices to create 3D arrays which were then processed using a CNN architecture.

The 3D CNN model achieved testing accuracy of 78.26% on a separate test set. While lower than the 2D model, this result is promising given the complexity of the 3D structure. The model's architecture included four convolutional blocks with progressively increasing filters, followed by fully connected layers. This design allowed the model to capture spatial relationships and contextual information across multiple slices, providing a holistic evaluation of the MPS.

4.2 Clinical Implications

Accurately assessing suture maturation stages is crucial for orthodontic and orthognathic surgery planning, enabling tailored interventions based on individual patient characteristics. By offering automated, objective classification, our models could improve diagnostic accuracy and streamline clinical workflows. Reducing the reliance on subjective manual classification can help standardize diagnosis and treatment planning across practitioners, leading to more consistent treatment outcomes.

Importantly, our model holds potential for guiding clinicians in selecting the most appropriate treatment options for both non-surgical and surgical cases. For non-surgical cases, the model can assist in determining the maturation stage of the midpalatal suture (MPS), enabling clinicians to choose interventions tailored to the specific stage of development. Similarly, for surgical cases, the model can aid in classifying the suture's maturation, ensuring that clinicians can make informed decisions about the necessity and type of surgical intervention. This automated approach may significantly improve treatment accuracy and outcomes in both contexts.

4.2.1 Scenarios with Small CNNs

Using small CNN models offers several key advantages:

1. Efficiency and Speed: Small CNNs are computationally efficient, which is crucial for applications in clinical settings where real-time processing may be required. This efficiency can lead to quicker diagnosis and treatment planning.
2. Resource Constraints: Smaller models require less computational power and memory, making them suitable for deployment on devices with limited resources, such as portable medical devices or older computers often found in clinical environments.
3. Generalization: Small CNNs, with fewer parameters, are less likely to overfit, especially on smaller datasets. This can be particularly beneficial in medical imaging where collecting large datasets can be challenging.

4. Interpretability and Maintenance: Smaller models are easier to interpret, debug, and maintain. This is an important factor in medical applications where understanding model decisions can be critical for gaining trust from medical professionals and ensuring patient safety.

4.2.2 Specific Use Cases: Real-World Applications

In real-world applications, small CNNs are particularly advantageous for initial screenings or routine check-ups where a quick assessment is needed. For example, in a busy orthodontic clinic, a small CNN can provide reliable results swiftly, enabling efficient patient management. Additionally, in rural or under-resourced medical facilities, deploying small CNNs on existing hardware can still provide accurate diagnostic capabilities without the need for expensive upgrades. These models can be used in portable medical devices for fieldwork, ensuring accessibility to advanced diagnostic tools even in remote locations.

4.3 Limitations and Future Directions

While our study demonstrates the potential of deep learning models in classifying MPS maturation stages, several limitations must be acknowledged. The relatively small sample size may limit the generalizability of our findings. Future studies should include larger and more diverse datasets to validate the robustness of our models. Additionally, the 3D CNN model's accuracy, while promising, indicates room for improvement.

4.4 Conclusion

This thesis highlights the potential of deep learning, particularly convolutional neural networks (CNNs), in automating the classification of midpalatal suture maturation stages on CBCT images. The developed 2D and 3D models offer promising tools that could assist orthodontic practitioners by improving diagnostic accuracy and efficiency. The 2D CNN model shows strengths in computational efficiency and accuracy, while the 3D CNN model provides a more comprehensive analysis, which may be useful in more complex cases where a 2D approach could be limited.

However, despite these encouraging results, further validation and refinement are necessary. Expanding the sample size, applying more advanced data augmentation techniques, and incorporating additional imaging modalities could improve model performance. Additionally, the practical integration of these models into clinical settings has yet to be evaluated, which is crucial to understanding their real-world impact on patient outcomes.

It is also important to note that this study relies heavily on Anglieri's staging system for midpalatal suture maturation. This presents a significant limitation, as the validity of our findings may be affected if future research calls into question the accuracy or relevance of Anglieri's classification. Acknowledging this, ongoing validation and the exploration of alternative classification systems will be critical to ensure the continued relevance and accuracy of AI-driven models in orthodontic diagnosis.

In conclusion, while this study lays a foundation for using advanced computational tools in craniofacial imaging analysis, much work remains. By harnessing the strengths of both 2D and 3D CNN models, we can continue to improve diagnostic tools, ultimately aiming to enhance patient care and treatment outcomes in orthodontics.

REFERENCES

Abadi, M., et al. (2016). "Tensorflow: Large-scale machine learning on heterogeneous distributed systems." arXiv preprint arXiv:1603.04467.

Abou Ghayda, R., et al. (2024). "Artificial intelligence in andrology: from semen analysis to image diagnostics." The World Journal of Men's Health **42**(1): 39.

Ackerman, J. L. and W. R. Proffit (1969). "The characteristics of malocclusion: a modern approach to classification and diagnosis." American journal of orthodontics **56**(5): 443-454.

Afzaal, U., et al. (2021). "An instance segmentation model for strawberry diseases based on mask R-CNN." Sensors **21**(19): 6565.

Agarwal, A. and R. Mathur (2010). "Maxillary expansion." International journal of clinical pediatric dentistry **3**(3): 139.

Akbulut, S., et al. (2020). "Prediction of rapid palatal expansion success via fractal analysis in hand-wrist radiographs." American Journal of Orthodontics and Dentofacial Orthopedics **158**(2): 192-198.

Al-Badri, A. H., et al. (2022). "Hybrid cnn model for classification of rumex obtusifolius in grassland." Ieee Access **10**: 90940-90957.

Al-Khuzaie, M. I. and W. A. M. Al-Jawher (2023). "Enhancing Medical Image Classification: A Deep Learning Perspective with Multi Wavelet Transform." Journal Port Science Research **6**(4): 365-373.

Al-Sarem, M., et al. (2022). "Enhanced tooth region detection using pretrained deep learning models." International Journal of Environmental Research and Public Health **19**(22): 15414.

Aljuaid, A. and M. Anwar (2022). "Survey of supervised learning for medical image processing." SN Computer Science **3**(4): 292.

Alloghani, M., et al. (2020). "A systematic review on supervised and unsupervised machine learning algorithms for data science." Supervised and unsupervised learning for data science: 3-21.

Alqazzaz, S., et al. (2022). "Combined features in region of interest for brain tumor segmentation." Journal of Digital Imaging **35**(4): 938-946.

Alzaid, N., et al. (2023). "Revolutionizing Dental Care: A Comprehensive Review of Artificial Intelligence Applications Among Various Dental Specialties." Cureus **15**(10).

Angelieri, F., et al. (2016). "Prediction of rapid maxillary expansion by assessing the maturation of the midpalatal suture on cone beam CT." Dental press journal of orthodontics **21**(06): 115-125.

Ashkani-Esfahani, S., Mojahed-Yazdi, R., et al. (2022). "Deep learning algorithms improve the detection of subtle Lisfranc malalignments on weightbearing radiographs." *Foot & Ankle International*, 43, 1118-1126.

Ayodele, T. O. (2010). "Types of machine learning algorithms." New advances in machine learning **3**(19-48): 5-1.

Barbosa, N. M. V., et al. (2019). "Reliability and reproducibility of the method of assessment of midpalatal suture maturation: A tomographic study." The Angle Orthodontist **89**(1): 71-77.

Baughan, N., et al. (2023). "Radiomic and deep learning characterization of breast parenchyma on full field digital mammograms and specimen radiographs: a pilot study of a potential cancer field effect." Journal of Medical Imaging **10**(4): 044501-044501.

Bloomquist, D. S. and D. R. Joondeph (2019). Orthognathic surgical procedures on non-growing patients with maxillary transverse deficiency. Seminars in Orthodontics, Elsevier.

Bocklet, M., et al. (2024). "Comparison of 3D-printed and laboratory-fabricated Hyrax on stress distribution and displacement of the maxillary complex: a 3D finite element study." Progress in orthodontics **25**(1): 11.

Bradski, G. and A. Kaehler (2000). "OpenCV." Dr. Dobb's journal of software tools **3**(2).

Brunetto, D. P., et al. (2017). "Non-surgical treatment of transverse deficiency in adults using Microimplant-assisted Rapid Palatal Expansion (MARPE)." Dental press journal of orthodontics **22**: 110-125.

Budach, L., et al. (2022). "The effects of data quality on machine learning performance." arXiv preprint arXiv:2207.14529.

Calvo-Henriquez, C., et al. (2021). "The impact of maxillary expansion on adults' nasal breathing: a systematic review and meta-analysis." American Journal of Rhinology & Allergy **35**(6): 923-934.

Cazzato, D., et al. (2020). "A survey of computer vision methods for 2d object detection from unmanned aerial vehicles." Journal of Imaging **6**(8): 78.

Celebi, F. and S. Akbulut (2020). "Relationship between the position of maxilla and rapid maxillary expansion failure." South European Journal of Orthodontics and Dentofacial Research **7**(2): 49-54.

Cha, Y. J., et al. (2017). "Deep learning-based crack damage detection using convolutional neural networks." Computer-Aided Civil and Infrastructure Engineering **32**(5): 361-378.

Cheng, Y., et al. (2017). "A survey of model compression and acceleration for deep neural networks." arXiv preprint arXiv:1710.09282.

Chiu, P.-F., et al. (2023). "Machine Learning Assisting the Prediction of Clinical Outcomes following Nucleoplasty for Lumbar Degenerative Disc Disease." Diagnostics **13**(11): 1863.

Chollet, F. (2019). "On the measure of intelligence." arXiv preprint arXiv:1911.01547.

Dakhil, N. B. and F. B. Salamah (2021). "The diagnosis methods and management modalities of maxillary transverse discrepancy." Cureus **13**(12).

Darawsheh, A. F., et al. (2023). "Applicability of fractal analysis for quantitative evaluation of midpalatal suture maturation." Journal of Clinical Medicine **12**(13): 4189.

de Melo Quintela, M., et al. (2021). "Miniscrew-assisted rapid palatal expansion (MARPE) protocols applied in different ages and stages of maturation of the midpalatal suture: cases report." Research, Society and Development **10**(11): e503101119480-e503101119480.

de Miranda Ladewig, V., et al. (2018). "Tomographic evaluation of the maturation stage of the midpalatal suture in postadolescents." American Journal of Orthodontics and Dentofacial Orthopedics **153**(6): 818-824.

Duque-Urióstegui, C., et al. (2023). "Treatment of maxillary transverse deficiency with rapid expansion of the palate with mini-implant: a review." International Journal of Research in Medical Sciences **11**(5): 1858.

El-Ghany, S. A., et al. (2023). "Robustness Fine-Tuning Deep Learning Model for Cancers Diagnosis Based on Histopathology Image Analysis." Diagnostics **13**(4): 699.

Elsken, T., et al. (2019). "Neural architecture search: A survey." Journal of Machine Learning Research **20**(55): 1-21.

Erten, O., & Yılmaz, B. (2018). “Three-dimensional imaging in orthodontics.” Turkish Journal of Orthodontics, 31(3), 86-94.

Esteva, A., et al. (2021). "Deep learning-enabled medical computer vision." NPJ digital medicine **4**(1): 5.

Esteva, A., et al. (2017). "Dermatologist-level classification of skin cancer with deep neural networks." nature **542**(7639): 115-118.

Feng, C., et al. (2021). Tood: Task-aligned one-stage object detection. 2021 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE Computer Society.

Forst, D., et al. (2014). "Radiographic assessment of external root resorption associated with jackscrew-based maxillary expansion therapies: a systematic review." European Journal of Orthodontics **36**(5): 576-585.

Franchi, L., et al. (2010). "Modifications of midpalatal sutural density induced by rapid maxillary expansion: A low-dose computed-tomography evaluation." American Journal of Orthodontics and Dentofacial Orthopedics **137**(4): 486-488.

Gao, L., et al. (2022). "Midpalatal suture CBCT image quantitive characteristics analysis based on machine learning algorithm construction and optimization." Bioengineering **9**(7): 316.

Ghosh, S., et al. (2019). "Understanding deep learning techniques for image segmentation." ACM computing surveys (CSUR) **52**(4): 1-35.

Gidaris, S. and N. Komodakis (2016). Locnet: Improving localization accuracy for object detection. Proceedings of the IEEE conference on computer vision and pattern recognition.

Glanz, V., et al. (2022). "Development and validation of a type 2 diabetes machine learning classification model for EHR-based diagnostics and clinical decision support." bioRxiv: 2022.2010. 2008.511400.

Goodfellow, I. (2016). Deep Learning, MIT Press.

Guo, J., et al. (2022). "TVGG dental implant identification system." Frontiers in Pharmacology **13**: 948283.

Guo, M.-H., et al. (2022). "Attention mechanisms in computer vision: A survey." Computational visual media **8**(3): 331-368.

Gupta, D. (2023). "Image segmentation keras: Implementation of segnet, fcn, unet, pspnet and other models in keras." arXiv preprint arXiv:2307.13215.

Hadi, Z. G., et al. (2021). "Comparison Between Convolutional Neural Network CNN and SVM in Skin Cancer Images Recognition." Journal of Techniques **3**(4): 15-22.

Haghanifar, S., et al. (2017). "Assessment of midpalatal suture ossification using cone-beam computed tomography." Electronic physician **9**(3): 4035.

Harris, C. R., et al. (2020). "Array programming with NumPy." nature **585**(7825): 357-362.

He, K., et al. (2016). Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition.

Hernandez-Alfaro, F., et al. (2010). "Minimally invasive surgically assisted rapid palatal expansion with limited approach under sedation: a report of 283 consecutive cases." Journal of oral and maxillofacial surgery **68**(9): 2154-2158.

Hernández-Alfaro, F. and A. Valls-Ontañón (2021). "SARPE and MARPE." Innovative Perspectives in Oral and Maxillofacial Surgery: 321-325.

Hoeser, T. and C. Kuenzer (2020). "Object detection and image segmentation with deep learning on earth observation data: A review-part i: Evolution and recent trends." Remote Sensing **12**(10): 1667.

Howard, A. G., et al. (2017). "Mobilenets: Efficient convolutional neural networks for mobile vision applications." arXiv preprint arXiv:1704.04861.

Howes, A. E. (1957). "Arch width in the premolar region—still the major problem in orthodontics." American journal of orthodontics **43**(1): 5-31.

Huang, S., et al. (2020). "A deep convolutional neural network architecture for interstitial lung disease pattern classification." Medical & biological engineering & computing **58**: 725-737.

Iandola, F. N., et al. (2016). "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size." arXiv preprint arXiv:1602.07360.

Iodice, G., et al. (2016). "Association between posterior crossbite, skeletal, and muscle asymmetry: a systematic review." European Journal of Orthodontics **38**(6): 638-651.

Isensee, F., et al. (2019). "Automated design of deep learning methods for biomedical image segmentation." arXiv preprint arXiv:1904.08128.

Isfeld, D., et al. (2019). "Evaluation of a novel palatal suture maturation classification as assessed by cone-beam computed tomography imaging of a pre-and postexpansion treatment cohort." The Angle Orthodontist **89**(2): 252-261.

Isfeld, D., et al. (2017). "Novel methodologies and technologies to assess mid-palatal suture maturation: a systematic review." Head & face medicine **13**: 1-15.

Janiesch, C., et al. (2021). "Machine learning and deep learning." Electronic Markets **31**(3): 685-695.

Jaswal, D., et al. (2014). "Image classification using convolutional neural networks." International Journal of Scientific and Engineering Research **5**(6): 1661-1668.

Jiang, B., et al. (2018). Acquisition of localization confidence for accurate object detection. Proceedings of the European conference on computer vision (ECCV).

Jimenez-Valdivia, L. M., et al. (2019). "Midpalatal suture maturation stage assessment in adolescents and young adults using cone-beam computed tomography." Prog Orthod **20**(1): 38.

Katti, G., et al. (2020). "Evaluation of midpalatal suture ossification using cone-beam computed tomography: a digital radiographic study." Acta Medica **63**(4): 188-193.

Kaya, N., et al. (2023). "Comparison of the effects of different rapid maxillary expansion techniques on craniofacial structures: a finite element analysis study." Progress in orthodontics **24**(1): 7.

Ker, J., et al. (2017). "Deep learning applications in medical image analysis." Ieee Access **6**: 9375-9389.

Keras, C. (2015). "Theano-based deep learning libraryCode: https://github. com/fchollet." Documentation: http://keras. io.

Khanagar, S. B., et al. (2021). "Scope and performance of artificial intelligence technology in orthodontic diagnosis, treatment planning, and clinical decision-making-a systematic review." Journal of dental sciences **16**(1): 482-492.

Khanehmasjedi, M., et al. (2020). "Characteristics and Dynamics of Full Arch Distalization Using Transpalatal Arches with Midpalatal and Interradicular Miniscrews as Temporary Anchorage Devices: A Preliminary Finite Element Analysis." International Journal of Dentistry **2020**(1): 6648526.

Kim, H., et al. (2021). "Nonsurgical maxillary expansion in a 60-year-old patient with gingival recession and crowding." Korean Journal of Orthodontics **51**(3): 217-227.

Kim, W., et al. (2020). "Unsupervised learning of image segmentation based on differentiable feature clustering." IEEE Transactions on Image Processing **29**: 8055-8068.

Kiyak, H. A. (2008). "Does orthodontic treatment affect patients’ quality of life?" Journal of dental education **72**(8): 886-894.

Knaak, C., et al. (2021). "A spatio-temporal ensemble deep learning architecture for real-time defect detection during laser welding on low power embedded computing boards." Sensors **21**(12): 4205.

Koo, Y.-J., et al. (2017). "Maxillomandibular arch width differences at estimated centers of resistance: comparison between normal occlusion and skeletal Class III malocclusion." The Korean journal of orthodontics **47**(3): 167-175.

Kwak, K. H., et al. (2016). "Quantitative evaluation of midpalatal suture maturation via fractal analysis." Korean Journal of Orthodontics **46**(5): 323.

Ladewig, V. D. M., et al. (2022). "Rapid Maxillary Expansion Based on the Correlation of the Maturation Between the Midpalatal and Zygomaticomaxillary Sutures through CBCT." Iranian Journal of Orthodontics **17**(2): 1-9.

Li, G., et al. (2021). "Practices and applications of convolutional neural network-based computer vision systems in animal farming: A review." Sensors **21**(4): 1492.

Li, S., et al. (2022). "Brain tumor segmentation based on region of interest-aided localization and segmentation U-Net." International Journal of Machine Learning and Cybernetics **13**(9): 2435-2445.

Li, Y.-J., et al. (2023). "A novel deep learning-based algorithm combining histopathological features with tissue areas to predict colorectal cancer survival from whole-slide images." Journal of Translational Medicine **21**(1): 731.

Liu, W., et al. (2016). Ssd: Single shot multibox detector. Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14, Springer.

Loddi, P. P., et al. (2008). "Transverse effects after surgically assisted rapid maxillary expansion in the midpalatal suture using computed tomography." Journal of Craniofacial Surgery **19**(2): 433-438.

Luo, J.-H., et al. (2017). Thinet: A filter level pruning method for deep neural network compression. Proceedings of the IEEE international conference on computer vision.

Mageet, A. O. (2016). "Classification of Skeletal and Dental Malocclusion: Revisited."

Mahdian, A., et al. (2019). "A systematic review of methods used to assess mid-palatal suture maturation." J. Dent. Sch **37**(1): 26-31.

Manjusha, K., et al. (2017). "Growth and development of dentofacial complex influenced by genetic and environmental factors using monozygotic twins."

McNamaraa, J. A. (2000). "Maxillary transverse deficiency." American Journal of Orthodontics and Dentofacial Orthopedics **117**(5): 567-570.

Mehta, C., et al. (2022). "Cardiovascular imaging databases: building machine learning algorithms for regenerative medicine." Current Stem Cell Reports **8**(4): 164-173.

Mehta, S. and K. S. Patnaik (2021). "Improved prediction of software defects using ensemble machine learning techniques." Neural Computing and Applications **33**(16): 10551-10562.

Minaee, S., et al. (2021). "Image segmentation using deep learning: A survey." IEEE transactions on pattern analysis and machine intelligence **44**(7): 3523-3542.

Mishra, S., et al. (2023). Prediction of Heart Disease using Machine Learning. 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), IEEE.

Moazen, M., et al. (2015). "Mechanical properties of calvarial bones in a mouse model for craniosynostosis." PloS one **10**(5): e0125757.

Mohammad, N., et al. (2022). "Applications of contemporary artificial intelligence technology in forensic odontology as primary forensic identifier: A scoping review." Frontiers in artificial intelligence **5**: 1049584.

Monsarrat, P., et al. (2022). "Systemic periodontal risk score using an innovative machine learning strategy: an observational study." Journal of Personalized Medicine **12**(2): 217.

Namdev, U., et al. (2022). "Object Detection Techniques based on Deep Learning: A Review." Computer Science & Engineering: An International Journal (CSEIJ) **12**: 125-134.

Naranjo-Torres, J., et al. (2020). "A review of convolutional neural network applied to fruit image processing." Applied Sciences **10**(10): 3443.

Nervina, J. M., et al. (2014). "Assessment of maxillary transverse deficiency and treatment outcomes by cone beam computed tomography." Cone beam computed tomography in orthodontics: indications, Insights, and Innovations: 383-409.

Nguyen, U. K. (2017). Retrospective Evaluation of Skeletal, Dentoalveolar, and Periodontal Changes of Microimplant Assisted Rapid Palatal Expansion (MARPE) In Skeletally Matured Patients, West Virginia University.

Nimkarn, Y., et al. (1995). "The validity of maxillary expansion indices." The Angle Orthodontist **65**(5): 321-326.

Nowak, R. and E. Zawiślak (2016). "An Analysis of Early Changes to the Craniofacial Skeleton Induced by Transpalatal Distraction in a Patient with Maxillary Constriction, Based on Computed Tomography Scans–Case Report." Dental and Medical Problems **53**(1): 134-141.

Nowak, R. M., et al. (2015). "Treatment options and limitations in transverse maxillary deficiency." Dent. Med. Probl **52**(4): 389-400.

O'Mahony, N., et al. (2020). Deep learning vs. traditional computer vision. Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Volume 1 1, Springer.

Pal, B. and S. Khaiyum (2019). "Low Memory Footprint CNN Models for end-to-end Driving of Autonomous Ground Vehicle and Custom Adaptation to Various Road Conditions." International Journal of Innovative Technology and Exploring Engineering **9**: 3252-3259.

Prajapati, S. A., et al. (2017). Classification of dental diseases using CNN and transfer learning. 2017 5th International Symposium on Computational and Business Intelligence (ISCBI), IEEE.

Rachmiel, A., et al. (2020). "Surgically assisted rapid palatal expansion to correct maxillary transverse deficiency." Annals of maxillofacial surgery **10**(1): 136-141.

Rebecq, H., et al. (2019). Events-to-video: Bringing modern computer vision to event cameras. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.

Redmon, J., et al. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE conference on computer vision and pattern recognition.

Rong, H. (2023). Exploration of Lightweight Neural Network Architectures for Sentiment Analysis. 2023 International Conference on Image Processing, Computer Vision and Machine Learning (ICICML), IEEE.

Roosenboom, J., et al. (2016). "Exploring the underlying genetics of craniofacial morphology through various sources of knowledge." BioMed research international **2016**(1): 3054578.

Roy, P., et al. (2018). "Numa-caffe: Numa-aware deep learning neural networks." ACM Transactions on Architecture and Code Optimization (TACO) **15**(2): 1-26.

Salunke, D., et al. (2022). Deep Learning Techniques for Dental Image Diagnostics: A Survey. 2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS), IEEE.

Sandler, M., et al. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. Proceedings of the IEEE conference on computer vision and pattern recognition.

Sawchuk, D., et al. (2016). "Diagnostic methods for assessing maxillary skeletal and dental transverse deficiencies: A systematic review." The Korean journal of orthodontics **46**(5): 331-342.

Sayar, G. and D. D. Kılınç (2019). "Rapid maxillary expansion outcomes according to midpalatal suture maturation levels." Progress in orthodontics **20**: 1-7.

Schmidhuber, J. (2015). "Deep learning." Scholarpedia **10**(11): 32832.

Schwarz, A. and I. Watzke (2009). "Orthognathic Maxillary Expansion (OME) Versus Surgical Assisted Rapid Palatal Expansion (SARPE): A Comparison." Informationen aus Orthodontie & Kieferorthopädie **41**(04): 253-258.

Schwendicke, F., et al. (2022). "Artificial intelligence for caries detection: value of data and information." Journal of dental research **101**(11): 1350-1356.

Shayani, A., et al. (2023). "Midpalatal Suture Maturation Stage in 10-to 25-Year-Olds Using Cone-Beam Computed Tomography—A Cross-Sectional Study." Diagnostics **13**(8): 1449.

Shih, S.-N., et al. (2022). "Management of Class III Malocclusion and Maxillary Transverse Deficiency with Microimplant-Assisted Rapid Palatal Expansion (MARPE): A Case Report." Medicina **58**(8): 1052.

Silva-Montero, J. C., et al. (2022). "Analysis of the Frequency and Correlated Factors of Midpalatal Suture Maturation Stages in Young Adults, Based on Cone Beam Computed Tomography Imaging." Journal of Clinical Medicine **11**(23): 6959.

Simonyan, K. (2014). "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556.

Singha, A., et al. (2021). "Deep learning applications in medical image analysis." Biomedical Data Mining for Information Retrieval: Methodologies, Techniques and Applications: 293-350.

Suri, L. and P. Taneja (2008). "Surgically assisted rapid palatal expansion: a literature review." American Journal of Orthodontics and Dentofacial Orthopedics **133**(2): 290-302.

Szeliski, R. (2022). Computer vision: algorithms and applications, Springer Nature.

Talwandi, N. S., et al. (2024). Object Detection Model for Tacking System Implemtation using Deep Learning Technique. 2024 IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT), IEEE.

Tasnim, S. and W. Qi (2023). "Progress in Object Detection: An In-Depth Analysis of Methods and Use Cases." European Journal of Electrical Engineering and Computer Science **7**(4): 39-45.

Top, A., et al. (2023). "Quantitative level determination of fixed restorations on panoramic radiographs using deep learning." International Journal of Computerized Dentistry **26**(4).

Tuzoff, D. V., et al. (2019). "Tooth detection and numbering in panoramic radiographs using convolutional neural networks." Dentomaxillofacial Radiology **48**(4): 20180051.

Venkatachalam, J. and S. Chandrabose (2023). "Optimizing Region Detection in Enhanced Infrared Images Using Deep Learning." Revue d'Intelligence Artificielle **37**(4).

Vijay, A., et al. (2021). "A Review of Brain Tumor Image Segmentation of MR Images Using Deep Learning Methods." AIJR Proceedings: 137-145.

Voulodimos, A., et al. (2018). "Deep learning for computer vision: A brief review." Computational intelligence and neuroscience **2018**(1): 7068349.

Wan, D., et al. (2018). Tbn: Convolutional neural network with ternary inputs and binary weights. Proceedings of the European Conference on Computer Vision (ECCV).

Wang, W. and J. Shen (2017). "Deep visual attention prediction." IEEE Transactions on Image Processing **27**(5): 2368-2378.

Wehrbein, H. and F. Yildizhan (2001). "The mid-palatal suture in young adults. A radiological-histological investigation." The European Journal of Orthodontics **23**(2): 105-114.

Yadav, S. S. and S. M. Jadhav (2019). "Deep convolutional neural network based medical image classification for disease diagnosis." Journal of Big data **6**(1): 1-18.

Yoo, C. H., et al. (2019). "Relevance regularization of convolutional neural network for interpretable classification." Network **50**: 5.

Zahra, S. E. and H. M. Samih (2017). "ABSENCE OR PRESENCE OF MID-PALATAL SUTURE IN PATIENTS WITH COMPLETE UNILATERAL CLEFT LIP AND PALATE,(A RETROSPECTIVE STUDY)." Egyptian Dental Journal **63**(2-April (Orthodontics, Pediatric & Preventive Dentistry)): 1155-1164.

Zhang, X., et al. (2018). Shufflenet: An extremely efficient convolutional neural network for mobile devices. Proceedings of the IEEE conference on computer vision and pattern recognition.

Zhu, M., et al. (2024). "Convolutional neural network-assisted diagnosis of midpalatal suture maturation stage in cone-beam computed tomography." Journal of Dentistry **141**: 104808.

Zong, C., et al. (2019). Skeletal and dentoalveolar changes in the transverse dimension using microimplant-assisted rapid palatal expansion (MARPE) appliances. Seminars in Orthodontics, Elsevier.

Brea, L. S., et al. (2020). Deep learning-based retinal vessel segmentation with cross-modal evaluation. Medical Imaging with Deep Learning, PMLR.

Chun, J.-H., et al. (2022). "Skeletal and alveolar changes in conventional rapid palatal expansion (RPE) and miniscrew-assisted RPE (MARPE): a prospective randomized clinical trial using low-dose CBCT." BMC oral health **22**(1): 114.

Jia, H., et al. (2021). "Comparison of skeletal maxillary transverse deficiency treated by microimplant-assisted rapid palatal expansion and tooth-borne expansion during the post-pubertal growth spurt stage: A prospective cone beam computed tomography study." The Angle Orthodontist **91**(1): 36-45.