

Gradient Based Novel Texture Feature Extraction Methods For Texture Classification

by

G M Mashrur E Elahi

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

Master of Science

Department of Computing Science

University of Alberta

© G M Mashrur E Elahi, 2016

Abstract

Texture analysis is a well-known research topic in computer vision and image processing and has many applications. For example, texture features in image classification have been shown to be useful. Texture features depend on representation, of which there are many methods. Among them, gradient-based methods have become popular in classification problems. One of the gradient based methods is Co-occurrence Histograms of Oriented Gradients (CoHOG) has been applied in many areas. CoHOG algorithm provides a unified description of both statistical and differential properties for a texture. But it discards some important texture information due to the use of sub-regions. In this thesis, based on the original CoHOG method, three novel feature extraction methods are proposed. All the methods use the whole image instead of sub-regions for feature calculation. Also we use a larger neighborhood size for the methods. In the first method, we use Sobel operators for gradient calculation named S-CoHOG. The second method uses Gaussian Derivative (GD) operators named GD-CoHOG and the third method named LFDG-CoHOG uses the Local Frequency Descriptor Gradient (LFDG) operators for gradient calculations. The extracted feature vector size is very large and classification using a large number of similar features does not provide the best results. In our proposed methods, only a minimum number of significant features are selected

using area under the receiver operator characteristic (ROC) curve (AUC) thresholds. The selected features are used in a linear support vector machine (SVM) classifier to determine the classification accuracy. The classification results of the proposed methods are compared with that of the original CoHOG method using three well-known texture datasets. The classification results show that the proposed methods achieve the best classification results in all the datasets. The proposed methods are also evaluated for medical image classification. Three different cohort datasets of 2D Magnetic Resonance Images (MRI) are used along with a multicenter dataset to compare the classification results of the proposed methods with that of the gray level co-occurrence matrix (GLCM) method. The experimental results show that the proposed methods outperform that of the GLCM method.

Acknowledgements

First and foremost, I offer my sincerest gratitude to my supervisors, Dr. Herbert Yang, Professor, Department of Computing Science, and Dr. Sanjay Kalra, Professor, Department of Medicine, Biomedical Engineering and Computing Science, who have supported me throughout my thesis with their patience, motivation, enthusiasm, and immense knowledge. The high standards that they hold in research are an invaluable inspiration that certainly shaped my academic pursuits and will continually encourage me all my life. I attribute the level of my Masters degree to their encouragement and effort and without them this thesis, would not have been completed or written. One simply could not wish for better or friendlier supervisors.

In my daily work I have been blessed with a friendly and cheerful group of fellow students. I am gratefully indebted to them for their very valuable comments and suggestions on this thesis.

I would like to thank the ALS Society of Canada, ALS Association of America, the MSI Foundation of Alberta, the University of Alberta Hospital Foundation, the Shelly Mrkonjic Foundation, Brain Canada, and NSERC for their generous financial support.

Finally, I must express my very profound gratitude to my parents and to my elder brother Md. Manjur E Elahee for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. At last but not the least, I must acknowledge my wife and friend, Anwara Khatun, without her love, encouragement and assistance I would not have finished this thesis. This accomplishment would not have been possible without them. Thank you.

Contents

Abstract	ii
Acknowledgements	iv
List of Tables	viii
List of Figures	xi
1 Introduction	1
1.1 Texture Analysis	1
1.2 The Thesis Work	3
1.3 Summery of Contributions	4
1.4 Thesis Outline	6
2 Related Works	7
2.1 2D Texture Analysis Methods	7
2.1.1 Local Binary Patterns (LBP)	8
2.1.2 Gray Level Co-occurrence Matrix (GLCM)	9
2.1.3 The Run Length Matrix (RLM)	11
2.1.4 Gradient Orientation based Texture Methods	11
3 Proposed Methodology	15
3.1 Overview	15
3.2 Pre-processing	16

3.2.1	Gradient Orientation and Quantization	17
3.3	Feature Extraction	19
3.3.1	Co-occurrence Matrix (CM) Calculation	19
3.3.2	Feature Vector Generation	22
3.4	Feature Selection	23
3.5	Classification	24
4	Experimental Results	26
4.1	Texture Datasets	26
4.1.1	INRIA Person Dataset	27
4.1.2	CUReT Dataset	27
4.1.3	UIUC Dataset	28
4.2	Human MRI Datasets	29
4.2.1	Amyotrophic Lateral Sclerosis (ALS)	30
4.2.2	ROI Selection	31
4.2.3	Downsampling	31
4.2.4	MR Dataset 1	32
4.2.5	MR Dataset 2	32
4.2.6	MR Dataset 3	35
4.3	Classification Results of Texture Datasets	36
4.3.1	Classification of INRIA Person Dataset	36
4.3.2	Classification of CUReT Dataset	38
4.3.3	Classification of the UIUC Dataset	39
4.3.4	Comparison with other CoHOG methods	40
4.4	Classification Results of MRI Datasets	41
4.4.1	ROC Analysis of MR Dataset 1	41
4.4.2	ROC Analysis of MR Dataset 2	42
4.4.3	ROC Analysis of MR Dataset 3	43
4.4.4	ROC Analysis using different Gradient Operators	46

4.4.5	Region Based Analysis	47
4.4.6	Comparison with the GLCM Method	48
4.4.7	ROC Analysis using Randomly Selected Slices	49
4.5	ROC Analysis of Multicenter Dataset	50
4.5.1	Classification using Different Centers for Training and Testing	53
5	Conclusion	55
5.1	Summary	55
5.2	Contributions	57
5.3	Future Work	58
	Bibliography	60

List of Tables

2.1	Texture features defined for the GLCM	10
4.1	Details of the subjects for MR Dataset 1	32
4.2	Details of the downsampled 2D MR images of the subjects for MR Dataset 1 and MR Dataset 3	33
4.3	Details of the subjects for MR Dataset 2 and MR Dataset 3	33
4.4	Details of the downsampled 2D MR images of the subjects for MR Dataset 2	34
4.5	Classification accuracy of the proposed methods using INRIA Person dataset. I means classification without feature selection and II mean classification with feature selection.	37
4.6	Number of features selected using AUC threshold for INRIA Person dataset using two neighborhood sizes.	37
4.7	Classification accuracy of the proposed methods using CURET dataset. I means classification without feature selection and II mean classifi- cation with feature selection.	38
4.8	Number of features selected using AUC threshold for CURET dataset using two neighborhood sizes.	39
4.9	Classification accuracy of the proposed methods using the UIUC dataset. I means classification without feature selection and II mean classification with feature selection.	39

4.10	Number of features selected using AUC threshold for UIUC dataset using two neighborhood sizes.	40
4.11	Comparison of the classification accuracies (CA) of the proposed methods with the original CoHOG and the Eig(Hess)-CoHOG methods.	41
4.12	ROC analysis of MR Dataset 1 using S-CoHOG features extracted for neighborhood radius of 4.	42
4.13	ROC analysis of MR Dataset 1 using S-CoHOG features extracted for neighborhood radius of 8.	42
4.14	ROC analysis of MR Dataset 2 using S-CoHOG features extracted for neighborhood radius of 4.	43
4.15	ROC analysis of MR Dataset 2 using S-CoHOG features extracted for neighborhood radius of 8.	43
4.16	ROC analysis of MR Dataset 3 using S-CoHOG features extracted for neighborhood radius of 4.	44
4.17	ROC analysis of MR Dataset 3 using S-CoHOG features extracted for neighborhood radius of 8.	44
4.18	ROC analysis of MR Dataset 1 using three proposed methods. CoHOG features extracted using a neighborhood radius of 8.	46
4.19	ROC analysis of MR Dataset 2 using three proposed methods. CoHOG features extracted using a neighborhood radius of 8.	47
4.20	Comparison of ROC analysis between GD-CoHOG and GLCM methods using MR dataset 1.	48
4.21	Comparison of ROC analysis between GD-CoHOG and GLCM methods using MR dataset 2.	49
4.22	MRI acquisition parameters for five different centers.	51
4.23	Multicenter (MC) dataset details.	52
4.24	ROC analysis of MC datasets using proposed methods with a neighboring radius of 8 and GLCM.	53

4.25 Classification accuracy using data from one center for training and
the other center for testing. 54

List of Figures

1.1	Texture patterns in various images.	1
2.1	LBP process ($R = 1, P = 8$). (a) A gray level image, (b) neighbors' values after thresholding, (c) Binary encoding and the corresponding decimal value.	8
2.2	GLCM computation process ($d = 1, \theta = 90^\circ$). (a) A gray level image, (b) corresponding GLCM.	9
2.3	Generation of (b) RLM matrix from a (a) gray level image using 0^0 direction of run.	11
2.4	Basic HOG calculation process. (a) Gradient orientation of the image pixels, (b) histogram of gradient orientation of each sub-region, (c) feature vector.	12
2.5	Basic CoHOG calculation process. (a) Gradient orientation of the image pixels, (b) Combination of of sub-regions and offsets for CM calculation, (c) CM for each sub-region, and (d) feature vector. . . .	13
3.1	Overview of the proposed approach.	16
3.2	$0^\circ - 360^\circ$ degree orientations are quantized into 8 bins.	19
3.3	(a) Offsets for different radii and (b) a specific offset at distance (x, y) from pixel (p, q)	20
3.4	Illustration of the Co-occurrence Histograms of Oriented Gradients (CoHOG) calculation for the proposed methods.	21

4.1	Three sample images of (a) human and (b) nonhuman classes from the INRIA Person dataset.	27
4.2	Three sample images of (a) class 1 (b) class 3 and (c) class 5 from the CURET dataset.	28
4.3	Three sample images of (a) class 1 (b) class 3 and (c) class 5 from the UIUC dataset.	29
4.4	(a) Sagittal, (b) Axial and (c) Coronal image slices. Coronal imaging is used in texture feature extraction.	31
4.5	ROI selection from a coronal image slice. The highlighted regions are selected as ROI.	32
4.6	Three sample image slices of (a) controls and (b) patients from MR Dataset 1. Patients and controls are not distinguishable by visual inspection.	33
4.7	Three sample image slices of (a) controls and (b) patients from MR Dataset 2. Patients and controls are not distinguishable by visual inspection.	34
4.8	Three sample image slices of (a) controls and (b) patients from MR Dataset 3. Patients and controls are not distinguishable by visual inspection.	35
4.9	Classification accuracy for selected features using different AUC thresholds for MR dataset 1.	45
4.10	Mean feature values with standard deviations of the mean feature values between patients and controls of ten selected features for (a) MR Dataset 1 and (b) MR Dataset 2.	45
4.11	Region based analysis of the subjects of MR dataset 1. Significant regions are marked by the colored boxes and classification accuracy of the corresponding boxes.	48

4.12 Classification accuracies for 10 random slice selection experiments and the mean and the standard deviation of the classification accuracies. 50

4.13 Sample image slices of (a) controls and (b) patients of each center from Multicenter dataset. Patients and controls are not distinguishable by visual inspection. 52

Chapter 1

Introduction

1.1 Texture Analysis

Texture analysis is a well-known and promising method in image processing and computer vision. Visual patterns appearing in images are called image textures and can be seen everywhere such as: carpet, wall, ultrasound images, fingerprint images, and medical images (see Fig.1.1).

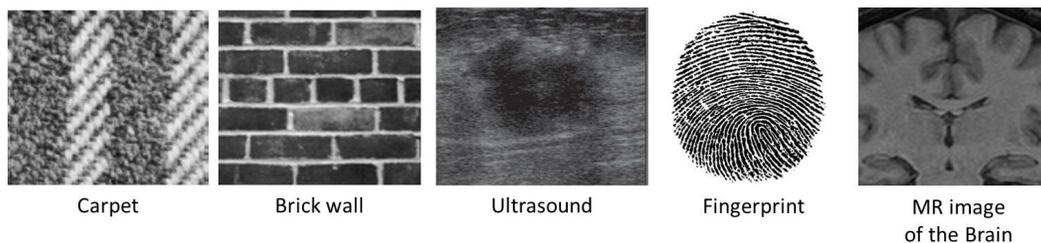


Figure 1.1: Texture patterns in various images.

In the real world, textures on the surface of objects can be classified as either micro structures or macro structures. The arrangement of bricks on a wall is an example of macro structures while the graininess pattern on a brick is an example of micro structures.

Texture analysis characterizes and quantifies pattern variations in images, including those that are imperceptible to the human eye. Textures are used as visual

cues to differentiate among different image regions or different images.

Texture analysis has been a major research topic for the last four decades. It has been used in different applications which include document processing [1], remote sensing [2], automated inspection [3], fingerprint identification [4], and medical image analysis [5], [6].

In texture analysis, texture features that are invariant to geometric transformation, noise, blurriness, and illumination changes are desired. Some well-known 2D texture methods are the Local Binary Patterns (LBP) [7] and its variants, gray level co-occurrence matrix (GLCM) [8], the gray level Aura matrix (GLAM) [9], the Run Length Matrix (RLM) [10], filter responses in the frequency [11] and spatial [12], [13] domains, Wavelets [14], Orientation Pyramid [15], Markov Random Fields (MRFs) [16], Gaussian MRFs [17], and spin images [18]. However, none of the above methods are insensitive to illumination changes.

Recently, gradient orientation based texture methods have become popular in computer vision and image processing. Among them Scale Invariant Feature Transform (SIFT) [19], Histograms of Oriented Gradients (HOG) [20] and Co-occurrence Histograms of Oriented gradients (CoHOG) [21] are commonly used in object detection. The CoHOG method and its variants [22], [23] have been successfully used in pedestrian detection [21], face recognition [24], fine-grained activity recognition [22], etc.

In recent years, texture analysis methods have also found applications in medical imaging that include texture analysis of MRI images. In particular, diagnosis of dementia using GLCM and Gabor filter responses [25], the study of pathological changes of the hippocampus in patients with Alzheimers disease and mild cognitive impairment using GLCM and RLM [26], brain tumor detection [27] and the study of epilepsy using Wavelet features [28] are some important contributions. Unfortunately, there is still no well-known feature descriptor for Amyotrophic Lateral Sclerosis (ALS) detection. A 2D region of interest based approach is used to find texture changes in ALS [29]. GLCM has been used for 3D texture analysis in ALS

[30]. Unfortunately, the sensitivity and specificity of these methods are sub-optimal for clinical utility. Therefore, a method with improved sensitivity and specificity as well as classification accuracy is much desired.

1.2 The Thesis Work

Gradient based texture feature extraction methods have become popular in image classification. These methods provide both statistical and differential properties of a texture. Intensity based methods that use intensity levels of the gray scale image will have unpredictable performance on images having varying intensities. To address this problem, instead of using intensity, gradient orientations are used for texture feature extraction in gradient based methods. Two of the well-known and popular methods are HOG and CoHOG. In both methods, a feature vector is formed based on the histogram of gradient orientations. The HOG method only counts the orientations for a local region. Inter-relationship between orientations is not used. The CoHOG method overcomes this limitation using the co-occurrence information of orientations to build the histogram. Also by using different length offsets for co-occurrence calculation, it can better represent local and global information.

In this thesis, based on the original CoHOG method, we propose three novel texture feature extraction methods. Three well-known methods such as Sobel [31], Gaussian Derivative (GD) [23] and Local Frequency Descriptor Gradient (LFDG) [32], [33] operators are used to calculate the gradient orientations of the image pixels. In the first method, called S-CoHOG, we use Sobel operators for gradient calculation. The second method, called GD-CoHOG, uses GD operators and the third method, called LFDG-CoHOG, uses the LFDG operators for gradient calculations. The original CoHOG method uses Sobel operators for gradient calculation and sub-regions of an image and calculates the sum of the co-occurrences of orientation pairs within each sub-region. The use of sub-region limits the accuracy of co-occurrence calculation for the boundary pixels and thus some information is

incomplete for each sub-region. For the first time in this thesis, we are applying CoHOG to the whole image for texture feature extraction without subdividing the image into sub-regions.

For each of the three proposed methods, the gradient orientations of each pixel of an image are calculated using the respective gradient operators. The gradient orientations of the pixels are then quantized into N bins. The co-occurrences of the gradients are summed for each offset and stored into an $N \times N$ co-occurrence matrix. An offset is defined by a distance and a direction. Offsets are limited to a radius specified by the distance from the pixel. All the co-occurrence matrices calculated from all the offsets are combined to create the feature vector (FV).

The size of the FV can be very large depending on the number of offsets and not all features are significant. Using this large number of similar features creates ambiguity in creation of an optimal hyperplane and leads to an incorrect classification by a classifier. So, we apply receiver operator characteristic (ROC) curve analysis to select significant features. The selected features vector (SFV) size is less than the original FV size. Most importantly, using the selected features, a significant difference between two classes is obtained. Particularly, for medical dataset, it is difficult to find differences between the patients and the controls without feature selection. We employ a linear support vector machine (SVM) [34] classifier to calculate the classification accuracy between two classes.

1.3 Summery of Contributions

The main contributions of this thesis are as follows:

- The proposed three methods use the whole image instead of subdividing it into sub-regions. The use of sub-regions limits the accuracy of co-occurrence matrix for boundary pixels and thus some information is incomplete for each sub-region. Also it increases the size of the feature vector. Thus, using the whole image not only reduces the boundary pixels problem in sub-regions but

also reduces the size of feature vector.

- The original CoHOG method uses Sobel operators for gradient calculation. For the first time, we adopt two gradient operators GD and LFDG for the proposed GD-CoHOG and LFDG-CoHOG methods, respectively. The proposed methods are compared to determine the impact of the gradient operators on classification accuracy using the whole image.
- Texture features are extracted using two different neighborhood sizes. The original CoHOG method uses a maximum neighborhood size of 4. We use a larger neighborhood size of 8 to evaluate the effect of using more global information for co-occurrence calculation on classification accuracy.
- The extracted feature vector size using the CoHOG method is very large with many similar features. Using this large number of similar features creates ambiguity in optimal hyperplane creation and leads to the wrong classification by an SVM classifier. We are the first to use a feature selection method to reduce the number of features used in CoHOG. In particular, we select significant features using area under the ROC curve (AUC) analysis for classification. Only features that contain significant differences between classes are selected using an AUC threshold. The experimental results show that the performance with feature selection outperforms that without feature selection.
- Three different datasets of 2D Magnetic Resonance Images (MRI) of Amyotrophic Lateral Sclerosis (ALS) are analyzed for the first time using the proposed methods. Every dataset uses different image resolutions and contrasts. Another multicenter ALS dataset of different image contrasts for each center is also used in this experiment to demonstrate the classification performance of the proposed methods. The experimental results show that the proposed methods can achieve very high specificity and sensitivity as well as classification accuracy in all datasets.

1.4 Thesis Outline

The rest of the thesis is organized as follows. Some important related works are discussed in Chapter 2. In Chapter 3, we explain the proposed approach of feature extraction and selection. The experimental results and discussions are presented in Chapter 4. Chapter 5 concludes the thesis.

Chapter 2

Related Works

Texture analysis is a promising topic in computer vision and image processing. Two dimensional texture analysis methods have been used for document processing [1], remote sensing [2], automated inspection [3], fingerprint identification [4], medical image analysis [6], etc. Some of the representative 2D methods are discussed below.

2.1 2D Texture Analysis Methods

The Gabor filter banks are used in filter responses in the frequency [11] and spatial [12], [13] domains for image texture recognition. Multiresolution Wavelets [14] are used in texture feature extraction and selection to segment textures. A Markov Random Fields (MRFs) [16] model is used to analyze textures in images. Some related methods include the Gaussian MRFs [17] and spin images [18]. Another multiresolution approach to gray-scale and rotation invariant texture classification based on Local Binary Patterns (LBP) is presented in [7] and in its variants. Gray level co-occurrence matrix (GLCM) [8] is a gray level intensity based texture analysis method that uses the co-occurrences of gray levels at different pixel locations. The Run Length Matrix (RLM) [10] is used to calculate features of different terrain types for classification. Some of the important related methods are discussed below.

2.1.1 Local Binary Patterns (LBP)

LBP uses the gray level differences between the center pixel and its neighbors and assigns either 0 or 1 to each of its neighbors depending on the difference as shown in Eq. 2.1 [7],

$$S(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}, \quad (2.1)$$

where $x = (g_i - g_c)$. Here g_c and g_i are the gray levels of the center pixel and its neighbor pixel (i), respectively. These values are used to form a binary local pattern. Then this binary pattern is converted into the corresponding decimal value using Eq 2.2 [7],

$$LBP_{P,R} = \sum_{p=0}^P S(g_p - g_c)2^p, \quad (2.2)$$

where P is the number of neighbors and R the radius of the neighboring pixels. An example of calculating the LBP decimal code for a neighboring radius $R = 1$ and number of neighbors $P = 8$ is shown in the Fig 2.1.

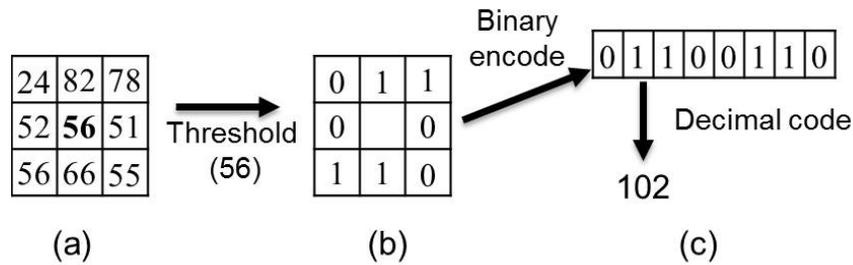


Figure 2.1: LBP process ($R = 1, P = 8$). (a) A gray level image, (b) neighbors' values after thresholding, (c) Binary encoding and the corresponding decimal value.

The center gray value of the window is compared with all the neighboring gray values (see Fig. 2.1(a)) and 1 or 0 is assigned to the corresponding neighbor based on Eq. 2.1 (see Fig. 2.1(b)). Finally these bits are encoded into a binary code and converted into the corresponding decimal code (see Fig. 2.1(c)).

After computing the LBP codes for all the pixels in the image a histogram is built from these decimal codes to represent the texture image. To achieve rotation invariant, the minimum right shifted binary pattern is used.

2.1.2 Gray Level Co-occurrence Matrix (GLCM)

In GLCM [8], image intensities are quantized into a fixed number of gray levels and a co-occurrence matrix is formed by summing the co-occurrences of a specific pair of gray levels. The process of GLCM can be divided into three steps. First, each pixel values of a given gray image is quantized into G number of gray levels. Then, using this gray level information a GLCM is formed. A GLCM is defined for a given direction (θ) and distance (d). A vector with distance d and direction angle θ connects image pixel $I(x1, y1)$ to $I(x2, y2)$ such that $x2 = x1 + d \cos(\theta)$ and $y2 = y1 + d \sin(\theta)$. $GLCM_{d,\theta}$ for distance d and direction angle θ is a $G \times G$ matrix where each entry $GLCM_{d,\theta}(i, j)$ shows the number of times that $I(x1, y1) = i$ and $I(x2, y2) = j$, where i and j are the gray levels at the corresponding locations. Simply, GLCM counts the number of times a particular gray level pair co-occurs. An example of the process of computing the GLCM is shown in Fig. 2.2.

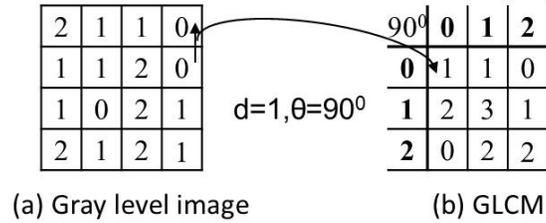


Figure 2.2: GLCM computation process ($d = 1, \theta = 90^\circ$). (a) A gray level image, (b) corresponding GLCM.

Usually, GLCM uses one of the eight directions ($0^\circ, \pm 45^\circ, \pm 90^\circ, \pm 135^\circ, 180^\circ$). Symmetric GLCM uses four directions instead of eight as diagonally opposite directions are symmetric. Finally, the GLCM is normalized to compute texture features. The normalization can be done using Eq. 2.3,

$$GLCM_{d,\theta}^{norm}(i, j) = \frac{GLCM_{d,\theta}(i, j)}{\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} GLCM_{d,\theta}(i, j)}. \quad (2.3)$$

Twelve well known features of GLCM are defined and used. The texture features are listed in Table 2.1. Here P is the normalized GLCM, G the number of gray levels. μ_x , μ_y , σ_x , and σ_y indicate means and standard deviations of the row and column sums of P . $P_{x+y}(k) = \sum_{i=1}^G \sum_{j=1}^G P(i, j)$ and $P_{x-y}(k) = \sum_{i=1}^G \sum_{j=1}^G P(i, j)$.

Table 2.1: Texture features defined for the GLCM

Texture Feature	Formula
Angular second moment	$f_1 = \sum_{i=1}^G \sum_{j=1}^G (P(i, j))^2$
Contrast	$f_2 = \sum_{i=1}^G \sum_{j=1}^G i - j ^2 P(i, j)$
Correlation	$f_3 = \frac{1}{\sigma_x \sigma_y} \sum_{i=1}^G \sum_{j=1}^G ij P(i, j) - \mu_x \mu_y$
Sum of Squares: Variance	$f_4 = \sum_{i=1}^G \sum_{j=1}^G (i - \mu)^2 P(i, j)$
Inverse difference moment normalized	$f_5 = \sum_{i=1}^G \sum_{j=1}^G \frac{1}{1+(i-j)^2/G^2} P(i, j)$
Sum average	$f_6 = \sum_{i=2}^{2G} iP_{x+y}(i)$
Sum variance	$f_7 = \sum_{i=2}^{2G} (1 - f_8)^2 P_{x+y}(i)$
Sum entropy	$f_8 = - \sum_{i=2}^{2G} P_{x+y}(i) \log(P_{x+y}(i))$
Entropy	$f_9 = - \sum_{i=1}^G \sum_{j=1}^G P(i, j) \log(P(i, j))$
Difference variance	$f_{10} = \text{variance of } P_{x-y}$
Difference entropy	$f_{11} = - \sum_{i=0}^{G-1} P_{x-y}(i) \log(P_{x-y}(i))$
Homogeneity	$f_{12} = \frac{1}{1+ i-j } \sum_{i=1}^G \sum_{j=1}^G P(i, j)$

Recently, GLCM has been used in medical imaging for the diagnosis of dementia [25], the study of pathological changes of hippocampus in patients with Alzheimer disease and mild cognitive impairment [26], and brain tumors detection [27]. A 3D variants of GLCM has been used for 3D texture analysis in Amyotrophic Lateral Sclerosis (ALS) [30]. The major limitation of GLCM is that it works with the intensity level of gray scale images, which will have unpredictable performance when the acquisition equipment or the scanning protocol changes.

2.1.3 The Run Length Matrix (RLM)

The RLM uses the gray level runs. A set of consecutive, co-linear pixels in an image having the same gray level value is called a gray level run [10]. The length of the run is defined by the number of pixels in the run. For a given run direction, the run length matrix of an image can be calculated. Fig. 2.3 shows an example of creating a RLM for 0^0 degree run direction.

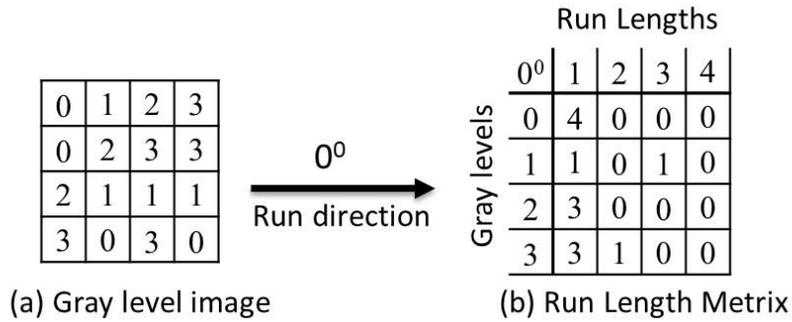


Figure 2.3: Generation of (b) RLM matrix from a (a) gray level image using 0^0 direction of run.

An RLM matrix element (i, j) is the number of times the gray level i appears in the image with run length j in specific run direction. The number of run lengths depends on the given gray level image size and the size of the RLM is equal to the number of run length \times number of gray levels of the image.

The numerical texture features are computed using some well-known functions that are used in the Gray Level Co-occurrence Matrix (GLCM) [8] method for feature calculation.

2.1.4 Gradient Orientation based Texture Methods

Gradient orientation based texture methods have become popular in recent years for their robustness against image intensity changes, blurriness and deformations. Moreover, gradient orientation based methods have better classification accuracy than LBP-like methods [35]. This is because LBP-like methods merely count the

number of patterns around pixels and lack gradient orientation related information [36].

Histograms of Oriented Gradients (HOG) [20] and Co-occurrence Histograms of Oriented gradients (CoHOG) are two such commonly used methods that have been used for objects detection[20], pedestrian detection [21], face recognition [24], fine-grained activity recognition [22], etc. We give a brief description of HOG and CoHOG below.

Histograms of Oriented Gradients (HOG)

The HOG method uses a gradient oriented image as input. The gradient orientations are quantized into N bins. Then the image is subdivided into M number of equal sub-regions. For each sub-region, a histogram of orientations is computed. The histogram is formed by simply counting different groups of orientations. The size of the histogram is N . For M sub-regions, there are M different histograms each of size N . Finally, these histograms are concatenated to form the feature vector histogram of size $M \times N$. An overview of HOG calculation process is shown in Fig. 2.4 [20].

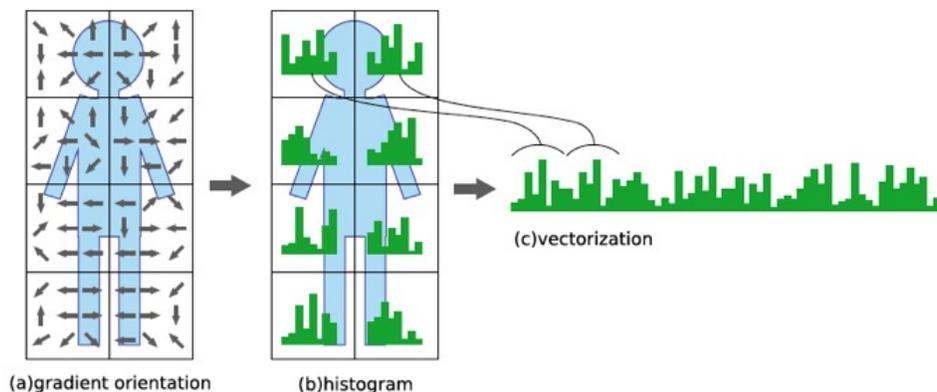


Figure 2.4: Basic HOG calculation process. (a) Gradient orientation of the image pixels, (b) histogram of gradient orientation of each sub-region, (c) feature vector.

The limitation of HOG is that it only counts the orientations for a local region. Inter-relationship information between orientations is not used. To overcome

this limitation, an improvement of HOG called the Co-occurrence HOG (CoHOG) method is proposed.

Co-occurrence Histograms of Oriented Gradients (CoHOG)

CoHOG is an extension of HOG. It also uses the quantized gradients as input and subdivides the image into a number of sub-regions. The CoHOG method uses a circular neighborhood with a given radius in which each pixel with the center pixel forms a pair called an offset. Now for each sub-region and for each offset, the co-occurrences of an orientation pair is computed by scanning all the pixels in the sub-region to form a co-occurrence matrix (CM). The size of the CM is $N \times N$, where N is the number of distinct orientations. The total number of CMs for a sub-region depends on the number of offsets. Finally, these CMs are concatenated to form the histogram for the sub-region. Then, the histograms of all the sub-regions are concatenated to form the CoHOG feature vector for the given image. The overview of CoHOG calculation process is shown in Fig. 2.5 [20].

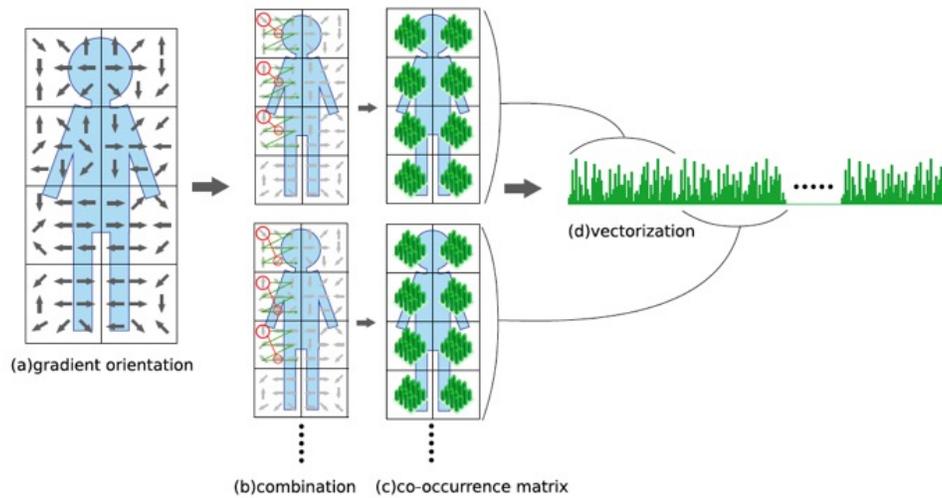


Figure 2.5: Basic CoHOG calculation process. (a) Gradient orientation of the image pixels, (b) Combination of of sub-regions and offsets for CM calculation, (c) CM for each sub-region, and (d) feature vector.

The feature vector size depends on the number of orientations, the number of offsets and the number of sub-regions. For example, if an image has M sub-

regions and K offsets with a CM size of $N \times N$, then the final feature vector size is $M \times K \times (N \times N)$.

The CoHOG method has the advantages over HOG in preserving the inter-relationship among the neighboring pixel orientations and, by using different offsets, the co-occurrence matrices can better represent the local and global orientation information. However, the use of sub-regions limits the accuracy of CM for boundary pixels and thus some information is incomplete for each sub-region.

In this thesis, we present a modified CoHOG method for texture feature extraction of the whole image which can overcome the sub-region issue mentioned above and can reduce the feature vector size.

Chapter 3

Proposed Methodology

3.1 Overview

In this chapter we discuss the proposed approaches of extracting texture features using the CoHOG method.

The original CoHOG method subdivides the original image and for each sub-region it calculates the co-occurrence matrices for all the offsets. Finally, all the co-occurrence matrices of each sub-region are combined to form the feature vector histogram. The histogram is very large depending on the number of sub-regions and the number of offsets. Image classes that contain very small changes in between the groups are almost similar in all other regions. Features extracted from these regions are also similar. Using this large number of similar features creates ambiguity in defining the optimal hyperplane and leads to an incorrect classification by a classifier.

In this thesis, based on the original CoHOG method, we propose three novel texture feature extraction methods. Since one of the key components in CoHOG is gradient calculation, three well-known methods such as Sobel [31], Gaussian Derivative (GD) [23] and Local Frequency Descriptor Gradient (LFDG) [32], [33] operators are used to calculate the gradient orientations of the image pixels. In the first method, we use Sobel operators for gradient calculation named S-CoHOG.

The second method uses GD operators named GD-CoHOG and the third method named LFDG-CoHOG uses the LFDG operators for gradient calculations.

The original CoHOG method uses sub-regions of an image and calculates the sum of the co-occurrences of orientation pairs. The use of sub-region limits the accuracy of co-occurrence calculation for the boundary pixels and thus some information is incomplete for each sub-region. While it is a simple idea, to the best knowledge of the author, it is the first time of applying CoHOG to the whole image using the three proposed methods for texture feature extraction.

The CoHOG features are extracted using two different neighborhood sizes. The original CoHOG method uses a maximum neighborhood size of 4. We use a larger neighborhood size of 8 to see the effect of using more distance information for co-occurrence calculation on classification accuracy.

Finally, we select significant features using area under the ROC curve (AUC) analysis for classification. Only features that contain significant differences in between the classes are selected using an AUC threshold.

The overview of the proposed approach is shown in Fig. 3.1. The proposed

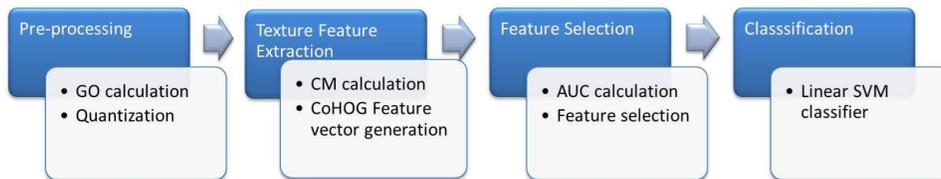


Figure 3.1: Overview of the proposed approach.

approach for all the three proposed methods consists of four steps: pre-processing, texture feature extraction, feature selection and classification. These steps are discussed below.

3.2 Pre-processing

Texture features are extracted from pre-processed images. Pre-processing involves gradient orientation (GO) calculation and quantization. The proposed S-CoHOG,

GD-CoHOG and LFDG-CoHOG methods use Sobel [31], Gaussian Derivative (GD) [23] and Local Frequency Descriptor Gradient (LFDG) [32] operators for gradient orientation calculation, respectively. The GO calculation and quantization steps are discussed below.

3.2.1 Gradient Orientation and Quantization

The gradient orientations of image pixels are computed by convolving the gradient operators with the image. Horizontal and vertical gradient operators are used to calculate the corresponding gradient images and then gradient orientations are calculated from the gradient images. The details of each of the three gradient operators are discussed below.

The Sobel Operators

Sobel uses two 3×3 kernels to estimate the horizontal and vertical derivatives. The two operators used in this method are shown in Eq. 3.1,

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}, \quad (3.1)$$

where G_x and G_y are the corresponding horizontal and vertical gradient operators.

The GD Operators

The GD operators that use two basic one-dimensional derivative filters are given in Eq. 3.2 [23], [37],

$$f_1(t) = \frac{-2t}{\sigma^2} e^{-\frac{t^2}{\sigma^2}}, f_2(t) = e^{-\frac{t^2}{\sigma^2}}, \quad (3.2)$$

where t is the width of the derivative filter and σ the standard deviation. These one-dimensional derivatives are used to calculate the two horizontal and vertical derivative filters as shown below [23], [37].

Basic filters	Filter in x	Filter in y
G_x	f_1	f_2
G_y	f_2	f_1

Here, f_1 and f_2 are two vectors defined in Eq. 3.2. For both G_x and G_y filters, filter in x and filter in y are convolved with each column and each row of image I , respectively, to form the corresponding gradient image.

The LFDG Operators

The LFDG operators can be calculated using the representation as shown in Eq. 3.3 and Eq. 3.4 [32],

$$G_x = \sum_{k=1}^p f_k \cos\left(\frac{2\pi(k-1)}{p}\right), \quad (3.3)$$

$$G_y = \sum_{k=1}^p f_k \cos\left(\frac{\pi}{2} + \frac{2\pi(k-1)}{p}\right), \quad (3.4)$$

where p is the number of neighboring points and f_k the corresponding gray level of the k th neighbor.

The kernel size depends on the specified radius. For a kernel with radius R , the LFDG operator has the kernel size of $N \times N$, where $N = 2R + 1$. In our experiments, we use $R = 1$ and 34 neighboring points to calculate the kernel operators.

For all of the operators discussed above, G_x and G_y are convolved with the original image to compute the horizontal and vertical gradient images, respectively. Gradient orientations are computed using Eq. 3.5,

$$GO = \arctan \frac{G_y}{G_x}. \quad (3.5)$$

Finally, the orientations are then quantized into 8 bins. In particular, $0^\circ - 360^\circ$ orientations are divided into eight bins of 45° each. Each pixel's orientation is assigned to the nearest bin. The orientation bins are shown in Fig. 3.2. The blue

lines are the boundary of the orientation bins.

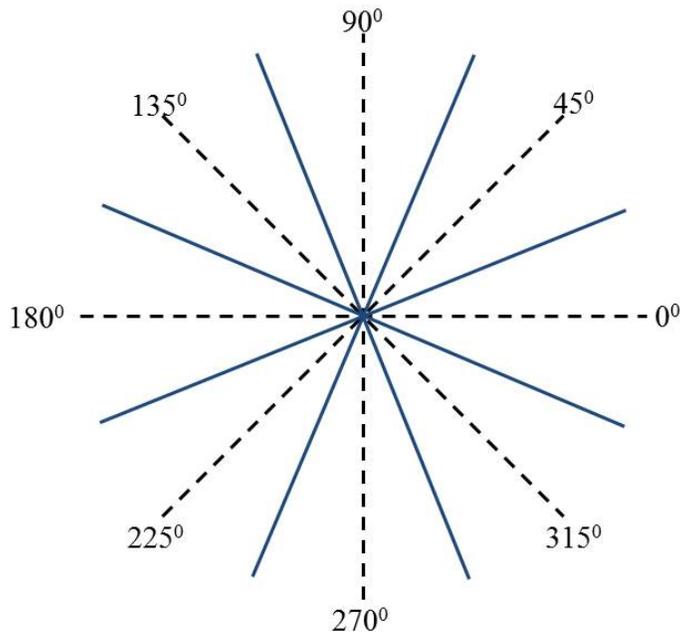


Figure 3.2: $0^\circ - 360^\circ$ degree orientations are quantized into 8 bins.

3.3 Feature Extraction

The quantized oriented image is used for feature extraction using the proposed methods. It is a two-step process. First, all the co-occurrence matrices for all the offsets are computed and then these matrices are combined to obtain the feature vector.

3.3.1 Co-occurrence Matrix (CM) Calculation

In CoHOG, an offset corresponds to the center pixel of the neighborhood to one of its neighbors (see Fig. 3.3 (b)). Fig. 3.3 (a) shows a neighborhood with a radius of size 4, 6, and 8 from the center green pixel. For a given radius, each neighbor within the radius is paired with the green pixel to form an offset. For example, using a neighborhood size of 4 we have a total of 31 offsets including the pair of the

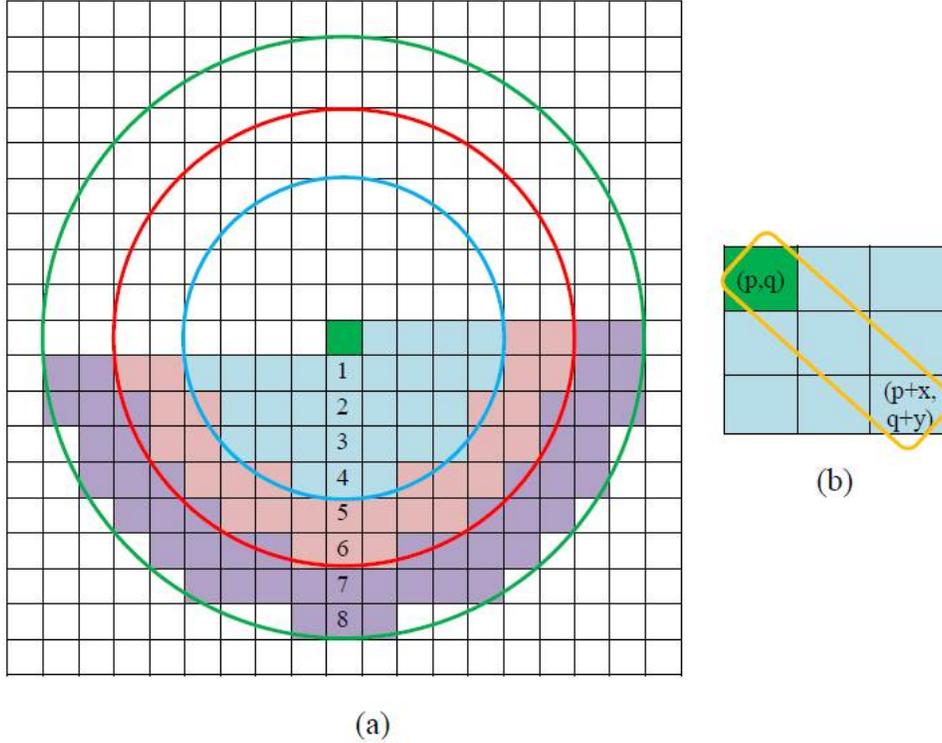


Figure 3.3: (a) Offsets for different radii and (b) a specific offset at distance (x, y) from pixel (p, q) .

green pixel with itself. Increasing the neighborhood size increases the number of offsets and thus the number of CMs. The upper half of the circular neighborhood is not considered because they are redundant since pixels in the top left corner of image are processed first.

In the proposed approach, we use the whole image for co-occurrence matrix calculation instead of using sub-regions as that used in the original CoHOG method. A neighborhood size of 4 and 8 are separately used for feature extraction. In CoHOG, the co-occurrence matrix is obtained by summing the co-occurrences of each pair of orientations for each offset. The size of the co-occurrence matrix is $N \times N$, where N is the number of distinct orientations which is pre-defined. For a specific offset (x, y) and a specific orientation at pixel $(p, q) = i$ and pixel $(p + x, q + y) = j$, the equation for calculating the CM is shown in Eq. 3.6 [21],

$$CM_{x,y}(i, j) = \sum_{p=1}^m \sum_{q=1}^n \begin{cases} 1 & \text{if } Q \text{ is True} \\ 0 & \text{Otherwise} \end{cases}, \quad (3.6)$$

where $Q = GO(p, q) = i$ and $GO(p + x, q + y) = j$ and $GM(p, q) \geq T$ and $GM(p + x, q + y) \geq T$. Here $m \times n$ is the size of the gradient oriented image I . GM is the gradient magnitude of the corresponding pixel and T the threshold magnitude to consider for the co-occurrence count.

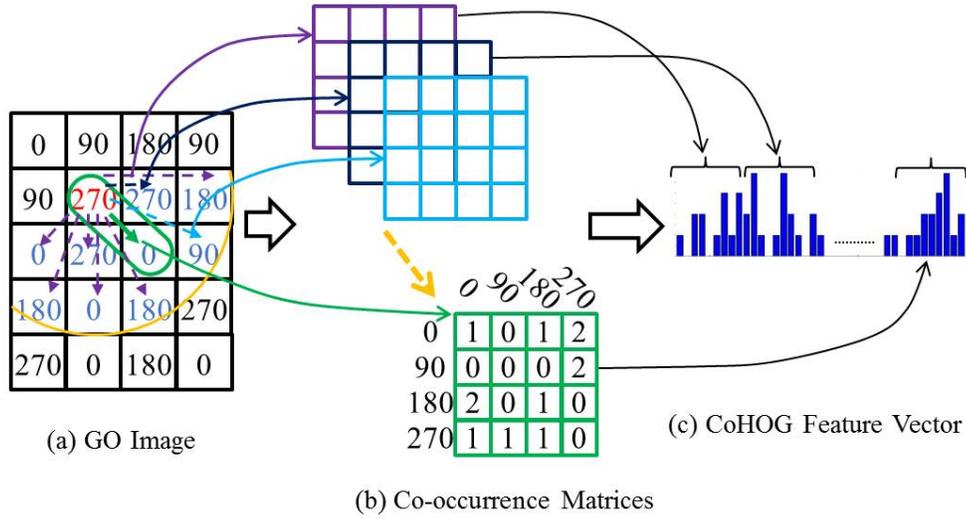


Figure 3.4: Illustration of the Co-occurrence Histograms of Oriented Gradients (CoHOG) calculation for the proposed methods.

Fig. 3.4 shows the workflow of CoHOG. It shows that using the gradient oriented image which is quantized into four different orientations (0^0 , 90^0 , 180^0 , and 270^0), a co-occurrence matrix of size 4×4 for each offset within the specified radius is created. The method scans each pixel for all the offsets and sums the co-occurrences of the orientations for that offset and stores the sums into the entry that corresponds to the pair of orientations of the specified co-occurrence matrix. After scanning all the pixels, it finishes in building up all the co-occurrence matrices. The algorithm

for computing the CMs are given in Algorithm 1.

<p>Algorithm 1: Algorithm for CM calculation of the proposed methods</p> <p>Given I: Gradient oriented Image;</p> <p>initialize: $CM \leftarrow 0$;</p> <p>for all positions (p, q) inside of the image do</p> <p style="padding-left: 1em;">$i \leftarrow I(p, q)$;</p> <p style="padding-left: 1em;">for all offsets (x, y) such that corresponds neighbors do</p> <p style="padding-left: 2em;">if $(p + x, q + y)$ is inside of the image then</p> <p style="padding-left: 3em;">$j \leftarrow I(p + x, q + y)$;</p> <p style="padding-left: 3em;">$CM(i, j, x, y) \leftarrow CM(i, j, x, y) + 1$;</p> <p style="padding-left: 2em;">end</p> <p style="padding-left: 1em;">end</p> <p>end</p>

3.3.2 Feature Vector Generation

Now all the created CMs are used to generate the feature vector for the selected image. A feature vector (FV) is generated by simply concatenating the CMs as shown in Eq. 3.7,

$$FV = \left\| \left\|_{i=1}^O \text{vec}(CM_i) \right. \right. \quad (3.7)$$

where, $\|$ is the concatenation operator, O is the number of offsets and vec is the vector representation of CM. The FV is a histogram of the co-occurrences of orientations of different offsets in the image (see Fig. 3.4(c)).

The size of the feature vector depends on the number of offsets used and the size of the CM as shown in Eq. 3.8,

$$FV \text{ size} = \text{number of offsets} \times \text{size of CM}. \quad (3.8)$$

One can see that the size of the feature vector is very large if the number of

offsets is large. For example, with a neighborhood size of 4, the total number of offsets is 31. Then the CM size is $8 \times 8 = 64$, and the FV size = $31 \times 64 = 1984$. With the same CM size, if the radius is increased to 6 with 61 offsets and 8 with 109 offsets then FV size = 3904 and FV size = 6976, respectively. When the FV size is large, it is difficult to distinguish between two classes of distinct categories, e.g. changes occur in a small portion of the images between two classes and the features of all other portions of the image classes are similar. Distinguishing between classes is difficult due to these similar features. Therefore, it is important to select significant features that are extracted from the changed portion of the images. These selected features have significant differences between the classes and are used for classification.

3.4 Feature Selection

The extracted feature vector size using the proposed methods is very large with many similar features. Changes that occur in a small portion of the images between two classes produce a large number of similar features. Hence, classification using this large number of similar features is very difficult for a classifier. In particular, using this large number of similar features creates ambiguity in defining the optimal hyperplane and leads to a wrong classification by a classifier. Therefore, it is necessary to select the significant features for better classification using a feature selection method.

ROC based methods [38], [39], [40], are well-known and promising in selecting important features. Some of the ROC based feature selection examples include feature ranking and significant feature selection using area under the ROC curve analysis [38], a regularized ROC method for disease classification and biomarker selection for microarray data [39], comparison of the ROC feature selection method with other popular methods [40] and feature selection using the ROC curve for small samples and imbalanced data classification problems [41]. These methods

demonstrate that better classification accuracy is obtained using the ROC feature selection approach.

In this thesis, feature selection is performed using ROC analysis. It is noteworthy that we are the first to use a feature selection method to extract significant CoHOG texture features to further improve classification accuracy. Significant features are selected using area under the ROC curve (AUC) analysis for classification. Only features that contain significant differences in between the classes are selected using an AUC threshold.

3.5 Classification

For classification we use a linear support vector machine (SVM) [34]. A two stage classification is used with the use of training and testing datasets.

For a two class classification, the SVM computes the optimal hyperplane to partition the feature space of the training samples into two halves. Samples from both classes are used for training. Each training sample consists of a feature vector and a label of its class.

Finally, the trained SVM is used to predict the class of a test sample using Eq. 3.9,

$$class(\vec{x}_t) = Sgn\left\{ \sum_{\forall k, l_k \in (p, c)} y(l_k) \alpha_k K(\vec{x}_t, \vec{x}_k) + b \right\}, \quad (3.9)$$

where $class(\vec{x}_t)$ is the class label of the test sample \vec{x}_t , \vec{x}_k is the feature vector of the k th training sample. $y(l_k)$ is the class label function of the k th sample which is either +1 or -1, α_k the Lagrangian multiplier for the training sample k , K the kernel function and b the bias parameter of the optimal hyperplane of the SVM. A linear kernel function is used to map data into higher dimensional spaces hoping that the data could be better separated. A linear kernel simply uses the dot product of two vectors.

The classification is performed using LIBSVM [42] version 3.20 package. The

SVM classifier is trained with a random selection of half of the dataset from each class and, then using the trained model, the classification accuracy is tested using the rest of the sample. The average classification accuracy was recorded over 1000 runs to reduce the effect of randomness.

Chapter 4

Experimental Results

In this chapter, we discuss the results using the proposed methods for different datasets. The proposed methods are implemented in Matlab. The program runs on a PC with an Intel Core i7 with 3.40GHz CPU with 24GB RAM running Windows 7 Professional. Three well-known texture datasets are used in this experiment to compare the classification performance of the proposed methods to other state of the art methods.

We also use another three datasets consisting of 2D MR images of ALS patients and healthy controls for classification. We compare the results of these MRI datasets with that of the GLCM method that has been used for texture classification in ALS. Another multicenter dataset in ALS of different image contrasts is also used to evaluate the classification performance of the proposed methods on datasets having various image resolutions and contrasts.

4.1 Texture Datasets

In our experiments, we use three well-known texture datasets, namely, INRIA Person [20], Columbia-Utrecht Reflectance and Texture (CURET) [43], and the UIUC [44]. The details of these texture datasets are discussed below.

4.1.1 INRIA Person Dataset

The INRIA Person [20] dataset is a widely used pedestrian detection benchmark dataset. The dataset contains two classes of human and nonhuman images of various sizes. We have used 200 images per class and human and nonhuman images are divided into equal halves in each class. Three image samples of each class of this dataset are shown in Fig. 4.1.



Figure 4.1: Three sample images of (a) human and (b) nonhuman classes from the INRIA Person dataset.

4.1.2 CURET Dataset

In Columbia-Utrecht Reflectance and Texture (CURET) [43] dataset, we use 10 different texture classes with 55 samples in each class. The image resolution is 640×480 for all the classes. The images are acquired from a physical texture sample

photographed under a range of viewing and illumination angles. Three sample images in class 1, 3 and 5 are shown in Fig. 4.2.

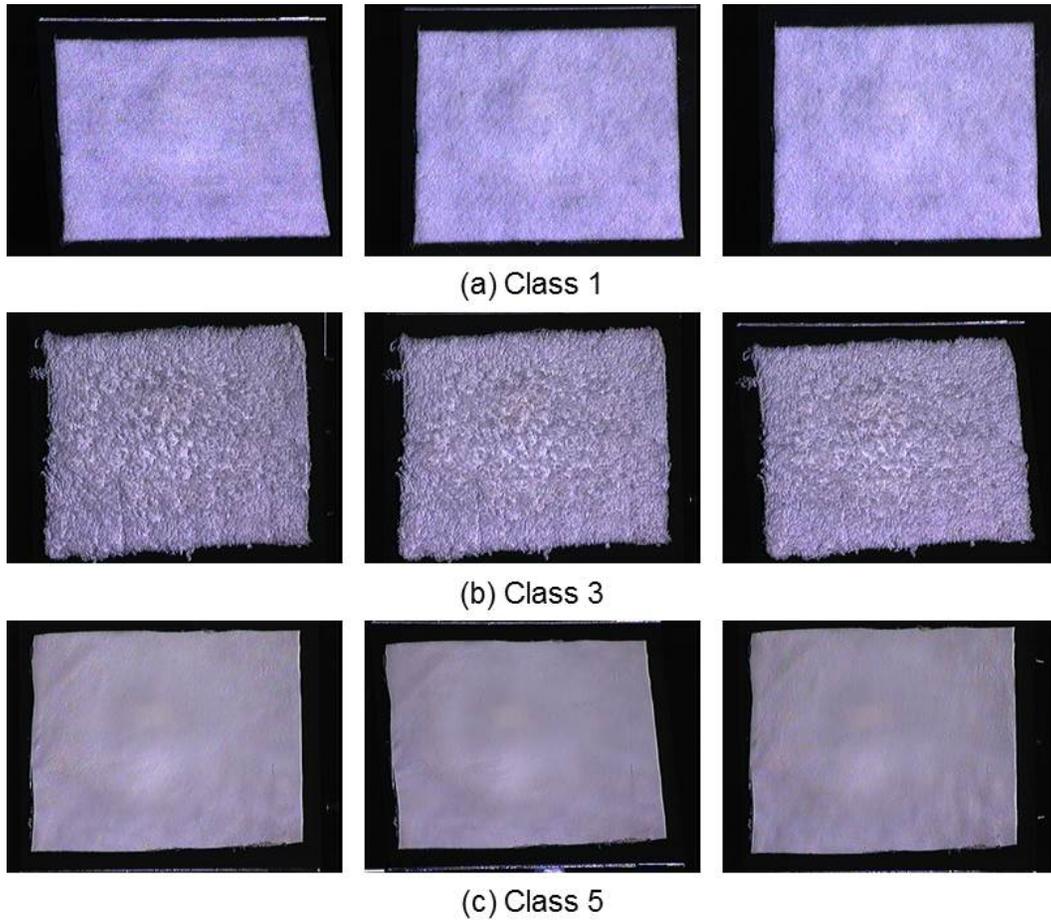


Figure 4.2: Three sample images of (a) class 1 (b) class 3 and (c) class 5 from the CURET dataset.

4.1.3 UIUC Dataset

The third texture dataset we use is UIUC [44]. In the UIUC dataset, we use 10 different texture classes with each class containing 40 image samples. All the classes have the same image resolution of 640×480 . The dataset includes materials imaged under significant view-point variations. Three sample images in class 1, 3 and 5 are shown in Fig. 4.3.

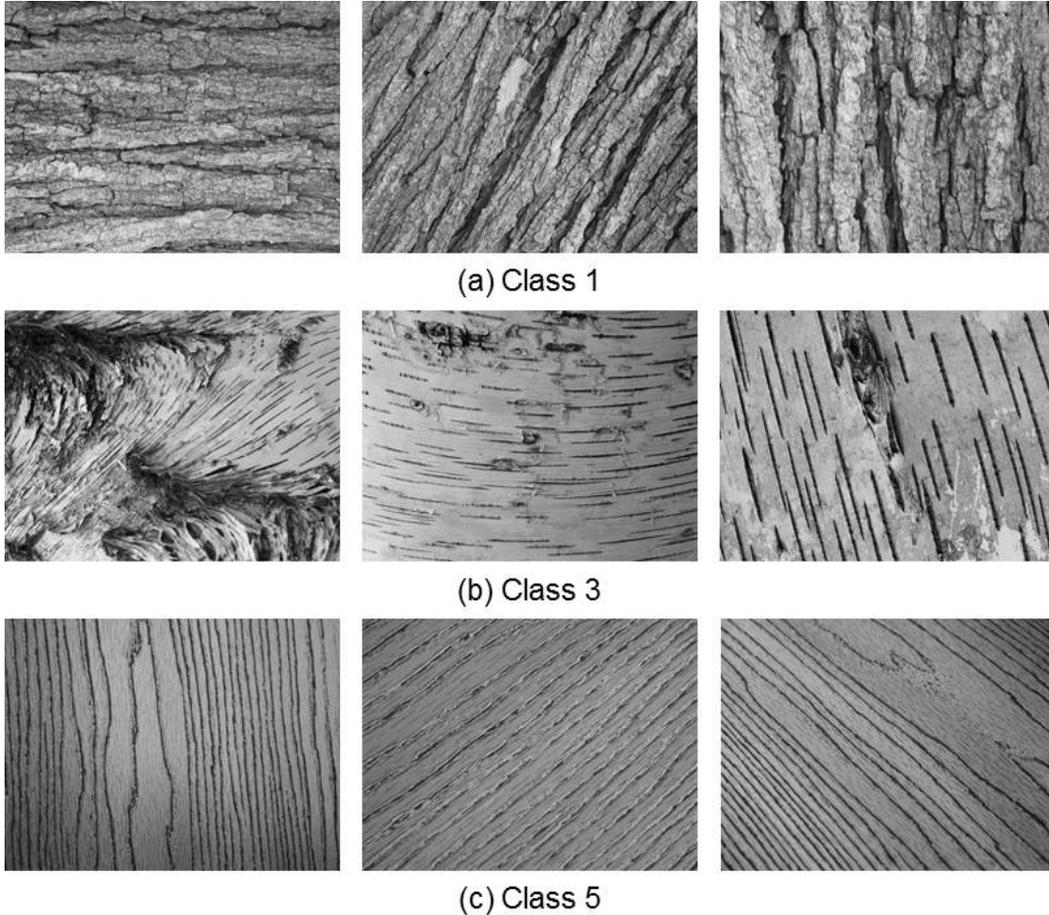


Figure 4.3: Three sample images of (a) class 1 (b) class 3 and (c) class 5 from the UIUC dataset.

4.2 Human MRI Datasets

We use three different datasets of Magnetic Resonance Images (MRI) of Amyotrophic Lateral Sclerosis (ALS) patients and healthy controls. These datasets are acquired using two different scanning machines with different scanning parameters for each dataset. The datasets are referred to as MR dataset 1, 2 and 3. The subjects in MR dataset 1 are different from that in MR dataset 2 and 3. Also the image resolution and contrast are different for each dataset. Another multicenter ALS dataset of different image contrasts is also used in this experiment.

All patients are clinically probable or definite sporadic ALS according to the revised El Escorial criteria [45] were recruited. By this criteria, all patients had

clinical evidence of upper motor neuron (UMN) and lower motor neuron (LMN) involvement.

In this experiment, we use the specified coronal slices of the the MRI scans of the whole brain. Then a region of interest (ROI) is selected for texture feature selection. We also use different downsampled version of the same image slice in the experiment. The details of ALS, ROI selection and downsampling of the subjects are discussed below.

4.2.1 Amyotrophic Lateral Sclerosis (ALS)

Amyotrophic Lateral Sclerosis (ALS) is a fatal progressive degenerating disorder of adulthood that causes rapid muscular weakness and disability. Several factors, including clinical presentation, rate of disease progression, early presence of respiratory failure, and the nutritional status of patients impact the survival of ALS [46].

ALS is an idiopathic disease of the human motor system. It affects both the UMN of the cerebral cortex and the LMN in the brainstem and spinal cord [46], [47], [48], [49]. UMN dysfunction leads to spasticity, weakness, and brisk deep tendon reflexes. By contrast, features of LMN impairment include fasciculations, muscle wasting, and weakness. Spastic dysarthria, which is characterized by slow, labored, and distorted speech, often with a nasal quality is caused by bulbar UMN dysfunction. Bulbar LMN dysfunction can be identified by tongue wasting, weakness, and fasciculations, accompanied by flaccid dysarthria and later dysphagia [46].

People are affected by ALS worldwide. Men have a higher incidence than do women, although the incidence is about the same in familial disease between men and women. About 90% of ALS patients have sporadic disease and rest of the 10% are familial. About 50% of patients die within 30 months and about 20% of them survive between 5 years to 10 years after the onset of symptom [46].

Currently, there is no reliable tool to provide a quantitative measure of cerebral degeneration in ALS and such a tool is desperately needed to aid in diagnosis and

to evaluate novel therapies.

4.2.2 ROI Selection

From the MRI scan of the whole brain, coronal slices with an angulation parallel to the corticospinal tract (CST) (see Fig. 4.4 (a)) are used for texture calculation (see Fig. 4.4 (c)). The image angulation is performed using Mango [50].

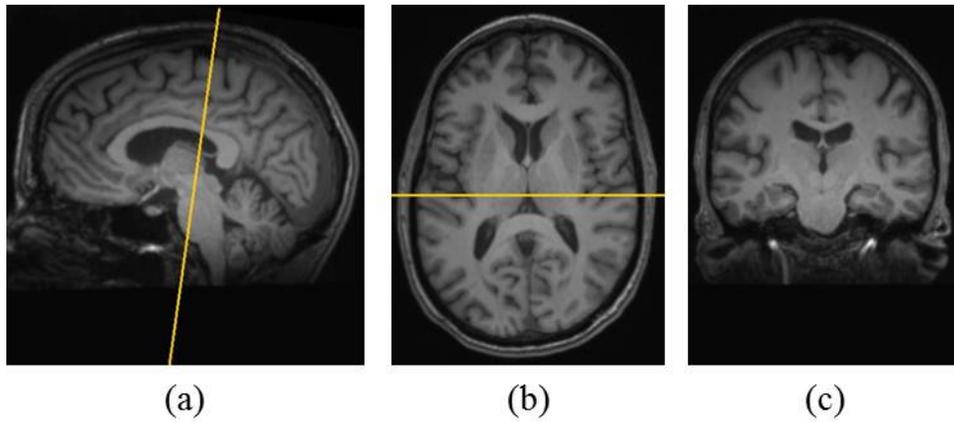


Figure 4.4: (a) Sagittal, (b) Axial and (c) Coronal image slices. Coronal imaging is used in texture feature extraction.

In particular, an ROI is manually defined that includes the region above the inferior horn of the lateral ventricles (see Fig. 4.5) and is specified by creating a mask to segment out the regions of interest. Masks for each subject are created separately using ITK-SNAP [51].

4.2.3 Downsampling

The selected ROI is downsampled to four different resolutions. Different scaling factors are chosen to downsample the ROI into $1 \times 1 \text{ mm}^2$, $2 \times 2 \text{ mm}^2$, $3 \times 3 \text{ mm}^2$ and $4 \times 4 \text{ mm}^2$ physical dimensions. The downsampling is done using ImageJ [52].

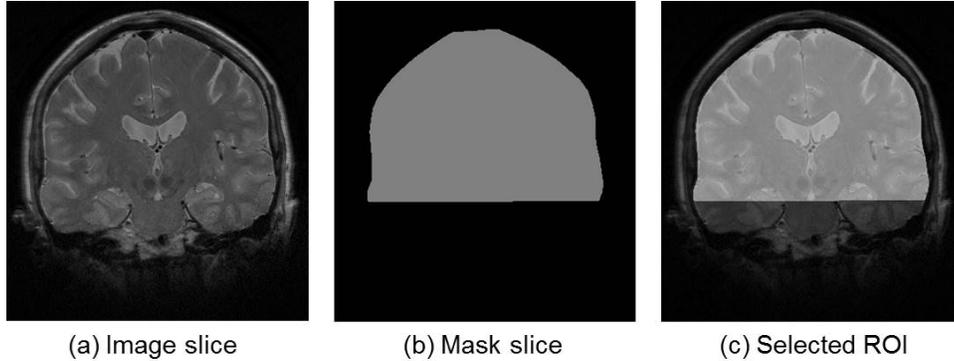


Figure 4.5: ROI selection from a coronal image slice. The highlighted regions are selected as ROI.

4.2.4 MR Dataset 1

Twelve patients and nineteen controls are in this datasets. Details of the patients and controls for this dataset are given in Table 4.1.

Table 4.1: Details of the subjects for MR Dataset 1

Subjects	No.	Average Age	Male	Female
Patients	12	57.4 ± 10.0	7	5
Controls	19	57.0 ± 10.5	8	11

MR imaging was performed on a 4.7 Tesla whole-body scanner (Varian Unity Inova console). High-resolution fast spin echo T2-weighted images were acquired in the coronal plane ($TR = 4000ms$, $TE = 33.3ms$, pixel size = $0.5 \times 0.5mm^2$, slice thickness = $2mm$). Three sample image slices of both patients and controls are shown in Fig. 4.6. We can see that patients and controls are not distinguishable by visual inspection. 2D MR images of the subjects are downsampled into four different resolutions. Details of the downsampled images are given in Table 4.2.

4.2.5 MR Dataset 2

Nineteen patients and twenty controls are in this dataset. Details of the patients and controls are given in Table 4.3.

MRI scans were done on a 1.5 Tesla system (Magnetom Sonata, Siemens Medical

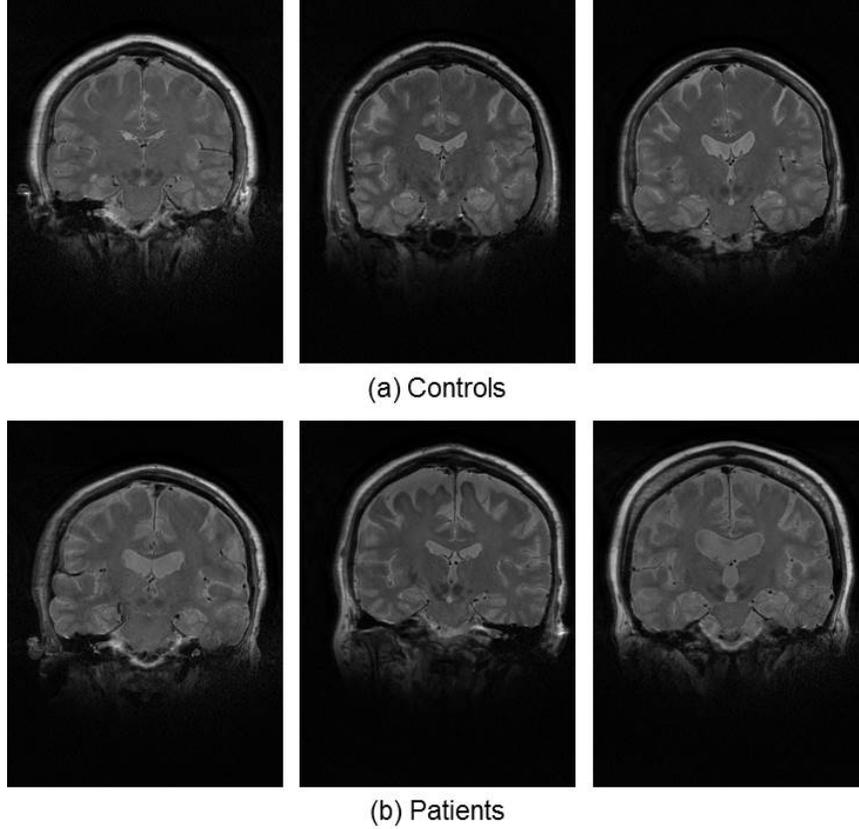


Figure 4.6: Three sample image slices of (a) controls and (b) patients from MR Dataset 1. Patients and controls are not distinguishable by visual inspection.

Table 4.2: Details of the downsampled 2D MR images of the subjects for MR Dataset 1 and MR Dataset 3

Scale Factor	Pixel size (mm^2)	Image size ($Pixel^2$)
1	0.5×0.5	385×512
0.5	1×1	192×256
0.25	2×2	96×128
0.167	3×3	64×85
0.125	4×4	48×64

Table 4.3: Details of the subjects for MR Dataset 2 and MR Dataset 3

Subjects	No.	Average Age	Male	Female
Patients	19	56.7 ± 13.7	10	9
Controls	20	56.8 ± 12.4	9	11

Systems). Coronal T2-weighted images were acquired ($TR=7510ms$, $TE=113ms$, pixel size = $1.1 \times 0.9mm^2$, slice thickness = $5mm$). Three sample image slices of

both patients and controls are shown in Fig. 4.7. We can see that patients and controls are not distinguishable by visual inspection. MR dataset 2 images are

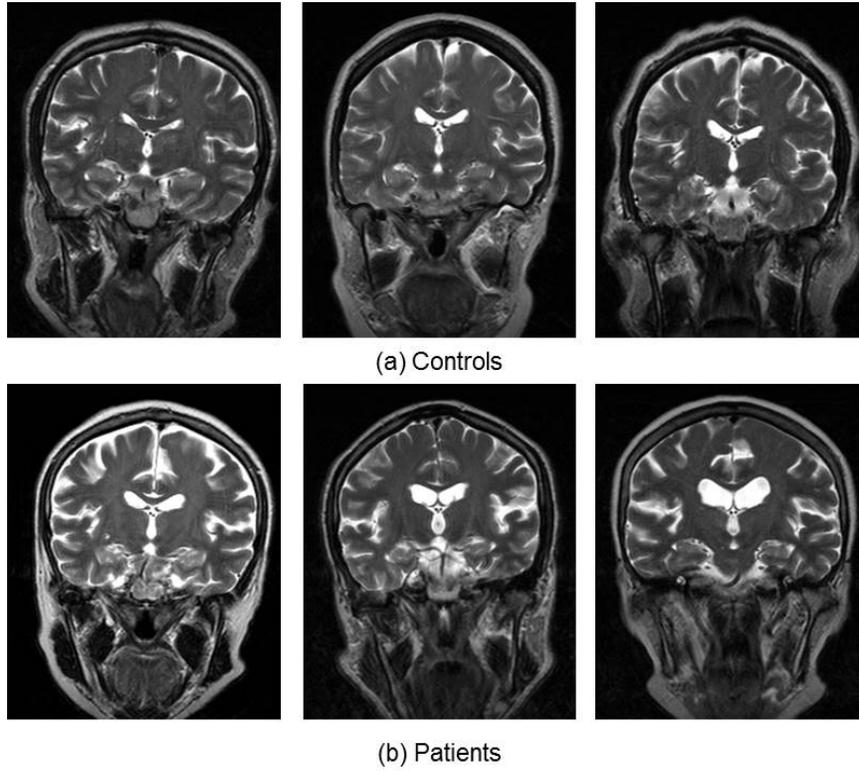


Figure 4.7: Three sample image slices of (a) controls and (b) patients from MR Dataset 2. Patients and controls are not distinguishable by visual inspection.

downsampled into four different resolutions. Details of the image resolutions for each scale are given in Table 4.4.

Table 4.4: Details of the downsampled 2D MR images of the subjects for MR Dataset 2

Scale Factor	Pixel size (mm^2)	Image size ($Pixel^2$)
1	0.86×0.86	208×256
0.86	1×1	178×220
0.43	2×2	89×110
0.285	3×3	59×72
0.215	4×4	44×55

4.2.6 MR Dataset 3

All the subjects are the same for MR dataset 2 and MR dataset 3 (see Table 4.3). But MR dataset 3 was acquired with a T1-weighted MPRAGE (TR=1600ms, TE=3.8ms, TI=1100ms, pixel size = $1.0 \times 1.0mm^2$, slice thickness = 1.5mm). MRI scanning were performed on a 1.5 Tesla system (Magnetom Sonata, Siemens Medical Systems). Three sample image slices of both patients and controls are shown in Fig. 4.8. We can see that patients and controls are not distinguishable by visual inspection. The resolutions of downsampled images are given in Table 4.2.

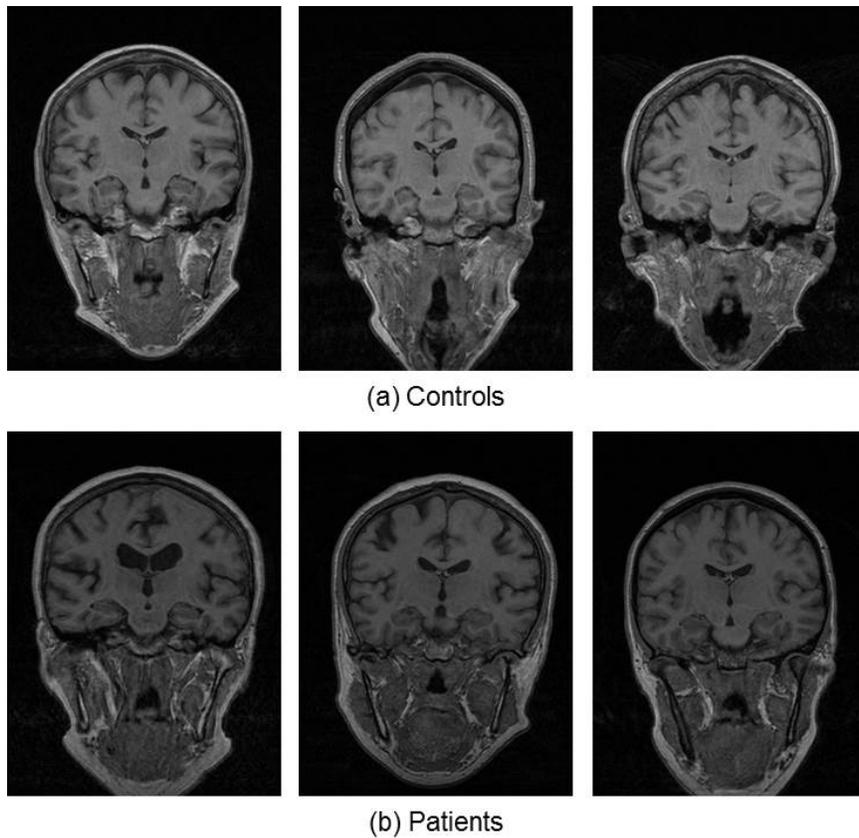


Figure 4.8: Three sample image slices of (a) controls and (b) patients from MR Dataset 3. Patients and controls are not distinguishable by visual inspection.

For all the MR datasets, coronal imaging was performed with an angulation parallel to the CST (see Fig. 4.4). ROI was manually selected for each subject that covers the region above the inferior horn of the lateral ventricles. One sample ROI

is shown in Fig. 4.5.

4.3 Classification Results of Texture Datasets

In this section, we discuss the classification accuracy of the proposed S-CoHOG, GD-CoHOG and LFDG-CoHOG methods using two neighborhood sizes of 4 and 8. The original CoHOG method uses a maximum neighborhood size of 4. In this thesis, we employ a larger neighborhood size of 8 to see the effect of using more global information for co-occurrence calculation on classification accuracy. The feature vector size of neighborhood size of 4 is 1984 and is 6976 for neighborhood size of 8. Classification using these very large feature vectors is difficult due to large number of similar features. A large number of similar features creates problem in defining the optimal hyperplane to separate the two classes and thus produces wrong classification.

For the first time, we employ a feature selection method with the CoHOG features in this thesis. In all of the reported experiments, we have applied the ROC feature selection method with a significance level of $p \leq 0.01$ using an AUC threshold to reduce the feature vector size and to select the significant features for classification. The AUC threshold is chosen to minimize the number of features such that the selected features can produce better classification results.

Classification accuracies with and without feature selection are compared. We also compare the results of the proposed methods with that of the other CoHOG methods. The detailed classification results of the three texture datasets using the proposed methods are discussed below.

4.3.1 Classification of INRIA Person Dataset

The INRIA database contains two classes of images, namely, human and nonhuman. We use the proposed methods for feature extraction and selected using ROC analysis with AUC thresholds of 0.95 and 0.99 for a neighborhood size of 4 and 8,

respectively. These thresholds are chosen to minimize the number of features such that selected features can produce better classification results. For classification, we use half of the images from each class for training and the remaining images from both classes for testing. The classification accuracies with and without feature selection for two neighborhood sizes are shown in Table 4.5.

Table 4.5: Classification accuracy of the proposed methods using INRIA Person dataset. I means classification without feature selection and II mean classification with feature selection.

Proposed Method	Neighborhood Size of 4		Neighborhood Size of 8	
	Classification Accuracy (I)	Classification Accuracy (II)	Classification Accuracy (I)	Classification Accuracy (II)
S-CoHOG	99.00%	99.30%	98.90%	99.90%
GD-CoHOG	99.40%	99.60%	99.00%	99.90%
LFDG-CoHOG	98.70%	99.30%	99.00%	99.50%

The classification results for this dataset are almost 100% for all the proposed methods. There is a small improvement in most cases in the classification accuracy between with and without feature selection. Also using a larger neighborhood size has little impact on the classification accuracy. The S-CoHOG and GD-CoHOG methods acquire a maximum classification accuracy of 99.90% with neighborhood size of 8 with feature selection.

Table 4.6: Number of features selected using AUC threshold for INRIA Person dataset using two neighborhood sizes.

Proposed Method	Neighborhood Size of 4		Neighborhood Size of 8	
	AUC Threshold	Selected Features	AUC Threshold	Selected Features
S-CoHOG	0.95	193	0.99	1403
GD-CoHOG	0.95	807	0.99	1080
LFDG-CoHOG	0.95	655	0.99	1245

The number of selected features using an AUC threshold for each of the method is shown in Table 4.6. The selected features are smaller than that of the total number of features for both the neighborhood sizes. The proposed methods achieve better classification accuracy using the selected features than using all the features.

4.3.2 Classification of CURET Dataset

The classification results of the proposed methods are calculated using 10 classes of CURET dataset. Each class contains 55 images and half of the images in each class are used to train the classifier and rests of the images in each class are used for testing. A two class classification is performed among the 10 classes and the average classification accuracy is recorded. Feature selection is performed using an AUC threshold of 0.70 and 0.80 for a neighborhood size of 4 and 8, respectively. These thresholds are chosen to minimize the number of features such that selected features can produce better classification results. The classification accuracies with and without feature selection for two neighborhood sizes are shown in Table 4.7.

Table 4.7: Classification accuracy of the proposed methods using CURET dataset. I means classification without feature selection and II mean classification with feature selection.

Proposed Method	Neighborhood Size of 4		Neighborhood Size of 8	
	Classification Accuracy (I)	Classification Accuracy (II)	Classification Accuracy (I)	Classification Accuracy (II)
S-CoHOG	96.70%	96.60%	96.30%	97.80%
GD-CoHOG	96.80%	97.40%	96.50%	98.30%
LFDG-CoHOG	96.80%	97.10%	96.40%	97.60%

From the classification results, we observe that the proposed methods have higher classification accuracy using feature selection than without feature selection when the neighborhood size is 8. For a neighborhood size of 4, we found an exception for S-CoHOG that it has a slightly lower classification accuracy with feature selection. Without feature selection, the classification accuracies are almost the same in all the methods for both neighborhood sizes, but using feature selection the proposed methods have a better classification accuracy for neighborhood size of 8. The GD-CoHOG method acquires a maximum classification accuracy of 98.30% for neighborhood size of 8 and using feature selection.

The number of selected features using an AUC threshold for each of the methods is shown in Table 4.8. The number of selected features are much smaller than that of

Table 4.8: Number of features selected using AUC threshold for CURET dataset using two neighborhood sizes.

Proposed Method	Neighborhood Size of 4		Neighborhood Size of 8	
	AUC Threshold	Selected Features	AUC Threshold	Selected Features
S-CoHOG	0.70	343	0.80	235
GD-CoHOG	0.70	371	0.80	403
LFDG-CoHOG	0.70	378	0.80	563

the (1984 for neighborhood size of 4 and 6976 for neighborhood size of 8) number of features for both neighborhood sizes. The proposed methods achieve better classification accuracy using the selected features than using all the features.

4.3.3 Classification of the UIUC Dataset

In this dataset, we also use 10 different classes of images and each of which contains 40 images. ROC feature selection is performed with an AUC threshold of 0.80 and 0.90 for a neighborhood size of 4 and 8, respectively. These thresholds are chosen to minimize the number of features such that selected features can produce better classification results. Half of the images in each class are used for training and the remaining images in each class are used for testing. A two class classification is performed among the 10 classes and the average classification accuracy is recorded. The classification accuracies with and without feature selection for two neighborhood sizes are shown in Table 4.9.

Table 4.9: Classification accuracy of the proposed methods using the UIUC dataset. I means classification without feature selection and II mean classification with feature selection.

Proposed Method	Neighborhood Size of 4		Neighborhood Size of 8	
	Classification Accuracy (I)	Classification Accuracy (II)	Classification Accuracy (I)	Classification Accuracy (II)
S-CoHOG	95.30%	95.20%	95.60%	97.00%
GD-CoHOG	93.70%	96.60%	95.00%	98.00%
LFDG-CoHOG	95.00%	95.60%	95.31%	97.50%

For this dataset, we observe that the proposed methods with feature selection

have similar or better classification results than without feature selection for both the neighborhood sizes. For a neighborhood size of 4, we found an exception for S-CoHOG method that it has a slightly lower classification accuracy with feature selection than that without feature selection. Using a larger neighborhood, we found better results than that of using a smaller neighborhood. The GD-CoHOG method acquires a maximum classification accuracy of 98.00% for neighborhood size of 8 using feature selection.

Table 4.10: Number of features selected using AUC threshold for UIUC dataset using two neighborhood sizes.

Proposed Method	Neighborhood Size of 4		Neighborhood Size of 8	
	AUC Threshold	Selected Features	AUC Threshold	Selected Features
S-CoHOG	0.80	436	0.90	1028
GD-CoHOG	0.80	188	0.90	743
LFDG-CoHOG	0.80	395	0.90	895

The AUC thresholds used for feature selection and number of selected features for the proposed methods using this dataset are shown in Table 4.10.

4.3.4 Comparison with other CoHOG methods

We compare the classification results of the proposed methods with that of the original CoHOG method [21] and the Eig(Hess)-CoHOG method [23] using the INRIA Person, CURET and the UIUC texture datasets.

The original CoHOG method uses 6 sub-regions and a neighborhood size of 4. Sobel operators are used for gradient calculations. The Eig(Hess)-CoHOG uses the Hessian matrix to calculate the eigen values of the image surface. These eigen values are used for pixel orientation calculation. This method uses 4 sub-regions with a neighborhood size of 4. The comparison results are shown in Table 4.11.

Here we compare the results found using feature selection and neighborhood size of 8 for the proposed methods with that of the original CoHOG and Eig(Hess)-CoHOG methods. The original CoHOG method uses the normalized images of zero

Table 4.11: Comparison of the classification accuracies (CA) of the proposed methods with the original CoHOG and the Eig(Hess)-CoHOG methods.

Method	INRIA Dataset CA	CUReT Dataset CA	UIUC Datasets CA
Original CoHOG	95.5%	94.94%	77.41%
Eig(Hess)-CoHOG	-	90.00%	91.66%
S-CoHOG	99.90%	97.80%	97.00%
GD-CoHOG	99.90%	98.30%	98.00%
LFDG-CoHOG	99.50%	97.60%	97.50%

mean and unit standard deviation. We do not use any normalization of the image dataset for the proposed methods. It is noteworthy that the results of the original CoHOG are worse for images without normalization. For the CUReT and UIUC datasets, GD-CoHOG has the best classification accuracies. S-CoHOG and GD-CoHOG have the best results for INRIA Person dataset. The Eig(Hess)-CoHOG method has better classification results than that of the original CoHOG method but are worse than that of our proposed methods.

4.4 Classification Results of MRI Datasets

For the three MR datasets, we use ROC analysis with a significance level of $p \leq 0.01$ to find the area under the ROC curve (AUC). Then classification of the subjects in the dataset is performed using linear SVM classifier. In each run, half of the patients and controls for each dataset are randomly selected for training and the rest of the patients and controls are used for testing. The average classification accuracy and the optimal sensitivity and specificity over 1000 runs are recorded. The details of the ROC analysis and classification results are discussed below.

4.4.1 ROC Analysis of MR Dataset 1

Six patients and ten controls are used to train the linear SVM classifier and the rest of the patients and controls are used for testing in this dataset. The maximum AUC is calculated for the selected features and then the classification accuracy is

calculated using the selected features. The results are shown in Table 4.12 and 4.13.

Table 4.12: ROC analysis of MR Dataset 1 using S-CoHOG features extracted for neighborhood radius of 4.

Image Pixel Size (mm^2)	Maximum AUC	Optimal Sensitivity	Optimal Specificity	Classification Accuracy
0.5×0.5	0.815	54%	67%	63.00%
1×1	0.895	57%	73%	67.30%
2×2	0.886	81%	84%	83.50%
3×3	0.895	81%	91%	87.30%
4×4	0.842	71%	84%	79.00%

Table 4.13: ROC analysis of MR Dataset 1 using S-CoHOG features extracted for neighborhood radius of 8.

Image Pixel Size (mm^2)	Maximum AUC	Optimal Sensitivity	Optimal Specificity	Classification Accuracy
0.5×0.5	0.831	50%	61%	57.00%
1×1	0.895	57%	74%	67.70%
2×2	0.906	74%	83%	79.30%
3×3	0.917	91%	95%	93.00%
4×4	0.921	90%	89%	90.30%

Four different downsampled images along with the original image are used in this experiment. From the results it is shown that features extracted (using both neighborhood size of 4 and 8) from downsampled images (image pixel size = $3 \times 3 mm^2$) have better classification accuracy with a higher maximum AUC than that of using the original image resolution. In particular, the best classification accuracy (93.00%), the maximum AUC (0.917) and the optimal sensitivity (91%) and specificity (95%) are obtained using features extracted with a neighborhood size of 8.

4.4.2 ROC Analysis of MR Dataset 2

In this dataset, we use 19 patients and 20 controls for classification and ROC analysis. Ten patients and 10 controls are used for training the linear SVM and

the other 9 patients and 10 controls are used for testing. The classification results for different downsampled images with two neighborhood sizes are shown in Table 4.14 (four neighbors) and Table 4.15 (eight neighbors).

Table 4.14: ROC analysis of MR Dataset 2 using S-CoHOG features extracted for neighborhood radius of 4.

Image Pixel Size (mm^2)	Maximum AUC	Optimal Sensitivity	Optimal Specificity	Classification Accuracy
0.86×0.86	0.783	48%	57%	53.30%
1×1	0.791	54%	59%	56.70%
2×2	0.810	63%	78%	71.00%
3×3	0.834	75%	78%	76.90%
4×4	0.856	84%	86%	85.30%

Table 4.15: ROC analysis of MR Dataset 2 using S-CoHOG features extracted for neighborhood radius of 8.

Image pixel Size (mm^2)	Maximum AUC	Optimal Sensitivity	Optimal Specificity	Classification Accuracy
0.86×0.86	0.834	55%	60%	57.70%
1×1	0.855	60%	61%	61.10%
2×2	0.850	66%	78%	72.70%
3×3	0.834	77%	82%	80.00%
4×4	0.867	92%	88%	90.40%

We observe from the results that downsampling increases the classification accuracy along with sensitivity and specificity. Here we found the best classification accuracy (90.40%), the maximum AUC (0.867) and with the best optimal sensitivity (92%) and specificity (88%) in downsampled images (image pixel size = $4 \times 4 mm^2$) with a neighborhood size of 8.

4.4.3 ROC Analysis of MR Dataset 3

MR Dataset 3 is a T1-weighted dataset of the same subjects as MR dataset 2. Similar experimental settings are used for this dataset as the other two. The observed ROC analysis and classification results are shown in Table 4.16 (four neighbors) and Table 4.17 (eight neighbors).

Table 4.16: ROC analysis of MR Dataset 3 using S-CoHOG features extracted for neighborhood radius of 4.

Image Pixel Size (mm^2)	Maximum AUC	Optimal Sensitivity	Optimal Specificity	Classification Accuracy
0.5×0.5	0.641	40%	35%	37.50%
1×1	0.753	53%	45%	49.30%
2×2	0.811	63%	65%	64.50%
3×3	0.818	68%	75%	72.70%
4×4	0.869	81%	75%	78.10%

Table 4.17: ROC analysis of MR Dataset 3 using S-CoHOG features extracted for neighborhood radius of 8.

Image Pixel Size (mm^2)	Maximum AUC	Optimal Sensitivity	Optimal Specificity	Classification Accuracy
0.5×0.5	0.659	31%	37%	34.20%
1×1	0.780	43%	39%	41.50%
2×2	0.821	75%	74%	74.20%
3×3	0.845	81%	83%	82.00%
4×4	0.895	94%	92%	93.50%

We also found the best results at a lower resolution for this dataset as well. In this case, downsampled images (image pixel size = $4 \times 4 \text{ mm}^2$) have the best results. We observed the best classification accuracy (93.50%), the maximum AUC (0.895) and the best optimal sensitivity (94%) and specificity (92%) when the neighborhood size is 8.

For all of the three MR datasets results, we apply the S-CoHOG method for texture feature extraction. MR Dataset 1 and 2 are both T2-weighted but are collected from different scanner parameters. MR dataset 3 is a T1-weighted dataset which is different from the other datasets. We select the features with $AUC \geq 0.8$ for all of the above results because we observe that the best classification results are obtained using $AUC \geq 0.8$. Fig. 4.9 shows the classification accuracy for the selected features using different AUC thresholds.

MR Dataset 1 and two neighborhood sizes are used in this experiment. For both neighborhoods, we found the best classification results at $AUC \geq 0.8$. Mean feature values and standard deviations of mean feature values of ten selected features of

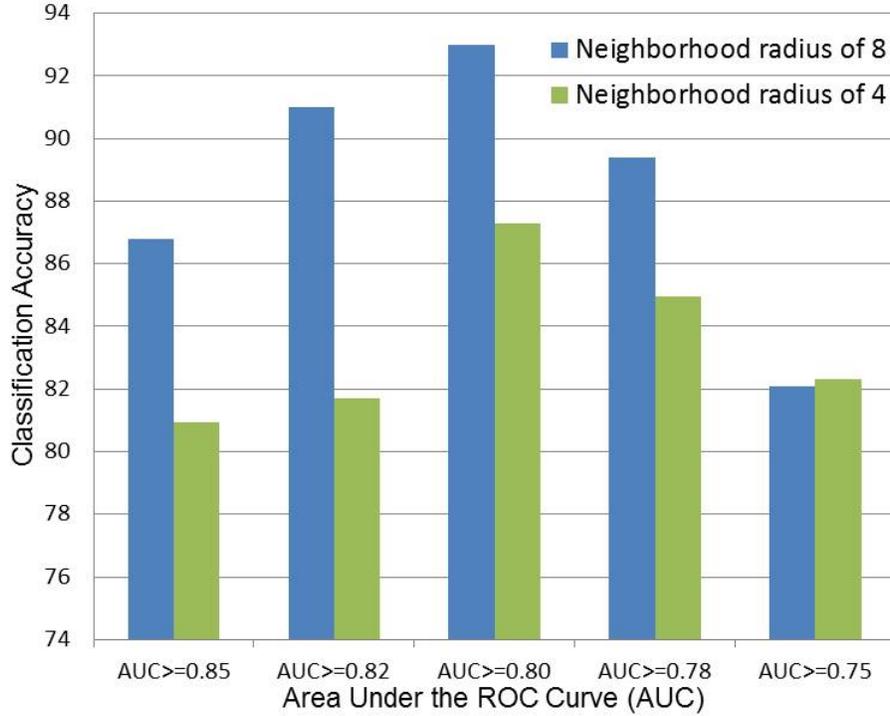


Figure 4.9: Classification accuracy for selected features using different AUC thresholds for MR dataset 1.

patients and controls using $AUC \geq 0.8$ for MR dataset 1 and MR dataset 2 are shown in Fig. 4.10. There is a significant difference between the patients and the controls in the mean feature values and also their standard deviations do not overlap.

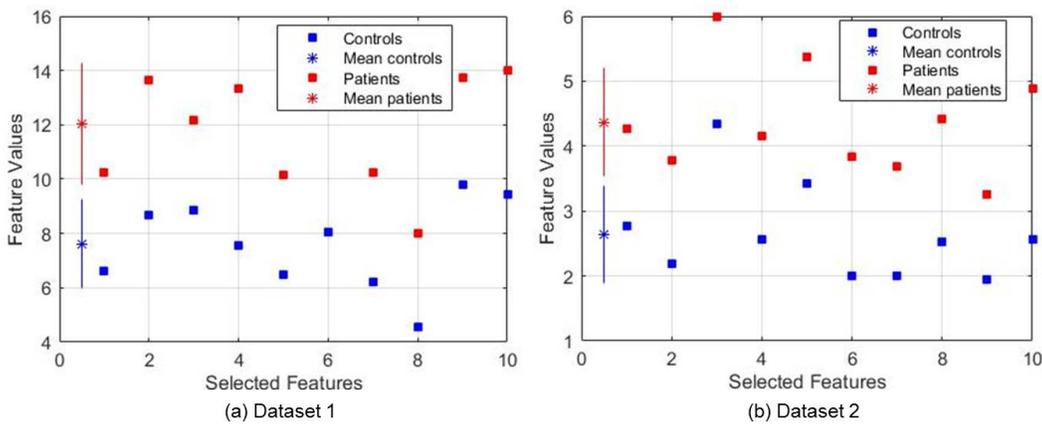


Figure 4.10: Mean feature values with standard deviations of the mean feature values between patients and controls of ten selected features for (a) MR Dataset 1 and (b) MR Dataset 2.

We observe that images with lower resolution give better results in classification. This is because the reduced resolution image contains more dense information than the original image for a fixed neighborhood size. With a fixed neighborhood size, CoHOG can cover more regions in feature extraction for downsampled image than for the original image. Thus the extracted features have more distance information for downsampled image than for the original image. For MR dataset 1, the best results are obtained using downsampled images with a pixel size of $3 \times 3 \text{ mm}^2$ and of $4 \times 4 \text{ mm}^2$ give the best results for MR dataset 2 and 3. For all of the MR datasets, the best results are observed using a neighborhood size of 8. Therefore, the rest of the experimental analysis is focused on using only downsampled images with a pixel size of $3 \times 3 \text{ mm}^2$ for MR dataset 1 and of $4 \times 4 \text{ mm}^2$ for MR dataset 2 and 3 with a neighborhood size of 8.

4.4.4 ROC Analysis using different Gradient Operators

Based on three different gradient operators, Sobel, Gaussian Derivative (GD) and Local Frequency Descriptor Gradient (LFDG), our three proposed methods S-CoHOG, GD-CoHOG and LFDG-CoHOG are used separately with the ROC curve to analyze the classification accuracy of MR dataset 1 and 2. The results of classification accuracy and optimal sensitivity and specificity for MR dataset 1 are shown in Table 4.18.

Table 4.18: ROC analysis of MR Dataset 1 using three proposed methods. CoHOG features extracted using a neighborhood radius of 8.

Proposed Methods	Maximum AUC	Optimal Sensitivity	Optimal Specificity	Classification Accuracy
S-CoHOG	0.917	91%	95%	93.00%
GD-CoHOG	0.954	98%	96%	97.30%
LFDG-CoHOG	0.897	88%	98%	94.75%

All of the three proposed methods have a high classification accuracy with excellent optimal sensitivity and specificity. Among them, GD-CoHOG method has the highest classification accuracy (97.30%) and the optimal sensitivity (98%) and

specificity (96%). Also, GD-CoHOG has the highest maximum AUC (0.954) among the all the operators.

Table 4.19: ROC analysis of MR Dataset 2 using three proposed methods. CoHOG features extracted using a neighborhood radius of 8.

Proposed Methods	Maximum AUC	Optimal Sensitivity	Optimal Specificity	Classification Accuracy
S-CoHOG	0.867	92%	88%	90.40%
GD-CoHOG	0.918	91%	93%	92.30%
LFDG-CoHOG	0.864	87%	95%	91.00%

For MR dataset 2, the results are shown in Table 4.19. In this case, GD-CoHOG method has a better performance than the other two methods too. GD-CoHOG acquires the highest classification accuracy of 92.30% along with the highest optimal sensitivity of 91% and specificity of 93%.

4.4.5 Region Based Analysis

In this experiment, we perform region based analysis of the subjects of MR dataset 1. We subdivided the downsampled image with an image pixel size of $3 \times 3 \text{ mm}^2$ into 15 equal sized square sub-regions, each of which is of size 10×10 pixels. Now for each sub-region, we apply our proposed S-CoHOG method with a neighborhood size of 8. Selected features using ROC analysis are applied to the SVM classifier. Based on the classification accuracy, we highlight seven regions that have the highest classification accuracy.

Fig. 4.11 shows the highlighted regions of the MRI image that have the highest classification accuracy. The classification accuracy of the corresponding colored box is shown in Fig. 4.11 as well. From the figure, we can see that regions with significant differences between patients and controls correspond to regions most severely affected by ALS, namely the motor cortex and the corticospinal tracts. The top left and right regions contains a little tissue in them. These small regions do not have enough texture information to extract features and thus have the lower classification accuracies.

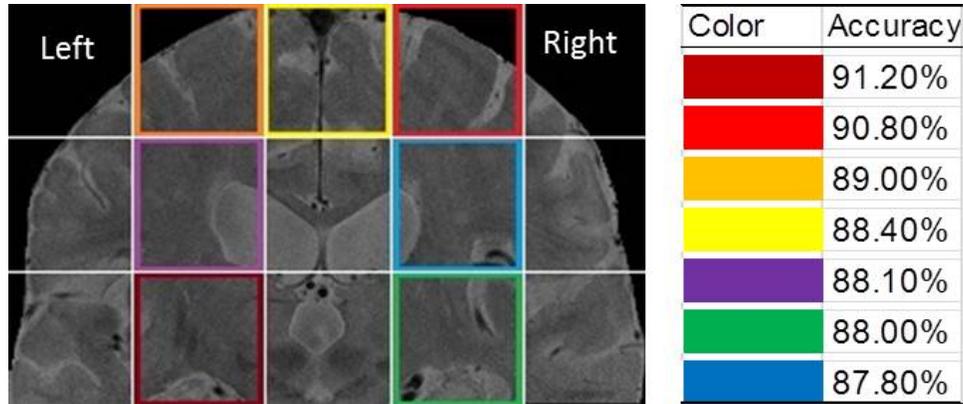


Figure 4.11: Region based analysis of the subjects of MR dataset 1. Significant regions are marked by the colored boxes and classification accuracy of the corresponding boxes.

4.4.6 Comparison with the GLCM Method

We compare the best results of our proposed methods with that of the well-known GLCM method. GLCM has been used in medical image analysis in many applications [25], [26], [27]. We implemented the GLCM method (Gray Labels = 32, Neighbor distance = 1, Neighbor direction = 0^0) in the same environment for MR datasets 1 and 2. In total, 22 features are calculated using well-known feature functions in GLCM (see Table 2.1). Among them only 3 features namely, Angular second moment, Entropy and Sum entropy are selected using ROC feature selection with an AUC threshold. The results are shown in Table 4.20 and Table 4.21 for MR datasets 1 and 2, respectively.

Table 4.20: Comparison of ROC analysis between GD-CoHOG and GLCM methods using MR dataset 1.

Methods	Maximum AUC	Optimal Sensitivity	Optimal Specificity	Classification Accuracy
GD-CoHOG	0.954	98%	96%	97.30%
GLCM	0.601	4%	95%	58.60%

We found the best results in downsampled images with a pixel size of $3 \times 3 \text{ mm}^2$ and of $4 \times 4 \text{ mm}^2$ for MR dataset 1 and 2, respectively. The proposed GD-CoHOG method outperforms the GLCM method for both the MR datasets. For MR dataset

Table 4.21: Comparison of ROC analysis between GD-CoHOG and GLCM methods using MR dataset 2.

Methods	Maximum AUC	Optimal Sensitivity	Optimal Specificity	Classification Accuracy
GD-CoHOG	0.918	91%	93%	92.30%
GLCM	0.805	68%	76%	72.30%

1, we observe that GLCM has very poor performance. In particular, it has a high specificity but a very low sensitivity. The overall classification accuracy is very low compared to that of GD-CoHOG. Using MR dataset 2, GLCM has a better performance than using MR dataset 1, but is still worse than that of GD-CoHOG. MR dataset 1 was acquired using a high resolution 4.7 Tesla MRI system and MR dataset 2 was acquired using a relatively low resolution 1.5 Tesla MRI system. So, we can see from the comparison results that GLCM has a poor performance using high resolution images than that of low resolution images. This is because of its sensitivity to changes in the intensity levels that GLCM uses for features. Such a finding is consistent with the observation that the proposed methods have very similar performance using either MR datasets.

Moreover, we also compare the results of the proposed methods with 3D GLCM that uses 3D texture analysis in ALS [30]. We compare the sensitivity and specificity in the CST using MR dataset 2. 3D GLCM achieves a sensitivity of 90% and specificity of 95% whereas our proposed 2D GD-CoHOG method acquires a sensitivity of 91% and specificity of 93%, which is comparable.

4.4.7 ROC Analysis using Randomly Selected Slices

In this section, we analyze the effect of selecting the wrong slice in the experiment to calculate the classification accuracy. Manual selection may cause error by selecting the wrong slice. In this experiment we randomly choose slices from each subjects to see how it affects the results.

MR Dataset 1 is used in this experiment. We use five slices of each subject.

These five slices include the manually selected slice along with two immediate slices from both right and left side of the manually selected slice. Texture features are calculated for one of the five slices of each subject. The experiment is done using S-CoHOG with a neighborhood size of 8.

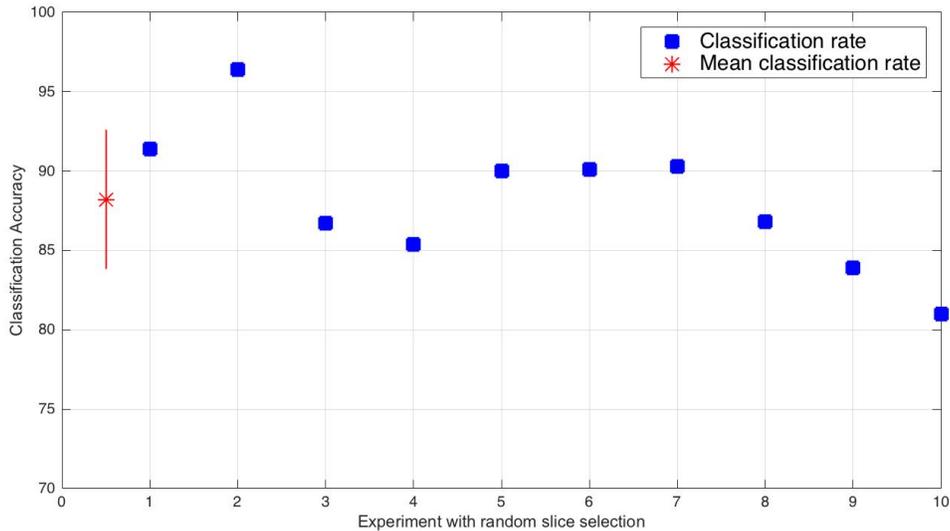


Figure 4.12: Classification accuracies for 10 random slice selection experiments and the mean and the standard deviation of the classification accuracies.

The experiment is performed 10 times where every time we randomly choose a slice for each subject. In each run, the classification accuracy is plotted in Fig. 4.12. The maximum and minimum classification accuracies are 96.40% and 81.00%, respectively. The mean classification accuracy is 88.20% with a standard deviation of 4.35. From the results, we observe that the classification accuracy decreases when we use random slices instead of using the selected slices which is 93.00%. This means that appropriate slice selection is important to get better results.

4.5 ROC Analysis of Multicenter Dataset

Multicenter structural MRI studies can have greater statistical power than single-center studies due to their ability to recruit a greater number of subjects than a single center can. However, across-center differences in contrast, spatial unifor-

mity, etc., may lead to tissue classification or image registration differences that could reduce or wholly offset the enhanced statistical power of multicenter data [53]. Therefore, using multicenter data for classification is still a major challenge due to the use of different scanning parameters as well as the inherent differences in image characteristics arising from different machines used in different centers. Several works have found differences in multicenter data. A multicenter Voxel-Based Morphometry (VBM) study was done using multicenter data in [54]. The same subjects were used in different scanners with the results showing differences in spatial patterns of the results between different scanners. Another VBM based multicenter MRI analysis is done to study reliability in multicenter data [53]. The study was to detect group differences and to estimate heritability when MRI scans from different scanners running different acquisition protocols in a multicenter setup. A study on multicenter data included subjects from three different countries to study gray matter changes with reading disability [55]. VBM analysis showed significant group differences.

In this experiment, we use multicenter data for classification using our proposed methods. We use data from five different centers. T1-weighted MRI scans of the subjects of different centers are performed using different MRI acquisition parameters. One sample image of the patients and controls for each center are shown in Fig. 4.13. The details of the parameters are shown in Table 4.22.

Table 4.22: MRI acquisition parameters for five different centers.

Centers	TR	TE	Voxel Size (mm^2)	Thickness (mm)	Image Size ($pixel^2$)	Scanner
C1	7.4	3.1	1.0×1.0	1	256×256	GE Medical Systems
C2	7.4	3.1	1.0×1.0	1	256×256	GE Medical Systems
C3	7.6	2.9	1.0×1.0	1	256×256	GE Medical Systems
C4	2300	3.4	1.0×1.0	1	256×256	Siemens Medical Systems
C5	2300	3.4	1.0×1.0	1	256×256	Siemens Medical Systems

Data from centers C1, C2 and C3 are acquired using a 3 Tesla GE Medical Systems scanner and data from centers C4 and C5 are acquired using a 3 Tesla Siemens Medical Systems scanner. As data from these two groups are acquired

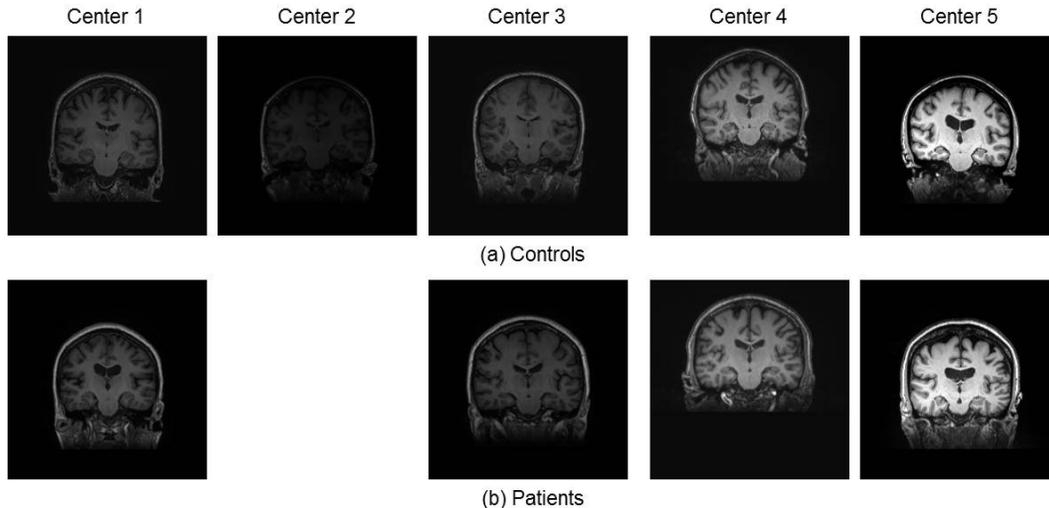


Figure 4.13: Sample image slices of (a) controls and (b) patients of each center from Multicenter dataset. Patients and controls are not distinguishable by visual inspection.

using two different scanners, we combine the subjects of C1, C2 and C3 and formed a multicenter (MC) dataset 1. Similarly, MC dataset 2 is formed using subjects of centers C4 and C5. This datasets are formed to study the scanner specific classification on multicenter data.

The details of the two MC datasets are shown in Table 4.23. MC dataset 1 contains 10 patients and 12 controls and MC dataset 2 contains 19 patients and 13 controls.

Table 4.23: Multicenter (MC) dataset details.

Datasets	Patients	Controls	Centers
MC Dataset 1	10	12	C1, C2, C3
MC Dataset 2	19	13	C4, C5
MC Full Dataset	29	25	C1, C2, C3, C4, C5

The selected 2D MRI images of the subjects are downsampled with an image pixel size of $3 \times 3 \text{ mm}^2$ for both the MC dataset 1 and MC dataset 2. We also form another dataset containing all the subjects from all the centers called MC Full dataset (see Table 4.23). This dataset contains 29 patients and 25 controls. we downsampled all the images to $3 \times 3 \text{ mm}^2$ physical resolution.

Table 4.24: ROC analysis of MC datasets using proposed methods with a neighboring radius of 8 and GLCM.

Datasets	Methods	Maximum AUC	Optimal Sensitivity	Optimal Specificity	Classification Accuracy
MC Dataset 1	S-CoHOG	0.950	81%	99%	91.00%
	LFDG-CoHOG	0.929	87%	98%	93.30%
	GD-CoHOG	0.912	75%	90%	83.70%
	GLCM	0.758	70%	67%	68.80%
MC Dataset 2	S-CoHOG	0.807	87%	87%	87.00%
	LFDG-CoHOG	0.802	88%	72%	81.80%
	GD-CoHOG	0.866	83%	84%	83.50%
	GLCM	0.729	42%	75%	61.70%
MC Full Dataset	S-CoHOG	0.846	86%	83%	85.10%
	LFDG-CoHOG	0.821	81%	78%	80.10%
	GD-CoHOG	0.817	80%	77%	78.60%
	GLCM	0.715	52%	66%	59.30%

For all of the datasets, texture features are extracted using the proposed methods with a neighborhood size of 8. The classification accuracy along with the sensitivity and specificity are shown in Table 4.24 for the three datasets. The results for MC dataset 1 is higher than that of the MC dataset 2. LFDG-CoHOG achieves the best classification accuracy of 93.30% for MC dataset 1. The best classification accuracy in MC Full dataset is achieved by S-CoHOG method. The classification accuracy, sensitivity and specificity of these datasets are comparable to that of using the datasets in section 4.2. Though these datasets have variations in intensities and illuminations, the proposed methods can still differentiate between the patients and controls.

We also compare our results with that of the GLCM method. For all the MC datasets, the proposed methods have much better classification accuracies than that of the GLCM as shown in Table 4.24.

4.5.1 Classification using Different Centers for Training and Testing

In this experiment, the classification accuracy of the proposed S-CoHOG method is analyzed using data from one center for training and data from the other center

for testing. Texture features are extracted using a neighboring radius of 8. We use centers C1 and C3 in one group and centers C4 and C5 in another group for this experiment to perform a scanner specific, between-center classification. Center C2 is not used in this experiment because it contains only two subjects. For a particular group, both centers are used for training and testing separately. The classification results are shown in Table 4.25.

Table 4.25: Classification accuracy using data from one center for training and the other center for testing.

Proposed Method	Train Center	Test Center	Classification Accuracy
S-CoHOG	C1	C3	93.00%
	C3	C1	90.00%
	C4	C5	91.50%
	C5	C4	80.00%

From the results, a maximum of 93.00% in classification accuracy is obtained using C1 data as training and C3 data as testing. A similar classification accuracy is also obtained using C3 data as training and C1 data as testing. Also the C4, C5 data group has good classification accuracy, except when C5 data is used for training. This is because the C5 center contains only few subjects for patients compared to controls.

Chapter 5

Conclusion

5.1 Summary

In this thesis, based on the original CoHOG method, three novel gradient-based methods are proposed. Gradient operators Sobel, GD and LFDG are used in the proposed S-CoHOG, GD-CoHOG and LFDG-CoHOG methods, respectively. For the first time, we apply the proposed methods to the whole image instead of to sub-regions for feature calculation to reduce the sub-region issue problem in the original CoHOG. The original CoHOG method uses a maximum neighborhood size of 4. We also use a larger neighborhood size of 8 for co-occurrence calculation. The extracted feature vector size is very large. Using this large number of similar features creates ambiguity in creation of an optimal hyperplane and leads to an erroneous classification by a classifier. For the first time, we apply the feature selection method on the extracted CoHOG features to select significant features using ROC analysis with a significance level of $p \leq 0.01$ and an AUC threshold. The selected features are used in a linear support vector machine (SVM) classifier to determine the classification accuracy.

Three well-known texture datasets, INRIA Person, CURET and the UIUC are used to evaluate the classification accuracy of the proposed methods. The proposed methods achieve the best classification results using a neighborhood size of 8 with

feature selection. The proposed S-CoHOG and GD-CoHOG methods achieve a maximum classification accuracy of 99.90% for the INRIA Person dataset. A maximum classification accuracy of 98.30% and 98.00% are achieved by GD-CoHOG method for the CURET and the UIUC datasets. The classification results of the proposed methods are compared with that of the original CoHOG method. The classification results show that the proposed methods achieve the best classification results on all the datasets that outperform that of the original CoHOG method.

Three different datasets of 2D MRI are used for classification. Each dataset has a different image resolution and contrast. MRI imaging of ALS patients and controls are classified using the proposed methods. To the best of our knowledge, we are the first to use the CoHOG-like methods to study cerebral degeneration in ALS. A multicenter ALS dataset with images having the same resolution but different contrasts is also used to demonstrate the classification performance of the proposed methods. The experimental results demonstrate that our methods have promising classification abilities with high sensitivity and specificity. In particular, the GD-CoHOG method achieves the maximum classification accuracy of 97.30% for MR dataset 1. For these datasets, we compare the results of the proposed methods with that of the GLCM method. The classification results show that the proposed methods outperform that of the GLCM method. Also the sensitivity and specificity of the proposed methods have higher than that of the GLCM method. Region based analysis is also performed and the result shows that areas most responsible for significant differences between the patients and controls are congruent with the spatial distribution of the pathology of ALS. For the multicenter dataset, classification is done using data from one center for training and that from the other center for testing. This experiment is done to address the issue when there is a lack of subjects in a center and whether data from another center can be used for training. The classification accuracy is promising for such a multicenter setting.

In summary the proposed CoHOG based methods show excellent classification accuracy in different texture datasets. As well, the proposed methods show excel-

lent classification accuracy in ALS datasets of different contrasts (T1 and T2) and data collected from different MRI machines. Thus, the proposed methods using texture show promise as a potential method to identify ALS. Future research using the proposed methods in a multicenter setting is much warranted in addition to determining the ability of the method to monitor disease progression.

5.2 Contributions

The main contributions of this thesis include:

1. The proposed three methods use the whole image instead of subdividing it into sub-regions. The use of sub-regions limits the accuracy of co-occurrence matrix for boundary pixels and thus some information is incomplete for each sub-region. Also it increases the size of the feature vector. Thus, using the whole image not only reduces the boundary pixels problem in sub-regions but also reduces the size of feature vector.
2. For the first time, we adopt two gradient operators GD and LFDG for the proposed GD-CoHOG and LFDG-CoHOG methods, respectively. The proposed methods are compared to see the impact of the gradient operators on classification accuracy using the whole image.
3. Texture features are extracted using two different neighborhood sizes. The original CoHOG method uses a maximum neighborhood size of 4. We use a larger neighborhood size of 8 to see the effect of using more spatial information for co-occurrence calculation on classification accuracy. Indeed, the experimental results confirm our expectation that a better classification accuracy is achieved using a neighborhood size of 8.
4. The extracted feature vector size using the CoHOG method is large with many similar features. Changes that occur in a small portion of the images

between two classes produce a large number of similar features. Using this large number of similar features not only creates ambiguity in creation of an optimal hyperplane but also leads to the wrong classification for a classifier. We are the first to use a feature selection method to extract significant CoHOG texture features using area under the ROC curve (AUC) analysis for classification. Only features that contain significant differences between classes are selected using an AUC threshold. Experimental results show that classification using feature selection has a better accuracy than that without using the feature selection.

5. Three different datasets of 2D Magnetic Resonance Images (MRI) are used for the first time in CoHOG-like methods for classification. Each dataset has different image resolutions and contrasts. MRI imaging data of Amyotrophic Lateral Sclerosis (ALS) patients and controls is used for classification. The experimental results show excellent classification accuracy with high sensitivity and high specificity using the proposed methods.

Another multicenter ALS dataset of different image contrast is also used in this experiment to demonstrate the classification performance of the proposed methods. We found comparable classification accuracy for this dataset, though multicenter data classification is a challenging task due to its variation in imaging parameters and qualities.

5.3 Future Work

The work in this thesis encourages future research in the following directions:

1. The proposed CoHOG based methods are applied to other areas of image classification and to study cerebral degeneration in ALS more extensively. Other areas that are applicable include document processing, remote sensing, automated inspection, fingerprint recognition, etc.

2. The proposed 2D CoHOG based methods can be extended to 3D methods to extract features from a 3D image. The 3D methods can be used to extract features from 3D MRI scans of the brain. For a 3D method we have to use a spherical neighborhood rather than a circular neighborhood that we have used for the proposed 2D CoHOG method. Also 3D gradient operators are needed for the calculation of gradient orientations of the 3D image.

The above are some of the many interesting problems in which the proposed methods may be useful.

Bibliography

- [1] Ming Zhao, Shutao Li, and James Kwok. Text detection in images using sparse representation with discriminative dictionaries. *Image and Vision Computing*, 28(12):1590–1599, 2010.
- [2] Jefersson Alex dos Santos, Philippe-Henri Gosselin, Sylvie Philipp-Foliguet, Ricardo da S Torres, and Alexandre Xavier Falao. Multiscale classification of remote sensing images. *Geoscience and Remote Sensing, IEEE Transactions on*, 50(10):3764–3775, 2012.
- [3] Wei-Chen Li and Du-Ming Tsai. Wavelet-based defect detection in solar wafer images with inhomogeneous texture. *Pattern Recognition*, 45(2):742–756, 2012.
- [4] Loris Nanni and Alessandra Lumini. Local binary patterns for a hybrid fingerprint matcher. *Pattern recognition*, 41(11):3461–3466, 2008.
- [5] A Kassner and RE Thornhill. Texture analysis: a review of neurologic mr imaging applications. *American Journal of Neuroradiology*, 31(5):809–816, 2010.
- [6] Parveen Lehana, Swapna Devi, Satnam Singh, Pawanesh Abrol, Saleem Khan, and Sandeep Arya. Investigations of the mri images using aura transformation. *Signal & Image Processing*, 3(1):95, 2012.
- [7] Timo Ojala, Matti Pietikäinen, and Topi Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7):971–987, 2002.

- [8] Robert M Haralick, Karthikeyan Shanmugam, and Its' Hak Dinstein. Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, (6):610–621, 1973.
- [9] Xuejie Qin and Yee-Hong Yang. Similarity measure and learning with gray level aura matrices (glam) for texture image retrieval. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 1, pages I–326. IEEE, 2004.
- [10] Mary M Galloway. Texture analysis using gray level run lengths. *Computer graphics and image processing*, 4(2):172–179, 1975.
- [11] Xinqi Chu and Kap Luk Chan. Rotation and scale invariant texture analysis with tunable gabor filter banks. In *Advances in Image and Video Technology*, pages 83–93. Springer, 2009.
- [12] Thomas Leung and Jitendra Malik. Representing and recognizing the visual appearance of materials using three-dimensional textons. *International journal of computer vision*, 43(1):29–44, 2001.
- [13] Manik Varma and Andrew Zisserman. A statistical approach to texture classification from single images. *International Journal of Computer Vision*, 62(1-2):61–81, 2005.
- [14] Michael Unser and Murray Eden. Multiresolution feature extraction and selection for texture segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 11(7):717–728, 1989.
- [15] Constantino Carlos Reyes-Aldasoro and Abhir Bhalerao. The bhattacharyya space for feature selection and its application to texture segmentation. *Pattern Recognition*, 39(5):812–826, 2006.
- [16] George R Cross and Anil K Jain. Markov random field texture models. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (1):25–39, 1983.

- [17] Fernand S. Cohen, Zhigang Fan, and Maqbool A Patel. Classification of rotated and scaled textured images using gaussian markov random field models. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (2):192–202, 1991.
- [18] Svetlana Lazebnik, Cordelia Schmid, and Jean Ponce. A sparse texture representation using local affine regions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(8):1265–1278, 2005.
- [19] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.
- [20] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893. IEEE, 2005.
- [21] Tomoki Watanabe, Satoshi Ito, and Kentaro Yokoi. Co-occurrence histograms of oriented gradients for pedestrian detection. In *Advances in Image and Video Technology*, pages 37–47. Springer, 2009.
- [22] Hirokatsu Kataoka, Kiyoshi Hashimoto, Kenji Iwata, Yutaka Satoh, Nassir Navab, Slobodan Ilic, and Yoshimitsu Aoki. Extended co-occurrence hog with dense trajectories for fine-grained activity recognition. In *Computer Vision–ACCV 2014*, pages 336–349. Springer, 2014.
- [23] Kazim Hanbay, Nuh Alpaslan, Muhammed Fatih Talu, Davut Hanbay, Ali Karci, and Adnan Fatih Kocamaz. Continuous rotation invariant features for gradient-based texture classification. *Computer Vision and Image Understanding*, 132:87–101, 2015.
- [24] Thanh-Toan Do and Ewa Kijak. Face recognition using co-occurrence histograms of oriented gradients. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 1301–1304. IEEE, 2012.

- [25] TR Sivapriya, V Saravanan, and P Ranjit Jeba Thangaiah. Texture analysis of brain mri and classification with bpn for the diagnosis of dementia. In *Trends in Computer Science, Engineering and Information Technology*, pages 553–563. Springer, 2011.
- [26] Xin Li, Hong Xia, Zhen Zhou, and Longzheng Tong. 3d texture analysis of hippocampus based on mr images in patients with alzheimer disease and mild cognitive impairment. In *Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on*, volume 1, pages 1–4. IEEE, 2010.
- [27] Ahmed Kharrat, Nacéra Benamrane, Mohamed Ben Messaoud, and Mohamed Abid. Detection of brain tumor in medical images. In *Signals, Circuits and Systems (SCS), 2009 3rd International Conference on*, pages 1–6. IEEE, 2009.
- [28] Kouros Jafari-Khouzani, Mohammad-Reza Siadat, Hamid Soltanian-Zadeh, and Kost Elisevich. Texture analysis of hippocampus for epilepsy. In *Medical Imaging 2003*, pages 279–288. International Society for Optics and Photonics, 2003.
- [29] Milena Albuquerque, Lara GV Anjos, Helen Maia Tavares de Andrade, Márcia S Oliveira, Gabriela Castellano, Thiago Junqueira Ribeiro de Rezende, Anamarli Nucci, and Marcondes Cavalcante França. Mri texture analysis reveals deep gray nuclei damage in amyotrophic lateral sclerosis. *Journal of Neuroimaging*, 2015.
- [30] Rouzbeh Maani, Yee-Hong Yang, Derek Emery, and Sanjay Kalra. Cerebral degeneration in amyotrophic lateral sclerosis revealed by 3-dimensional texture analysis. *Frontiers in Neuroscience*, 10, 2016.
- [31] Richard O Duda, Peter E Hart, et al. *Pattern classification and scene analysis*, volume 3. Wiley New York, 1973.
- [32] Rouzbeh Maani, Sanjay Kalra, and Yee-Hong Yang. Robust volumetric texture

- classification of magnetic resonance images of the brain using local frequency descriptor. *Image Processing, IEEE Transactions on*, 23(10):4625–4636, 2014.
- [33] Rouzbeh Maani, Sanjay Kalra, and Yee-Hong Yang. Rotation invariant local frequency descriptors for texture classification. *Image Processing, IEEE Transactions on*, 22(6):2409–2419, 2013.
- [34] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.
- [35] Yang Zhao, Rong-Gang Wang, Wen-Min Wang, and Wen Gao. Local quantization code histogram for texture classification. *Neurocomputing*, 2016.
- [36] Yang Zhao, Rong-Gang Wang, Wen-Min Wang, and Wen Gao. Local quantization code histogram for texture classification. *Neurocomputing*, 2016.
- [37] Jun Zhang, Heng Zhao, and Jimin Liang. Continuous rotation invariant local descriptors for texton dictionary-based texture classification. *Computer Vision and Image Understanding*, 117(1):56–75, 2013.
- [38] Antonio J Serrano, Emilio Soria-Olivas, José David Martín-Guerrero, Rafael Magdalena, and Juan Gomez. Feature selection using roc curves on classification problems. In *IJCNN*, pages 1–6, 2010.
- [39] Shuangge Ma and Jian Huang. Regularized roc method for disease classification and biomarker selection with microarray data. *Bioinformatics*, 21(24):4356–4362, 2005.
- [40] D Lorente, J Blasco, AJ Serrano, E Soria-Olivas, N Aleixos, and J Gómez-Sanchis. Comparison of roc feature selection method for the detection of decay in citrus fruit using hyperspectral images. *Food and Bioprocess Technology*, 6(12):3613–3619, 2013.

- [41] Malak Alshawabkeh, Javed A Aslam, Jennifer Dy, and David Kaeli. Feature selection metric using auc margin for small samples and imbalanced data classification problems. In *Machine Learning and Applications and Workshops (ICMLA), 2011 10th International Conference on*, volume 1, pages 145–150. IEEE, 2011.
- [42] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [43] Kristin J Dana, Bram Van Ginneken, Shree K Nayar, and Jan J Koenderink. Reflectance and texture of real-world surfaces. *ACM Transactions on Graphics (TOG)*, 18(1):1–34, 1999. The Dataset is available at <http://www.cs.columbia.edu/CAVE/software/curet/>.
- [44] Svetlana Lazebnik, Cordelia Schmid, and Jean Ponce. A sparse texture representation using local affine regions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1265–1278, 2005. The Dataset is available at http://www-cvr.ai.uiuc.edu/ponce_grp/data/.
- [45] Benjamin Rix Brooks, Robert G Miller, Michael Swash, and Theodore L Munsat. El escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotrophic lateral sclerosis and other motor neuron disorders*, 1(5):293–299, 2000.
- [46] Matthew C Kiernan, Steve Vucic, Benjamin C Cheah, Martin R Turner, Andrew Eisen, Orla Hardiman, James R Burrell, and Margaret C Zoing. Amyotrophic lateral sclerosis. *The Lancet*, 377(9769):942–955, 2011.
- [47] F Agosta, E Pagani, MA Rocca, D Caputo, M Perini, F Salvi, A Prella, and M Filippi. Voxel-based morphometry study of brain volumetry and diffusivity

- in amyotrophic lateral sclerosis patients with mild disability. *Human brain mapping*, 28(12):1430–1438, 2007.
- [48] Julian Grosskreutz, Jörn Kaufmann, Julia Frädrieh, Reinhard Dengler, Hans-Jochen Heinze, and Thomas Peschel. Widespread sensorimotor and frontal cortical atrophy in amyotrophic lateral sclerosis. *BMC neurology*, 6(1):1, 2006.
- [49] DM Mezzapesa, A Ceccarelli, F Dicuonzo, A Carella, MF De Caro, M Lopez, V Samarelli, P Livrea, and IL Simone. Whole-brain and regional brain atrophy in amyotrophic lateral sclerosis. *American Journal of Neuroradiology*, 28(2):255–259, 2007.
- [50] Jack L. Lancaster and Michael J. Martinez. *Mango Software*. URL <http://rii.uthscsa.edu/mango/mango.html>.
- [51] Paul A. Yushkevich, Joseph Piven, Heather Cody Hazlett, Rachel Gimpel Smith, Sean Ho, James C. Gee, and Guido Gerig. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*, 31(3):1116–1128, 2006.
- [52] Caroline A Schneider, Wayne S Rasband, Kevin W Eliceiri, et al. Nih image to imagej: 25 years of image analysis. *Nat methods*, 9(7):671–675, 2012.
- [53] Hugo G Schnack, Neeltje EM van Haren, Rachel M Brouwer, G Caroline M van Baal, Marco Picchioni, Matthias Weisbrod, Heinrich Sauer, Tyrone D Cannon, Matti Huttunen, Claude Lepage, et al. Mapping reliability in multicenter mri: Voxel-based morphometry and cortical thickness. *Human brain mapping*, 31(12):1967–1982, 2010.
- [54] Niels K Focke, Gunther Helms, Susanne Kaspar, Christine Diederich, Vera Tóth, Peter Dechent, Alexander Mohr, and Walter Paulus. Multi-site voxel-based morphometry not quite there yet. *Neuroimage*, 56(3):1164–1170, 2011.

- [55] Katarzyna Jednorog, Artur Marchewka, Irene Altarelli, Ana Karla Monzalvo Lopez, Muna van Ermingen-Marbach, Marion Grande, Anna Grabowska, Stefan Heim, and Franck Ramus. How reliable are gray matter disruptions in specific reading disability across multiple countries and languages? insights from a large-scale voxel-based morphometry study. *Human brain mapping*, 36(5):1741–1754, 2015.