REPUBLIC OF TÜRKİYE

ALTINBAŞ UNIVERSITY Institute

of Graduate Studies Electrical and

Computer Engineering

SKIN CANCER DETECTION AND CLASSIFICATION USING DEEP LEARNING

Alzahraa Yahya Haider HAIDER

Master's Thesis

Supervisor

Prof. Dr. Galip CANSEVER

İstanbul, 2024

SKIN CANCER DETECTION AND CLASSIFICATION USING DEEP LEARNING

Alzahraa Yahya Haider HAIDER

Electrical and Computer Engineering

Master's Thesis

ALTINBAŞ UNIVERSITY

2024

The thesis/dissertation titled (SKIN CANCER DETECTION AND CLASSIFICATION USING DEEP LEARNING) prepared by (ALZAHRAA YAHYA HAIDER HAIDER) and submitted on (2024) has been **accepted unanimously/by majority of votes** for the degree of Master of Arts in Electrical and Computer Engineering.

Academic Title First and LAST NAME of Co - Supervisor	Academic Title First and LAST NAME of Supervisor

Thesis Defense Committee Members:

Academic Title First LAST NAME	Department of …, University	
Academic Title First LAST NAME	Department of …, University	
Academic Title First LAST NAME	Department of …, University	
Academic Title First LAST NAME	Department of …, University	
Academic Title First LAST NAME	Department of …, University	

I hereby declare that this thesis/dissertation meets all format and submission requirements for a …… (Master's/PhD) thesis/dissertation.

Submission date of the thesis/dissertation to Institute of Graduate Studies: ___/___/___

I hereby declare that all information/data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct.

Alzahraa Yahya Haider HAIDER

Signature

DEDICATION

I thank God Almighty first and foremost for the great grace that He has bestowed upon me, then I thank those who favored them. My beloved parents do not cease to me for all their efforts from the moment of my birth to these moments. For everyone who advised me, guided me, contributed, or directed me in preparing this research and connecting me to the required references and sources at any of the stages it went through, and I especially thank the distinguished professor for helping me Supporting me, and guiding me with advice, education, correction, and all that he did with me.

In the end, I want to say that this thesis is dedicated to the souls of my father, my dear brother & his wife, and my sisters. From those who supported me during my studies and from whom I learned knowledge, diligence, and perseverance to achieve success in all aspects of life.

PREFACE

At the outset, I would like to thank my supervisor, Prof. Dr. Galip CANSEVER, for helping me while writing the thesis, and answering my inquiries regarding the thesis. his suggestions, directives, and suggestions as well as continuous encouragement to me, I thank my family and friends.

ABSTRACT

SKIN CANCER DETECTION AND CLASSIFICATION USING DEEP LEARNING

Yahya Haider HAIDER, ALZAHRAA
M.Sc., Electrical and Computer Engineering, Altınbaş University,
Supervisor: Prof. Dr. Galip CANSEVER
Date: Januery / 2024
Pages: 65

One of the worst malignancies is skin cancer. It is expected to spread to other parts of the body if it is not identified and treated at the outset. The method for successfully treating skin cancer uses both image processing and deep learning. In this research, three different kinds of skin cancer, including melanoma, pigmented benign keratoses, and basal cell carcinoma are introduced for detection and classification using efficient methods in Machine Learning (ML) such as K-Mean clustering and Multi-class support vector machine (M-SVM) algorithm, and then Deep Learning (DL) techniques such as AlexNet and ResNet are used. Additionally, deep convolutional neural networks (CNN) effectiveness and capacity are seen. The data set contains 1176 images of skin cancer with different classes of disease, this data set is used in both ML and DL. In ML the data set is divided into training and testing sets, these sets are pass through a few steps of enhancement using a canny edge detection filter and extracting the features using the Gray-Level Co-occurrence Matrix (GLCM) method. Then, these images are segmented into three clustering using K-mean clustering algorithms. In DL two models are used AlexNet and ResNet, in these models the data sets are divided into training and testing sets. Then, an augmentation technique has been proposed, it's very useful for small data sets, and the results show changes in accuracy result. The results show an accuracy of 99.03%, and 97.02% for AlexNet, and ResNet respectively.

Keywords: ML, DL, M-SVM, CNN, GLCM, AlexNet, ResNet.

TABLE OF CONTENTS

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

LIST OF CHARTS

Pages

ABBREVIATIONS

CV	:	Computer Vision
AI	:	Artificial Intelligence
ML	:	Machine Learning
DL	:	Deep Learning
CNN	:	Convolutional Neural Network
BBC	:	Basal Cell Carcinoma
AIP	:	Analog Image Processing
DIP	:	Digital Image Processing
PCA	:	Principal Component Analysis
GLCM	:	Gray Level Co-Occurrence Matrix
ANN	:	Artificial Neural Network
CNN	:	Video Processing
ReLU	:	Rectified Linear Unit
ROI	:	Area Of Interest
M-SVM	:	Multiclass Support Vector Machine
GUI	:	Graphical User Interface

1. INTRODUCTION

1.1 SKIN CANCER DISEASES

The most common disease in the world has historically been skin cancer. In the following decades, both non-melanoma and melanoma skin cancer incidence have increased. According to the World Health Organization (WHO), skin cancer can be found in one of every three cancer cases, and one in five Americans will develop skin cancer at some point in their lives, according to statistics from the Skin Cancer Foundation. Skin conditions appear to have a significant negative impact on population health worldwide, particularly in countries like the United States. According to a 2017 study, skin cancer accounts for 1.79 percent of the global illness burden measured in years of life with a handicap. [1]. Skin cancer accounts for about 7% of all new cancer cases worldwide [2], costing the US Medicare project over eight billion dollars in 2011. Clinical data support racial disparities in skin cancer outcomes: those with darker skin are about 20–30% less likely to acquire melanoma than people with lighter skin, and they have a reduced or greater chance of dying for certain melanoma forms.

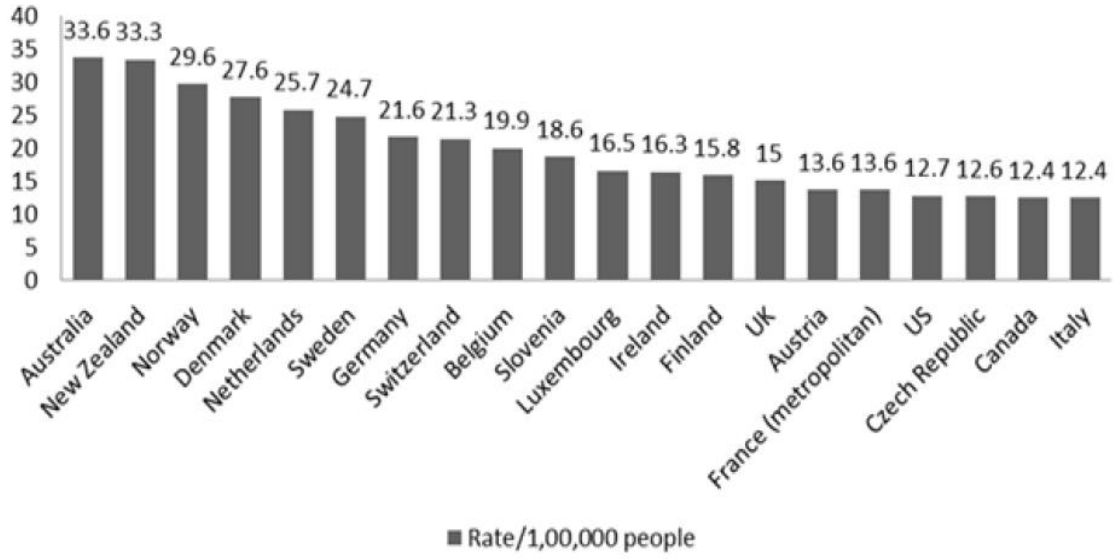


Figure 1.1: According to the World Cancer Research Foundation, skin cancer rates vary by nation.

1.2 COMPUTER VISION

Computer vision (CV) is a branch of modern computing science and a type of artificial intelligence that aims to give machines the ability to see and aids in seeing the outside world,

analyzing visual data, and then drawing conclusions from it, as well as identifying and processing objects like images and videos in a way that is comparable to how humans do.
The huge number of digital data available today is one of the primary forces influencing CV development because, as the field of CV has advanced and become more compatible with contemporary devices thanks to algorithms and recent advancements in DL and neural networks, the accuracy rates in object identification and object analysis have increased. This allows for the extraction of symbolic information from the images using models created with assistance from engineering, physics, statistics, and learning theory.
Autonomous vehicles, Google Translate, face recognition, healthcare, sports tracking, industry, and agriculture are a few examples of CV uses. [3].

1.3 MACHINE LEARNING

A form of artificial intelligence (AI), machine learning (ML) aims to provide computers the ability to carry out their tasks skillfully utilizing intelligent software [4]. The four categories of machine learning (ML) are reinforcement learning when there is no data, supervised learning when there is classified (labeled) data, unsupervised learning when there is unclassified (unlabeled) data, and semi-supervised learning when there is a mixture of classified and unclassified data [5].

1.4 DEEP LEARNING

One kind of DL is ML, which uses deep neural networks with many hidden layers. It is one of the most recent models studied in ML and AI [6]. DL is one of the most significant methods of detection and classification used today. The fundamental DL tool is the convolutional neural network (CNN), which improves accuracy by utilizing an ML model with programming a large amount of input for extracting features and several hidden layers [7]. In the area of machine learning and pattern recognition, deep learning is an exciting topic whose effectiveness has been shown in a variety of applications. DL is only an ML method that classifies images, text, or even sound by learning from examples. In the great majority of its applied assignments, DL can outflank opponents. Deep learning is based on applying hierarchical methodologies with deep layers, which is different from previous machine learning approaches. On any image recognition task, DL techniques perform better than any other technique. In all prior scientific publications, convolutional neural is regarded

as the most well-liked effective DL strategy [8]. Convolutional layers, a pooling layer, Fully Connected layers, and the SoftMax layer make up the conventional CNN model. Various CNN designs, including AlexNet, ResNet, GoogleNet, VGG, and others, are used for classification.[9][10].

AlexNet is a CNN model created by Alex Krizhevsky[11], and it has been used for many different benchmarks across a wide range of industries. This model architecture is therefore used in a wide range of picture categorization experiments. It is a deep CNN with 25 layers, five convolutional layers, and two fully connected layers. The first convolutional layer contains 96 different (11x11) filters, while the max-pooling layer uses (3x3) filters in the lower layers while using (5x5) filters in the upper layers. The third, fourth, and fifth layers also use (3x3) filters. The size of the image is (227x227) [8]. ResNet is a deep CNN that uses a specially created residual structure to create a very deep network. Traditional deep CNN is limited in its depth and loses accuracy as depth grows. ResNet's author surmises that identity mapping is challenging to learn. The deep residual learning framework is suggested as a solution to this issue, and the network learns the residual rather than the direct mapping. The winning model in the 2015 ImageNet contest was this one. The input image is (224x224) in size [12].

1.5 PROBLEM STATEMENT

Although deep learning outperforms conventional ML algorithms and allows for autonomous feature extraction and classification, the main hurdles are data availability and legal considerations. Deep learning works better with a large amount of data, but collecting large amounts of medical data is currently difficult. The main challenges are a lack of data and having to deal with legal issues about protecting the confidentiality of patients. For example, if we are constructing an algorithm to detect skin cancer, we would need hundreds or millions of CT scans from clients with their permission; As a result, the algorithm must handle the confidentiality of patients and data security. Data shortages, private data confidentiality, and other legal constraints limit deep learning model training.

To solve the issue of data shortage, many data augmentation strategies are available. Many ways for increasing data volume and quality to help in deep learning algorithm training are included in these augmentation strategies. To magnify samples, medical imaging technologies that have been available for a long use transformation such as scaling, rotation,

flipping, and translation. These adjustments are effective up to a limit, beyond which they could limit deep learning's capacity to generalize to fresh data. One approach is to use traditional augmentation procedures carefully before mixing them with other augmentation approaches. Another exciting method is the use of generative models for image enhancement.

Additionally, an ensemble method, in which a single CNN is substituted by an ensemble of numerous CNNs for final classification, can be used to improve the performance of a deep learning model on limited data. Ensemble techniques can enhance performance by 1% to 5% despite the large computing hardware requirements. Researchers are applying ensemble techniques in CNN to enhance classifier performance recently due to the availability of complicated hardware.

This thesis tries to accomplish the intended goal of an effective deep-learning model that can give superior classification results on a small training data set by fusing augmentation approaches with an ensemble of CNNs.

1.6 OBJECTIVES OF THIS MESSAGE

a. Design identification and classification model for skin cancer using deep learning technique.
b. Achieve the highest accuracy by selecting the best method of enhancement, segmentation, and classification of images. Also, decreasing the training and testing time for the proposed model.
c. Design a system that helps to result in high-accuracy decisions in the case of small databases for learning and testing.
d. Calculating the affected area of the skin cancer disease after classifying their type.

1.7 LITERATURE REVIEW

In this section, we will offer studies from prior years on the same subject that we will suggest a solution to, which is the challenge of skin cancer detection and classification. Where you will offer an overview of other researchers' findings and approaches for detecting and classifying skin cancer.

CNNs have already been widely employed in the classification of medical images, among

other applications [13]. It has also produced some enlightening results, including the identification of skin cancers [16] [17] [18] [19] and clinical epidermal 2 cells [14], diabetic retinopathy fundus [15], cervical cell image classification, etc.

Yuet al. [20] created an extremely deep CNN and several kinds of learning frameworks with little training data. Esteva et al. [21] developed and acquired a dermatologist-level diagnosis from over 120.000 photos using a pre-trained CNN approach. Dermatologists have provided CNN models that have been more effective or superior to them, according to Haenssle et al. [22]. Deep learning was utilized to generate other strategies, like the ensemble model [23], which combines characteristics from many models, to identify skin cancer [24] [25].

Several approaches for segmenting and classifying skin cancer have been proposed throughout the last few decades. In the International Skin Imaging Challenge (ISIC) 2016, various methods for feature extraction and classification of skin cancer were proposed. The classification's conclusions were based on the discovery of just two cancer types: benign and aggressive [26]. Numerous methodologies were proposed and demonstrate excellent recognition performance using various DCNNs and Support Vector Machine (LSVM) approaches [26].

The same year, a Complete Convolutional Residual Network (FCRN) was created to apply it to the classification of certain lesions, and it was introduced to the ISBI (2016) dataset [26]. Following the separation of the lesion areas from the input images, a Deep Residual Network (DRN) was employed to distinguish between melanoma and non-melanoma lesions. In contrast, many DCNN techniques, such as VGG-16, GoogleNet, FCRN-38, RCRN-50, and FCRN-101 [26], have been tested. The ISIC 2016 skin lesion segmentation datasets, which yield the greatest results of 61.3% and 69.3% test accuracy, respectively, were used to estimate the effectiveness of VGG-16 and Inception-v3. [27].

However, an updated and contemporary architecture known as "U-Net" was proposed in 2015, specifically for medical image segmentation operations. From that point on, U-Net spread rapidly and was successfully applied in many computerized pathology and medical imaging modalities. U-Net produces trustworthy segmentation data while allowing for a reduced training sample size [28]. R2U-Net, sometimes referred to as the Recurrent Residual Convolutional (RRN) U-Net, was first introduced in 2018 [29]. A retinal blood vessel, ISIC 2017 skin cancer categorization, and lung classification datasets were used to evaluate an updated version of the U-Net model for clinical picture classification.

The integrated diagnostic methodology suggested by AI-Masni et al. [71] encompasses both the skin lesion border segmentation step and the multiple skin lesion categorization stage. A full-resolution convolutional network was used to accurately delineate the borders of skin lesions from whole dermoscopic pictures. Furthermore, Inception-v3, ResNet-50, Inception-ResNet-v2, and DenseNet-201 architectures were utilized to categorize the segmented skin lesions. The accuracy of each classifier model in terms of classification performance was 88.05%, 89.28%, 87.74%, and 88.70%, respectively. The final results were achieved by aggregating the accuracy scores of these classifiers. Additionally, both the training and inference procedures are characterized by a significant expenditure of time.

Barata et al. [72] used a hierarchical classification approach to categorize skin lesions into two main groups: melanocytic and nonmelanocytic lesions. These groups were then further subdivided into benign or malignant lesions based on the level of malignancy shown. The classification methodology used in this study has a resemblance to the approach often utilized for binary tree issues. The method of deriving a hierarchical diagnosis from pictures that underwent feature extraction using the image encoder included the use of a long short-term memory (LSTM) network for the image decoder. The ultimate level of accuracy achieved was 78.0%, with a classification accuracy of 92.2% specifically for distinguishing between melanocytic and nonmelanocytic tumors. In this work, two models were used as the image encoder, while a decoder based on RNN-based LSTM was utilized. The LSTM model requires a lengthier duration for both training and inference compared to models based on CNNs, mostly because of its recurrent design.

1.8 THESIS LAYOUT

Six chapters comprise the thesis report. The first chapter contains an introduction to the issue that explains the problem that inspired us to conduct the study and the objective that we hoped to attain via the subsequent phases of this research. The second chapter provides a detailed explanation of skin cancer diseases and methods of image processing and also includes the main theory of ML and DL with different CNN categories. The third chapter describes the proposed methodology and its implementation using the MATLAB program. The fourth chapter is the practical model, illustrates the experimental result, and analysis, and discusses the measurement and tests that are explained in chapter three. The fifth chapter

comprises the conclusions and suggestions for future works. The sixth chapter includes the references used in writing the research.

2. METHODOLOGY

2.1 INTRODUCTION

The basic concepts of image processing using ML, such as the enhancement technique, segmentation method, and classification method, as well as DL models like AlexNet and ResNet models, are covered in this chapter, along with an overview of popular skin cancer diseases.

2.2 COMMON SKIN CANCER DISEASES

With incidence steadily rising throughout the world, skin cancer is one of the most prevalent and deadly cancer kinds. It can cause metastases and have significant death rates if it is not discovered in the early stages. If discovered early enough, skin cancer is curable. Consequently, an important research goal at the moment is the prompt and precise detection of such tumors.

Skin cancers develop when abnormal cells proliferate, which may invade or spread to other parts of the body depending on their kind and aggressiveness. Our skin is the organ that is most sensitive to environmental poisons since it is the most exposed to them. Thus, skin cancer is the most prevalent type of cancer.

Below is an explanation of several types of diseases affecting skin cancer:

2.2.1 Melanoma

Melanoma is a tumor that typically affects the skin, eyes, nerve centers, and meninges. It develops in melanin-producing cells. Despite having the greatest rate of growing incidence of any kind of skin cancer, melanoma has a higher chance of being curable if detected early [30]. According to one study, when the initial stages of skin cancer are discovered, the chance of death is reduced by 90% [31]. Those with stage I cancer, for example, have an average 94% to 98% chance of surviving after 10 years, whereas those with stage IV illness have an estimated 10% to 15% chance of survival after 10 years. Melanoma is more common in certain groups than others. These high-risk scenarios can be avoided by being aware of these groups [32].

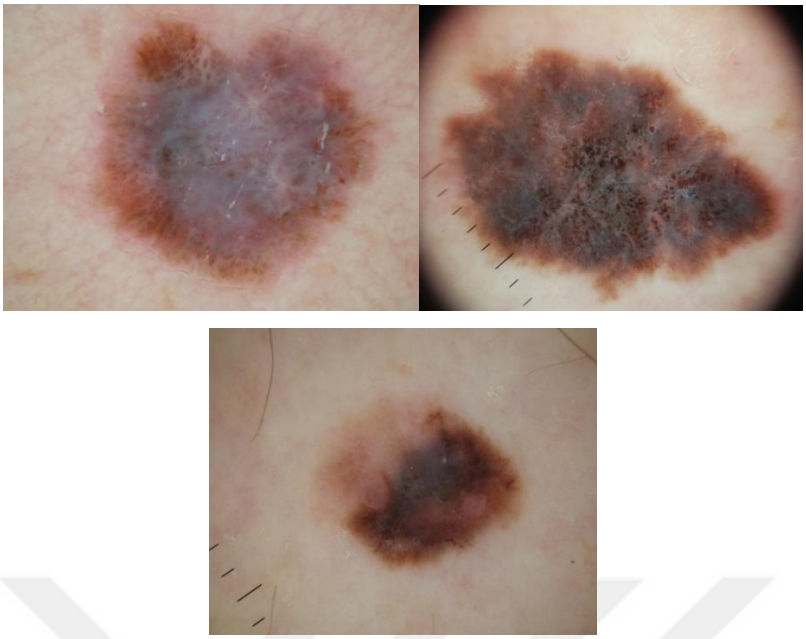


Figure 2.1: Melanoma Disease

2.2.2 Basal Cell Carcinoma

Basal cell carcinoma (BCC) is both affected by and the source of this form of cancer. Keratinocytes, which can be found in the epidermis, are the source of basal cell carcinoma. These might penetrate all of the epidermal thickness [33]. With more than 4 million cases identified each year in the United States, basal cell carcinoma (BCC) is the most prevalent kind of skin cancer [34]. Even though BCC seldom metastasizes, when it is advanced, it can be incredibly disfiguring and damaging to the underlying tissue. Advanced BCC can significantly harm a patient's physical health and burden them psychologically because of the apparent deformity. Additionally, the health care system must pay substantial surgical expenditures for the treatment of advanced BCC [35].

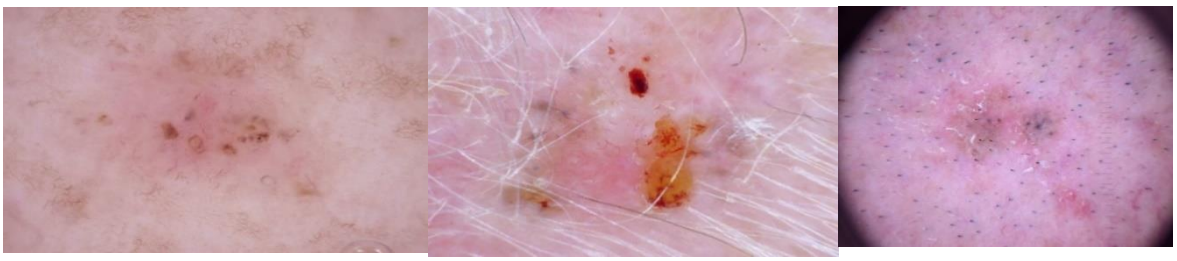


Figure 2.2: Basal Cell Carcinoma Disease

2.2.3 Pigmented Benign Keratosis

One of the most prevalent epidermal tumors, pigmented benign keratosis, is commonly observed in clinical practice. The head and neck area is where it most frequently manifests itself, followed by the body, upper, and lower extremities [36]. In a comprehensive, retrospective investigation that included 639 patients, 85 of the cases had additional lesions in addition to pigmented benign keratosis. Premalignant lesions, malignancies, melanocytic lesions, and other lesions were among these concurrent lesions. Haemangiomas, fibroepithelial polyps, chondrodermatitis nodularis helices, and keratoacanthoma were among the several lesions [37].

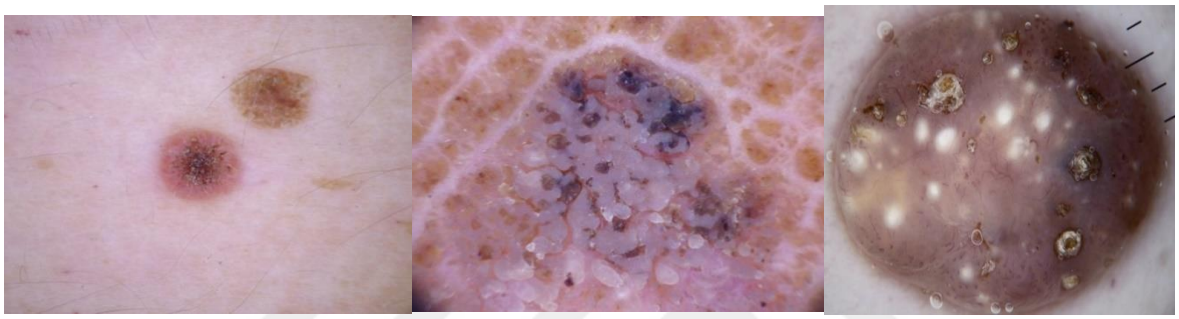


Figure 2.3: Pigmented Benign Keratosis Disease.

2.3 IMAGE PROCESSING

Image processing is a method for converting an image into a digital format and applying specific operations to it to create an improved image or extract some important data. It is a form of signal release in which an image, such as a video frame or photograph, serves as the input, while the output might either be an image or attributes associated with images. The image processing may be divided into two categories: analog image processing (AIP) and digital image processing (DIP). Analog image processing (AIP) works with electrical waves and 2D analog signals, such as those from television broadcasts using dish antennas, pictures taken with a camera, and other media. Because it is ongoing, this kind of processing takes a long time and costs money. Additionally, it takes a long time to fix the signal distortion during processing. On the other hand, the DIP deals with discrete zeros and ones in a pixel matrix when working with digital pictures. Digital image manipulation is simple because of the variety of processing methods that are accessible. DIP is regarded as one of the areas of ongoing study and has helped academics find solutions to a variety of real-world issues requiring visual observations, including image processing and analysis [38].

2.3.1 Digital Image Processing (DIP)

DIP is a method that uses a few computer algorithms to manipulate digital photos. Images or a collection of representational aspects or attributes of the source photographs are the possible results of this method. Additionally, these technologies are frequently used in medical imaging, remote sensing, robotics, and intelligent devices. DIP directly works with an image made up of several image points. These picture points, also known as pixels, are spatial coordinates that show the locations of the image's points as well as their intensities or levels of gray. Since red, green, and blue values are frequently employed to imitate image colors in various combinations in the actual world, colorful images have better dimensional features than grey images [39].

DIP's primary goal is to enhance important picture properties and improve image data (functions) by eliminating pointless distortions. Additionally, it enables humans to receive a detailed or high-quality version of the original image. Furthermore, imaging devices or sensors are hesitant to automatically capture "meaningful" targets, in contrast to the human visual system, which can adjust to different settings.

2.3.2 Image Enhancement

The main goal of image enhancement is to manipulate a given image in a way that makes the final product appear better than the original image for a certain application. It emphasizes or sharpens picture elements like edges, limits, and contrast to make a graphic display more effective for presenting and research. The improvement increases the dynamic range of the chosen characteristics, making them easier to identify, but it does not improve the intrinsic information quality of the data. The measurement of the enhancement criteria presents the largest obstacle in picture enhancement, hence many image enhancement approaches are empirical and depend on interactive procedures to get good results [40].

2.3.2.1 Principal component analysis (pca)

Is a well-known unsupervised feature extractor that converts RGB images into grayscale images and determines an orthogonal projection to ensure that the few dimensions of the new feature space adequately describe the variation in the data. Medical HSI applications frequently use PCA for categorization and reconstruction tasks. [41] Although it has not yet been used on medical HSI, super pixel-based PCA (SuperPCA) has been proposed as an

enhancement for hyperspectral image data that contain inhomogeneous regions [42].

2.3.2.2 Morphological closing filter

An essential operator in the study of mathematical morphology is closing. It, like its dual operator opening, can be created by the basic operations of erosion and dilation. Although grayscale variations exist, it is most commonly employed in binary images like those operators. Closing is comparable to dilatation in some aspects, such as the tendency to extend the limits of foreground (bright) parts in a picture (and to reduce the size of background color gaps in such places), However, it is less damaging to the original border shape. A structural component, like other morphological operators, determines the precise operation [43]. The operator has the effect of deleting all other parts of the background pixels while keeping the areas in the background that are similar in form to this structuring element or can completely encompass it [44]. Figure (2.4) depicts the impact of a closure on a binary picture using a 3*3 square structuring element.

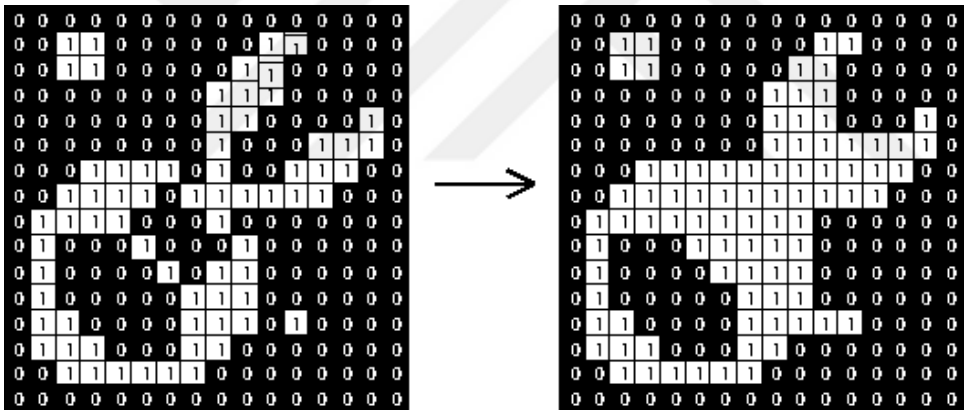


Figure 2.4: Closing With A 3*3 Square Structural Element.

2.3.3 Image Segmentation

The method of segmentation is used to divide a picture into items or parts. A segment is a divide or partition of a picture into various pieces. It is not a good idea to process the full image at once since there will be areas where there are no details. We may utilize the suitable segments for image processing by segmenting the picture. Since each image is made up of a different collection of pixels, image segmentation groups the pixels with the same characteristics. While image segmentation creates a pixel mask for each item in the picture,

image detection builds a boundary box for each object in the image without providing any more information about it. We can understand the object(s) in the image considerably more precisely thanks to this technique [45]. Figure (2.5) shows the difference clearly.

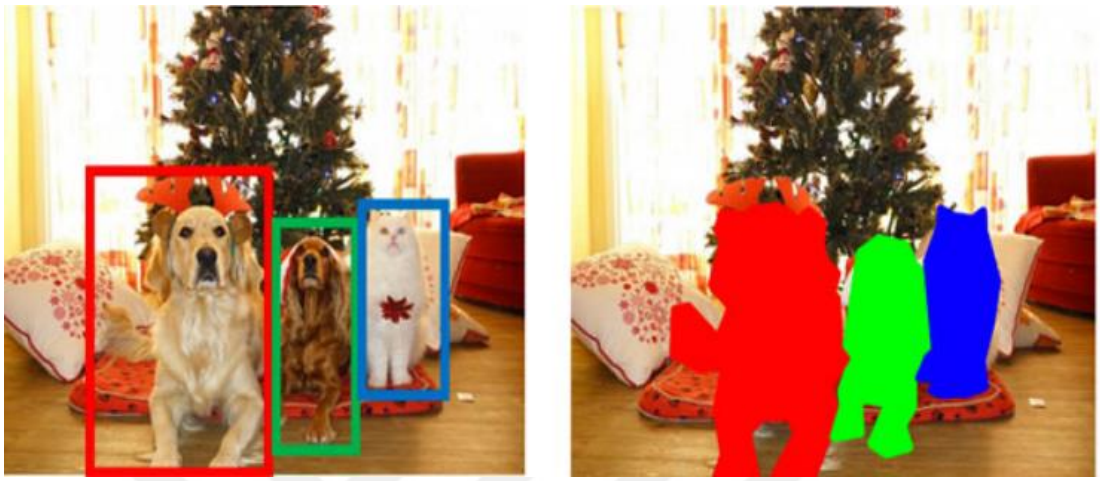


Figure 2.5: Different Between Object Detection and Segmentation
Images Can Be Segmented Using A Variety of Techniques, Including Thresholding, Clustering, Compression-Based Techniques, Etc.

2.3.3.1 K-Mean clustering algorithm

K-Means The goal of the clustering method is to divide the provided data set into K clusters; the value of K (the number of clusters) is determined by the user and is fixed. The mean of related locations can be used to calculate the centroid. The data points with the shortest distance from the specified cluster are assigned to that specific cluster once the centroid of each cluster has been picked for clustering. Data points are measured in terms of their distance from a certain centroid using the Euclidean Distance (ED) formula. The following is an expression of the objective fu nction:

$$J(V) = \sum_{i=1}^{C} \sum_{j=1}^{c_i} \left(\left\| x_i - v_j \right\| \right)^2 \quad (2.1)$$

Where ' c_i ' is the number of data points in the i^{th} cluster and $\left\| x_i - v_j \right\|$ is the ED between x_i and v_j. ' c_i 'stands for several cluster centers. The K-Mean method offers benefits like speedy access to massive databases and simplicity [46] [47]. As follows is how the K-Mean clustering technique operates:

a. Choose the clustering factor (K).

b. Choose K random data points without replacing the centroids to initialize the centroids by

first rearranging the data set.

c. Continue iterating till centroids stay the same.

d. Compute the squared sum of all the centroids' distances from the data points.

e. Assign each data point to the centroid of the nearest cluster.

f. Calculate the centroids for the clusters by adding together all the data points in each cluster.

The K-Mean clustering process is depicted in Figure (2.6) with three clusters.

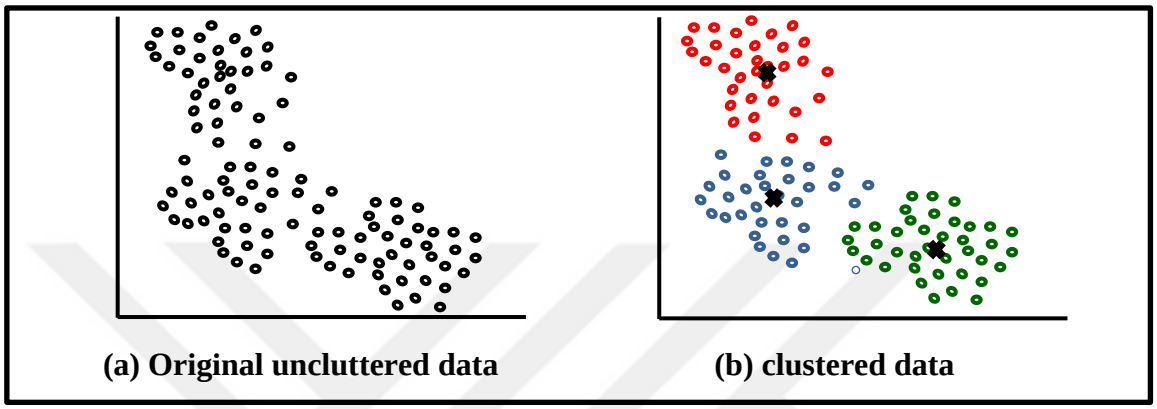


Figure 2.6: Clustering Data With (K=3), Centroid Denoted By X Sign.

2.4 IMAGE FEATURE EXTRACTION

A specific example of dimensionality reduction is feature extraction is the process of generating a collection of features from a set of data. The fundamental objective of the feature extraction approach is to extract the most crucial information from the original data and represent that information in a reduced dimensional space. We must carefully choose the feature extraction to carry out the intended function, utilizing this condensed representation rather than the full-size input, to extract the crucial information from the input data. The high dimensional features in ML and CV applications require more processing power and memory, which increases the execution time on the device. To extract the most significant characteristics from irrelevant data, where irrelevant refers to duplicate information, selection techniques are used [48].

The Gray Level Co-occurrence Matrix (GLCM) is a technique used in the image processing step to extract information from grayscale pictures. The brightness values of the pixels in the picture can be combined in various ways. The pixel location data, which has comparable grey level values, is contained in GLCM. It can be used to determine the degree of spatial dependency in an image's grey scale. The number of columns and rows in the GLCM matrix, which has four orientations (0, 45, 90, and 135), perfectly matches the number of grey levels

in the picture. The average of the earlier matrices is used to construct a new matrix. The relationship between "two neighboring pixels" is taken into account by GLCM; the first pixel is referred to as a reference and the second as a neighbor pixel [49].

Let the matrix's size be $(N \times N)$ and the GLCM be $P(i,j)$. (i,j) can be used to express the frequency of each grey level (i), which is spatially connected to a GLCM scale grey level pixel (j). The creation of the GLCM matrix from the picture is shown in Figure (2.7).

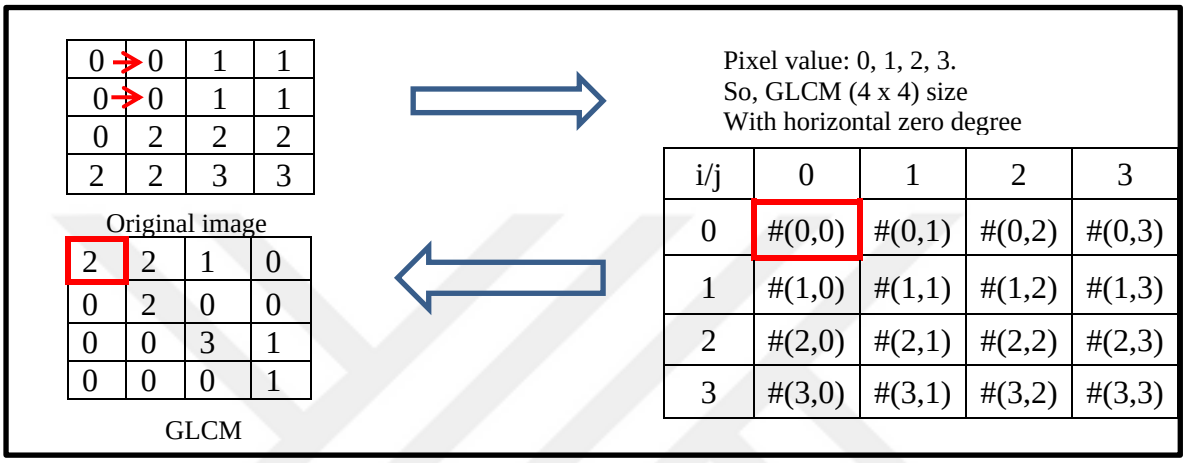


Figure 2.7: Construction GLCM from an Image.

Several features may be explained by the following [50]:

a. Contrast: the difference in pixel intensity between its minimum and maximum. It may be stated as follows:

$$Contrast = \sum_{i,j=0}^{N-1} P(i,j) \times (i-j)^2 \tag{2.2}$$

b. Correlation: A correlation demonstrates the relationship between a pixel and its surroundings. Its value is between -1 and 1.

$$Correlation = \sum_{i,j=0}^{N-1} P(i,j) \times \frac{(i-\mu)\times(j-\mu)}{\sigma^2} \tag{2.3}$$

Where μ is the GLCM mean, which is an estimate of the intensity of all the pixels involved in the connection and who made contributions to the GLCM, it is calculable as:

$$\mu = \sum_{i,j=0}^{N-1} i\, P(i,j) \tag{2.4}$$

And σ^2 is the variance of all reference pixels' intensities across all associations that contributed to the GLCM, as determined by:

$$\sigma^2 = \sum_{i,j=0}^{N-1} P(i,j)(i-\mu)^2 \quad (2.5)$$

c. Energy: Energy, which ranges from 0 to 1 and has a value of 1 for the constant image, determines uniformity with squared elements in GLCM.

$$Energy = \sum_{i,j=0}^{N-1} P(i,j)^2 \quad (2.6)$$

d. Entropy: which quantifies the disorder in a picture, has its highest value when all p matrices are equal.

$$Entropy = \sum_{i}^{N-1} \sum_{j}^{N-1} P(i,j) \times \log[P(i,j)] \quad (2.7)$$

e. Homogeneity: Homogeneity provides a number that indicates how closely an element's distribution is to the GLCM diagonal.

$$Homogeneity = \sum_{i,j=0}^{N-1} P(i,j)/R \quad (2.8)$$

f. Variance: It measures the dispersion of the gray-level distribution.

$$Variance = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i-\mu)^2 \times P(i,j) \quad (2.9)$$

2.5 MACHINE LEARNING ALGORITHM

ML is a computer science subfield that evolved from the study of AI recognize patterns and practical learning theory. It investigates the design and research of algorithms that can learn and predict data. Rather than following strictly fixed instructions for the program, such algorithms function by constructing a model from example inputs to make predictions or assessments based on facts [60]. There are several algorithms in ML. Several aspects must be considered when selecting an algorithm to tackle a given issue, including the kind of method, parametrization, memory size, overfitting propensity, learning time, and prediction time [61].

Machine learning methods can be supervised or unsupervised and reinforced, depending on the quantity and kind of supervision required during training, because such approaches study data and discover patterns to react to an environment [62].

2.5.1 Supervised Learning

Supervised learning is often used in classification issues since the goal is always to teach the computer our categorization system. Another frequent example of how to learn

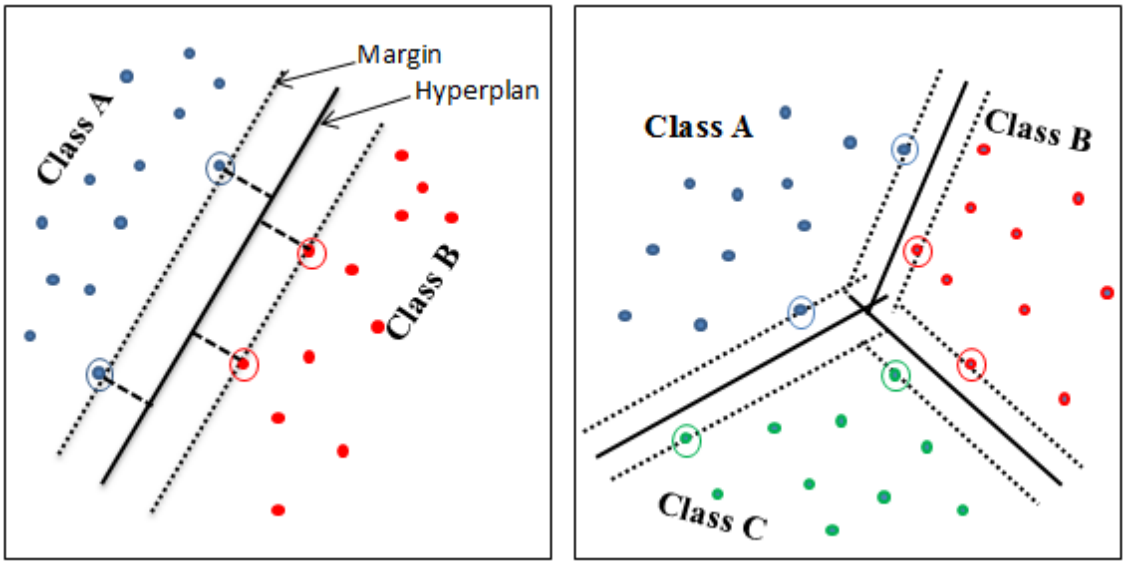
categorization is digit recognition [63]. The supervised ML algorithms require external assistance. The input data set is divided into train and test datasets. The performance of the training dataset is varied and must be predicted or categorized. All algorithms learn some form of pattern from the training dataset and use it in the test dataset for prediction or classification. Many supervised algorithms exist, including Decision Tree, Nave Bayes, Support Vector Machine (SVM), and others [64].

2.5.1.1 Multi-class support vector machine (m-svm)

The data points closest to the selection surface (or hyperplane) are referred to as "support vectors." SVM is a supervised ML method that is utilized in regression as well as classification issues, however, it is more commonly employed in classification problems. Each data item is plotted in m-dimensional space with the value of each feature in a specific coordinate (where m refers to the number of features). Finally, the classification works effectively by locating the hyperplane that divides the two groups. SVM creates margins as the boundary between groups in the provided dataset.

The notion is that the margins should be constructed in such a way that the distance between each class and the nearest margin is maximized, resulting in the lowest possible classification error.

The margins are defined as the distance between two supporting vectors separated by a hyperplane [65]. SVMs were initially intended for binary classification. However, by dividing the multiclass classification issue into a sequence of binary classification problems, it may be successfully expanded to multiclass classification. A multi-class support vector machine (M-SVM) is an SVM that handles all categories at once. Figures (2.8) depict the categorization of SVM and M-SVM [66].



(a) SVM **(b) M-SVM**

Figure 2.8: Classification Of Two Classes In SVM And Three Classes In M-SVM.

2.5.2 Unsupervised Learning

Unsupervised learning is a type of machine learning that looks for previously unnoticed patterns in a data collection with no prior labeling and as little human supervision as possible. Unsupervised learning, often known as self-organization, differs from supervised learning in that it does not employ human-labeled data. The principal component analysis and analysis of clusters are two of the most often utilized methodologies in unsupervised learning. In unsupervised learning, the analysis of clusters is used to group or divide datasets with similar properties to deduce algorithmic links. Cluster analysis is a field of machine learning that clusters data that has not been labeled, classed, or categorized.

Rather than ignoring feedback, cluster analysis finds similarities in data and responds depending on the presence or absence of these patterns in each new piece of data. This method aids in the detection of abnormal data points that do not fit into any of the groups. One of the data clustering strategies is the K-Mean algorithm [67].

2.5.3 Reinforcement Learning

Reinforcement learning (RL) is a branch of machine learning that studies how software agents should behave in a given environment to maximize the concept of cumulative reward. In contrast to supervised learning, it does not need labeled input/output pairings and does not require explicit correction of suboptimal behaviors. Instead, the emphasis is on striking

a balance between discovery (of unexplored territory) and utilization (of existing knowledge) [64].

2.6 DEEP LEARNING ALGORITHM

DL is an ML technique that aids computers in understanding the environment in terms of concept hierarchy and learning from experience. There is no particular necessity for a human-computer user to ascertain all of the machine's essential knowledge since the computer gathers information from experience. By building complicated concepts out of smaller ones, a structure of ideas aids with machine understanding [7].

With time, DL has expanded to encompass a larger idea of algorithms that can comprehend several levels of data representation that correlate to various complexity hierarchies. In other words, algorithms can learn different levels of complexity in addition to their ability to predict and categorize. This may be seen in how a neural network learns to recognize objects in images, such as faces, eyelashes, people, etc. The advantage of this is clear: we can reach the necessary level of sophistication to create intelligent software. We may presently observe this in functions like autocorrect, which simulate the suggested alterations to perceptible speech patterns that are particular to each person's language. The different layers in these models accomplish varying degrees of data abstraction. DL models are always constructed with layers of nonlinear units processing data, or neurons. The input layer, the hidden layers (h1, h2, and h3), and the output layer are the three layers of neural networks that are illustrated in Figure (2.8). [51].

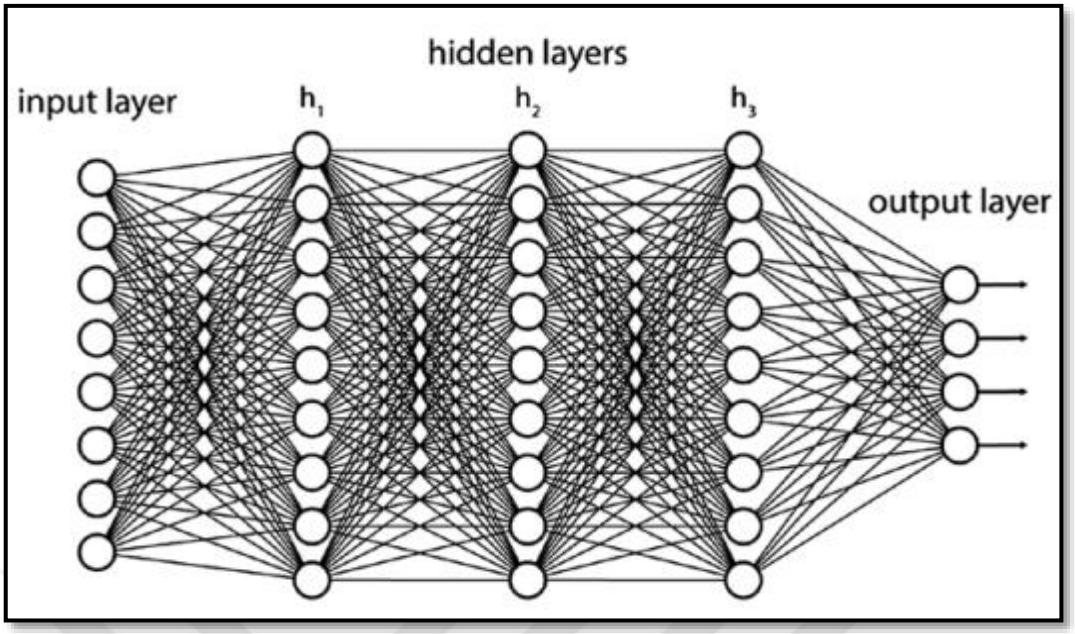


Figure 2.9: Deep Neural Network.

2.6.1 Convolutional Neural Network

The most popular and efficient DL method for handling the image classification issue is CNN, one of the DL architectures. CNNs are a development of conventional artificial neural networks (ANNs), concentrating mostly on applications with repeated patterns in different regions of the modeling space in a specific picture. Their primary distinguishing feature is that, according to the technology utilized in their layering, they significantly lower the number of structural components (number of artificial neurons) required in comparison to typical feedforward neural networks. CNN is a feed-forward, very effective detection technique. The network has an easy structure and fewer training parameters.

CNN is an abbreviation that stands for a very efficient detection method. On the other hand, the complexity and weight values of the network model are reduced. [52] [53]. The primary structure of CNN is seen in Figure (2.10) and consists of four layers: a convolution layer with an activation function, a pooling layer, a fully connected layer, and a softmax layer. It will be covered in more detail in the section after this.

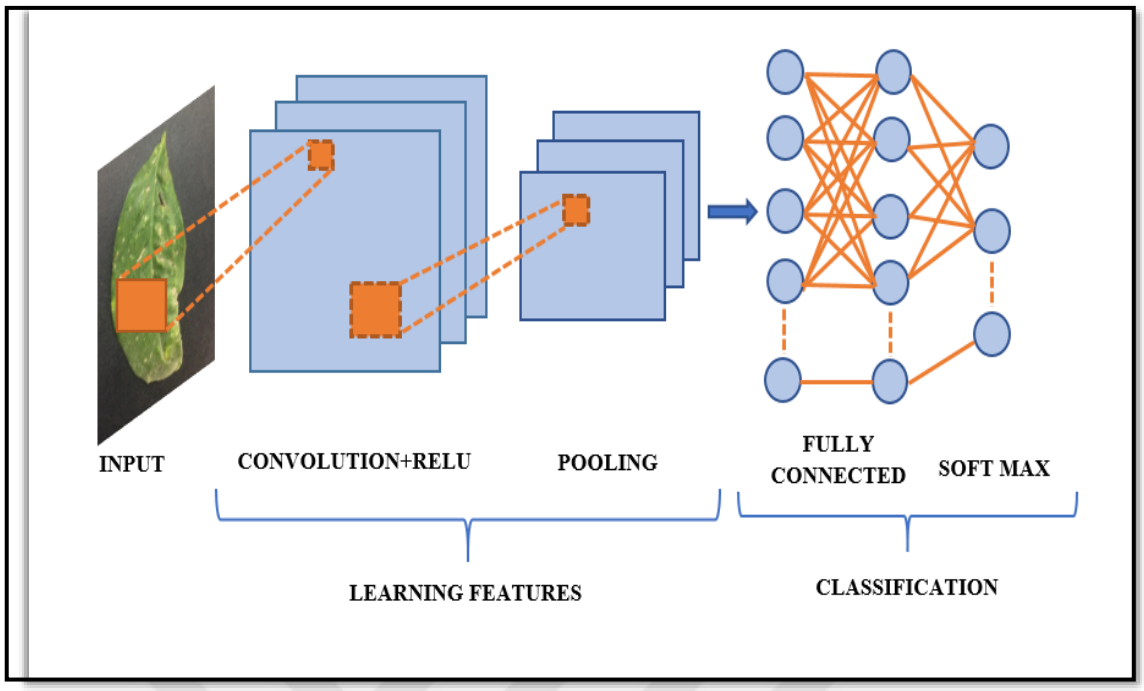


Figure 2.10: Main Structure of Convolutional Neural Network

a. Convolutional layer

The simplest description of a convolution operation is an operation between tensors (multidimensional arrays). When the input is a pair of tensors and the output is a single tensor, the operator * is typically used to denote the convolution operation. Additionally, filters are among the most crucial elements of CNNs. Another name for the filters is kernels, which are square matrices with tiny dimensions like 3x3 or 5x5. Here, the stride determines how the row and column values in the convolution process will change [54].

The fundamental components of CNN designs are regarded as convolutional layers. CNN creates new feature maps by using different kernels in the convolutional layers to converge over the full picture and the intermediate feature maps. Convolutional processing offers three main benefits. The weight-sharing approach minimizes the number of parameters and, therefore, the number of operations in the same feature map. It is possible to analyze correlations between nearby pixels thanks to local connectivity. Finally, regardless of the object's position in the picture, invariance to the object's location aids in object localization [55]. Figure (2.10) depicts the convolution process for an RGB picture using a (3x3xC) kernel size, where C is the minimum number of kernel channels required to match the number of image channels. The convolution

operation will be performed by each channel of the kernel in the same channel of the image, and the resulting three-channel feature map will then be added to create the final feature map.

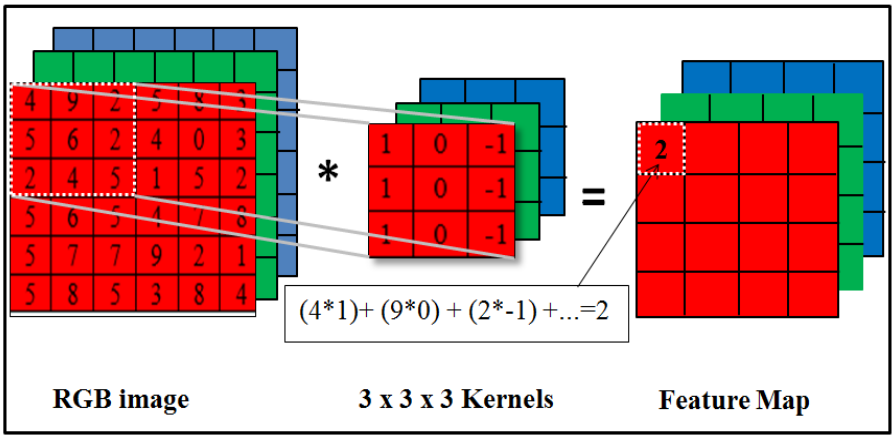


Figure 2.11: Convolution Using A (3 X 3) Kernel On An RGB Picture.

The Rectified Linear Unit (ReLU) activation function is often applied after each convolutional layer. It permits the network's insertion of nonlinearity. ReLU outperformed tanh or sigmoid functions in terms of computing efficiency while exhibiting little change in accuracy. So we can write the ReLU equation as follows:

$$ReLU(x) = \max(0, x) \tag{2.10}$$

If x is less than 0, the result will be equal to zero, and if x is more than 0, the result will be equal to one [56].

b. Pooling layer

The second process that is essential to CNNs is pooling. Contrary to convolution, this technique is much simpler to comprehend. The feature maps' and the network parameters' sizes are continuously reduced using the pooling layer. Since pooling layers' factor in nearby pixels when computing, they are also translation-invariant. Max-pooling layers and average-pooling layers are the two primary categories of pooling layers. The majority of implementations use max-pooling because it can enhance generalization, choose superior invariant features, and accelerate convergence. By only preserving a group of pixels' maximum values, the max-pooling layer chooses a set of pixels to combine into a single

pixel. A mask of size 2*2 will begin at the top-left 4 pixels if one is utilized. then choose the output, which is the highest value. Max pooling will move the mask to work on an extra 4 pixels, much like convolution; therefore, it requires a stride. The pooling layer's stride value is, at the very least, 2. Because a stride of 1 duplicates values without output, it is useless [57] and should not be used. The maximum and average polling for a (4x4) matrix with a (2x2) filter and a stride of 2 are shown in Figure (2.11).

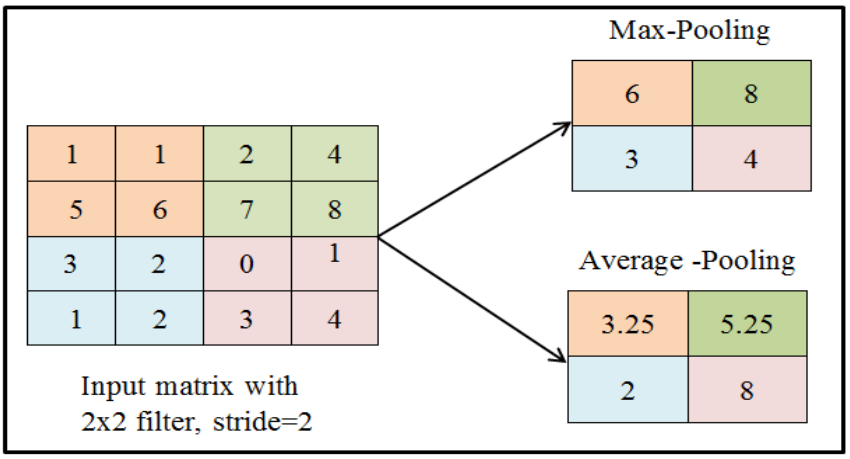


Figure 2.12: Pooling using a 2x2 filter and a stride of two

c. Fully connected layers (FC)

About 90% of CNN's parameters are FC layers. Using this, a vector with a predetermined length is sent forward into the neural network. For image classification, we could either feed the vector forward into a fixed number of categories or use it as a feature vector for later processing. Although modifying the FC layer's structure is unusual, some work has been done to improve it. The FC layer, which is the product of the previously added convolution, activation, and pooling layers, is a higher-level representation of the input picture. These layers are not intended to offer classification estimations. The input image is now defined using the FC layer by looking at the characteristics of the training set [58.].

2.6.2 CNN Model

a. AlexNet Model

The AlexNet model, a state-of-the-art CNN that has been pre-trained, was created by Alex Krizhevsky. It has been utilized for several benchmarks across a wide range of industries. With three clear benefits: better efficiency, fewer training parameters, and strong robustness,

AlexNet is CNN's most often used model. When simulating dual-channel visual communication and examining image features using two channels, as shown in figure (2.13). it is influenced by notions of human vision. It is a deep CNN of twenty-five layers, comprising an input layer, five convolutional layers, seven ReLU layers, two cross-channel normalization levels, a SoftMax layer, and finally an output layer with 1000 labels. The learning feature is executed across the whole model in the two-channel convolutional layer, and the channels are only crossed in the third extraction layer. The characteristics of the two groups are mixed in the first FC layer in the appropriate ways, and the process is repeated until the last FC layer, which then combines the features of the two groups to produce a 4096-dimensional feature vector [59].

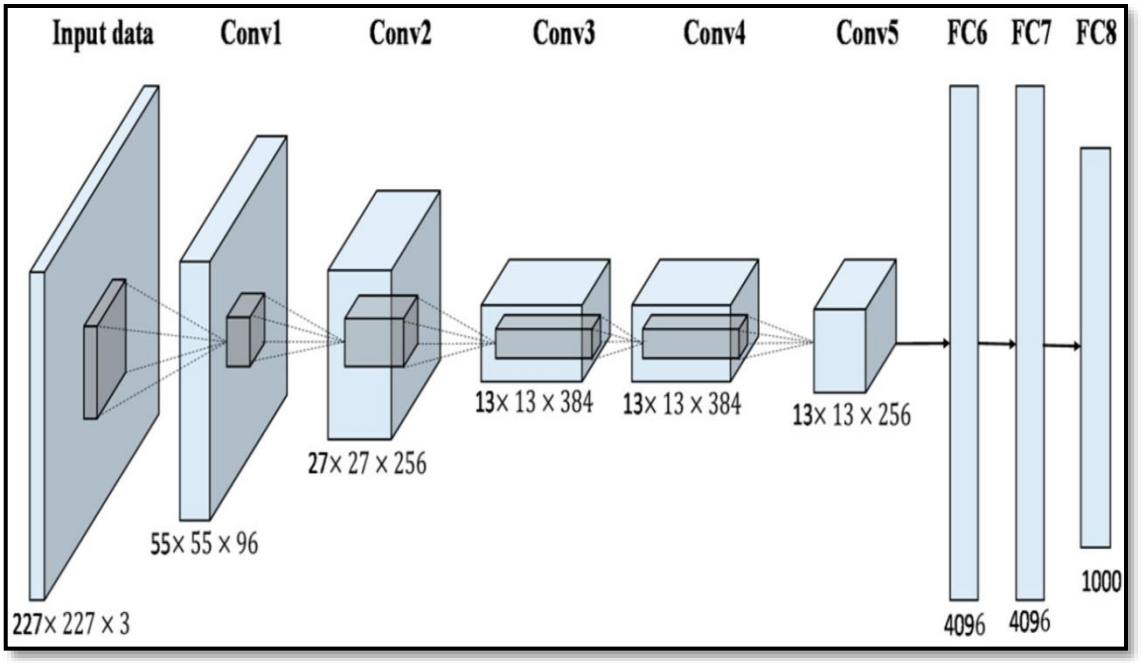


Figure 2.13: the Framework of the Alexnet Model.

Due to its dynamic structure, numerous training parameters, and a huge quantity of training data, the AlexNet model has to adopt relevant strategies to make training simpler and prevent overfitting. To make the initialized system more consistent with the theory, a model may be created by only applying the ReLU's nonlinearity to the data structures and starting the network's training process from scratch.

b. ResNet Model

ResNet is a deep CNN that has a residual structure that was specifically designed to achieve an extremely deep network. Traditional deep CNN is limited in its depth and loses accuracy

as it grows. The inventor of ResNet speculates that identity mapping is challenging to learn. This issue is recommended to be resolved using the Deep Residual Learning System, where the network learns the residual instead of the direct mapping [12]. This model won the ImageNet competition in 2015. ResNet's key breakthrough was its ability to efficiently train incredibly deep neural networks with over 150 layers. Before ResNet testing, very deep neural networks were challenging because of the problem of vanishing gradients. The ResNet-50 model's primary structure is seen in Figure (2.13).

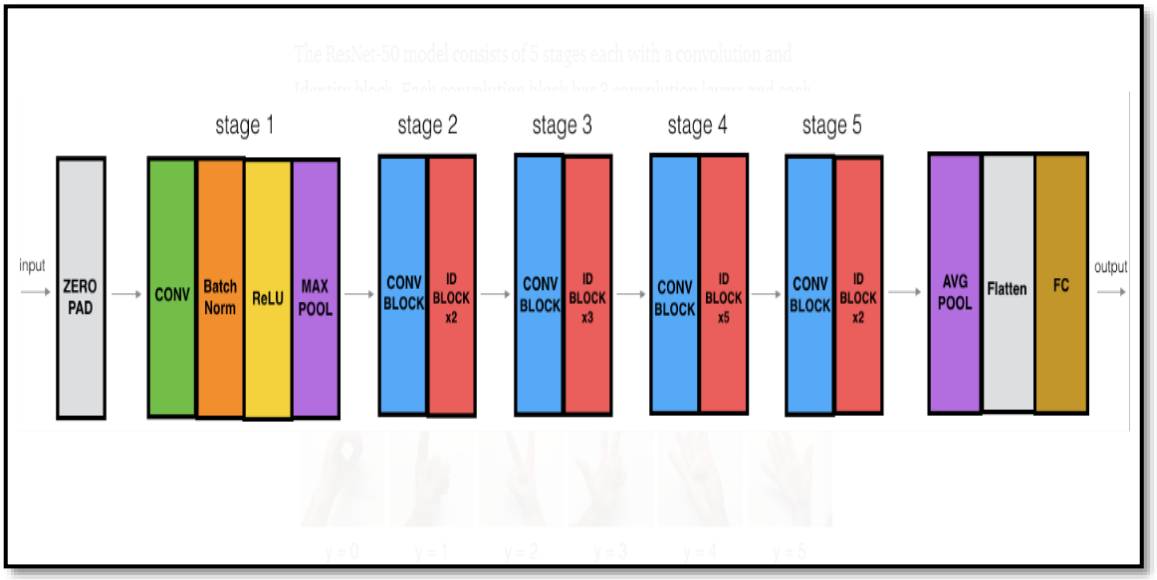


Figure 2.14: the Main Structure of the Resnet -50 Model.

Five stages make up the ResNet-50 model, each comprising a block of convolution and identity. Three levels of convolution are present in every block of convolution, and three layers of convolution are also present in every block of identity. There are around 23 million trainable parameters in the ResNet-50.

2.7 COLOR MODEL

Color models are a human-interpretable color measurement system and a means of integrating diverse values as a collection of basic colors. Color models often include three or four-color components, with distinct color spaces accessible for each application. There are three color model groups according to image processing applications [68];

a. Device-oriented color models: Also known as device-dependent color models, these models relate to and impact the device's signal, as well as the resultant color as it is changed by display tools. Such devices are often utilized in many situations where the color must

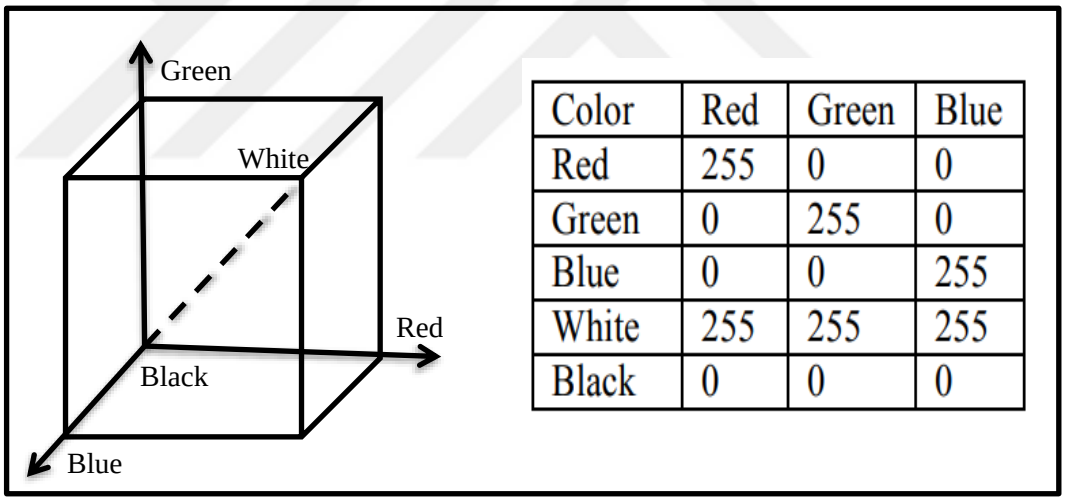
match hardware instruments such as televisions.

b. User-oriented color models: These models, which are viewed as a link between the viewer and the equipment that processes color information, allow the user to specify and approximate what he expects from the color display.

c. Device-independent color models: the color model is unaffected by the device's characteristics, and the same color would arise from a set of parameters with no regard for the equipment's output. This sort of color model is important in network information transfer since visual data must be passed by many hardware devices.

2.7.1 RGB Color Model

It is a three-color additive color model that adds red, green, and blue light waves to replicate a broad spectrum of colors. Because RGB is device-dependent, the quality of white relies on the principal light sources. The red, blue, and green color components all have values in the [0-255] [69] range.



Color	Red	Green	Blue
Red	255	0	0
Green	0	255	0
Blue	0	0	255
White	255	255	255
Black	0	0	0

Figure 2.15: RGB Color Space Fundamental.

2.7.2 Hue, Saturation, and Value (HSV)

HSV is a cylindrical geometry with hue as its angular length, beginning at 0 °, traveling through 120 ° of primary red, 240 ° of primary green, and 360 ° of primary blue, then returning to red. Each geometry's center vertical axis is made up of neutral, achromatic, or gray hues that range from top to bottom, from white at brightness 1 (value 1) to black at lightness 0 (value 0). It is based on three variables [70];

• Hue: from 0 to 360 degrees (or zero to 2 radians);

• Saturation: from 0.0 to 1.0 (0-100%, inclusive);

• Value: from 0.0 to 1.0 (0-100%, like Saturation).

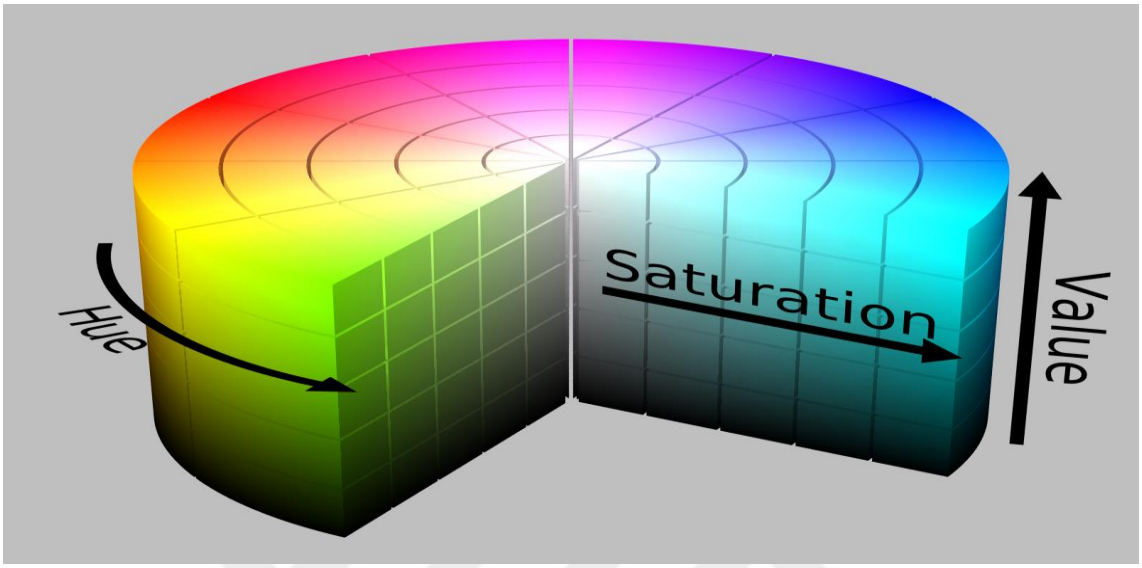


Figure 2.16: HSV Color Space Representation.

3. MACHINE AND DEEP LEARNING IMPLEMENTATION IN SKIN CANCER DISEASE DETECTION

3.1 INTRODUCTION

To detect and classify the disease for skin cancer, the key processes of DL with area computation are given in this chapter.

3.2 PROPOSED METHODOLOGY FOR ML ALGORITHM

Four processes make up the suggested approach that is used to identify and classify skin cancer diseases, as indicated in chart (3.1). The ML algorithm's flowchart is displayed in chart (3.2).

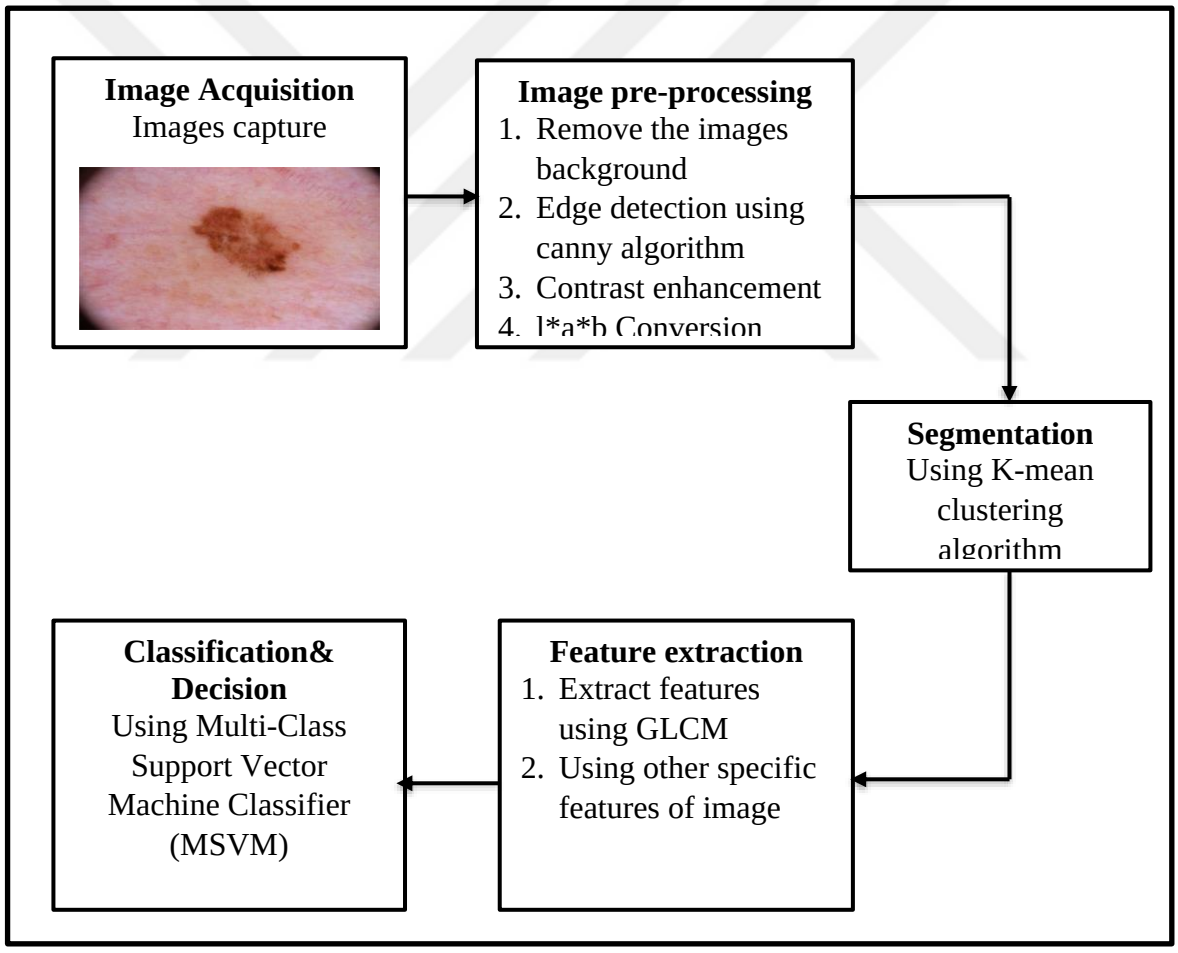


Chart 3.1: General Structure of the Proposed ML Algorithm.

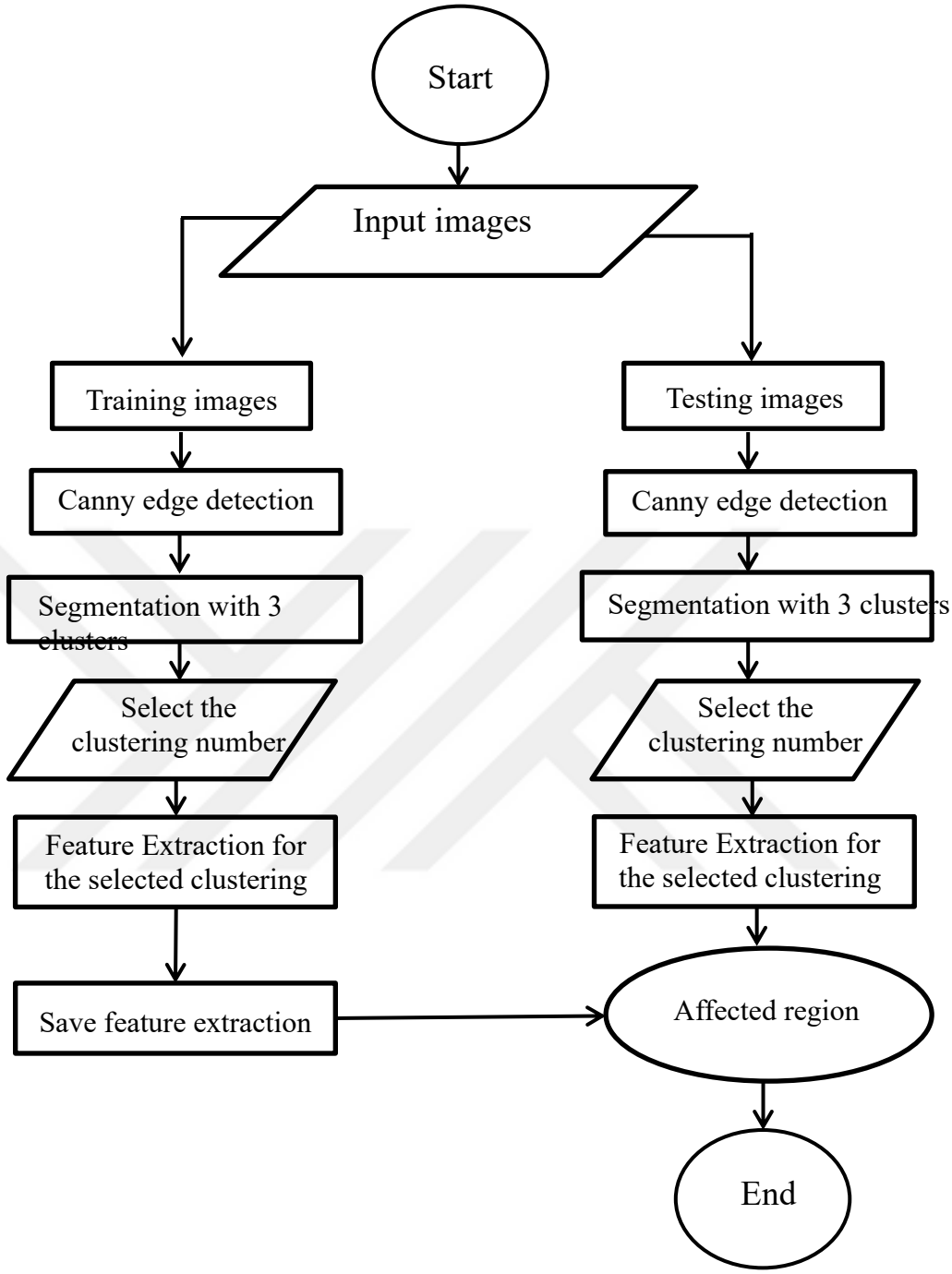


Chart 3.2: Flowchart of the Proposed ML Algorithm.

3.2.1 Image Acquisition

The system's first stage is image acquisition, during which the picture is taken and saved in a format that is compatible with image processing, such as JPEG, JPG, BMP, or PNG. Dermatoscopic pictures are used to capture a variety of skin cancer disease classes, including

pigmented benign keratosis, basal cell carcinoma, and melanoma.

1176 photos of skin cancer diseases are utilized in this study; they are split into training set and testing set images. Examples of the captured photos utilized in this study are shown in the figure below. Samples from the compiled dataset of skin cancer are shown in Figure (3.1).

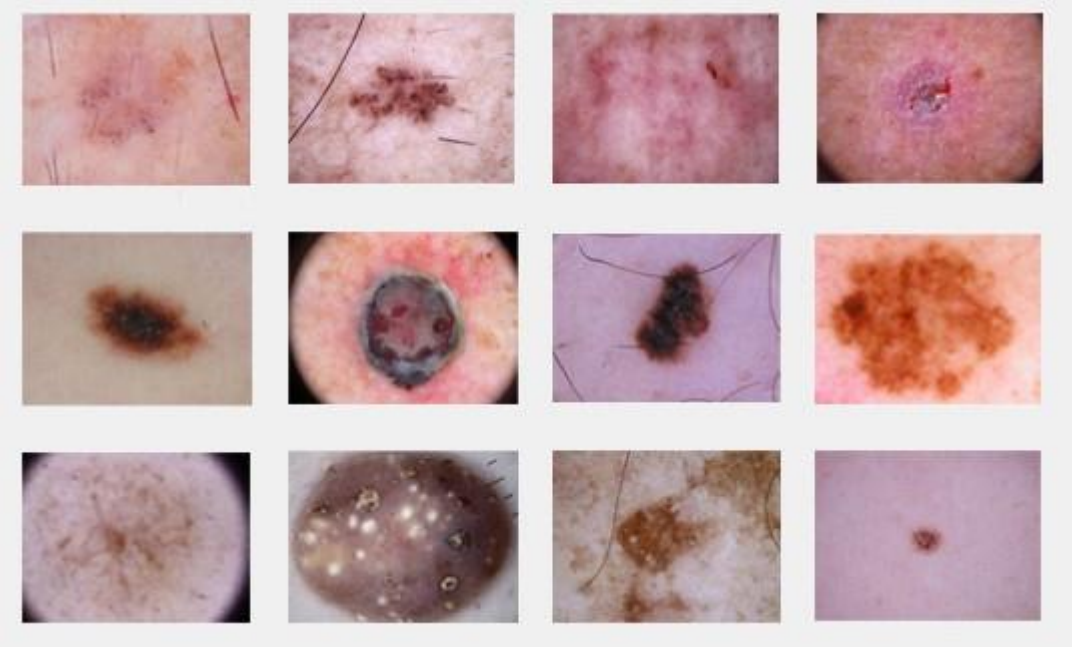


Figure 3.1: Image Acquisition for Different Class Skin Cancer.

3.2.2 Image Pre-Processing

The data gathered is frequently disorganized and originates from several sources. To be input into the ML model, they must be cleaned up and standardized. Pre-processing is typically used to carry out steps that simplify the algorithm and improve its accuracy. We are unable to create a specific method for each circumstance in which a picture is obtained; therefore, once we have an image, we attempt to change it into a shape that can be solved by a general approach. Because of light source reflection, which causes local brightness and contrast variance, skin cancer image lighting is frequently not uniform. Additionally, the color, size, and quality of skin cancer images are all over the place. This may have an impact on how accurately skin cancer abnormalities are found and provide false-positive findings.

In this study, the enhancement was applied using a hybrid contrast stretching approach, which is an edge detection and enhancement routine, and contrast enhancement. The noisy and brighter image produced distortions of the proper class of skin diseases. The process for enhancement is as follows:

a. Removing the background from the image to get rid of any undesired areas.

b. Convert the color of the given image to grayscale.

When it comes to image detection and classification, it is preferable to transform an RGB image to a grayscale form since this allows for faster edge detection and feature extraction.

c. Create the morphological structuring element: After using the morphological operators, the original image's mask is moved to reveal a new value for each pixel. Each mask's value of 1 denotes the final image's efficacy, whereas its value of 0 denotes ineffectiveness. One can choose from a variety of forms to create a mask. In our study, the (23x23) matrix-shaped mask with a radius of 12 is utilized.

d. After morphological ways to improve and identify the edge for the images, apply the canny edge.

e. Change an image's color space from RGB to L*A*B*. It is made up of three layers: layer L for brightness, layer A for chromaticity, and layer B for color information.

The flow chart for image processing is shown in chart (3.3).

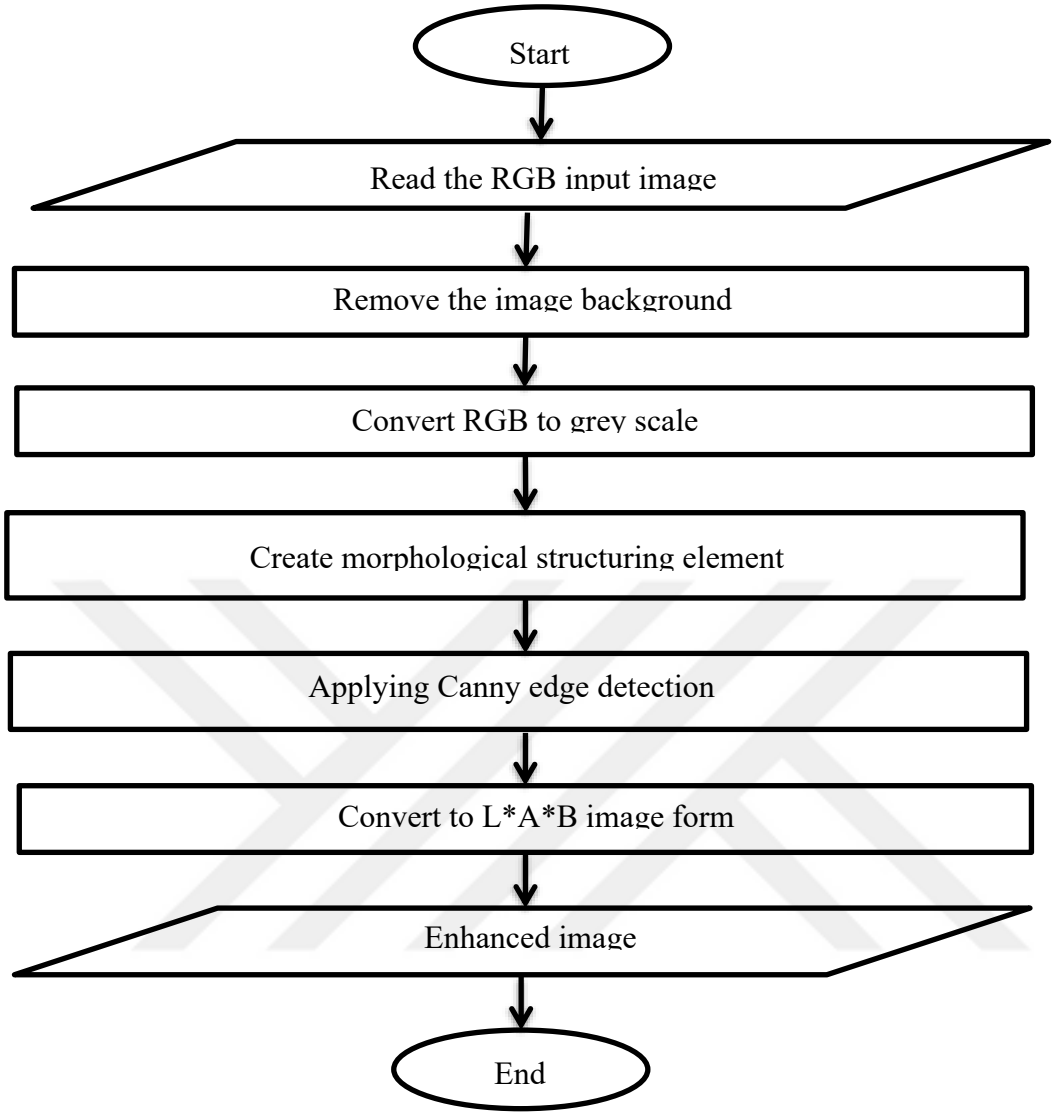


Chart 3.3: Flowchart of the Pre-Processing Image.

3.2.3 Image Segmentation

The method used to divide an image into items or parts is called segmentation. To simplify the representation of images in a more relevant fashion, the division of images is done using the k-mean clustering algorithms.

The next stages include applying image segmentation using the enhanced image that was produced by the previous procedure.

a. Using the l*a*b enhanced image, classify the colors in a*b.

b. Using the technique for K-mean clustering By using equation (2.1).

c. Select the clustering factor (K).

The process flow chart for K-mean clustering is shown in chart (3.4).

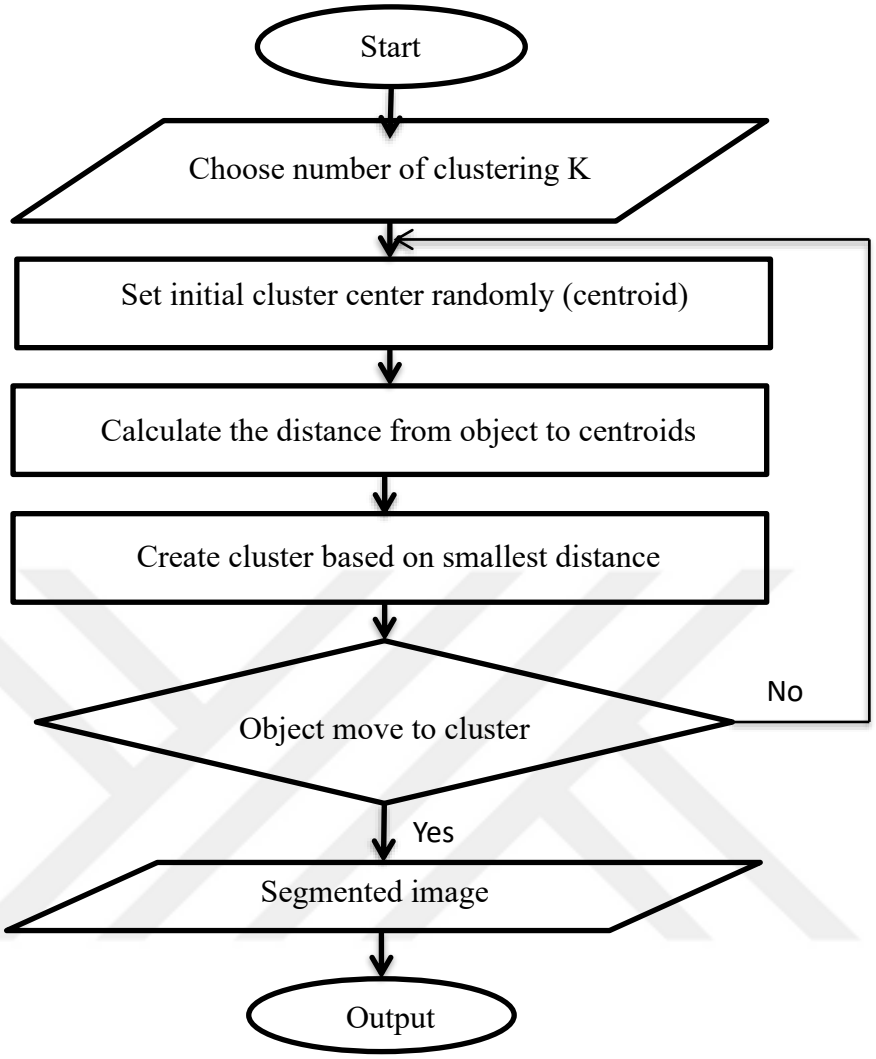


Chart 3.4: Flowchart of K-Mean Clustering Algorithm.

3.2.4 Feature Extraction

High-dimensional features in CV and ML applications increase processing space and system execution time requirements. As a result, using feature selection approaches, a subset of the most important features are extracted from irrelevant data, where irrelevant refers to duplicated information. Apply equations (2.2) to (2.6) to the gray-scale clustered image to extract the contrast, correlation, energy, entropy, homogeneity, and variance features, respectively. Then, choose the clustered images that include the area of interest (ROI) to extract features using GLCM. Additionally, several other characteristics including Mean, Standard Deviation, Root Mean Square (RMS), Skewness, and Kurtosis, were recovered

from the RGB clustered image. Chart (3.5) shows the process flowchart for feature extraction.

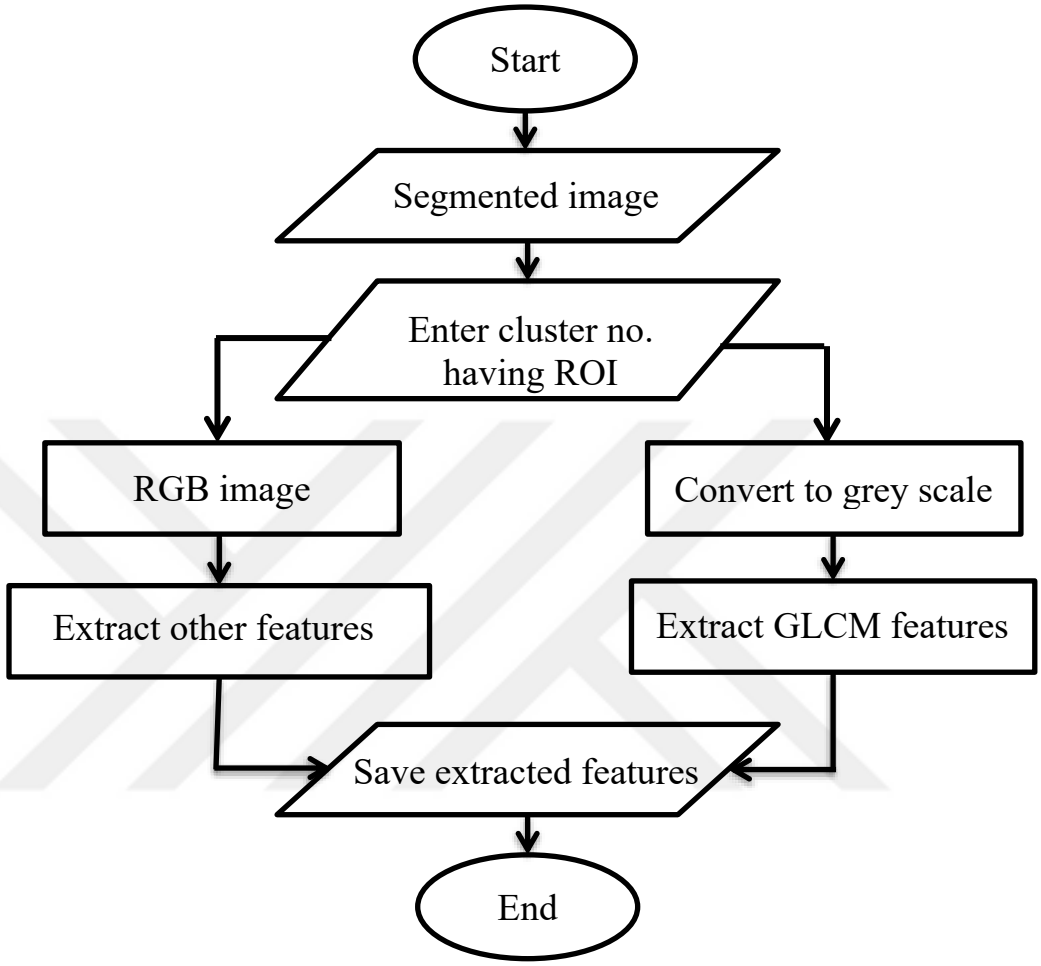


Chart 3.5: Flow Chart of the Feature Extraction Process

3.2.5 Classification

At this stage, classification is done using the features that were extracted from the testing images. To place the tests in the appropriate class, the classifier is provided with a set of skin cancer disease features (training features) along with the target class (training labels).

The suggested system, which is depicted in figure (3.2), uses M-SVM as a classifier approach. (SVMs) are supervised learning models that evaluate data and identify patterns. They also have related learning methods. They are utilized in regression and classification analyses.

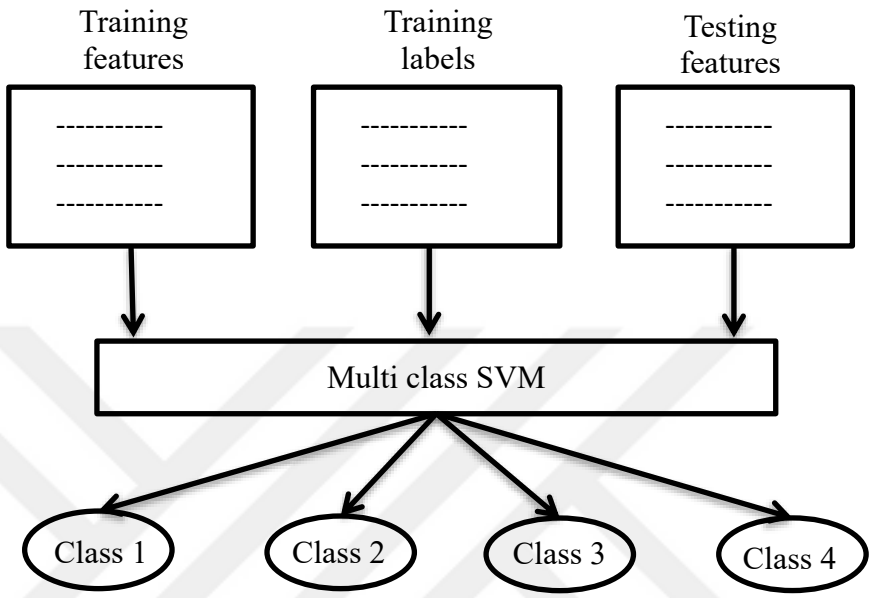


Figure 3.2: Main Structure of M-SVM Classifier.

3.3 PROPOSED METHODOLOGY FOR DEEP LEARNING ALGORITHM

Chart (3.6) depicts the suggested approach for employing CNN models to detect and classify skin cancer diseases:

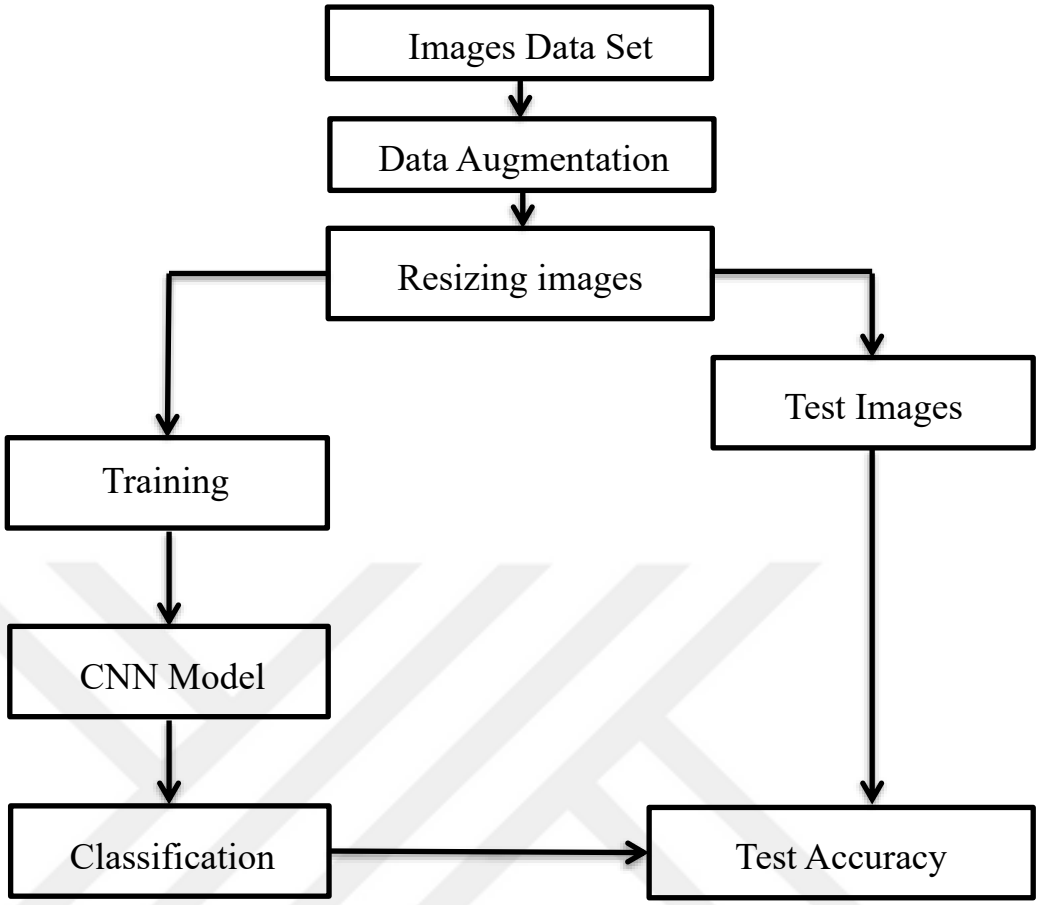


Chart 3.6: The Flowchart of Image Classification With Data Augmentation Using CNN.

In this research, we combine the data augmentation technique with two CNN models, Alex Net and Res Net.

Data augmentation is the process of altering the original data to produce new data points. This method, which works well with small datasets, increases the number of training images for deep learning without requiring the addition of new pictures. Despite being extremely strong, CNN may end up overfitting and failing to produce the desired results because there are insufficient pictures utilized, thus artificially increasing the dataset through label-preserving changes.

The following is how the data augmentation is used:

a. A seemingly random left-to-right reflection.

b. The input image's horizontal translation range is measured in pixels, with a pixel range of [30–30].

c. The input image's vertical translation's applied range is measured in pixels, with a pixel range of [30–30].

3.3.1 AlexNet and ResNet CNN Models

These models were used to detect and classify the images that follow skin cancer diseases.

a. Launch the training image collection.

b. Resize photos into the following dimensions: (227x227x3) for AlexNet and (224x224x3) for ResNet, where 3 is the number of colors.

c. Determine the number of images per label and save the class numbers.

Label	**Count**
basal cell carcinoma	392
melanoma	392
pigmented benign keratosis	392

d. Split the images into 70% training images and 30% test images.

e. Load ResNet and the already-trained neural network AlexNet, which can categorize photos into different item categories. Figures (3.9), (3.10), and (3.11), respectively, exhibit the AlexNet and ResNet CNN-trained layers. The scaling in the preceding stage was necessary since the first input layer of the networks needed a picture of a certain size.

```
ans =
  25x1 Layer array with layers:

     1   'data'     Image Input                 227x227x3 images with 'zerocenter' normalization
     2   'conv1'    Convolution                 96 11x11x3 convolutions with stride [4  4] and padding [0  0  0  0]
     3   'relu1'    ReLU                        ReLU
     4   'norm1'    Cross Channel Normalization   cross channel normalization with 5 channels per element
     5   'pool1'    Max Pooling                 3x3 max pooling with stride [2  2] and padding [0  0  0  0]
     6   'conv2'    Grouped Convolution         2 groups of 128 5x5x48 convolutions with stride [1  1] and padding [2  2  2  2]
     7   'relu2'    ReLU                        ReLU
     8   'norm2'    Cross Channel Normalization   cross channel normalization with 5 channels per element
     9   'pool2'    Max Pooling                 3x3 max pooling with stride [2  2] and padding [0  0  0  0]
    10   'conv3'    Convolution                 384 3x3x256 convolutions with stride [1  1] and padding [1  1  1  1]
    11   'relu3'    ReLU                        ReLU
    12   'conv4'    Grouped Convolution         2 groups of 192 3x3x192 convolutions with stride [1  1] and padding [1  1  1  1]
    13   'relu4'    ReLU                        ReLU
    14   'conv5'    Grouped Convolution         2 groups of 128 3x3x192 convolutions with stride [1  1] and padding [1  1  1  1]
    15   'relu5'    ReLU                        ReLU
    16   'pool5'    Max Pooling                 3x3 max pooling with stride [2  2] and padding [0  0  0  0]
    17   'fc6'      Fully Connected             4096 fully connected layer
    18   'relu6'    ReLU                        ReLU
    19   'drop6'    Dropout                     50% dropout
    20   'fc7'      Fully Connected             4096 fully connected layer
    21   'relu7'    ReLU                        ReLU
    22   'drop7'    Dropout                     50% dropout
    23   'fc8'      Fully Connected             1000 fully connected layer
    24   'prob'     Softmax                     softmax
    25   'output'   Classification Output       crossentropyex with 'tench' and 999 other classes
```

Figure 3.3: AlexNet CNN Layers.

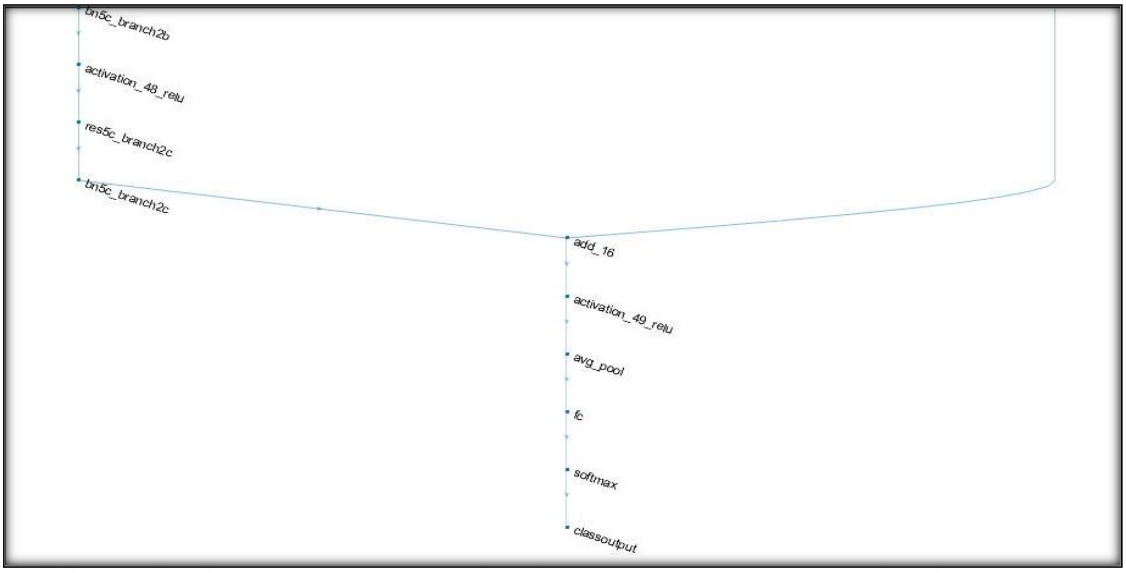


Figure 3.4: ResNet CNN Final Layer.

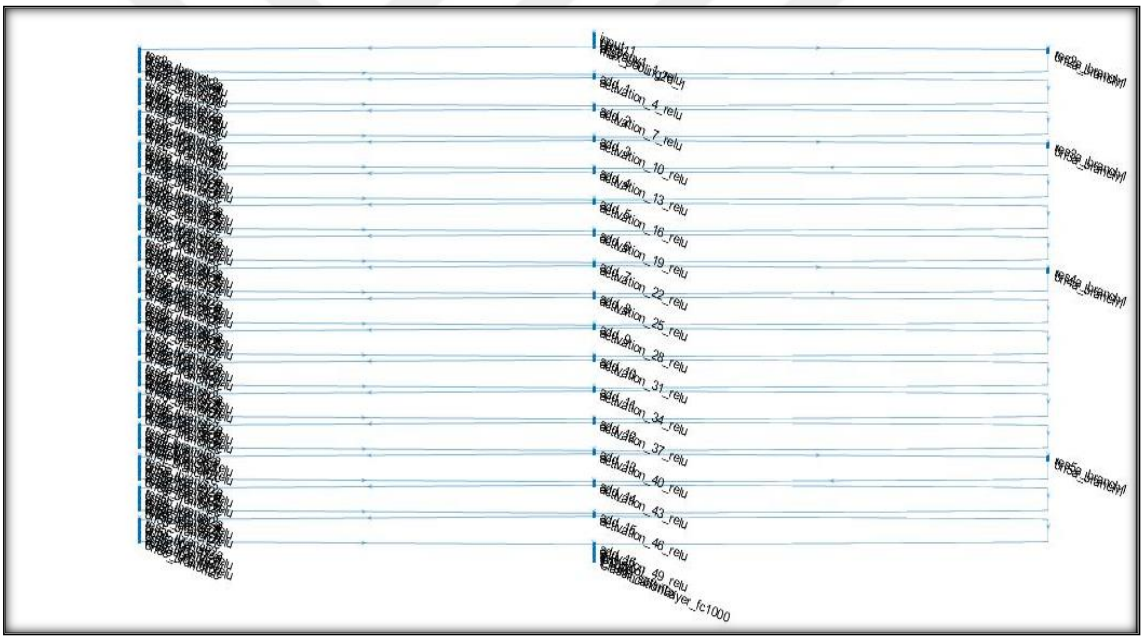


Figure 3.5: ResNet CNN Layers.

f. Swap out the final three levels to create three classifications for diseases.

j. Image augmentation is employed to prevent the network from overfitting due to the short dataset.

h. Specify the training option for each model as Stochastic Gradient Descent with Momentum (SGDM), with a maximum epoch of 100, a minimum batch size of 6, and an initial learning rate of 1e-4.

4. SIMULATION RESULTS

4.1 INTRODUCTION

The simulation results of the suggested work utilizing deep learning and machine learning are described in this chapter. Additionally, the area's results are computed. In the end, based on the dataset and following the three disease types, skin cancer was detected and classified. MATLAB R2019b was used to create the work.

4.2 MACHINE LEARNING IMPLEMENTATION RESULTS

The ML operations are completed in a limited number of stages, as stated in Section 3.2. The dataset's first component includes 1176 dermatoscopic photos of skin cancer diseases that were gathered in JPG format. The outcomes of the dataset's picture enhancement utilizing canny edge detection are then displayed in figure (4.1), along with the image enhancement process that was used.

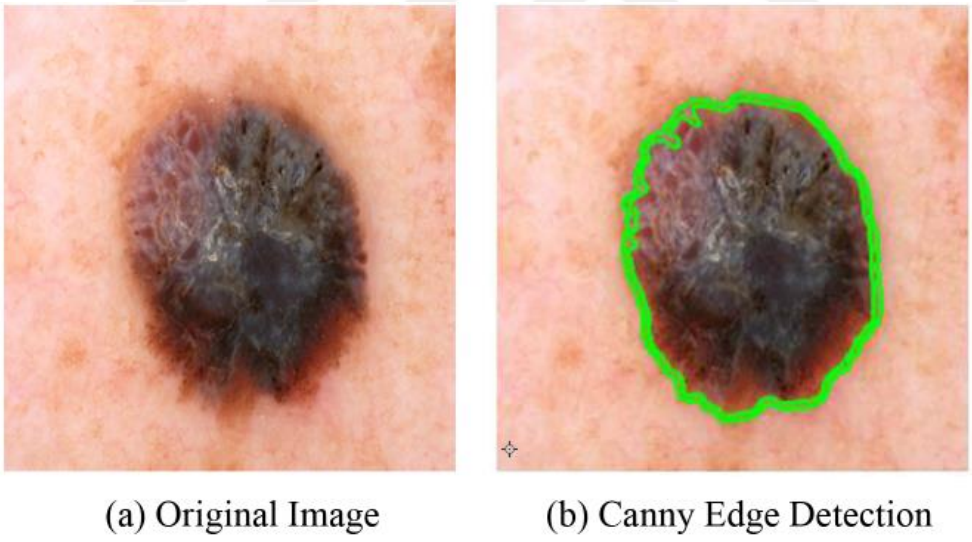


Figure 4.1: Image Enhancement Process For Skin Cancer Disease Image.

The improved image was divided into three clusters in the third stage using the K-mean clustering technique. The clustering for skin cancer diseases is depicted in Figure (4.2).

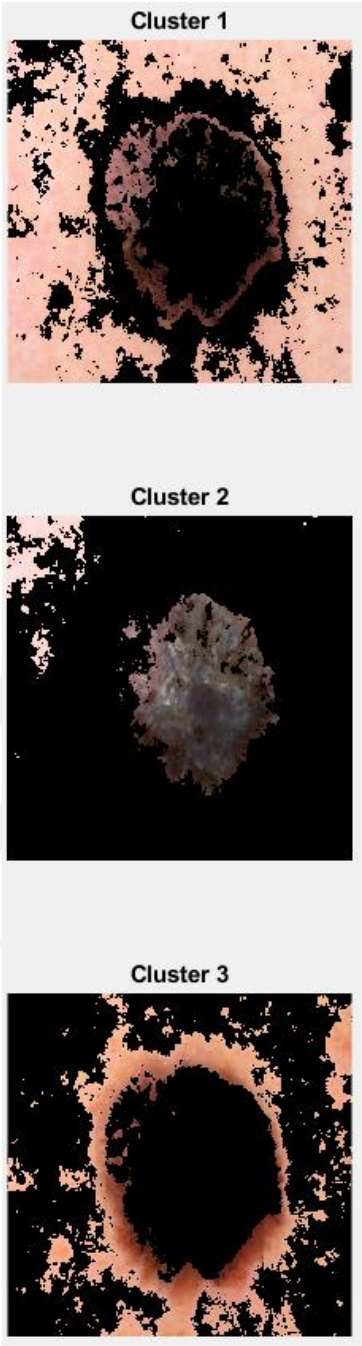


Figure 4.2: Applying the K-Mean Clustering Algorithm For Skin Cancer Disease.

Then, using the GLCM technique and other statistical characteristics, extract the features from clusters 1, 2, and 3. In tables (4-1), (4-2) and (4-3), the average for each characteristic in each cluster is displayed.

Table 4.1: Feature Extraction Average for Cluster 1.

	Basal cell carcinoma	Melanoma	Pigmented benign keratosis
Mean	58.0097	10.627	57.1139
S.D	86.3912	38.2256	69.8863
Entropy	2.69569	1.27845	3.52286
RMS	8.23025	4.47533	10.076
Variance	4967.09	1403.37	4505.62
Kurtosis	1.81583	22.65	1.30954
Skewness	0.856172	4.35361	0.460306
Contrast	1.73785	0.782093	1.93555
Correlation	0.842416	0.663345	0.730263
Energy	0.487629	0.759316	0.33063
Homogeneity	0.934048	0.956059	0.878983

Table 4.2: Feature Extraction Average for Cluster 2.

	Basal cell carcinoma	Melanoma	Pigmented benign keratosis
Mean	29.6198	147.83	63.7817
S.D	65.2266	108.181	70.5472
Entropy	1.8093	4.40538	3.89267
RMS	5.5575	14.4495	10.4651
Variance	3346.12	6376.39	4205.32
Kurtosis	4.77481	1.20958	1.19143
Skewness	1.86396	4.222691	0.266763
Contrast	0.968995	0.642019	1.51918
Correlation	0.845554	0.966552	0.787413
Energy	0.647406	0.305115	0.313532
Homogeneity	0.952344	0.939217	0.902377

Table 4.3: Feature Extraction Average for Cluster 3.

	Basal cell carcinoma	Melanoma	Pigmented benign keratosis
Mean	42.1851	7.90844	64.8241
S.D	78.4251	32.591	70.7439
Entropy	2.04598	0.90853	3.93738
RMS	6.56724	3.4474	10.5548
Variance	3896.45	1003.86	4233.07
Kurtosis	2.91249	25.362	1.17882
Skewness	1.35413	4.65114	0.240369
Contrast	1.48366	0.252022	1.53315
Correlation	0.836205	0.82777	0.786831
Energy	0.582879	0.845924	0.309151
Homogeneity	0.943418	0.975533	0.901865

4.3 GRAPHICAL USER INTERFACE (GUI)

Figure (4.3), Figure (4.4), and Figure (4.5) depict a screenshot of the Graphical User Interface (GUI) for the proposed System that was created using MATLAB R2019b. It consists of three buttons that each perform a specific task for the application and allow for

both image display and area calculation results to be displayed in the same window.

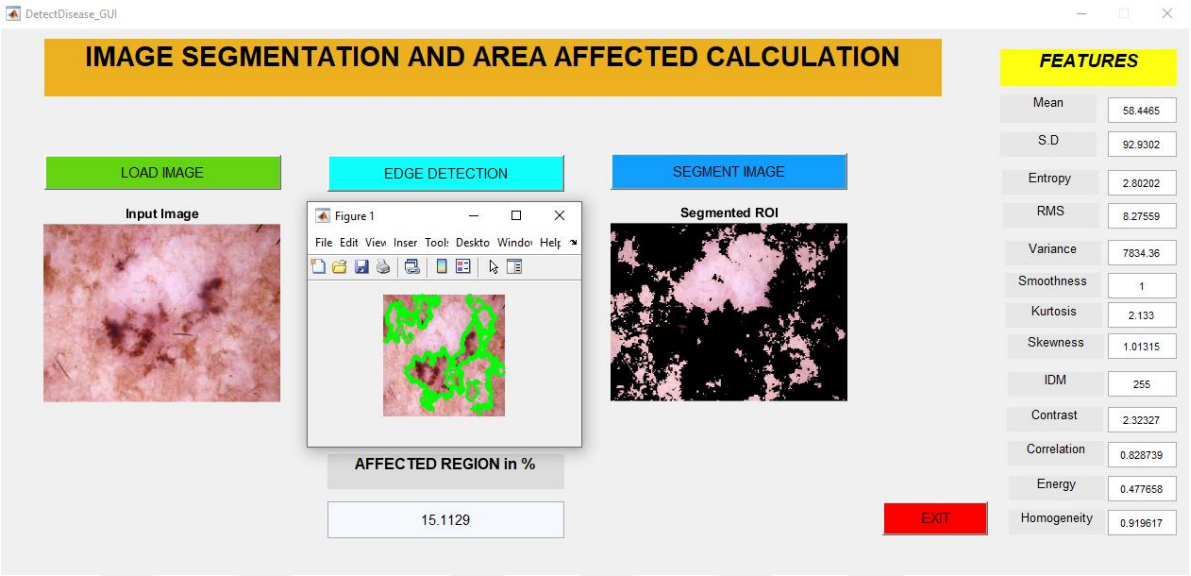


Figure 4.3: The Snapshot Of (GUI) For Basal Cell Carcinoma Disease.

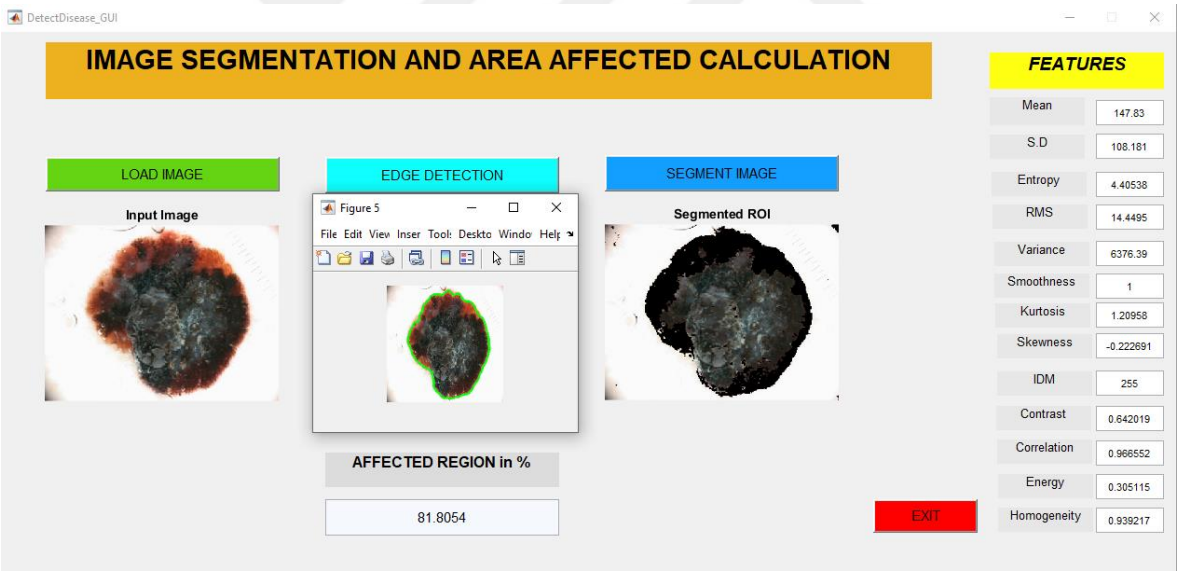


Figure 4.4: The Snapshot of (GUI) For Melanoma Disease.

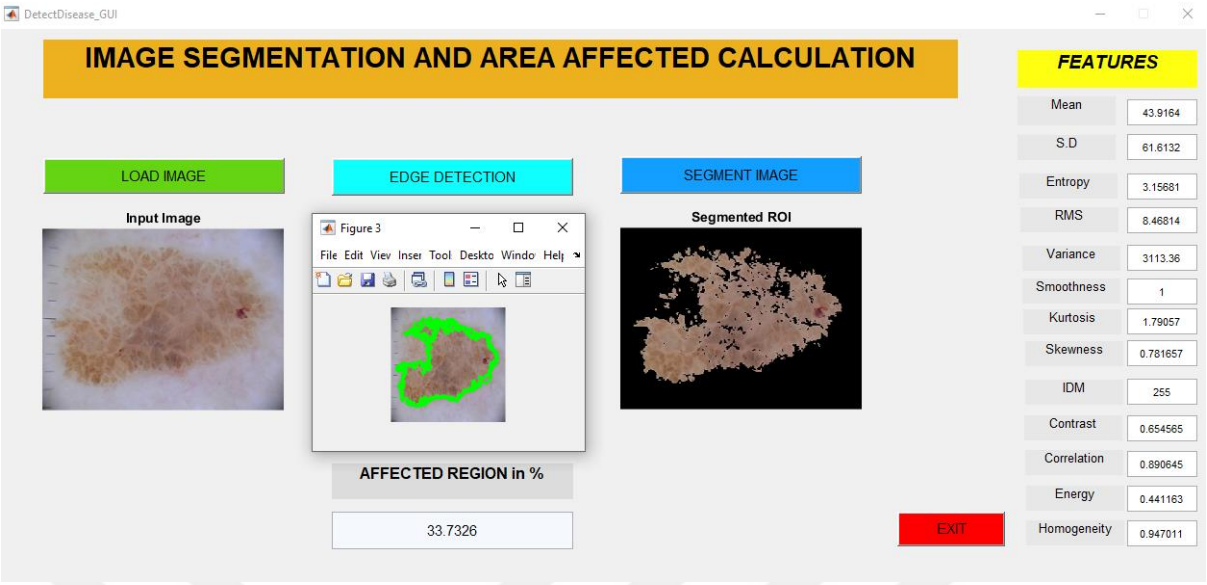


Figure 4.5: The Snapshot of (GUI) for Pigmented Benign Keratosis Disease.

4.4 EXPERIMENTAL RESULTS FOR DEEP LEARNING CLASSIFICATION

The algorithms DL AlexNet and ResNet are used to detect and classify skin cancer. The method of classification is described in section 3.3.

The first option used for the technique is stochastic gradient descent with momentum (SGDM), with a maximum epoch of 100, a minimum batch size of 6, and an initial learning rate of 1e-4. Figures (4.6) and (4.7) depict the AlexNet and ResNet training processes, with the blue line denoting training accuracy and the red line denoting training loss, respectively.

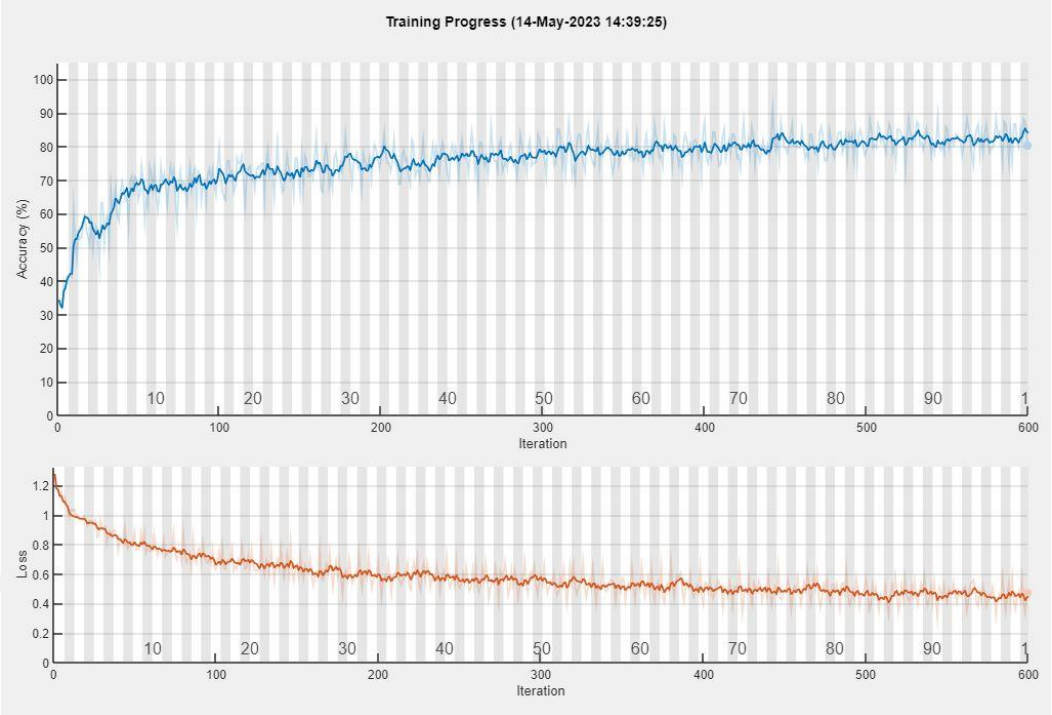


Figure 4.6: Alexnet Training Process.

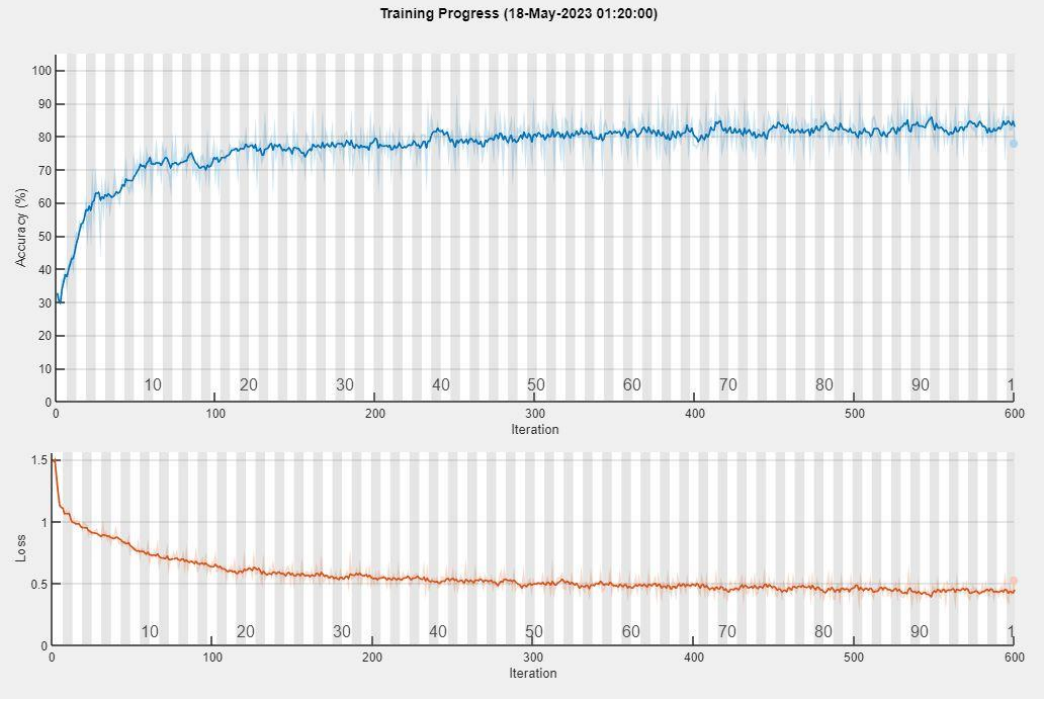


Figure 4.7: ResNet Training Process.

According to the two models' findings (table 4-4), AlexNet has the best training accuracy (99.3 %) and the shortest training time (14 min. 9 sec.).

Table 4.4: Training Accuracy for Each Model

Train / Model	Training Accuracy	Elapsed Time
AlexNet	99.3 %	14min 9sec
ResNet	97.2 %	31min 12sec

Figure (4.8) and (4.9) show the training accuracy of each model (AlexNet and ResNet).

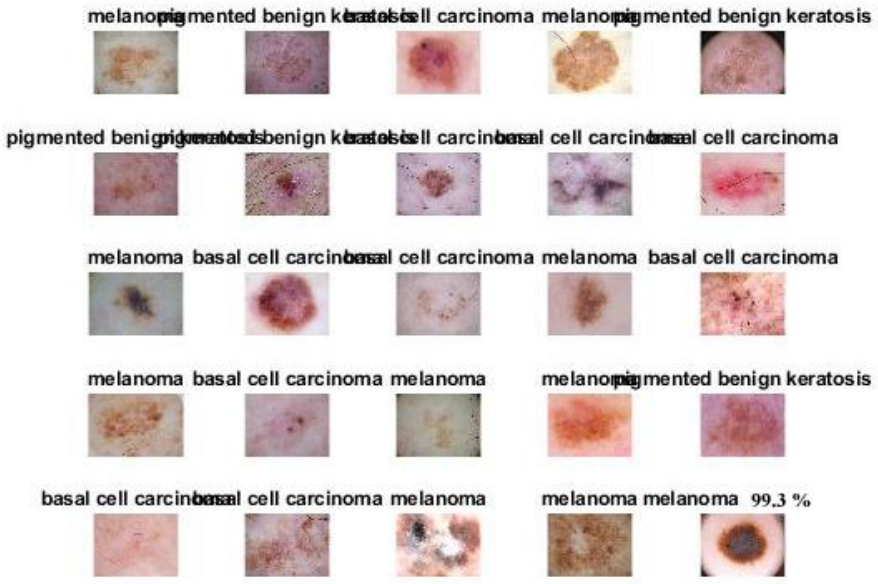


Figure 4.8: Training Accuracy of AlexNet.

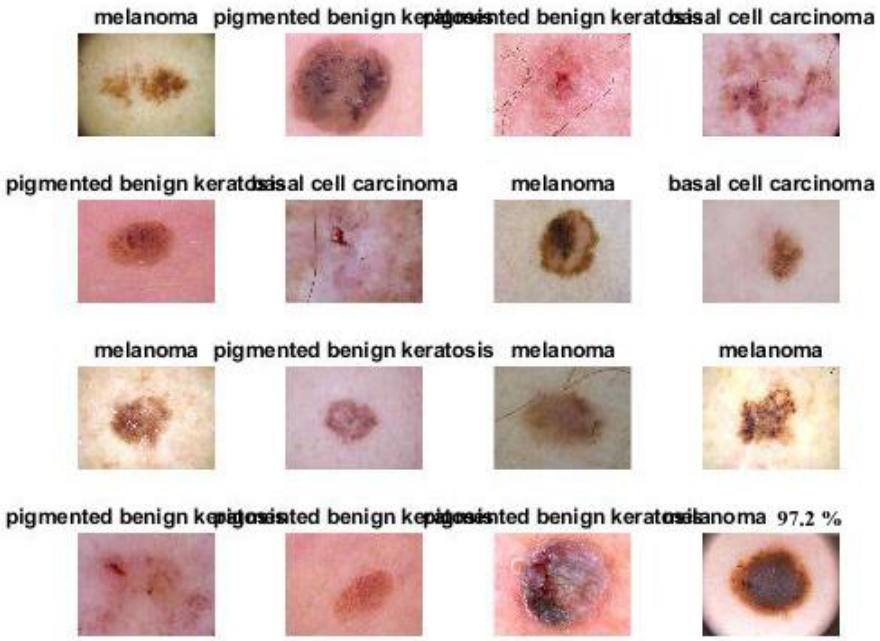


Figure 4.9: Training Accuracy of ResNet.

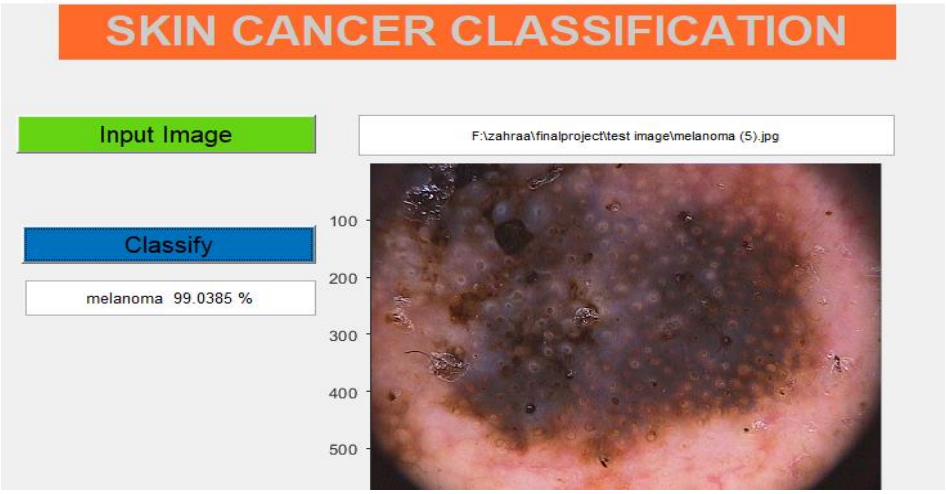


Figure 4.10: Classification Result for Melanoma.

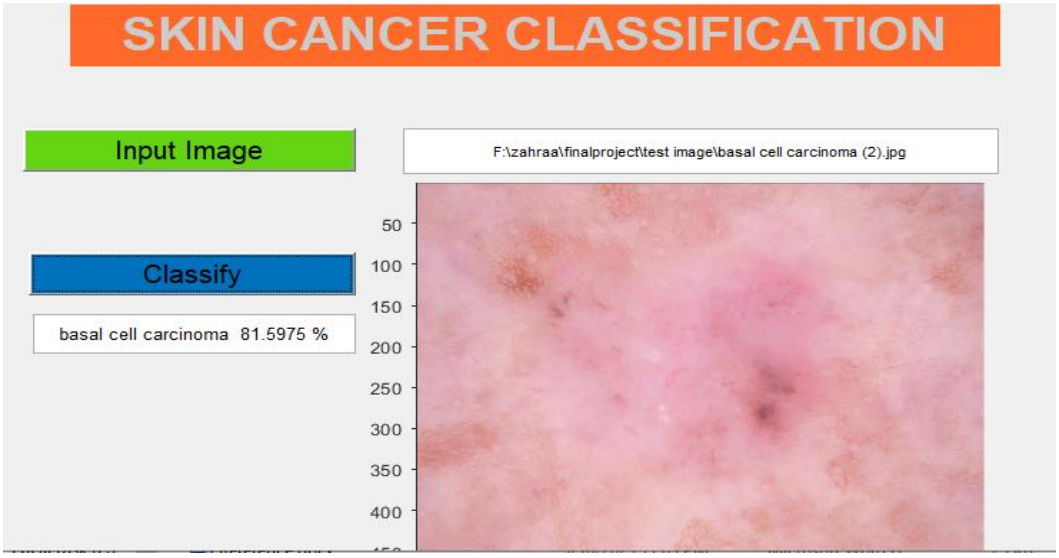


Figure 4.11: Classification Result For Basal Cell Carcinoma.

Figure 4.12: Classification Result for Pigmented Benign Keratosis.

5. CONCLUSION AND FUTURE WORK

5.1 CONCLUSION

As the world's death toll from skin cancer rises daily, this worldwide public health crisis must be addressed. A skin cancer diagnosis can benefit from deep convolution models' excellent picture dataset performance. However, a new problem requires a different approach when employing deep learning models.

The photos of skin cancer were detected and classified using machine learning and CNN approaches in the suggested work. The HAM10000 dataset was used for studies. Following the studies, the modified CNN and machine learning algorithms were assessed based on accuracy. The photos were pre-processed, divided into feature and target values, and then data augmentation before the training and testing stages. This indicates that the suggested CNN performs better in terms of classification for the HAM10000 data set. On the same data set, the suggested work has been compared to previous existing work, and it has been found to produce superior accuracy with fewer errors and losses.

With the ultimate objective of producing an AI-powered device that can identify skin cancers in real-time, deep learning algorithm-based algorithms are being created to help dermatologists in the prompt and accurate detection of skin malignancies. We spoke about several deep learning architectures for skin cancer diagnosis and especially focused on classifying skin tumors using deep learning algorithms. The performance and computational costs of many deep learning techniques covered by this research were compared in this survey study.

There are two algorithms. ML and DL are used to detect and classify three types of skin cancer known as Basal cell carcinoma, Melanoma, and Pigmented benign keratosis, as well as evaluate the area of disease detection in the skin. First, ML is used to improve and extract features using the K-Mean clustering and M-SVM classification algorithms. Then, using MATLAB R2019b, two DL models named Alex Net and Res Net are employed in different ways to calculate the accuracy of detecting and classifying diseases in the skin. The following conclusions can be drawn from simulation and testing results:

a. The Machine Learning system was effectively improved by applying the canny edge detection technique, segmented by the k-mean clustering approach, and classified by the MSVM method for the diseases Basal cell carcinoma, Melanoma, and Pigmented benign keratosis.

b. In ML, the classification is done via three-cluster segmentation.

c. Image processing techniques are used to efficiently determine the area of skin cancer diseases.

d. In ML, features are extracted manually by the user, but in DL, they are extracted automatically.

e. The AlexNet and ResNet models were used effectively in this project.

f. Increasing the maximum number of epochs improves model accuracy while increasing the predicted time for the model.

j. DL produces the greatest accuracy results for AlexNet and ResNet, with 99.3% and 97.2%, respectively.

5.2 FUTURE WORK

A future idea for developing the work is as follows:

a. Extending the data set images and skin cancer disease classifications, as well as extending the technology to different types of cancer.

b. Use transfer learning CNN to improve the performance of the suggested models.

c. Real-time skin disease identification using AlexNet and ResNet in Raspberry Pi with Python and OpenCV.

REFERENCES

[1] C. Karimkhani, R. P. Dellavalle, L. E. Coffeng, C. Flohr, R. J. Hay, S. M. Langan, E. O. Nsoesie, A. J. Ferrari, H. E. Erskine, J. I. Silverberg et al., "Global skin disease morbidity and mortality: an update from the global burden of disease study 2013," JAMA dermatology, vol. 153, no. 5, pp. 406–412, 2017.

[2] T. Gansler, P. A. Ganz, M. Grant, F. L. Greene, P. Johnstone, M. Ma-honey, L. A. Newman, W. K. Oh, C. R. Thomas Jr, M. J. Thun, et al., "Sixty years of ca: a cancer journal for clinicians," CA: a cancer journal for clinicians, vol. 60, no. 6, pp. 345–350, 2010.

[3] David A. Forsyth; and J. Ponce, COMPUTER VISION A MODERN APPROACH. University of Illinois at Urbana-Champaign, 2003.

[4] M. Mohammed, M. B. Khan, and E. B. M. Bashier, Machine Learning, no. December. Boca Raton : CRC Press, 2017.: CRC Press, 2016.

[5] M. Kang and N. J. Jameson, "Machine Learning: Fundamentals," in Prognostics and Health Management of Electronics, Chichester, UK: John Wiley and Sons Ltd, 2018, pp. 85–109.

[6] M. R. Minar and J. Naher, "Recent Advances in Deep Learning: An Overview," no. February, 2018.

[7] K. Li, J. Lin, J. Liu, and Y. Zhao, "Using deep learning for Image-Based different degrees of ginkgo leaf disease classification," Inf., vol. 11, no. 2, 2020.

[8] A. Y. A. Salama, A. E. Hassanien, and A. L. Y. Fahmy, "Sheep Identification Using a Hybrid Deep Learning and Bayesian Optimization Approach," IEEE Access, vol. 7, pp. 31681–31687, 2019.

[9] S. Vishvanathan, A. V. Vidyapeetham, S. Kp, and A. V. Vidyapeetham, "Breast Cancer Classification using Capsule Network with Preprocessed Histology Images," no. April, 2019.

[10] I. Namatēvs, "Deep Convolutional Neural Networks: Structure, Feature Extraction, and Training," Inf. Technol. Manag. Sci., vol. 20, no. 1, pp. 40–47, 2018.

[11] N. N. A. A. Hamid, R. A. Razali, and Z. Ibrahim, "Comparing bags of features, conventional convolutional neural network and alexnet for fruit recognition," Indones. J. Electr. Eng. Comput. Sci., vol. 14, no. 1, pp. 333–339, 2019.

[12] Y. Li, Z. Ding, C. Zhang, Y. Wang, and J. Chen, “SAR Ship Detection Based on Resnet and Transfer Learning,” IGARSS 2019 - 2019 IEEE Int. Geosci. Remote Sens. Symp., pp. 1188–1191, 2019.

[13] G. Kasinathan, S. Jayakumar, A. H. Gandomi, M. Ramachandran, S. J. Fong, and R. Patan, “Automated 3-d lung tumor detection and classification by an active contour model and cnn classifier,” Expert Systems with Applications, vol. 134, pp. 112–119, 2019.

[14] Z. Gao, L. Wang, L. Zhou, and J. Zhang, “Hep-2 cell image classification with deep convolutional neural networks,” IEEE Journal of biomedical and health informatics, vol. 21, no. 2, pp. 416–428, 2016.

[15] P. Wang, L. Wang, Y. Li, Q. Song, S. Lv, and X. Hu, “Automatic cell nuclei segmentation and classification of cervical pap smear images,” Biomedical Signal Processing and Control, vol. 48, pp. 93–103, 2019.

[16] K. M. Hosny, M. A. Kassem, and M. M. Foaud, “Classification of skin lesions using transfer learning and augmentation with alex-net,” PloS one, vol. 14, no. 5, p. e0217293, 2019.

[17] X. He, Z. Yu, T. Wang, B. Lei, and Y. Shi, “Dense deconvolution net: Multipath fusion and dense deconvolution for high-resolution skin lesion segmentation,” Technology and Health Care, vol. 26, no. S1, pp. 307–316, 2018.

[18] B. Harangi, “Skin lesion classification with ensembles of deep convolutional neural networks,” Journal of Biomedical Informatics, vol. 86, pp.25–32, 2018.

[19] T. J. Brinker, A. Hekler, A. H. Enk, J. Klode, A. Hauschild, C. Berking, B. Schilling, S. Haferkamp, D. Schadendorf, T. Holland-Letz, et al., “Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task,” European Journal of Cancer, vol. 113, pp. 47–54, 2019.

[20] L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. Heng, “Automated melanoma recognition in dermoscopy images via very deep residual networks,” IEEE Transactions on medical imaging, vol. 36, no. 4, pp. 994–1004, 2016.

[21] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S.Thrun, “Dermatologist-level classification of skin cancer with deep neural networks,” Nature, vol. 542, no. 7639, pp. 115–118, 2017.

[22] H. A. Haenssle, C. Fink, R. Schneiderbauer, F. Toberer, T. Buhl, A. Blum, A. Kalloo,

A. B. H. Hassen, L. Thomas, A. Enk et al., “Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists,” Annals of Oncology, vol. 29, no. 8, pp. 1836–1842, 2018.

[23] N. C. Codella, Q.-B. Nguyen, S. Pankanti, D. A. Gutman, B. Helba, A. C. Halpern, and J. R. Smith, “Deep learning ensembles for melanoma recognition in dermoscopy images,” IBM Journal of Research and Development, vol. 61, no. 4/5, pp. 5–1, 2017.

[24] M. Attia, M. Hossny, S. Nahavandi, and A. Yazdabadi, “Skin melanoma segmentation using recurrent and convolutional neural networks,” in 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI2017). IEEE, 2017, pp. 292–296.

[25] N. Nida, A. Irtaza, A. Javed, M. H. Yousaf, and M. T. Mahmood, “Melanoma lesion detection and segmentation using deep region-based convolutional neural network and fuzzy c-means clustering,” International journal of medical informatics, vol. 124, pp. 37–48, 2019.

[26] N. C. Codella, D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti, S. W. Dusza, A. Kalloo, K. Liopyris, N. Mishra, H. Kittler, et al., “Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (ISIC),” in 2018 IEEE 15th.

[27] J. Burdick, O. Marques, J. Weinthal, and B. Furht, “Rethinking skin lesion segmentation in a convolutional classifier,” Journal of digital imaging, vol. 31, no. 4, pp. 435–440, 2018.

[28] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in International Conference on Medical image computing and computer-assisted intervention. Springer, 2015, pp. 234–241.

[29] M. Z. Alom, C. Yakopcic, M. Hasan, T. M. Taha, and V. K. Asari, “Recurrent residual u-net for medical image segmentation,” Journal of Medical Imaging, vol. 6, no. 1, p. 014006, 2019.

[30] I Pirnog, C. Oprea, Proceedings of 11th International Symposium on Electronics and Telecommunications, Institute of Electrical and Electronics Engineers Inc (2015), 10.1109/ISETC.2014.7010796

[31] MQ Khan, A Hussain, SU Rehman, et al., “Classification of Melanoma and Nevus in

Digital Images for Diagnosis of Skin Cancer, ” IEEE Access, 7 (2019), pp. 90132-90144, 10.1109/ACCESS.2019.2926837

[32] CH O'Neill, CR. Scoggins, Melanoma, J Surg Oncol, 120 (5) (2019), pp. 873-881, 10.1002/jso.25604

[33] Korhonen N., Ylitalo L., Luukkaala T., Itkonen J., Häihälä H., Jernman J., Snellman E., Palve J. Premalignant lesions, basal cell carcinoma and melanoma in patients with cutaneous squamous cell carcinoma. *Arch. Dermatol. Res.* 2020;313:879–884. doi: 10.1007/s00403-020-02114-w.

[34] Rogers HW, Weinstock MA, Feldman SR, Coldiron BM. Incidence Estimate of Nonmelanoma Skin Cancer (Keratinocyte Carcinomas) in the U.S. Population, 2012. *JAMA Dermatol*. 2015; **151**: 1081- 1086.

[35] Medical and Editorial Content Team ACS. Key Statistics for Basal and Squamous Cell Skin Cancers: American Cancer Society; 2016. https://www.cancer.org/cancer/basal-and-squamous-cellskin-cancer/about/key-statistics.html. Accessed October 12, 2017.

[36] M Ishida, N Ohsato, H Okabe. Basal cell carcinoma arising within a seborrheic keratosis with respect to immunohistochemical characteristics. *Oncol Lett.* 2011;2:625–27.

[37] C Lim. Seborrhoeic keratoses with associated lesions: a retrospective analysis of 85 lesions. *Australas J Dermatol.* 2006;47:109–13.

[38] V. Hlavac, “Fundamentals of Image Processing,” Opt. Digit. Image Process. Fundam. Appl., no. July, pp. 71–96, 2011.

[39] J. Z. & V. Huiyu Zhou, Jiahua Wu, Digital Image Processing – Part I, vol. 137, no. 4. 2010.

[40] H. Hassanpour, N. Samadiani, and S. M. Mahdi Salehi, “Using morphological transforms to enhance the contrast of medical images,” Egypt. J. Radiol. Nucl. Med., vol. 46, no. 2, pp. 481–489, 2015.

[41] Zherdeva LA, Bratchenko IA, Myakinin OO, Moryatov AA, Kozlov SV, Zakharov VP. In vivo hyperspectral imaging and differentiation of skin cancer. Optics in Health Care and Biomedical Optics VII. Vol 10024. 2016:100244G. doi:10.1117/12.2246433

[42] Jiang J, Ma J, Chen C, Wang Z, Cai Z, Wang L. SuperPCA: a superpixelwise PCA approach for unsupervised feature extraction of hyperspectral imagery. IEEE Trans Geosci Remote Sens. 2018; 56(8): 4581- 4593. doi:10.1109/TGRS.2018.2828029

[43] R. Gonzalez and R. Woods Digital Image Processing, Addison-Wesley Publishing

Company, 1992, pp 524, 552.

[44] R. Haralick and L. Shapiro Computer and Robot Vision, Vol. 1, Addison-Wesley Publishing Company, 1992, pp 174 - 185.

[45] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once : Unified, Real-Time Object Detection," 2016.

[46] Q. S. Kalra M., Lal N., K-Mean Clustering Algorithm Approach for Data Mining of Heterogeneous Data. Information and Communication Technology for Sustainable Development., Springer, Singapore, 2018.

[47] J. Yadav and M. Sharma, "A Review of K-mean Algorithm," Int. J. Eng. Trends Technol., vol. 4, no. 7, pp. 2972–2976, 2013.

[48] G. Kumar, "A Detailed Review of Feature Extraction in Image Processing Systems," pp. 5–12, 2014.

[49] Ş. Öztürk and B. Akdemir, "Application of Feature Extraction and Classification Methods for and GLCM , Histopathological Image using a SFTA," Procedia Comput. Sci., vol. 132, no. Iccids, pp. 40–46, 2018.

[50] P. S. Veling, "Mango Disease Detection by using Image Processing," Int. J. Res. Appl. Sci. Eng. Technol., vol. 7, no. 4, pp. 3717–3726, 2019.

[51] T. Beysolow II, Introduction to Deep Learning Using R: A Step-by-Step Guide to Learning and Implementing Deep Learning Models Using R. 2017.

[52] A. Y. A. Salama, A. E. Hassanien, and A. L. Y. Fahmy, "Sheep Identification Using a Hybrid Deep Learning and Bayesian Optimization Approach," IEEE Access, vol. 7, pp. 31681–31687, 2019.

[53] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis Deep learning models for plant disease detection and diagnosis," Comput. Electron. Agric., vol. 145, no. February, pp. 311–318, 2018.

[54] U. Michelucci, Applied Deep Learning. Berkeley, CA: A Case-Based Approach to Understanding Deep Neural Networks Apress, Umberto Michelucci toelt.ai, Dübendorf, Switzerland 2018.

[55] A. Ahmad, "Object Recognition in 3D data using Capsules," thesis at Syracuse University 2018.

[56] A. Sacré, "Efficient Interactive Annotation for Cytomine," Université de Liège, Liège, Belgique, 2019.

[57] A. F. Gad, Practical Computer Vision Applications Using Deep Learning with CNNs Applications Using Deep Learning with CNNs. Springer Science+Business Media New York, 2018.

[58] E. Cengİl, “Multiple Classification of Flower Images Using Transfer Learning,” 2019 Int. Artif. Intell. Data Process. Symp. (IDAP)., 2019.

[59] F. S. Jing Sun, Jianguo Zhang, “Scene image classification method based on Alex-Net model,” 2016 3rd Int. Conf. Inf. Cybern. Comput. Soc. Syst., 2016.

[60] P. Dönmez, “Introduction to Machine Learning, 2nd ed.,” Nat. Lang. Eng., vol. 19, no. 2, pp. 285–288, Apr. 2013.

[61] N. Dhanda, S. S. Datta, and M. Dhanda, “Machine Learning Algorithms,” no. May, 2019, pp. 210–233.

[62] Michael W. Berry, A. Mohamed, and B. W. Yap, “Supervised and Unsupervised Learning for Data Science,” Springer Nat. Switz., no. January, 2020.

[63] T. O. Ayodele, “Types of Machine Learning Algorithms,” IntechOpen, 2010, p. 366.

[64] A. Dey and A. S. Learning, “Machine Learning Algorithms : A Review,” vol. 7, no. 3, pp. 1174–1179, 2016.

[65] M. Kamalahasan, “Review on Classification Algorithms in Image,” Int. J. Innov. Trends Eng. Res., vol. 2, no. November, 2017.

[66] R. Don and D. P. W, “Multiclass Classification Using Support Vector Machines,” 2018.

[67] S. R. Krishnan, A. ArulS, and Sivakumari, “Study on Machine Learning Techniques for Anomaly Detection,” Conf. Mach. Learn. its Appl. Coimbatore, 2020.

[68] N. a Ibraheem, M. M. Hasan, R. Z. Khan, and P. K. Mishra, “Understanding Color Models : A Review,” ARPN J. Sci. Technol., vol. 2, no. 3, pp. 265–275, 2012.

[69] H. Kour, “Analysis on Image Color Model,” Int. J. Adv. Res. Comput. Commun. Eng., vol. 4, no. 12, pp. 233–235, 2015.

[70] I. Rynning, “Hue, Saturation and Value (HSV) Image Manipulation and Evaluation Through Numerical Techniques: A Review of the Tools Thereby,” 1993.

[71] Al-Masni, M.A.; Kim, D.H.; Kim, T.S. Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification. Comput. Methods Programs Biomed. 2020, 190, 105351.

[72] Barata, C.; Marques, J.S.; Emre Celebi, M. Deep attention model for the hierarchical

diagnosis of skin lesions. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Workshops (CVPRW), Long Beach, CA, USA, 16–17 June 2019; pp. 2757–2765.