



**Department of Computer Science and Engineering**

**Dhaka University of Engineering & Technology, Gazipur**

**SKIN CANCER DETECTION UTILIZING INTENSITY  
VALUE ESTIMATION MODEL WITH A DEEP NEURAL  
NETWORK**

**AUTHOR**

N. I. Md. Ashafuddula

Roll No. 18204016

Department of Computer Science and Engineering  
Dhaka University of Engineering & Technology, Gazipur  
Gazipur, Bangladesh

**SUPERVISED BY**

Dr. Rafiqul Islam

Professor

Department of Computer Science and Engineering  
Dhaka University of Engineering & Technology, Gazipur  
Gazipur, Bangladesh

# **SKIN CANCER DETECTION UTILIZING INTENSITY VALUE ESTIMATION MODEL WITH A DEEP NEURAL NETWORK**

N. I. Md. Ashafuddula

A thesis in fulfilment of the requirements for the degree  
of  
Master of Engineering in Computer Science and Engineering



**Department of Computer Science and Engineering  
Dhaka University of Engineering & Technology, Gazipur  
Gazipur, Bangladesh**

## **SUPERVISED BY**

Dr. Rafiqul Islam  
Professor

Department of Computer Science and Engineering  
Dhaka University of Engineering & Technology, Gazipur  
Gazipur, Bangladesh

## **DECLARATION**

It is herewith proclaimed that this thesis work is my original work. The exploration presented in this thesis or any part of this thesis has not been submitted elsewhere for the award of any degree or diploma. Knowledge derived from the other works has been acknowledged with appropriate referencing.

Signatures:



.....

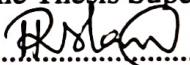
(N. I. Md. Ashafuddula)

Student ID: 18204016

Department of Computer Science and Engineering

Dhaka University of Engineering & Technology, Gazipur

Signature of the Thesis Supervisor



.....

**DR. RAFIQUL ISLAM**

Professor

Department of Computer Science and Engineering

Dhaka University of Engineering & Technology, Gazipur.

*This is to certify that the thesis entitled "SKIN CANCER DETECTION UTILIZING INTENSITY VALUE ESTIMATION MODEL WITH A DEEP NEURAL NETWORK" submitted by N. I. Md. Ashafuddula, Roll 18204016 has been accepted as satisfactory in fulfillment of the requirement for the award of Master of Science in Computer Science and Engineering from Dhaka University of Engineering & Technology, Gazipur, Bangladesh on January 30, 2023.*

**BOARD OF EXAMINERS**

(Dr. Rafiqul Islam)

1. Professor  
Department of Computer Science and Engineering  
Dhaka University of Engineering & Technology, Gazipur  
Gazipur-1707, Bangladesh.

Chairman  
(Supervisor)

2. Department of Computer Science and Engineering  
Dhaka University of Engineering & Technology, Gazipur  
Gazipur-1707, Bangladesh.

Member  
(Ex-Officio)

(Dr. Fazlul Hasan Siddiqui)

3. Professor  
Department of Computer Science and Engineering  
Dhaka University of Engineering & Technology, Gazipur  
Gazipur-1707, Bangladesh.

Member

(Dr. Md. Shafiqul Islam)

4. Associate Professor  
Department of Computer Science and Engineering  
Dhaka University of Engineering & Technology, Gazipur  
Gazipur-1707, Bangladesh.

Member

(Dr. Muhammad Shahin Uddin )

5. Professor  
Department of Information and Communication Technology  
Mawlana Bhashani Science and Technology University  
Tangail-1902, Bangladesh.

External Member

## **ACKNOWLEDGEMENT**

The real spirit of achieving the goal is through the way of excellence and discipline. I would never succeed to complete my task without the cooperation, encouragement, and help provided to me by various personalities.

First of all, I render my humble gratitude to my almighty Allah for his immense grace and blessings to me to successfully complete this work with noticeable results.

With a deep sense of gratitude, I express my sincere thanks to my supervisor, **Rafiqul Islam, Ph.D., Professor**, Department of Computer Science and Engineering, Dhaka University of Engineering & Technology (DUET), Gazipur, Bangladesh, for his invaluable supervision, continuous support, and patience during my study. With all his profound knowledge and extensive experience, he guided and encouraged me throughout my MSc. Engg. research. His evaluation and editorial observations were also invaluable in formulating this thesis.

I further extend my thanks to all my teachers, friends, and classmates in my department, who have shared their valuable opinions and experiences that enabled me to enhance my knowledge regarding this work. Finally, yet importantly, I would like to express my heartfelt gratitude to my beloved parents and siblings for their understanding and endless support, encouragement and blessings in accomplishing this work.

**January 30, 2023  
DUET, Gazipur**

**N. I. Md. Ashafuddula**

## ABSTRACT

Melanoma skin cancer is among the most hazardous and life-intimidation cancers in the midst of many skin diseases. Exposure to ultraviolet (UV) rays may impair the skin's DNA cells, which could cause possible Melanoma skin cancer. Nevertheless, the detection and classification of Melanoma skin cancer at a rudimentary stage is complicated due to its irregular nature. To aid the physicians here, we developed an automatic deep learning-based diagnosis system by incorporating the intensity value estimation (IVE) with a deep neural network model (DNN) to detect and classify Melanoma and Nevus mole more precisely. Since the diverse level of intensities is the most attributable feature for the region of interest (ROI) or object identification, the high-intensity pixel weights are imposed from the segmented lesion images. Embedding those high-intensity features into the DNN intensifies the overall performance of the suggested model than the state-of-the-art methods in Melanoma skin cancer detection. Various methodologies have been studied for Melanoma skin cancer detection and categorizing, particularly at an early stage. The problems and scopes in conventional techniques with their proposed methodology are identified. Then an Intensity Value Estimation (IVE) model is developed to retain important features in the detection and classification of Melanoma skin cancer which is embedded with Deep Neural Network architecture. In this work, we proposed an artificially intelligent model by embedding the novel IVE method with a DNN model and named it the IVEwDNN model. We have taken a widely known MED-NODE macroscopic images dataset contributed by the digital image library of the Dermatology of the University Medical Center Groningen (UMCG) department. Then the classification model was enriched by embedding the IVE model with different transfer learning approaches. We concluded that our newly developed method, IVE with VGG-16 transfer learning (IVEwVGG-16 TFL) based deep learning model performed best in Melanoma skin cancer identification and classifying from Nevus mole. The model achieved the best average test accuracy of 94.93% compared to the convolutional neural network (CNN) model accuracy of 90.66% and inception v3 transfer learning models accuracy of 93.75%. Moreover, considering the popular model evaluation metrics, all the DNN models with intensity value estimation surpassed the state-of-the-art methods.

# Contents

<b>Declaration</b>	<b>i</b>
<b>Declaration</b>	<b>ii</b>
<b>Acknowledgement</b>	<b>iii</b>
<b>Abstract</b>	<b>iv</b>
<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Problem statement and motivation . . . . .	2
1.3 Objectives . . . . .	2
1.4 Contributions . . . . .	3
1.5 Organization of the thesis . . . . .	3
1.6 Publications . . . . .	4
<b>2 Literature Study</b>	<b>5</b>
2.1 Introduction . . . . .	5
2.2 Imaging in Dermatology . . . . .	5
2.2.1 Digital photography . . . . .	7
2.2.2 Total body photography . . . . .	7
2.2.3 Dermoscopy . . . . .	7
2.2.4 Reflectance Confocal Microscopy . . . . .	9
2.2.5 Optical Coherence Tomography . . . . .	9

2.3	Classification . . . . .	10
2.3.1	Supervised Learning . . . . .	10
2.3.2	Pros and Cons supervised learning . . . . .	12
2.3.3	Unsupervised Learning . . . . .	12
2.3.4	Pros and Cons unsupervised learning . . . . .	12
2.4	Pre-processing . . . . .	13
2.4.1	Image Enhancement . . . . .	13
2.4.2	Grayscale Conversion . . . . .	14
2.4.3	Image Resizing . . . . .	14
2.4.4	Data Augmentation . . . . .	14
2.5	Segmentation . . . . .	15
2.5.1	Cluster based Segmentation . . . . .	16
2.5.2	Thresholding based Segmentation . . . . .	16
2.5.3	Edge Detection . . . . .	17
2.5.4	Watershed Segmentation . . . . .	17
2.6	Image Features . . . . .	18
2.7	Histogram . . . . .	19
2.8	Deep Neural Network . . . . .	19
2.8.1	Convolution Neural Network . . . . .	20
2.8.2	Transfer Learning . . . . .	21
2.8.2.1	DenseNet201 . . . . .	22
2.8.2.2	ResNet . . . . .	22
2.8.2.3	Inception-v3 . . . . .	23
2.8.2.4	VGG . . . . .	23
2.9	Deep Learning for Medical Imaging Analysis . . . . .	23
2.10	Melanoma Skin Cancer Detection . . . . .	24
2.11	Discussion . . . . .	27
<b>3</b>	<b>Proposed Classification using Convolutional Neural Network</b>	<b>29</b>
3.1	Introduction . . . . .	29
3.2	Proposed Methodology . . . . .	29
3.2.1	Pre-processing . . . . .	30
3.2.2	Segmentation . . . . .	31

3.2.3	Data augmentation . . . . .	34
3.2.4	Lesion Feature Analysis . . . . .	35
3.2.4.1	Canny Edge Detection . . . . .	35
3.2.4.2	Proposed Intensity Value Estimation . . . . .	37
3.2.5	Deep Neural Network . . . . .	38
3.2.5.1	Convolutional Neural Network . . . . .	39
3.3	Dataset . . . . .	41
3.4	Experimental Analysis . . . . .	43
3.4.1	Training and Testing . . . . .	44
3.4.2	Performance evaluation metrics . . . . .	44
3.5	Experimental Results . . . . .	46
3.6	Discussion . . . . .	48
3.7	Conclusion . . . . .	48
<b>4</b>	<b>Enhanced Classification using Transfer Learning</b>	<b>50</b>
4.1	Introduction . . . . .	50
4.2	Transfer Learning Model . . . . .	50
4.2.1	Transfer Learning using VGG-16 . . . . .	51
4.2.2	Transfer Learning using Inception v3 . . . . .	52
4.3	Performance Evaluation . . . . .	53
4.3.1	IVE with VGG-16 based Transfer Learning Algorithm . . . . .	53
4.3.2	IVE with Inception V3 Transfer Learning Algorithm . . . . .	55
4.3.3	Discussion . . . . .	57
4.4	Conclusion . . . . .	57
<b>5</b>	<b>Experimental Analysis</b>	<b>59</b>
5.1	Introduction . . . . .	59
5.2	Analysis . . . . .	59
5.3	Scope and data sensitivity . . . . .	61
5.4	Conclusion . . . . .	65
<b>6</b>	<b>Conclusion and Future works</b>	<b>66</b>
6.1	Conclusion . . . . .	66
6.2	Limitations . . . . .	67

6.3 Future Works . . . . .	67
<b>References</b>	<b>68</b>

# List of Figures

2.1	Sample dermatology imaging technology in skin cancer diagnosis . . . . .	6
2.2	Attached mobile dermatoscope allows to simultaneous dermoscopic assessment and photograph capture . . . . .	8
2.3	Vivascope Reflectance Confocal Microscope System . . . . .	10
2.4	Deep Neural Network-based image classification model . . . . .	11
2.5	The working of supervised learning algorithm . . . . .	11
2.6	The working of unsupervised learning algorithm . . . . .	12
2.7	Sample augmentations of a sample image: rotation, zooming and flipping .	15
2.8	General image segmentation techniques . . . . .	16
2.9	Difference between Machine Learning model to Deep Neural Network . .	20
2.10	Convolutional Neural Network (CNN) general architecture . . . . .	21
2.11	Transfer Learning process . . . . .	22
2.12	Melanoma skin cancer diagnostic process . . . . .	27
3.1	Proposed model to detect and classify Melanoma skin cancer and Nevus mole	30
3.2	Image analysis on BSC corrected gray-image using Otsu segmentation . .	32
3.3	Skin lesion data augmentation sample image . . . . .	35
3.4	Edge Tracking by Hysteresis . . . . .	37
3.5	Melanoma and Nevus lesion Histogram comparison . . . . .	38
3.6	Proposed IVEwCNN model workflow . . . . .	40
3.7	Understanding convolution with $(3 \times 3)$ image and $(2 \times 2)$ filter . . . .	40
3.8	Process of convolution with $(3 \times 3)$ image and $(2 \times 2)$ filter . . . . .	41
3.9	Dataset sample images (Melanoma skin cancer) . . . . .	43
3.10	Dataset sample images (Nevus mole) . . . . .	43

3.11	Average performance analysis of the proposed IVE and CED model embedded with CNN for Melanoma detection and classification . . . . .	48
4.1	Architecture of the Proposed IVEwVGG-16 Transfer Learning Model . . . . .	51
4.2	Architecture of the Proposed IVEwIV3 Transfer Learning Model . . . . .	53
4.3	Average performance analysis of the proposed IVE and CED model embedded with VGG-16 transfer learning for Melanoma detection and classification . . . . .	55
4.4	Average performance analysis of the proposed IVE and CED model embedded with Inception V3 transfer learning for Melanoma detection and classification . . . . .	57
5.1	Melanoma Skin-Cancer images (a) Melanoma skin lesion (b) Resized segmented lesion (c) Canny Edge detection (d) IVE model output . . . . .	60
5.2	Nevus Mole Images (a) Nevus skin lesion (b) Resized segmented lesion (c) Canny Edge detection (d) IVE model output . . . . .	60
5.3	Average performance analysis for CNN, IV3 transfer learning and VGG-16 transfer learning for Melanoma detection and classification . . . . .	62
5.4	Proposed IVEwVGG-16 TFL model's (a) training-validation loss (b) training-validation accuracy (c) AUC-ROC curve for 4056 data . . . . .	64

# List of Tables

3.1	Architecture of CNN Layers. . . . .	42
3.2	5-Fold cross-validation experimental result(%) analysis using CEDwCNN model (best accuracy shown in <b>bold</b> ) . . . . .	47
3.3	5-Fold cross-validation experimental result(%) analysis using IVEwCNN model (best accuracy shown in <b>bold</b> ) . . . . .	47
4.1	5-Fold cross-validation experimental result(%) analysis using using CEDwVGG-16 transfer learning model (best accuracy shown in <b>bold</b> ) . . . . .	54
4.2	5-Fold cross-validation experimental result(%) analysis using using IVEwVGG-16 transfer learning model (best accuracy shown in <b>bold</b> ) . . . . .	54
4.3	5-Fold cross-validation experimental result(%) analysis using using CED-wIV3 transfer learning model (best accuracy shown in <b>bold</b> ) . . . . .	56
4.4	5-Fold cross-validation experimental result(%) analysis using using IVEwIV3 transfer learning model (best accuracy shown in <b>bold</b> ) . . . . .	56
5.1	Experimental result(%) evaluation of the Proposed Methodology with the State-of-the-art Methods . . . . .	61
5.2	Average experimental result(%) comparison between CNN, VGG-16 TFL and Inception V3 TFL model (best results are shown in <b>bold</b> ) . . . . .	62
5.3	Result (%) sensitivity through the amount of data variation for the best-proposed model IVEwVGG-16 TFL (best results are indicated as bold) . .	63

# **Chapter 1**

## **Introduction**

---

### **1.1 Introduction**

The skin encloses the inner body organs and defends them from any exterior injury. However, due to its locality, it often suffers from some diseases that could lead to life-threatening conditions. Mainly three categories of skin cancer (i.e., Basal Cell Carcinoma, Squamous Cell Carcinoma, and Melanoma) are considered the most dangerous. Among them, Melanoma is the most dangerous and life-threatening one [1]. It rapidly develops and extends to the other body areas. The melanocyte cells could turn into Melanoma. These cells could also aid in developing benign skin tumors. Though nearly all moles (nevi) are harmless, however, some moles can still rise the risk of Melanoma. On the report of the American Cancer Society, [2], in 2023, about 97,610 new Melanomas will be discovered, and almost 7,990 people are anticipated to be passed away from Melanoma. The enormous number of Melanoma infected patients could burden the healthcare system. Nevertheless, researchers [3] conclude that if the Melanoma category is identified at an early stage, the five years of patient survival rate could be ameliorated.

## 1.2 Problem statement and motivation

Owing to the intricate nature of the derm lesions, dermatologists use several visual observations, such as symmetry of the lesion area, size, shape, color, and border, to diagnose malignant Melanoma [4]. The ugly duckling lesion evidence is another caution of Melanoma [5]. Medical specialists use popular scoring methods to identify malignant Melanoma, such as a 7-point checklist (7PCL), the Menzies method, a 3-point checklist, and ABCDE rules (lesion area, border, color, diameter, and evolving) [6, 5]. This conventional detection process is time-consuming and needs specialist doctors to diagnose [7]. Since this detection procedure highly depends on the medico specialization, that may influence certain inconsistent diagnosing outcomes. Studies found Melanoma mole differs from Nevus one in rough and uneven edges. Hence, researchers utilized the border (B) detection approach for Melanoma identification [8, 9, 10]. However, they lack in addressing other Melanoma detection features in color (C), and evolving (E) form and the significance of resizing the high-resolution images to low resolution while keeping the lesion shape as the origin.

In this thesis, we aim to overcome these shortcomings by utilizing intensity values in lesion areas as Intensity levels are crucial for identifying objects or regions of interest [11]. Finally, an automatic intelligent system is developed to predict and classify Melanoma from the Nevus mole early, considering the complicated circumstance. Our novel approach employs pre-processing steps, an intensity value estimation (IVE) technique, and a Deep Neural Network (DNN) model.

## 1.3 Objectives

Our objective is to design an algorithm for identifying and categorizing Melanoma skin cancer and Nevus mole. The key goals of our thesis are as follows:

- To develop an Intensity Value Estimation (IVE) model
- To embed the IVE model with deep neural network architecture
- To apply the proposed algorithm for early-stage skin cancer detection

## **1.4 Contributions**

To sum up, the noticeable contributions of this thesis can be enlisted as follows:

- Developing an Intensity Value Estimation (IVE) model.
- Presenting an effective method for resizing images while preserving lesion shape and minimizing the possible data loss.
- Designing the proposed methodology by combining the image pre-processing techniques, resizing process, IVE model, and a Deep Neural Network (IVEwDNN) for automatic categorization and detection of Melanoma skin cancer and Nevus mole.
- Performing a comparative histogram analysis between our newly developed method and other state-of-the-art strategies to determine the efficacy of our newly developed model.

## **1.5 Organization of the thesis**

This thesis report consists of six chapters and those are organized as follows:

### **Chapter 1: Introduction**

This chapter introduces life-threatening skin cancer, and finally, the problem statement and motivation, objectives, and contributions of this work are discussed.

### **Chapter 2: Background**

This chapter discusses the background of Melanoma skin cancer, imaging in Dermatology, pre-processing, segmentation, lesion features, histogram analysis, existing skin image classification models and methods, Deep Neural Network models, and some basic knowledge of this thesis.

### **Chapter 3: Proposed Classification using Convolutional Neural Network**

This chapter briefly discusses our proposed novel method (i.e., Intensity Value Estimation (IVE) model) to efficiently predict Melanoma at a primary phase with a DNN model (i.e., CNN). Then the experimental setups, such as datasets, training-testing data organization, result acquisition, and model evaluation metrics, are discussed.

## **Chapter 4: Enhanced Classification using Transfer Learning**

Here, we discuss the scopes and conveniences of a transfer learning model in skin cancer detection by incorporating our proposed IVE model with a pre-trained DNN model such as VGG-16 or Inception V3.

## **Chapter 5: Experimental Analysis**

This chapter discusses the experimental results of the IVEwDNN models and compares them in order to find the top-performed model in skin cancer detection and classification. Then the best DNN model is assimilated with the state-of-the-art methodologies to analyze the significance of the work. We also analyze the scope and sensitivity of data diversity in the proposed deep learning model.

## **Chapter 6: Conclusion and Future work**

Here we discuss the conclusion, limitations of this thesis, and possible future works.

## **1.6 Publications**

1. N. I. Md. Ashafuddula and Rafiqul Islam, "Melanoma skin cancer and Nevus mole classification using intensity value estimation with convolutional neural network." Computer Science, AGH University of Science and Technology, ISSN: 1508-2806, 2022, forthcoming.
2. N. I. Md. Ashafuddula and Rafiqul Islam, "Transfer Learning in Deep Neural Network for Skin Cancer Detection and Classification", Submitted.

# **Chapter 2**

## **Literature Study**

---

### **2.1 Introduction**

Various methods have been suggested in recent days for the identification and categorization of skin cancer. The impressive idea has evolved from these proposed methods. Before starting our proposed method in this chapter, we demonstrated the inevitable elements of this thesis work, such as skin imaging techniques, image classification models, feature extractions, and the DNN models (i.e., Convolutional Neural Network, Transfer Learning), along with general image classification methodology.

### **2.2 Imaging in Dermatology**

Medical imaging technology has considerably reorganized the practice of medicine, particularly dermatology, to diagnose, counsel, monitor, and treat skin diseases. The highly visual characteristics of dermal diseases make digital imaging the utmost convenience in everyday operations to simplify the data transfer between providers, estimate response to therapy, and play a crucial part in observing and examining skin cancer [12]. Imaging accelerates the alteration of data between medics as well as patients. Imaging technologies

make it feasible to record and circulate an enormous amount of clinical information in a single image. Technological progress in imaging has enhanced our capacity to image and document clinical discoveries digitally. Diagnosis of skin cancer early and more accurately is one critical application of these technologies [12]. Focusing on imaging technology for skin cancer diagnosis, we provided an overview of contemporary imaging technologies employed in dermatology.

Dermatologists operate imaging in their daily exercise to establish the presence of dermal disease. The dermal photographs may be used to differentiate disease phases before and after treatment to assess effectiveness, recall the specific position of lesions after the examination (i.e., biopsy or excision) and or monitor the development of the disease state over time. Imaging in dermatology is precious for identifying benign and malignant lesions that are crucial in resolving medicolegal issues and practice development. Some popular imaging techniques are shown in sample imaging Fig. 2.1.

The significance of an alteration lesion (evolution [E]) was portrayed by [13] in 2004, underrating the importance of precisely recording variations in pigmented lesions. Highlighting the necessity for ameliorated imaging technology to recognize and observe questionable lesions, the number of biopsies needed may be as high as 80:1 in particular populations to capture a Melanoma [12]. Imaging could be at the plane of complete physique photography to identify changes in the shape, size, or color of specific lesions or at the subcellular level with reflectance confocal microscopy to see abnormal cells. Photography helps to determine sites of prior incisions and progress our ability to monitor for repetition.



Figure 2.1: Sample dermatology imaging technology in skin cancer diagnosis (Toronto dermatology center)

### **2.2.1 Digital photography**

Dermatologists widely use Digital photography. About 82.2% of practicing US dermatologists use digital photography as shown in a recent survey [14] 90.1% of young dermatologists (under 40 years old), and 93.9% of academic researchers use digital photography more often. Recap and contiguity images can be adopted at baseline and follow-up visits to monitor lesions periodically. Particularly in the issue of Melanoma diagnosis, rudimentary-stage detection is crucial and can remarkably deteriorate the mortality rate. About 44% of Melanoma was identified *in situ* with a lower thickness in the class with baseline photography, compared to 35% for the general people [15]. Another investigation [16] revealed that baseline photography assists the specialist in detecting premature Melanoma among high-risk patients. The limitation of this procedure is that it only captures skin surface morphology.

### **2.2.2 Total body photography**

Total-body photography (TBP) is a family of digital photography operated to identify new or changing lesions. Generally, the physicians or technicians take 24 photos on average (range, 4-50) in different patient positions [17]. Close-up images might be considered at regular intervals for long-term follow-up for a selected lesion. In early-stage Melanoma detection, TBP to skin shown self-examination (SSE) enriches the sensitivity for identifying new or changed nevi from 60% to 72% and increases specificity from 96% to 98% [18]. TBP is a time-consuming procedure where patients must be photoed in diverse locations to apprehend all body externals.

### **2.2.3 Dermoscopy**

The derm disease that cannot be detected by the naked eye there, dermoscopy is used to visualize subsurface structures. The dermatoscopy is an optical hand-held instrument with (10x) magnification and a transilluminating illumination origin. The sample mobile dermatoscopy is shown in Fig. 2.2. Applications of dermoscopy include evaluating inflammation, infection, autoimmune diseases, and scalp and hair disorders. Polarized dermoscopy delivers adequate visualization of the exhibitions general in skin cancer, involving vascular and crystalline structures [19]. Dermoscopic images can be obtained, letting for improved

viewing of doubtful lesions and scrutinizing subtle transitions over time. Examinations have discovered that dermoscopy enhances the sensitivity and accuracy of identifying Melanoma corresponded to unaided eyes. A contemporary meta-investigation of 9 prospective reflections in clinical settings revealed that the comparative diagnostic odds proportion for Melanoma of dermoscopy compared to unassisted eyes was 9.0 ( $P = .03$ ; 95% CI 1.5-54.6) after the deduction of 2 outlier examinations [20]. For 7 of the 9 examinations, the positive predictive value (PPV) was more significant for dermoscopy than for the unassisted eye, and dermoscopy provided an assessed 18% improvement in sensitivity (95% CI 9-27;  $P = .002$ ) [20]. Aforementioned meta-analyses also revealed extremely improved diagnostic accurateness for Melanoma employing dermoscopy corresponded to unassisted eyes [21, 22].



Figure 2.2: Attached mobile dermatoscope allows simultaneous dermoscopic assessment and photograph capture (Canfield Scientific Inc, Fairfield, NJ)

The benefit of dermoscopy yields more acceptable specificity, as demonstrated by the 42% deduction of patients guided to biopsy in a randomized trial[23]. Also, it induces a decreased benign/malignant proportion of excised melanocytic lesions from 18:1 in the dermoscopy period to 4:1 in the post-dermoscopy period ( $P = .04$ ) [24]. Dermoscopy also enhances the precision of detecting non-melanocytic lesions, including basal cell carcinoma [25]. In addition, besides diagnosing the skin lesions, dermoscopy is also employed in monitoring. Although widely operated, dermoscopy demands appropriate training. Ex-

aminations have demonstrated that proper diagnosis can be extremely affected based on the medic's understanding; hence incapable or less experienced individuals were no better than assessment without dermoscopy [21]. In addition, anchoring bias and search satisfaction can result in an erroneous diagnosis.

#### **2.2.4 Reflectance Confocal Microscopy**

Reflectance Confocal Microscopy (RCM) accommodates noninvasive, high-resolution (30x) imaging of derm to a profoundness of  $200\mu m$ , reaching the papillary dermis [26]. A low-power laser acted as a ray source to emit near-infrared light, which scans across the specimen Fig. 2.3. From a preferred focal point, ray is reflected, passes via a pinhole, and then penetrates the detector. Giving contrast to the photograph, different cells and structures have different reflection indexes. Melanin and keratin have high distinctions with reflection indexes of 1.7 and 1.5, correspondingly [26]. High-contrast configurations emerge bright/white, and output photographs are horizontal sections of the skin. RCM keeps the natural form of the tissue. Its elevated resolution allows imaging of cellular, nuclear, and tissue model, and its noninvasive character avoids excessive biopsy. One analysis demonstrated the number needed to excise a Melanoma dropped from 14.6 to 6.8 using RCM [27]. RCM can be a beneficial supplement tool to analyze skin lesions with questionable features. It is also used to illustrate surgical margins and to monitor patients' responses to non-surgical treatments [28, 29]. RCM has been typically operated to estimate tumors in sensitive regions (i.e., eyelid [30], oral and genital mucosa) [30, 31] The developed algorithms with RCM performed high accuracy in skin cancer detection.

#### **2.2.5 Optical Coherence Tomography**

Optical Coherence Tomography (OCT) is a group of non-invasive imaging techniques that uses infrared light to operate analogously to ultrasound. It has a 3 to  $15\mu m$  resolution at a profoundness of utmost of 1.5 mm [31] and can produce real-time 2D and 3D photographs of the tissue. It is especially used to expose skin morphology, including arrangements like blood vessels, hair follicles, dermo-epidermal junction, epidermis, glands, and dermis [32]. It also assists the diagnosis of skin lesions, including inflammatory diseases, and cancer and also monitors patients' response to therapy. Examination of skin moistness, skin atrophy, burn depth, UV damage, wound healing, and engineered tissues are some other applica-



Figure 2.3: Vivascope Reflectance Confocal Microscope System (Lucid, Inc, Rochester, NY)

tions [33]. OCT estimates the depth of invasion more satisfactorily than a general scanning microscopy laser. In contrast, it still has a narrow depth of 1.5 mm due to substantial scattering and thus can only image moderately thin tumors. It also has restricted resolution and cannot effectively discriminate between malignant and benign lesions [34].

## 2.3 Classification

Classification is the approach for categorizing the input patterns into equivalent classes. Diverse classification techniques are available but the detection of skin cancer rate still needs to be adequate. Also, precise image segmentation into meaningful areas is an fundamental key to the success or failure of image classification [35]. Most classification methods use pre-processing, enhancement, segmentation, and classification processes to complete the task. Picking of a appropriate classifier demands consideration of many aspects: i) Classification accurateness, ii) Algorithm performance iii) Computational resources. There are two kinds of classification approaches. One is known as unsupervised classification, and the other type is known as supervised classification.

### 2.3.1 Supervised Learning

The understanding is where the algorithm produces a function that maps inputs to expected outputs. Supervised learning is the machine learning procedure of implying a function

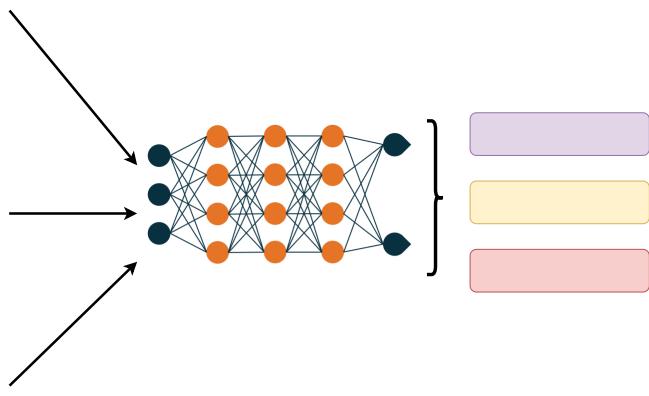


Figure 2.4: Deep Neural Network-based image classification model

from labeled training sample. Originally, supervised learning was when a machine learned or trained using well-labeled data, as shown in Fig. 2.5.

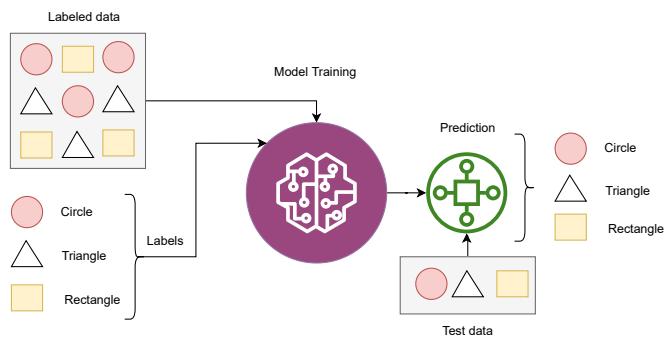


Figure 2.5: The working of supervised learning algorithm

A supervised learning algorithm investigates the training sample and generates an implied function, which can be operated for mapping unknown instances. To precisely determine the class labels for unseen instances demands the learning algorithm to generalize from the training data to unseen environments in a “reasonable” way. Here we used DNN architecture as a supervised learning algorithm for Melanoma prediction.

### 2.3.2 Pros and Cons supervised learning

Supervised learning enables a machine to collect and generate output data from previous experiences. With the help of its past experiences, it aids in performance optimization. Though many real-world computational problems are solved using supervised machine-learning techniques, it consumes a long training time a machine-learning model because computational complexity and classifying big data are challenging.

### 2.3.3 Unsupervised Learning

An unsupervised learning algorithm is training a machine or model using unlabeled data and permitting the algorithm to operate on that data without guidance, delineated in Fig. 2.6. Here the machine groups the unsorted data according to patterns, similarities, and differences without prior knowledge of the data. Unsupervised learning is classified into two categories of algorithms: 1) Clustering and 2) Association. Cluster or group analysis is the most common unsupervised learning method used for exploratory sample exploration to discover obscure patterns or grouping in samples. An association rule learning situation is where machines desire to uncover rules that describe large sections of the data; for instance, individuals who purchase X items also tend to purchase Y items.

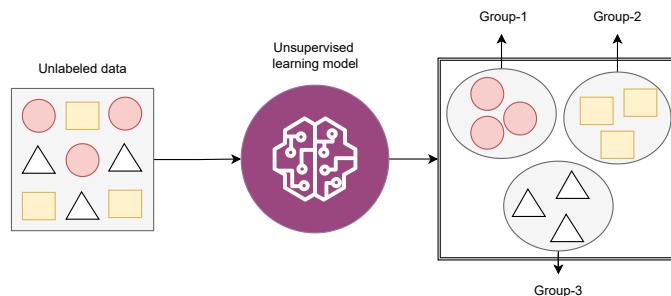


Figure 2.6: The working of unsupervised learning algorithm

### 2.3.4 Pros and Cons unsupervised learning

This learning procedure is more complex than supervised learning and less accurate as class labels are unknown. However, it is useful for data that are unlabeled.

## 2.4 Pre-processing

Image enhancement can be deemed one of the rudimentary methodologies in image examination. Until today, numerous photograph enhancement strategies have been proposed for several applications, and actions have been directed to improve the grade of the enhancement outcomes further and minimize computational complexity and memory usage. Images must undergo many image processing tasks before feeding them into a Deep Neural Network (DNN) model. The model includes, however, is not limited to resizing the images, grayscale conversion, image enhancement (i.e., brightness, contrast, sharpness), opening or closing, and image augmentation. Pre-processing aims to enhance the image's quality so we can examine it more effectively. Pre-processing eliminates artifacts and improves specific qualities essential for a particular application. Depending on the application, those characteristics could be changed.

### 2.4.1 Image Enhancement

Miscellaneous images are inadequate quality and difficult to detect and extract knowledge. Therefore, the picture has to get under a procedure called image enhancement which includes an aggregation of strategies that look to ameliorate the visible aspect of an image for enhancing attributes and gain better attributes of medical images. Medical images are one of the fundamental images because they are operated in a additional sensitive field which is the medical field. The raw data acquired straight from devices of medical acquisition may afford a comparatively poor picture quality illustration and may demolish by several types of noises. Image noise lessens the accuracy of the detection of abnormalities. Especially in skin mole identification pre-processing is the most essential step so that the segmentation algorithm functions appropriately. Hence the photograph quality should be enhanced before the segmentation procedure takes place. While clearing noise and artifacts, it is necessary to preserve its images without degradation of the original image. To ensure image quality, sharpness has a significant contribution. Sharpness describes the clarity of detail in an image that can be a worthwhile creative tool for highlighting texture. Image brightness (or luminous brightness) is a measure of intensity after the image has been acquired. Modifying the brightness of a photograph impacts all pixels equally. Advancing it will make the shady spots lighter and the light regions lighter too. Similarly, lowering the brightness will

make light regions darker and dark locations darker. Contrast distinguishes intensity levels in an image.

### **2.4.2 Grayscale Conversion**

A grayscale image is more accessible than an RGB image. A grayscale image performs more complex operations in less time for various image processing tasks such as morphological operation segmentation.

### **2.4.3 Image Resizing**

Resizing an image is an crucial component of pre-processing, which enhance or diminish the total number of pixels. Photograph resizing on the conventional approach followed by scaling and cropping. Content-aware resizing techniques seam carving, object carving, include warping, and multi-operator. A homogeneous map between the actual image's pixels and the target image's pixels defines the scaling. The most typical approach for scaling embraces the interpolation of actual image pixels. The cropping process extracts a rectangular window with the preferred size from the actual image. The classic cropping process crops the preferred rectangle from the center of the image as its output resizing result. The pivotal concern of image resizing is to maintain the most attractive parts and valuable information, achieve real-time resizing, minimize visual distortion, and satisfy user choices under the constraint of topological relations and the global context. As in Melanoma skin cancer identification shape, the mole size is the most distinguishable feature, so retaining its original size is challenging after resizing the image.

### **2.4.4 Data Augmentation**

Data enlargement is a strategy that enormously increases the diversity and amount of data from available data to train a model without gathering unexplored data. Padding, Cropping, Zooming, Flipping, and Altering the image size are the popular techniques used to augment the data. Sample data augmentation is depicted in Fig. 2.7.

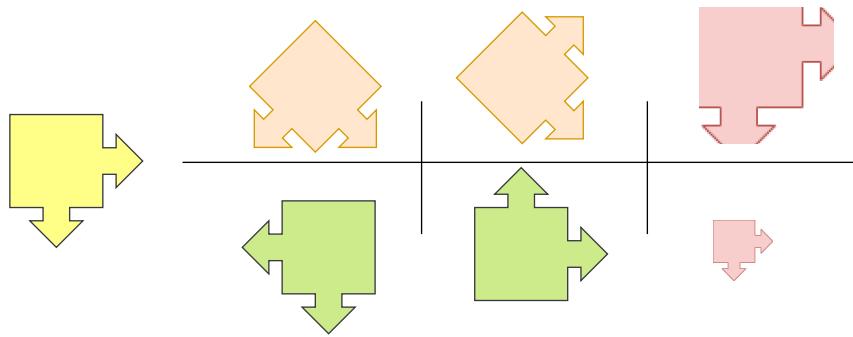


Figure 2.7: Augmentations of a sample image: rotation, zooming and flipping

## 2.5 Segmentation

Strong noise, poor gray-scale contrast, and blurred margins of tissue are features of medical images. Image segmentation is an ordinarily operated procedure in digital image processing and investigation. It is a procedure of partitioning an image into considerable regions or components established on the characteristics of the image pixels, such as gray level, color, texture, illumination, and contrast. Image segmentation could involve clustering or grouping regions of pixels based on analogous color or shape Fig. 2.8. Medical imaging, self-driving, video surveillance, and computer vision are possible image segmentation applications. Image segmentation involves the following objectives [36],

- To assess ROI (Region of Interest)
- To explore the anatomical form of body parts
- To estimate the scope of the lesion location in the skin appearance
- It also helps the radiologist in preparing the quantity of radiation preparatory to the radiation treatment.

The intensity of equivalence, presence of artifacts, and distinct soft tissue are the performance estimating facts in image segmentation [35]. At present, gray level and texture-based image segmentation methods are mostly used techniques. Some commonly used image segmentation techniques are delineated in Fig. 2.8.

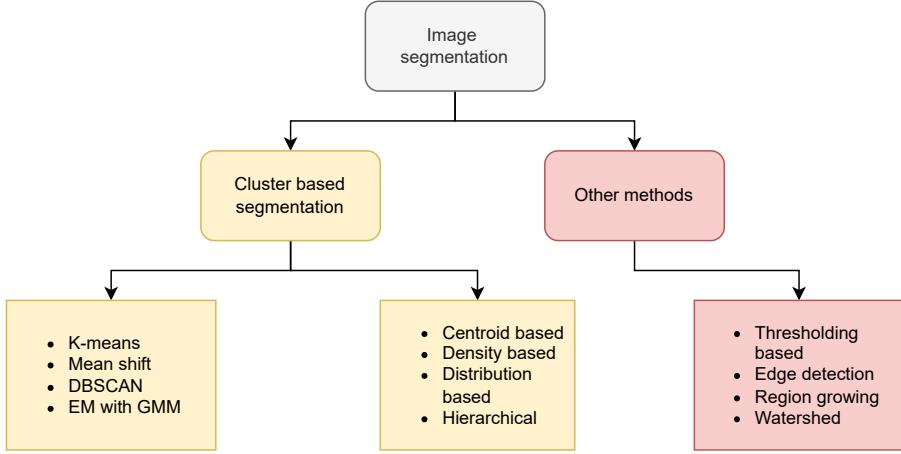


Figure 2.8: General image segmentation techniques

### 2.5.1 Cluster based Segmentation

Clustering is one of the most sensible unsupervised approaches in image Segmentation, where it categorizes pixels, meaning without containing prior knowledge. It organizes pixels with maximum likelihood into the analogous class. Training is done by using pixel features with each class's belongings for suitable grouping. K-means cluster, Mean shift, DBSCAN, EM with GMM and centroid, Density, Distribution, and Hierarchical based segmentation techniques are included in the clustering technique.

### 2.5.2 Thresholding based Segmentation

Thresholding is considered a general approach for picture segmentation that can be used for complex backgrounds and multiple objects. This strategy is remarkable for images dealing with diverse intensities. Employing this process, the image is separated directly into different parts based on the intensity weights [[35]]. Threshold computing methods are most effective as it uses both local properties and gray-level information for the thresholding process. Thresholding-based segmentation is two varieties: i) Global Thresholding and ii) Local Thresholding.

The global thresholding approach is the effortless and easiest segmentation method and operates the threshold value for segmentation, where the value is computed by calculating the broad intensity distribution of an photograph. Designated pixel values are compared with a threshold value; if the pixel value is higher than the threshold, then the pixel is

assessed as foreground; otherwise, consider it as background and set to be black. Most available thresholding employs two values: bi-level thresholding or global thresholding. Global thresholding supplies good outcomes only for images that have a constant background. However, it performs adequate segmentation for one portion, but it fails in another portion.

The local thresholding separates the entire image into smaller images horizontally and vertically and assesses the 15 thresholds for every sub-image. Finally, the outcome for all sub-images is aggregated. The corresponded result is developed using interpolation process. Unlike the global threshold, it desires extra time to perform segmentation. This technique's threshold value is calculated using various statistical procedures like standard deviation, mean and standard deviation together, mean, and mean of maximum and minimum. It is the agreeably appropriate technique for images that have differing backgrounds.

### **2.5.3 Edge Detection**

Edge detection is an image processing procedure for discovering the borders of entities within images by observing discontinued points with more or less sharp modifications in gray level [35]. Image boundary detection has been operated particularly for object segmentation and data extraction in image processing, machine vision, and computer vision scopes. The edge-detecting approach is how humans perceive objects and work well for images with good contrast between regions. However, it only performs well with images in which the boundaries are ill-defined, or there are too many boundaries and Less resistant to noise than other procedures (i.e., Thresholding and clustering) [35].

### **2.5.4 Watershed Segmentation**

Watershed transformation is a general approach for image segmentation. However, its use for computerized medical image segmentation has been restricted, particularly due to over-segmentation and sensitivity to noise. The process floods basins from the markers until basins featured to various markers meet on watershed stripes. In many circumstances, markers are taken as regional minima of an image from which basins are flooded. The initial shape intuition has significantly improved medical image segmentation algorithms. Employing initial shape understanding has revealed robust progress in medical image seg-

mentation algorithms. The watershed algorithm treats pixel values as a regional topography (elevation).

## 2.6 Image Features

A feature is a unique assessable property or unit of an empirical phenomenon that acts as an input to a machine-learning model. In image processing applications (i.e., computer vision, pattern recognition, and medical image investigation), the detection and description of medical photograph features play an essential function. The diverse features of an image comprise shape, color, texture, or domain-specific features. The texture is believed one of the noteworthy features of an image. The Gray level co-occurrence matrix (GLCM) is used to compute the global texture properties for an image, including energy, entropy, homogeneity, correlation, contrast, dissimilarity, and maximum likelihood. The regional texture features for an image are extracted using a texture feature descriptor named Speeded Up Robust Features (SURF). The feature acquisition performances are based on the classification results. Overall, two major types of features, local and global features, can be derived from an image. The key points within an image are detected in local features and describe regions around the points; In contrast, the global features express the image as a whole to interpret the definite property of an image considering all pixels.

Melanoma is a disease or condition that affects the melanocyte cells, thereby impeding the synthesis of melanin [26]. The disease is illustrated by the development of skin lesions and changes in color, shape, size, and texture. The texture features are regarded to be important in the detection of Melanoma. The spatial distribution of pixels in the neighborhood of an image characterizes the texture. The spatial dependency of gray levels is depicted by a two-dimensional matrix known as GLCM, which is operated for an image's global texture study. The GLCM matrix specifies an image's texture and how often the pairs of pixels with distinctive values appear in an image. The statistical measure is then calculated from the GLCM matrix. The textural attributes represent the spatial distribution of gray tonal discrepancy within a distinct portion. The adjacent pixel is correlated in images, and spatial values are acquired by the redundancy between the neighboring pixel values [37]. The local texture features are calculated using the SURF descriptor. In SURF, there are three steps:

interest point description, interest point detection, and feature descriptor matching. The interest points are detected using factor repeatability, and a descriptor vector is generated and established on these interest points. The descriptor vector is acquired by calculating the Haar wavelet responses (HWR). The feature vectors are compared against the adjacent using Euclidean distance [38]. The Melanoma affected skin areas are predominated by texture.

## 2.7 Histogram

An image histogram is a graphical representation of a digital image where gray-scale value distribution reveals the frequency of incidence of particular gray-level values. For an image dimension of  $1024 \times 1024 \times 8$  bits, the abscissa varies from 0 to 255; the total number of pixels in the image equals  $1024 \times 1024$ . Tonal variations are indicated by the horizontal axis, whereas the vertical axis denotes the number of pixels in that particular pixel.

## 2.8 Deep Neural Network

A Deep Neural Network (DNN) is an Artificial Neural Network (ANN) having numerous hidden layers between the input and output layers Fig. 2.9. Deep learning is a Machine Learning (ML) class based on ANN in people's fields of computer vision. It is a data processing strategy that operates considerable layers of complex networks or multiple processing layers composed of many nonlinear modifications. The actual data is abstracted layer by layer by incorporating multiple nonlinear processing layers. Different abstract features are acquired from the data and operated for target segmentation, detection, and/or categorization [39]. The benefit of deep learning is to alter the manual acquisition feature with unsupervised or semi-supervised feature learning and hierarchical feature extraction efficient algorithms. Although the quantity of medical data is immense, due to different equipment used, the quality of data varies greatly; data present fluctuating characteristics over time, and specific events change; due to dissimilarities in individuals, the ordinance of the disease has no ubiquitous applicability.

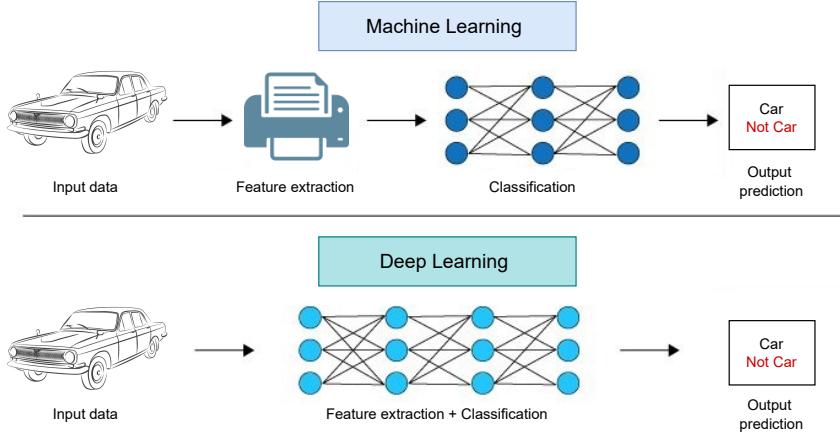


Figure 2.9: Difference between Machine Learning model to Deep Neural Network

### 2.8.1 Convolution Neural Network

Employing deep learning for image categorization is the earliest rise, and it is also a subject of prosperity. The convolutional neural network (CNN) is the broadest structure among them. CNN is beneficial for image categorization and recognition because of its ability to acquire high accuracy in a concise time [40, 41]. Compared to its predecessors, the main advantage of CNN is that it identifies essential features automatically without any human supervision [42]. As stated in Fig. 2.10 the typical CNN has four layers: i) an input layer ii) a convolutional layer, iii) a pooling layer, and iv) a fully connected layer. The layers are responsible for the following tasks,

- As seen in other structures of ANN, the input layer will keep the image's pixel values.
- The convolutional layer will determine the output of neurons that are connected to local areas of the input via the calculation of the scalar product between their weights and the area connected to the input volume. The layer's parameters focus on the use of learnable kernels. The rectified linear unit (generally shortened to ReLu) aims to implement.
- The pooling layers are responsible for lowering the number of dimensions of the features by simply conducting downsampling along the spatial dimensionality of the provided input; thus the overall computation cost will be minimized.
- The fully-connected layers operate the identical duties as an artificial neural network

(ANN) accomplishes; for instance, for a classification task, each neuron executes a linear transformation to the input vector through a weight matrix to produce category scores from the activation functions.

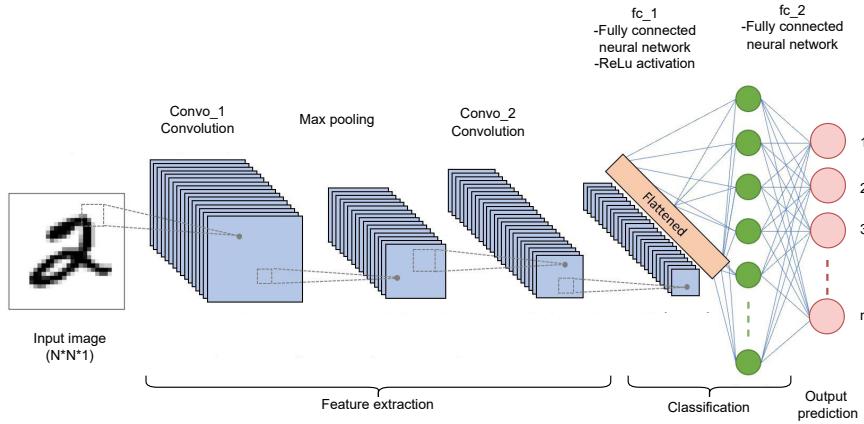


Figure 2.10: Convolutional Neural Network (CNN) general architecture

### 2.8.2 Transfer Learning

Transfer learning solves problems by applying previously learned knowledge to a new task. Rather than creating a model from scratch, deep transfer learning employs a pre-trained model trained on a massive amount of data, such as ImageNet for image classification weights, to transfer prior knowledge from an authority domain to a target domain [43]. The transfer learning strategy is illustrated in Fig. 2.11. This positively impacts many difficult domains to improve due to a shortage of training data [43]. Transfer learning has the advantage of reducing overall training time because it uses the weights of a pre-trained model. As the weights of a previously trained model with a massive amount of data yield more consistent and high-performance outcomes, transfer learning is a more beneficial classification procedure on a small dataset [44]. In transfer learning, retraining distinctive characteristics on a unique target dataset play a crucial role in improving overall performance. Furthermore, the second vital characteristic is a proper network complexity that matches the level of the dataset.

Many recent studies have discovered that DenseNet, ResNet, Inception-v3, and VGG are the most popular object recognition algorithms. One motivation is to avoid having to train the models from start to accomplish jobs concurrently, saving time and resources.

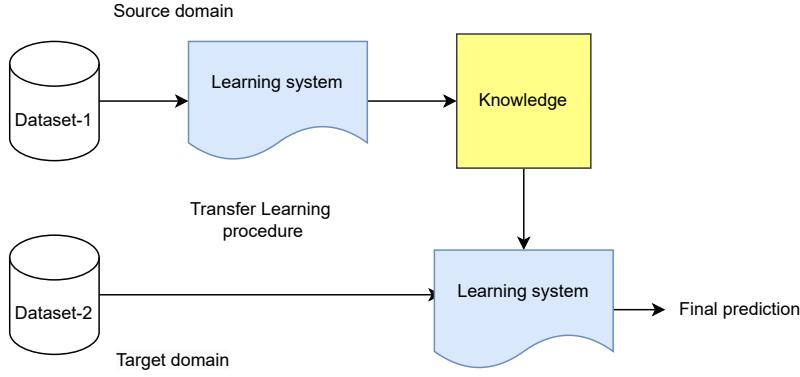


Figure 2.11: Transfer Learning process

### 2.8.2.1 DenseNet201

This model is well recognized for its excellent accuracy, and it has been used in a variety of image classification applications, as described in [45]. DenseNet201 features 201 layers in total, is relatively sophisticated in comparison to other models and has a significant number of parameters. As a result, it is considered as a computationally expensive model that necessitates powerful technology to train and run. When the dataset is vast and the job is complex, it performs better for transfer learning. DenseNet architecture employs a technique known as dense connectivity, in which each layer is feed-forward connected to every other layer. This permits information and features to flow across the network rather than merely through a linear sequence of levels.

### 2.8.2.2 ResNet

The ResNet-50, ResNet-101 and ResNet-152 models are based on the “ResNet architecture” and have been trained on the ImageNet dataset. They utilize skip connections to mitigate the vanishing gradient problem in deep networks [46]. The building block of residual learning is a key element that enables the network to learn identity functions and have deeper architectures. Compared to other versions of the ResNet model, ResNet50 has fewer parameters, making it ideal for transfer learning in complex tasks with limited computational resources [46]. The models have been used in several computer vision applications.

### **2.8.2.3 Inception-v3**

This model is developed by Google based on the innovative architecture “Inception module”. The module reduces computational power by introducing parallel branches with different filter sizes and then concatenating their results, allowing the network to learn a variety of features in a more computationally efficient manner [47]. The model has been trained on the ImageNet dataset and It has been used in several medical imaging tasks such as retinal disease classification, breast cancer detection, and skin lesion classification.

### **2.8.2.4 VGG**

The VGG-16 and VGG-19 models are based on the Visual Geometry Group (VGG) architecture [48] which contains a 16-layer convolutional neural network for VGG-16 and a 19-layer convolutional neural network for VGG-19 respectively. Both models are known for their simple structure and for using only 3x3 convolutional filters, which is unique compared to other popular architectures. They have been trained on the ImageNet dataset. Despite their simplicity, VGG-16 and VGG-19 have been very successful in several image classification tasks and have become widely used in various computer vision applications including lung nodule detection, breast cancer classification, and brain tumor detection.

## **2.9 Deep Learning for Medical Imaging Analysis**

With the expansion of deep learning, computer vision operates a lot of deep learning to deal with miscellaneous image problems. Medical photograph as special visual images has attracted the attention of many researchers. Miscellaneous medical image processing and recognition strategies have recently adopted deep learning methods. The method comprises macroscopic images, Reflectance Confocal Microscopy (RCT), Optical Coherence Tomography (OCM), endoscopic images, CT/MRI images, ultrasound images, and many more [39]. In modern times, deep learning technology is particularly operated to classify and segment medical images.

## 2.10 Melanoma Skin Cancer Detection

To identify Melanoma skin cancer, recently MED-NODE dataset has achieved high popularity among other datasets due to the simplicity of the macroscopic images. The researchers [49, 50, 51] used the MED-NODE dataset to analyze their models' performances. Another well-known dataset, HAM-1000, is utilized by the author [52]. The DERMIS dataset (total 397 images) is used in [53]. The authors [36] collaborated with the international skin imaging collaboration (ISIC) dataset with the MED-NODE dataset to train and assess their suggested model.

Before segmentation of the image lesion, researchers usually employ various pre-processing techniques to handle the artifacts. For removing the overall image noise and other noisy features, authors [49, 54, 53] used Gaussian smoothing and Kuwahara smoothing filters which additionally preserve the edges. Increasing the systems discrimination capability through illumination and noise correction are discussed in [54, 50]. A Multi-Parameter Extraction and Classification System (MPECS) is proposed in [55] where the authors defined the skin lesion images concerning the distinctive parameters, which are additionally operated for classification. They performed the Sobel edge detection method on the region of interest (ROI) components to extract the features. To enhance the image texture features and the darkened perimeter of the lesion, author [36] used hair removal, glare removal, and shading removal methods. Which further aids in identifying the color, shape, and size of the lesion. After contrast enhancement, noise removal is accomplished using an anisotropic diffusion filter (ADF).

Authors [52] focused on removing the noise from an image and working with lesser-resolution images. Their proposed system utilized Histogram Equalization to increase the skin lesion contrast. For an efficient image classification model, the effect of the image texture analysis and contrast-enhancing considering intense pixel values is discussed in [56]. The authors discovered that contrast-enhancing is designed to increase the separation between the image intensity values. So, they can be easily determined by humans and computer vision. They also found that the rudimentary information of an image is stored in the form of pixel intensity value. It can be three values for a color image or a single point

of a gray-scale image. The significance of image texture analysis for pattern recognition due to its separation capability is also discussed. Image normalization is used in [50] to increase skin lesion pixel intensity which carries skin cancer-related information. Unlike Nevus moles, Melanoma has rough and uneven edges [57]. The ADF filter helps the authors to remove noise from images to keep the edges [50]. Rescaled intensity values are utilized to ensure better convergence in the classifier. The canny edge detection (CED) technique has the optimal ability for edge detection [8]. So the authors [8] employed the CED model on the lesion image that was priorly segmented by using the watershed method. They extracted ABDC features from the images and the calculated total dermoscopy score (TDS) to identify Melanoma. Authors in [49, 54, 53] segmented the ROI area using the K-means clustering algorithm using ( $k=2$ ). The image mask was further enhanced by the authors [54]. A multi-level segmentation is introduced in [36], where the authors collaborated on three segmentation methods, watershed, Otsu, and Modified Otsu, for lesion segmentation. Whereas the authors [51] used only the Otsu segmentation method. Three segmentation methods (i.e., region, edge, and cluster-based segmentation) are checked in [50].

In order to increase the learnability of a model for better predicting the image augmentation methods are used [52, 36, 50, 57]. This technique helped the authors to avoid the overfitting problem. The system produces several copies of the images from the existing images by using various factors by applying rotation and, zooming, translation on images. To identify and categorize Melanoma and Nevus mole efficiently, the authors used different techniques to extract the features from the segmented lesion image. To extract 675 features per image, the authors [49] developed a cluster-based adaptive metric classifier. Three different image features are extracted by [53] using Grey Level Co-occurrence Matrix (GLCM), Local Binary Pattern (LBP), and three-color channel features from an RGB skin lesions region of interest (ROI). They created a distinctive hybrid super feature vector using the lesion's extracted textural and color features. For the classification purpose, the utilized SVM, K-nearest neighbor (KNN), Naïve Bayes (NB), and Decision Tree (DT) classifiers were used. They achieved the best performance by combining the texture, LBP, and GLCM features with the color features. A total of 1900 features are extracted from the segmented lesion image in [51]. Firstly, they excluded rank-wise 25 less relevant features to train and test on the multi-layer neural network (MLP) classifier. Then they used principal component

analysis (PCA) and achieved the best result using only 25 principal components.

In the analysis of medical images, deep neural network (DNN) has achieved significant performances. In [58], authors used CNN to detect and classify Melanoma from Nevus mole. The CNN model has the high capability to consume essential features from an image [59]. Particularly in Melanoma identification, DNN has shown a massive impact [54]. A two-stage model is proposed for spontaneous image segmentation and Melanoma detection in [60]. Authors [54] used the enhanced images from pre-processing to feed into a pre-trained CNN model. In their work, they utilized two convolutional layers; the first generates 20 feature maps, and the second generates 50 feature maps. Two pooling layers are used after each convolutional layer. The classification and output prediction is made in the fully connected layer. With no overlapping, they split the dataset into 20% for testing and the rest 80% for training the model. The researchers in [36] proposed another system in which a support vector machine (SVM) and back-propagation neural network (BPNN) were utilized. Also, a convolutional neural network (CNN) is a deep neural network to train the model and predicts Melanoma. Four features (color, shape, size, and texture) are extracted after lesion segmentation. They used a 0.5 ratio in the dropout layer to avoid overfitting issues. A decision support system is developed to detect and classify Melanoma skin cancer in [52]. The authors used balanced and resized images to train and test the model; along with some ML models such as XBoost, SVM, and Random Forest (SVM) classifiers, they also used CNN and Transfer learning approaches using a pre-trained VGG-16 model. Working with medical data is often challenging due to data shortages. To handle this shortcoming CNN model is improved further by increasing the network depth [61]. However, researchers utilize this improved model by using pre-trained model weight, and smaller medical data, also known as transfer learning [62]. A mobile device familiar model MobileNetV2 network is used in [63] for Melanoma identification effectively. The bottom-level layers of a deep learning model tend to be relatively generic; as we move top in the network, specialization ameliorates [64]. Hence, a transfer learning model is generally used easily: freezing the weights of a pre-trained DNN model till a chosen tier and then training the other layers for the new assessment. The researchers modeled the transfer learning techniques by employing pre-trained models in combination with focal loss and class weight. In the study, authors [57, 65, 54] used image resizing methods on all the input images before feeding

them to a DNN model. The general process of detecting and diagnosing Melanoma skin cancer has been summarized into key operating procedures such as image pre-processing, image segmentation, feature extraction and analysis, and classification of lesions images [66, 67] as described in Fig. 2.12.

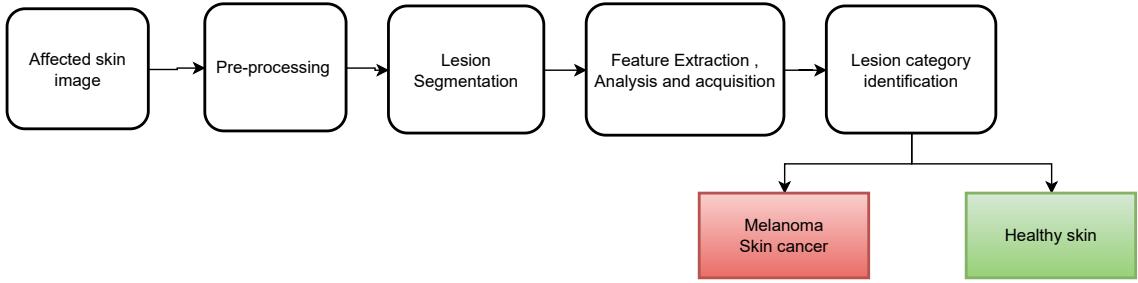


Figure 2.12: Melanoma skin cancer diagnostic process

In this work, a Deep Neural Network (DNN) architecture (i.e., CNN, VGG-16 transfer learning, Inception V3 transfer learning) based model is performed to identify the best model for efficiently detecting and classifying malignant Melanoma from benign one at an early stage. Before feeding the model, skin images are pre-processed to improve model performance. Following the visual feature identification rules ABCDE (asymmetry, border, color, diameter, evolving) for Melanoma, we found from the studies that the lesion's shape and size keep distinguishable features. So we developed an improved image-resizing algorithm for skin cancer detection to keep the lesion shape identical to the original. Finally, The proposed model's outcome is compared with the state-of-the-art methods. We also investigated data sensitivity toward the DNN model performance and analyzed the performance evaluation metrics on the data variation.

## 2.11 Discussion

In this chapter, we discussed the rudimentary knowledge about image acquisition, image processing methods, and techniques, segmentation, supervised and unsupervised machine learning, study on the histogram in image analysis, Deep Neural Networks in image feature extraction and classification, and study on Melanoma skin cancer.

Study on Melanoma shows that the researchers usually follows three types of method (i)

feature extraction from ROI using feature extraction methods (i.e., GLCM, LBP, etc.) or convolutional layer in CNN model then used PCA to optimize the large feature spaces to get better prediction (ii) a few researchers focused on the visual observation for Melanoma detection, i.e., ABCD features are extracted (iii) others used edge detection techniques particularly canny edge detection (CED) model on the segmented lesion image to distinguish Melanoma form Nevus mole. However, no study discusses the image resizing issue, as the size and shape of the lesion in visual Melanoma detection could highly influence the ABCDE feature extraction. So there could be a high possibility of changing the lesion area using a conventional image resizing technique before feeding them in a DNN model. Another complication was noticed in extracting the ABCDE features from the CED model. It contains only limited information about the skin lesion as CED outputs the edges with high-intensity value only (experimentally, one on a scale of 0 to 1). Considering the above study, we organize the proposed classification methodology using pre-processing, segmentation, and deep learning for Melanoma skin cancer detection and classification, described in the following Chapters.

# **Chapter 3**

## **Proposed Classification using Convolutional Neural Network**

---

### **3.1 Introduction**

Detection of Melanoma from skin lesion images at an early stage plays an essential role in saving a life. Physicians may miss the abnormality due to inexperience in skin lesion analysis. Hence, a precise computer aid intelligence system is required to identify and classify Melanoma from benign moles. This chapter briefly discussed our proposed innovative intensity value estimation model embedded with a DNN-based system to detect Melanoma and classify it from a benign mole by utilizing previously gathered knowledge such as image pre-processing, segmentation, resizing, and data augmentation.

### **3.2 Proposed Methodology**

Medical image classification plays a crucial function in clinical treatment and teaching lessons. However, traditional approaches have achieved the apex of their performances. Moreover, using them requires much time span and effort to be spent on extracting and determining categorization features. Our proposed methodology is divided into pre-processing, intensity value estimation, and deep neural network (DNN) steps. A few sub-parts are uti-

lized in the pre-processing step to reduce artifacts that could mislead the DNN model. The proposed methodology is illustrated by the block diagram shown in Fig. 3.1.

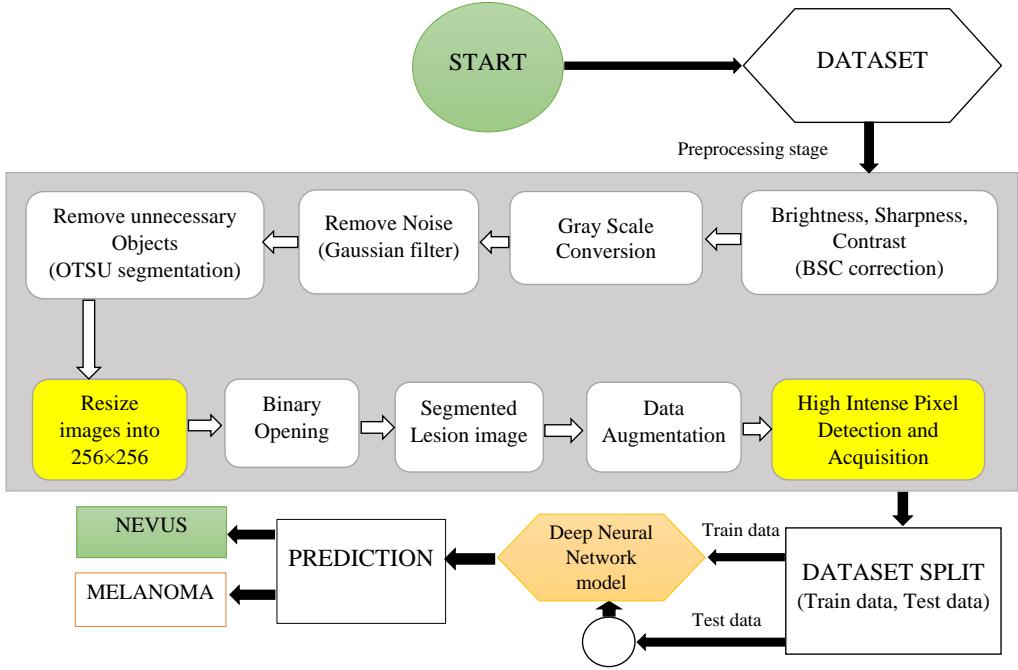


Figure 3.1: Proposed model to detect and classify Melanoma skin cancer and Nevus mole

### 3.2.1 Pre-processing

In the pre-processing phase, images from the collected MED-NODE dataset [49] are passed through a series of basic image processing methods to reduce the effects of misleading factors on the CNN model. As images have both the sound skin and the lesion parts, hence we correct image brightness (B), sharpness (S), and contrast (C) in order to articulate the distinguishable features. The pixel's intensity refers to an image's brightness, and the histogram clarifies the brightness. The higher end in the histogram indicates a brighter image, and the image is darker when the histogram is limited to a small region towards the lower end [68]. The number of details presented in an image refers to its sharpness. Lossy compression, denoising, motion blur, and out-of-focus filtering are some of the reasons that influence image sharpness recognition [68]. Commonly, contrast guides split up the bright and dark areas in a photograph [69]. The contrast enhancement technique eliminates the anomaly that

would otherwise occur between different regions in an image.

An RGB photograph carries a three-channel red (R), green (G), and blue (B). However, the grayscale photograph has only one color channel. The RGB image with three color channels contains diverse intensity scales to convey that image in a color form. For various image processing tasks, a grayscale image is more accessible than an RGB image because it has only one channel. So, we convert the images from RGB to grayscale images. The grayscale pixel value is estimated as the weighted sum of the corresponding R (red), G (green), and B (blue) pixels as in the equation 3.1 to transform an RGB channel image into a single grayscale channel image. Cathode-ray Tubes (CRT) phosphors use these weights to illustrate a more satisfactory human understanding of RGB images than equal weights [70]. The photograph texture suggests information about the spatial structure of color or intensities in that photograph or preferred photograph area [71]. Since the grayscale image is formed from the three-channel RGB image using equation 3.1, the image color and texture could be represented by the different pixel intensity levels in the grayscale image [72, 73].

$$G_{im} = 0.2125R + 0.7154G + 0.0721B \quad (3.1)$$

Though high-resolution professional cameras take images, non-uniform lights create noise effects. We used the Gaussian filter to remove noise from grayscale lesion images ( $\sigma = 1.35$ ).

### 3.2.2 Segmentation

Segmentation is a procedure of extracting desired objects from an image in which thresholding techniques are widely used. Otsu segmentation is one of the thresholding techniques proposed by Nobuyuki Otsu. This method is repeatedly used in numerous fields, particularly in image segmentation. However, it produces competent developments only when the numbers of pixels in individual classes are nearer to each other [74]. The original notion of this method is to iterate via all the potential threshold values and calculate the expansion of background and foreground pixels. Then find the threshold from the least spread area. The healthy skin part of the images is irrelevant to our model. So, before providing the images into the CNN model, we segmented the lesion part from the healthy skin part using Otsu segmentation as illustrated in Fig. 3.2.

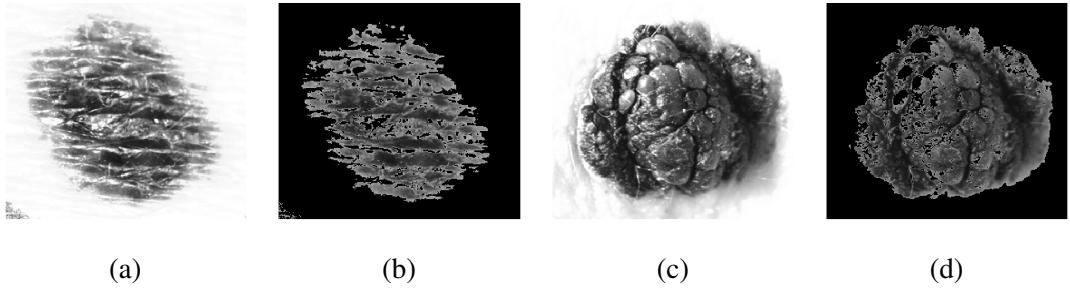


Figure 3.2: Image analysis on BSC corrected gray-image using Otsu segmentation (a) Nevi sample (b) Segmented lesion (c) Melanoma sample (d) Segmented lesion

The algorithm iteratively investigates a threshold point for minimizing the within-class variance. It is a weighted aggregate of variances of the two categories (i.e., foreground and background). Usually, in a gray-level image, colors are between 0-255 (0-1 in the case of float). Hence, If a threshold value of 100 is chosen, all the pixels with values greater than or equal to 100 become the foreground, and the remaining pixels are considered to be the image's background. The within-class variance at any threshold “t” can be discovered by equation 3.2.

$$\sigma^2(t) = \omega_{bg}(t)\sigma_{bg}^2(t) + \omega_{fg}(t)\sigma_{fg}^2(t) \quad (3.2)$$

where the  $\omega_{bg}(t)$  and  $\omega_{fg}(t)$  illustrate the likelihood of a number of pixels for each class at threshold “t” and  $\sigma^2$  carries the variance of color values. Here, the probability denotes,

- $P_{all}$  be the entire computation of pixels in an image,
- $P_{BG}(t)$  be the aggregate of background pixels at threshold “t”,
- $P_{FG}(t)$  be the total of foreground pixels at threshold “t”

Hence, the weights are provided by equation 3.3 and 3.4,

$$\omega_{fg}(t) = \frac{P_{BG}(t)}{P_{all}} \quad (3.3)$$

$$\omega_{bg}(t) = \frac{P_{BG}(t)}{P_{all}} \quad (3.4)$$

And the variance can be formulated by the below equation:

$$\sigma^2(t) = \frac{\sum(x_i - \bar{x})^2}{N-1} \quad (3.5)$$

where,  $x_i$  is the value of the pixel at  $i$  in the class of bg or fg,  $\bar{x}$  stands for pixel values in the class (i.e. bg or fg) and  $N$  denotes the amount of pixels.

---

### **Algorithm 1: Resize Image**

---

**Input:** An 2D image  $[oldIm_{r:c}]$ , where  $r \leq N_r, c \leq N_c$

**Output:** Resized 2D Image of size  $(N_r \times N_c)$

1. Initialization

- (a)  $reqSize = (N_r; N_c)$
- (b) An 2D image  $[newIm_{r:c}]$ ,  $r:c = reqSize$ , where each element initialized with integer 0.
- (c)  $hr \leftarrow \lfloor (reqSize - len(oldIm_{rowsize})) / 2 \rfloor$
- (d)  $hc \leftarrow \lfloor (reqSize - len(oldIm_{colsizes})) / 2 \rfloor$
- (e)  $hcUp \leftarrow hc$

2. **For**  $r \leftarrow 0$  **to**  $len(oldIm_{rowsize})$  **do**

- (a) **For**  $c \leftarrow 0$  **to**  $len(oldIm_{colsizes})$  **do**
  - i.  $newIm[hr][hcUp] \leftarrow oldIm[r][c]$
  - ii.  $hcUp \leftarrow hcUp + 1$

(b) **End For**

- (c)  $hr \leftarrow hr + 1$

- (d)  $hcUp \leftarrow hc$

3. **End For**

4. **Return**  $newIm$

---

The use of a conventional resizing algorithm on the grayscale image would directly distort the skin session shape [75]. So for convenience, keeping all the lesion shapes the same as the origin, we resize all The segmented masks (SM), and Gaussian filtered images (GFI) into a unique size ( $N_r \times N_c$ ) (where  $N_r = 256$  and  $N_c = 256$ ) using Algorithm 1. We apply a binary opening on the resized mask images to remove hair, small objects, and unnecessary things to enhance the lesion area. The segmented mask is used to segment the lesion area (SLA) from the image following equation 3.6 where “r” and “c” represents spatial (plane) coordinates of a 2D image, and the amplitude (r, c) is called the intensity or gray level at the point for that function.

$$SLA_{r;c} = GFI_{r;c} * SM_{r;c} \quad \text{where } (r;c) = 1;2;\dots;(N_c;N_r) \quad (3.6)$$

### 3.2.3 Data augmentation

The data augmentation strategy is a methodology of expanding the amount of data in a dataset [57, 76]. Different image variations are created using data augmentation techniques to reduce model over-fitting during training and increase detection accuracy for unseen data. The basic techniques for increasing the amount of data are rotation, flipping, and zooming (i.e., zoom in, zoom out), as shown in Fig. 3.3. The rotation is carried out using the equations 3.7 and 3.8. The image is also mirrored from the vertical and horizontal directions, as described in equations 3.9 and 3.10.

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad (3.7)$$

$$Img_{rotate}(X;Y) = Img_{R\theta} \quad (3.8)$$

$$Img_{flip-v}(X;Y) = Img(X;-Y) \quad (3.9)$$

$$Img_{flip-h}(X;Y) = Img(-X;Y) \quad (3.10)$$

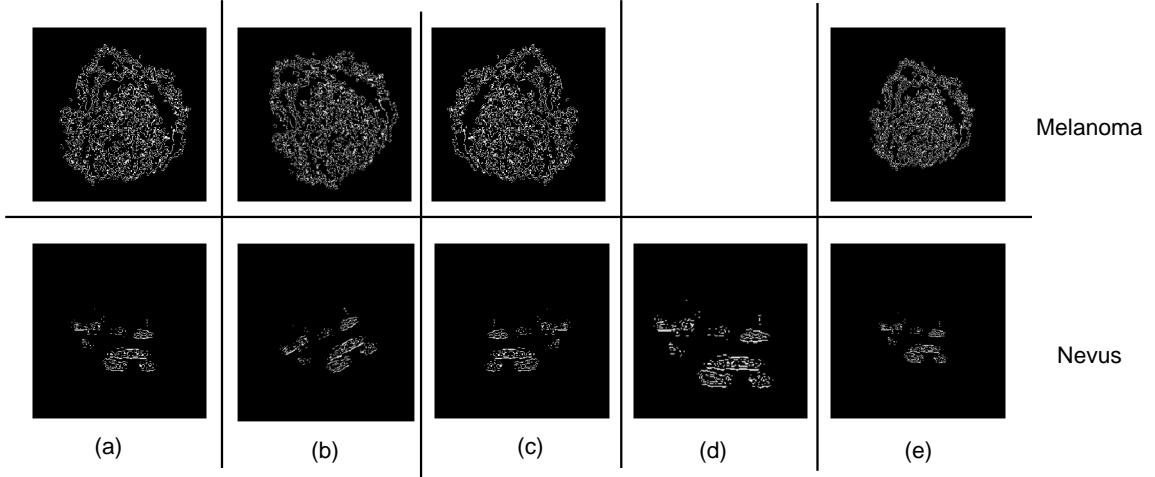


Figure 3.3: Skin lesion data augmentation sample image, (a) Actual image (b) Rotation (c) flip (d) zoom-in (e) zoom out

### 3.2.4 Lesion Feature Analysis

Before feeding the images into the proposed DNN models, skin images' lesions are analyzed so artifacts cannot influence DNN. The analysis is divided into two sub-steps: edge detection and high-intense pixel value estimation and acquisition. For edge detection, we chose the popularly used Canny Edge Detection (CED). In the CED model, to avoid losing precision while converting the image pixel values to unsigned 8-bit, we multiplied the segmented lesion image by 255. The resized segmented lesion images are fed into the proposed IVE model to analyze the high-intense pixels. The IVE model produces a highly intense pixel value articulated in Fig. (3.5g) and Fig. (3.5h).

#### 3.2.4.1 Canny Edge Detection

The Canny Edge Detection (CED) algorithm operates a multi-stage algorithm to detect an extended range of boundaries in images. The algorithm is based on a grayscale image and is a collection of 5 steps as noted below [77],

- Noise lessening
- Gradient computation
- Non-maximum suppression
- Dual threshold

- Boundary Tracking by Hysteresis

As the CED developments are highly sensitive to image noise, it firstly eliminates the image noise by smoothing using Gaussian blur. The image convolution method uses a Gaussian kernel ( $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$ , etc.). The kernel size relies on the anticipated blurring effect. Basically, the larger the kernel, the more the blur is visible. In a 2D space, the general equation of the Gaussian filter can be expressed as equation 3.11:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3.11)$$

where the vertical axis  $y$  and the horizontal  $x$  is the distance from the origin and  $\sigma$  is the standard deviation of the Gaussian distribution. The next step, Gradient calculation, discovers the image gradient to emphasize portions with high-spatial derivatives [77]. It identifies the border intensity and direction by computing the gradient of the picture using border detection operators. The gradient magnitude  $G$  and gradient angle  $\theta$  are formulated as follows:

$$G = \sqrt{(G_x^2 + G_y^2)} \quad (3.12)$$

$$\text{Angle}(\theta) = \arctan\left(\frac{G_y}{G_x}\right) \quad (3.13)$$

Non-maxima suppression (NMS) strategy is based on one of the two approaches commonly utilized for edge identification. Where one is to examine borders as the zero-crossings of the Laplacian of image intensity, and the other is to suppress the local non-maxima of the magnitude of the gradient of image intensity in the direction of this gradient; this procedure is called NMS [77]. NMS is also based on the gradient magnitudes in which the identifier transforms the solid edges of the image to relatively thin and sharp edges for identification purposes [77]. In double thresholding, the threshold value is based on two stages, T1=high threshold and T2=low threshold. When the pixel values of a grayscale level are higher than T1 are denoted as a strong edge pixel, hence the edge area. The outcome is the non-edge portion if the pixels contain a grayscale value slighter than the T2 level. The hysteresis comprises of altering weak pixels into strong ones based on the threshold outcomes, if and only if the pixels around the one being processed have at least one strong pixel, as delineated in Fig. 3.4:

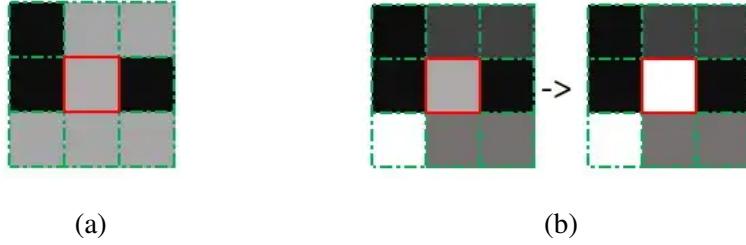


Figure 3.4: Edge Tracking by Hysteresis (a)No adjacent strong pixels (b) Has one adjacent strong pixels

### 3.2.4.2 Proposed Intensity Value Estimation

The intensity value estimation (IVE) model is one of the significant contributions of the work, as intensity ranks act as one of the considerable distinctive characteristics for object or region of interest recognition [11]. In the IVE model, a multiplication between every pixel of the segmented lesion image and a constant value is performed using equation 3.14. Here,  $f$  is a 2D light intensity function, and  $(x;y)$  denotes the spatial coordinates of the image. Then the normalization on the calculated image is used to distribute the intensity of the lesion region pixels using equation 3.15. The normalization procedure transforms the gray-level intensity of the photograph by the range of 0 to 255 Fig. (3.5e) and Fig. (3.5f). The histogram shows the pixel value between 0 and 255, whereas the edge detection of segmented lesion image histogram shows the pixel intensity in 0 or 255 Fig. (3.5c) and Fig. (3.5d). Hence, in the edge detection analysis, no intensity values were found between 0 and 255, which denotes the loss of a few features from the lesion image. To mitigate this issue, our proposed IVE model aids in retaining different intensity levels in the skin lesion images. These varieties of the intensity level of lesion images contain more information (color and texture) in the intensity level form than normal edge detection techniques.

Since using the CED model, the improvement of the effect is not significant; hence we resolved the model to distribute the gray-level intensity of the image using equation 3.15. It is observed that this modification can select the pixel values in the segmented lesion images that are greater or equal to a threshold value equation 3.16.

$$f(x;y) = f(x;y) \times \text{Constant}; \text{ for every } x;y \quad (3.14)$$

$$f(x;y) = \text{normalize}(f(x;y)) \quad (3.15)$$

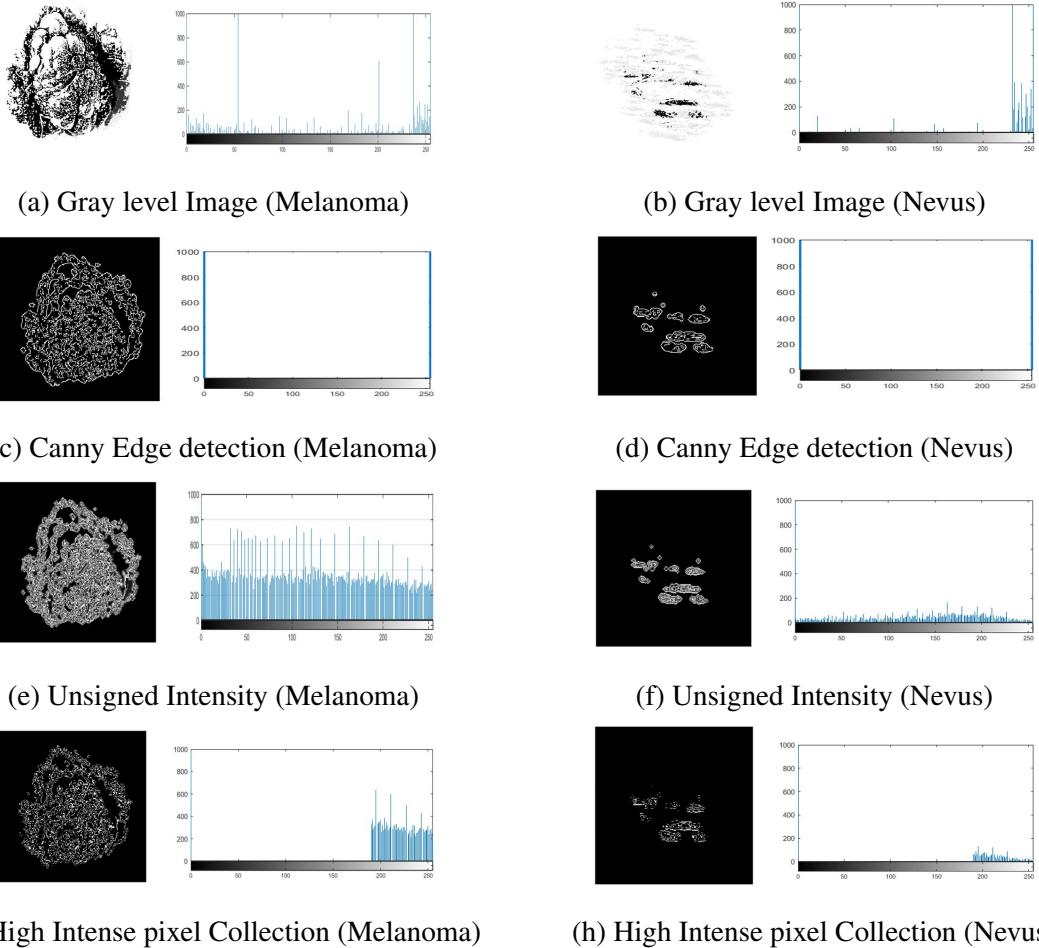


Figure 3.5: Melanoma and Nevus lesion Histogram comparison

$$f(x;y) = \begin{cases} f(x;y); & \text{if } f(x;y) > Threshold \\ 0; & \text{otherwise} \end{cases} \quad (3.16)$$

Using the proposed high-intensity value estimation model produced more features in different intensity levels than the only edge detection of the segmented lesion Fig. (3.5g) and Fig. (3.5h). The proposed model extracts texture and color features in the arrangement of the diverse intensity levels of the pigmentary lesion. In this aspect, choosing intense high pixels helps our model to differentiate Melanoma cells from Nevus cells. This decision experiment is depicted in Fig. 3.5.

### 3.2.5 Deep Neural Network

The Deep Neural Network (DNN) is an emerging machine learning procedure that has established its potential for various classification tasks. The convolutional neural network

dominates significantly in achieving the best outcomes on varying image classification tasks. In this chapter, we embedded the proposed intensity value estimation (IVE) with the Convolutional Neural Network (CNN) model; in short, IVEwCNN to detect and classify Melanoma skin cancer at an early stage.

### 3.2.5.1 Convolutional Neural Network

The researchers [78] found some limitations in using a straightforward artificial neural network (ANN) that often fails at a particular point; specifically an over-fitting could arise due to data size. The Convolutional Neural Network (CNN) has immense benefits over ANN in image classification problems. CNN is a subset of Deep Neural Networks (DNN) that makes an enormous advancement in image pattern recognition, classification, object identification, face recognition, and many more [79]. CNN model aids in self-detecting significant features and extracting these features from the images, which may be beneficial in the image category assessment [57]. It uses filters to extract image features and diminishes the number of total learnable parameters using the max-pooling function.

The CNN model comprises five stages of neural layers in its structure: the input layers, convolutional layer (Convo + ReLu), pooling layer, fully connected layer, and output layer. We illustrated the proposed CNN model architecture in Fig 3.6. CNN model accepts the input as a 3-dimensional matrix (*Width*  $\times$  *Height*  $\times$  *Dimension*) to the input layer. The dimension of the input layer denotes the number of image channels. For a grayscale image, the dimension is 1, and for an RGB image, it is defined as 3. The first layer of the CNN model is the input layer, where the input image with size  $256 \times 256 \times 1$  is given. Then input image is passed to the convolutional layer, where filters are applied to the original images. The filters slide over the receptive fields of the same input image by a stride and continue through the whole image. The convolutional layer uses the ReLu activation function (equation 4.2) to zero all negative values. Moreover, this layer is responsible for feature extraction from the input image. Mathematically, convolution is usually narrated with an asterisk\* symbol. Suppose an input image has a  $(N \times N)$  square neuron layer followed by a convolutional layer. If we utilize an  $(m \times m)$  filter denoted as  $\omega$ , then the output of the convolutional layer will be of size  $(N - m + 1) \times (N - m + 1)$  following equation 3.17.

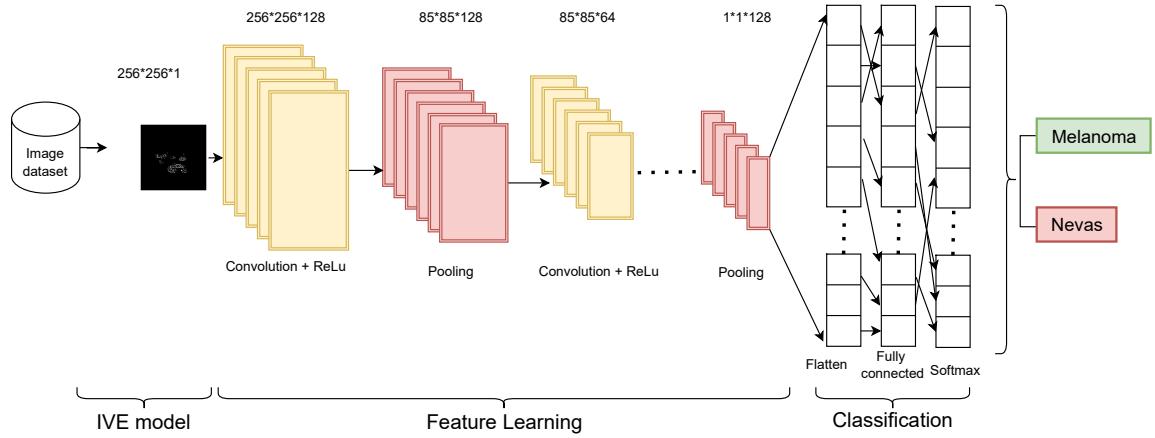


Figure 3.6: Proposed IVEwCNN model workflow

$$x_{ij}^l = \sum_{a=0}^{m-1} \sum_{b=0}^{m-1} \omega_{ab} y_{(i+a)(j+b)}^{l-1} \quad (3.17)$$

The approach of the simple convolution strategy is illustrated in Fig 3.7 and 3.8 where a  $(3 \times 3)$  size image and a  $(2 \times 2)$  size of the filter are taken.

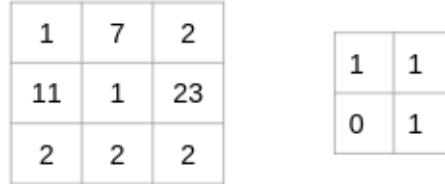


Figure 3.7: Understanding convolution with  $(3 \times 3)$  image and  $(2 \times 2)$  filter

The filter  $(2 \times 2)$  slides through the  $(3 \times 3)$  image using stride 1, then an element-wise multiplication is performed, and finally, the outcomes are aggregated together as delineated below,

- $(1 \times 1 + 7 \times 1 + 11 \times 0 + 1 \times 1) = 9$
- $(7 \times 1 + 2 \times 1 + 1 \times 0 + 23 \times 1) = 32$
- $(11 \times 1 + 1 \times 1 + 2 \times 0 + 2 \times 1) = 14$
- $(1 \times 1 + 23 \times 1 + 2 \times 0 + 2 \times 1) = 26$

A pooling layer is used after each convolutional layer to reduce the spatial volume of the input image. It simply takes some  $(k \times k)$  part of an image and produces a single value that

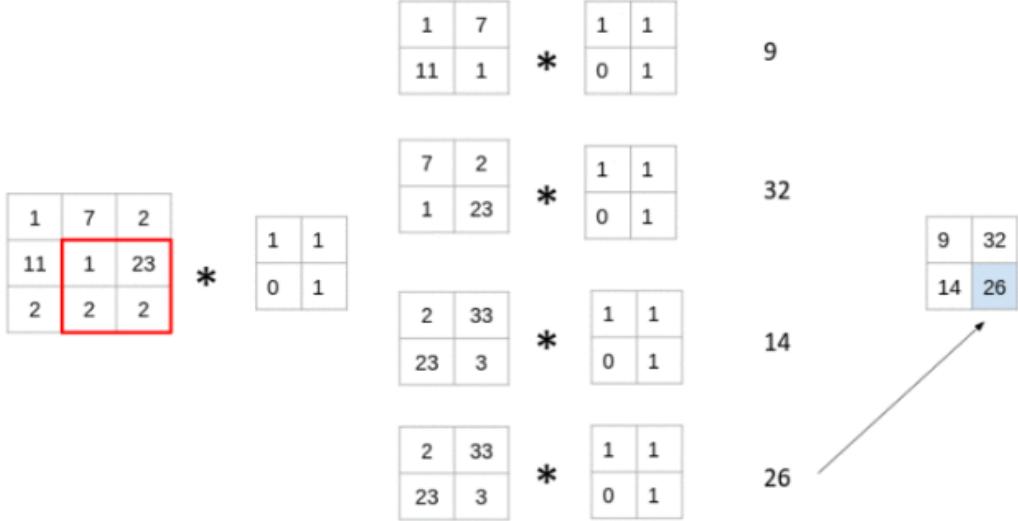


Figure 3.8: Process of convolution with  $(3 \times 3)$  image and  $(2 \times 2)$  filter

is a maximum in that part. For instance, from an  $(N \times N)$  input layer, it will produce an  $(\frac{N}{K} \times \frac{N}{K})$  layer, as utilizing the max function each  $(k \times k)$  region is reduced to just a single value.

The outputs from the earlier layer are then fed into the fully connected layer, flattened the inputs, and converted into a single vector. This layer involves weights, biases, neurons, and activation functions and updates weights in the training session. Here, we utilize the rectified linear activation unit (ReLU activation) function to propagate the outputs from each layer. The output layer is the final layer that uses the Softmax activation function (equation 4.1) to estimate Melanoma skin cancer and Nevus mole probability. The entire CNN model has illustrated in Table 3.1.

### 3.3 Dataset

The proposed model is estimated on a publicly known standard Melanoma dataset, i.e., MED-NODE. The dataset has macroscopic high-resolution skin lesion images [49]. MED-NODE image dataset has a total of 170 images, where 70 images are Melanoma affected and 100 nevi classes. The Department of Dermatology of the University Medical Center Groningen (UMCG) gathered 50,000 images in the digital image archives (DIA). The MED-NODE dataset is the subset of DIA. All of these pictures have been examined and

Table 3.1: Architecture of CNN Layers.

Layers	Output Size	Kernel Size	Activation Function	Rate
Input Layer	256×256×1	-	-	-
Conv2D	256×256×128	7×7	ReLU Activation	-
MaxPooling2D	85×85×128	3×3 (strides = 3)	-	-
Conv2D	85×85×64	4×4	ReLU Activation	-
MaxPooling2D	28×28×64	3×3 (strides = 3)	-	-
Conv2D	28×28×32	5×5	ReLU Activation	-
MaxPooling2D	14×14×32	2×2 (strides = 2)	-	-
Conv2D	14×14×128	6×6	ReLU Activation	-
MaxPooling2D	4×4×128	3×3 (strides = 3)	-	-
Conv2D	4×4×32	5×5	ReLU Activation	-
MaxPooling2D	2×2×32	2×2 (strides = 2)	-	-
Conv2D	2×2×128	6×6	ReLU Activation	-
MaxPooling2D	1×1×128	2×2 (strides = 2)	-	-
Flatten	128	-	-	-
Dense182	512	-	ReLU Activation	-
Dropout	-	-	-	0.4
Dense183	128	-	ReLU Activation	-
Dropout	-	-	-	0.3
Dense184	64	-	ReLU Activation	-
Dense185	512	-	ReLU Activation	-
Dropout	-	-	-	0.6
Dense186	512	-	ReLU Activation	-
Dropout	-	-	-	0.3
Dense187	64	-	ReLU Activation	-
Dense188	2	-	Softmax Activation	-

assessed by a dermatologist as being of the utmost cognition [49]. All of these images have been investigated and evaluated by a specialist dermatologist [49]. These images were taken in JPEG format with a Nikkor lens on a Nikon D3 and Nikon D1x camera from around 33 cm from the skin lesion area. Some example images from the MED-NODE dataset are

portrayed in Fig. 3.9 and 3.10. The stability of the dataset has been assured by examining the following characteristics [49].

- MED-NODE dataset is created by loosely selecting images.
- The dataset included superficial spreading Melanoma and nevi cases.
- Each photograph is acquired from a diverse patient (aside from the picture, which illustrates how the condition remarkably varies in different physical regions).
- Each image is properly sharp and well exposed to annotate precisely.
- Each image represents the class it is associated with.

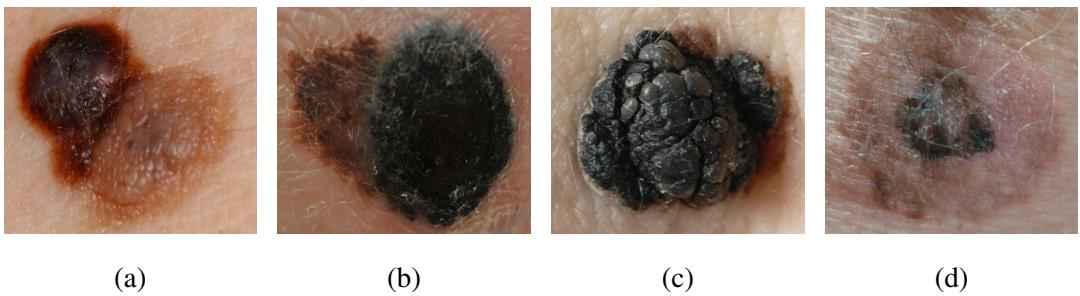


Figure 3.9: Sample Melanoma skin cancer images

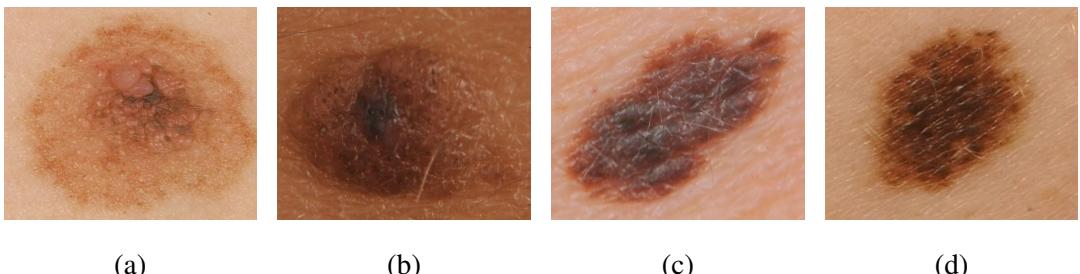


Figure 3.10: Sample Nevus mole images

## 3.4 Experimental Analysis

Some of the dataset image resolutions are above  $(N_r \times N_c)$  (where  $N_r = 256$  and  $N_c = 256$ ), which needs a high cost of computing power. As the rescaled lesion images would be a better fit for a deep-learning model [75], we used pre-processing steps to resize the images into

$(N_r \times N_c)$  using our improved resizing algorithm 3.2.2 and remove unwanted artifacts. The limited amount of data could misguide and overfit a deep neural network model. Hence, to expand the amount of data and variations, the image augmentation technique is performed [80] to reduce the data misleading and overfitting. Then the proposed IVE model is conducted to obtain highly intense pixel values from every segmented lesion image, ensuring significant information about Melanoma and Nevus mole. Feeding the substantial feature value to the DNN model achieves a better result. Authors [81] discovered that for the training/ validation set, the number of available adjustable parameters for the dataset should be inversely proportional to the square of the significant portion of patterns. As a result, the 80/20 split was determined to avoid overtraining the DNN models. Most of the state-of-the-art approaches [82, 83, 49, 54] that use the identical or smaller dataset operated on 80% data for training and 20% data for testing purposes.

### 3.4.1 Training and Testing

With the collaboration of the image augmentation technique, the system first generates a total of 2028 images. Then it works with a much larger dataset than the previous one. For the experiment, 80% of the images are utilized for training, and the rest 20% is used for testing purposes. Our system split the dataset such that no overlapping occurs in the train and test data set. We set 40 as the number of epochs to train the CNN model. For optimization purposes, we used “Adam Optimization” and “Sparse Categorical Cross entropy” to calculate the model’s loss. We chose average results from the 5-fold cross-validation to avoid biased performance results of our proposed model. We implemented our proposed model in Colab Notebook, a product of the google cloud platform. This allows the utilization of a 12 GB NVIDIA Tesla K80 GPU that can be conducted for up to 12 hours successively and is highly integrated with Google Drive.

### 3.4.2 Performance evaluation metrics

The right metrics are required to estimate a machine learning (ML) model based on application and purposes. The most vital aspect of the evaluation metrics is the capability of discriminating among model results. To evaluate our proposed model in this work, we choose the seven most popular metrics (i.e., Sensitivity or Recall, Specificity, positive predicted value or PPV, negative predicted value or NPV, Accuracy, F1-Score, and Region of

curve and area under curve or ROC-AUC). Sensitivity or Recall is another critical metric. For instance, suppose an investigation for a condition in a medical diagnosis. Sensitivity denotes the correct test's ability to recognize sick patients who carry the condition. This can be characterized as the fraction of examples from a class that is truly predicted by the ML model as denoted in (3.18) [84].

$$\text{Sensitivity} = \left( \frac{\text{true detected melanoma cases}}{\text{all Melanoma cases}} \right) \times 100 \quad (3.18)$$

An optimistic outcome in an investigation with high specificity is useful for governing disease. A true result signifies a high probability of the existence of the disease. The specificity of an investigation is the ratio of those who truly test negative for the condition or do not have the condition. The specificity can also be defined mathematically as (3.19):

$$\text{Specificity} = \left( \frac{\text{true detected non melanoma cases}}{\text{all non melanoma cases}} \right) \times 100 \quad (3.19)$$

There may be a situation where the classification accuracy may not be a good indicator of a model's performance. For instance, when the class distribution is imbalanced i.e. if one category is significantly higher than others. The precision can be defined as (3.20):

$$PPV = \left( \frac{\text{true detected melanoma cases}}{\text{detected melanoma cases}} \right) \times 100 \quad (3.20)$$

The greater the negative predictive value (NPV) (3.21), the less likely to be an individual with a negative test will have. The NPV is the proportion of subjects correctly diagnosed as negative to all the others who also had negative test results (embracing patients who were investigated wrongly as healthy). This characteristic can predict how likely someone will be healthy in case of a negative test result.

$$NPV = \left( \frac{\text{true detected non melanoma cases}}{\text{detected non melanoma cases}} \right) \times 100 \quad (3.21)$$

The accuracy metric is perhaps one of the simplest metrics in the classification using an ML model, and this can be defined as the total number of true predictions divided by the total number of predictions and then reproduced by 100; denoted as (3.22).

$$Accuracy = \left( \frac{\text{true detected cases}}{\text{all cases}} \right) \times 100 \quad (3.22)$$

Both recall and precision can be significant in many applications. Hence, a mathematical equation (3.23) could be developed to incorporate these two into a single metric. The F1-scoring technique is one of the popular metrics which combines both, and this states a

harmonic mean of precision and recall outcomes. F1-score is mostly used in comparing the performance of classifiers [84].

$$F1score = 2 \times \left( \frac{Precision \times Recall}{(Precision + Recall)} \right) \times 100 \quad (3.23)$$

The receiver operating characteristic (ROC) essentially shows the true positive rate (TPR) against the false positive rate (FPR) for different threshold points. The area under the curve (AUC) computes the area under the ROC curve, which is between 0 and 1. The AUC is an accumulated estimate of the performance of a binary classifier on all potential threshold points(hence it is threshold invariant). In an ML model evaluation, the higher the AUC value, the more satisfactory it is [84].

### 3.5 Experimental Results

We incorporated the Canny edge detection (CED) and Intensity value estimation (IVE) model with CNN to analyze the results set described in Tables 3.2 and 3.3. Finally, the average results are compared in Fig. 3.11. Table 3.2 holds the results of the CED model with CNN (CEDwCNN). Four folds Fold-1, Fold-2, Fold-4, and Fold-5 contain high performances. Though in Fold-1, two metrics, sensitivity (91.60%) and NPV (93.83%), show top performances. However, the other five carry the least performance results comparing other folds. The worst performance of sensitivity (82.52%) and NPV (85.88%) are found in Fold-4. The Fold-4 also contains the best performances for two metrics, specificity (94.41%) and PPV (92.91%). F1-score shows the highest value at the Fold-5. The best accuracy and ROC-AUC value are found in Fold-2 as 90.46% and 90.35%, respectively. Overall, The CEDwCNN model reached 88.63% average accuracy.

The exploratory results of IVEwCNN are expressed in table 3.3. Noticeable performances are seen in the Fold-2. Five metrics, specificity (97.62%), PPV (96.75%), accuracy (93.09%), F1-score (91.89%), and ROC-AUC (92.56%), achieved the best results in the Fold-2 among other folds. Though the highest marks were achieved by sensitivity (96.24%) and NPV (96.71%) in Fold-4, this fold also contains the least result of specificity (85.96%) and PPV (84.21%). The least NPV (89.27%) and accuracy (88.16%) are found in the Fold-3. Fold-5 contains the three least performed metrics sensitivity (83.90%), F1-score (84.98%),

and ROC-AUC (87.65%). The average accuracy of the IVEwCNN model is calculated as 90.66% which is almost 2% better than the CEDwCNN model.

Table 3.2: 5-Fold cross-validation experimental result(%) analysis using CEDwCNN model (best accuracy shown in **bold**)

Metrics	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Average
Sensitivity	<b>91.60</b>	89.47	83.78	82.52	86.89	86.85
Specificity	82.16	<b>91.22</b>	93.37	<b>94.41</b>	90.66	90.36
PPV	76.76	88.81	91.40	<b>92.91</b>	86.18	87.21
NPV	<b>93.83</b>	91.76	88.07	85.88	91.16	90.14
Accuracy	85.86	<b>90.46</b>	89.47	88.20	89.15	88.63
F1-score	83.53	89.14	87.97	87.41	<b>89.53</b>	87.52
ROC-AUC	86.88	<b>90.35</b>	89.07	88.46	88.77	88.71

Table 3.3: 5-Fold cross-validation experimental result(%) analysis using IVEwCNN model (best accuracy shown in **bold**)

Metrics	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Average
Sensitivity	88.50	87.50	85.27	<b>96.24</b>	83.90	88.28
Specificity	95.82	<b>97.62</b>	90.29	85.97	91.40	92.22
PPV	92.59	<b>96.75</b>	86.61	84.21	86.09	89.25
NPV	93.38	90.61	89.27	<b>96.71</b>	89.95	91.98
Accuracy	<b>93.09</b>	<b>93.09</b>	88.16	90.46	88.49	90.66
F1-score	90.50	<b>91.89</b>	85.94	89.82	84.98	88.63
ROC-AUC	92.15	<b>92.56</b>	87.78	91.10	87.65	90.25

The CNN model performance is briefed in Tables 3.2 and 3.3. A total of 3 folds, Fold-1, Fold-2, and Fold-4, show better accuracy in our proposed IVE model compared to the CED model, whereas the other two folds, Fold-3 and Fold-5, performed lesser in both models. However, between the models, the best performance for the F1-score is achieved in the Fold-5. Comparing both models, the best accuracy of 93.09% is achieved by Fold-1 and Fold-2 using the IVE model. Overall, the average performance of the CNN model in Fig. 3.11 illustrates the best outcomes using the IVE model. Whereas the IVE model achieved

almost 1.5% higher performances considering all the mentioned metrics.

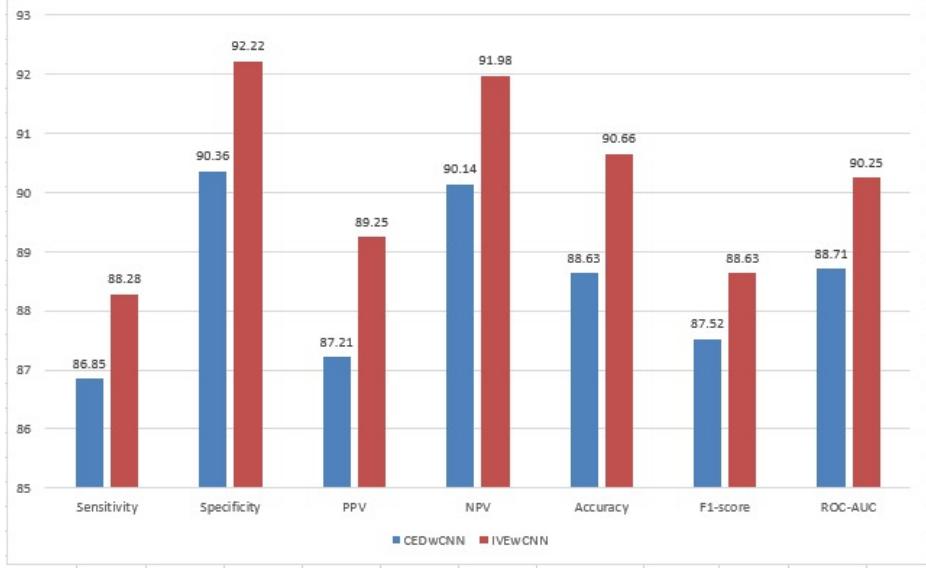


Figure 3.11: Average performance analysis of the proposed IVE and CED model embedded with CNN for Melanoma detection and classification

## 3.6 Discussion

It is demonstrated in the above analysis that pixel intensity levels serve as one of the distinctive features for identifying an object or region of interest, particularly when considering Melanoma skin cancer. Textural features are regularly used in image classification because they enhance the classification of Nevus and Melanoma by computing the abnormality of their arrangement [11]. Since choosing the high-intense pixels gives more information, the model can easily differentiate Melanoma cells from Nevus cells more precisely. Therefore the proposed IVE model shows a better result when compared with the CED model.

## 3.7 Conclusion

This chapter discusses an ingenious IVEwCNN approach to detect and categorize two types of mole (i.e., Melanoma and nevi) from macroscopic skin lesion images. Here, we analyzed the experimental result acquisition by collaborating the proposed IVE model and the popularly used CED model with the CNN model by examining the seven model evaluation metrics. Finally, we compared the average performances of the proposed IVE model with

the generally used CED model to conclude our contribution. This strategy works successfully in feature obtaining as well as in classification. Hence, we conclude that the intensity value-based proposed IVE model aids the DNN models in extracting features efficiently in the convolutional layers.

Though with other pre-trained transfer learning models, VGG-16 and inception v3 is the most popular choice in medical image classification; however, VGG-16 has been seen to outperform the inception v3 model [44]. In the next chapter, we embedded the innovative IVE model with other pre-trained deep learning models to investigate the IVE performance analysis in skin cancer detection by transfer learning (TFL) approach.

# **Chapter 4**

## **Enhanced Classification using Transfer Learning**

---

### **4.1 Introduction**

The previous chapter showed that the proposed IVE-based algorithm exhibited more promising results than the traditional edge detection-based algorithm. This chapter presents an enhanced classification algorithm in which the IVE model is integrated with the different transfer learning models. Here we discussed the experimental outcomes by embedding the IVE and CED models with two transfer learning strategies (i.e., VGG-16 and Inception V3 model) to detect Melanoma skin cancer efficiently.

### **4.2 Transfer Learning Model**

Training a new Deep Neural Network (DNN) model for efficient classification, massive data, and hardware supports [85] are required. In this approach, the model consumes much time while training the model on a massive dataset. Transfer Learning (TFL) model commonly transfers its prior knowledge to a second network that is already trained on the target dataset.

### 4.2.1 Transfer Learning using VGG-16

This section has prepared a DNN by employing the VGG-16 pre-trained model. The VGG-16 model contains ten pre-trained convolution layers and four max-pooling layers in which each convolution layer uses filtering of size  $(3 \times 3)$ . The pooling layers reduce the extracted feature in convolution layers. As prior knowledge, the popular “ImageNet” dataset is used in the pre-trained VGG-16 model. Then a new sub-layer model is designed, which comprises with seven layers (three convolution layers, three dropout layers, and one softmax layer) Fig. 4.1. To develop the proposed TFL model, the pre-trained VGG-16 model is incorporated with the new sub-model. The ReLu activation function (4.2) is incorporated in the convolution layers, and the final layer operates the softmax activation function (4.1) to predict the output. During the training phase, the proposed VGG-16 TFL model used a “sparse categorical cross-entropy” loss function and “RMSprop( $lr=2e-5$ )” optimizer to compute the volume that the VGG-16 model should aim to minimize.

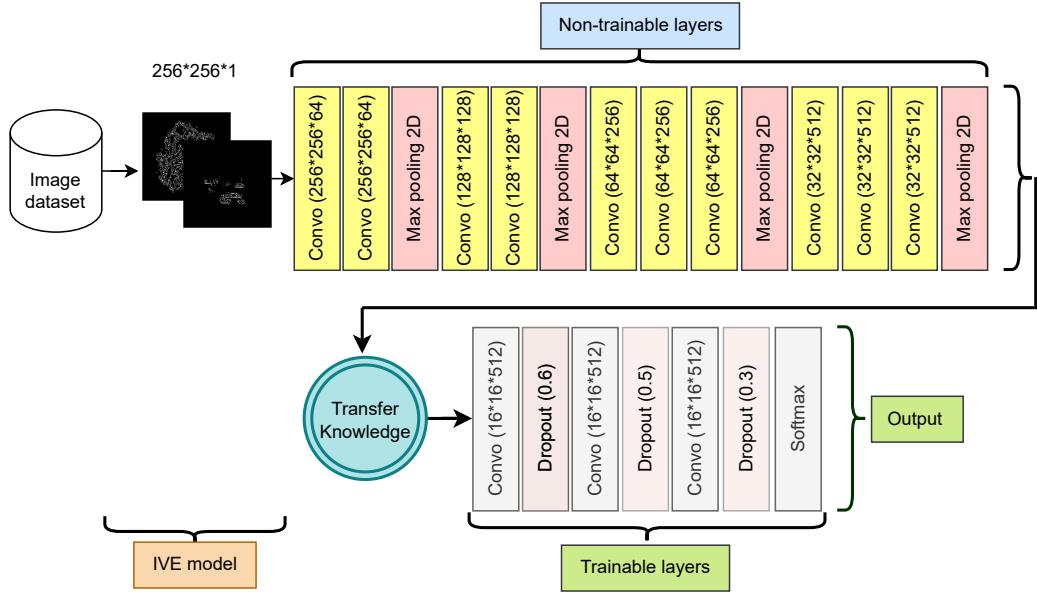


Figure 4.1: Architecture of the Proposed IVEwVGG-16 Transfer Learning Model

$$S(z_i) = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}} \quad (4.1)$$

where,  $z$  = output layer's value from the neuron.

$$ReLU(x) = \max(0; x) \quad (4.2)$$

where,  $x$  = input to a neuron.

### 4.2.2 Transfer Learning using Inception v3

Inception V3 (IV3) primarily focuses on reducing exhaustive computing power by utilizing the past Inception model. This concept was introduced in a paper entitled “Rethinking the Inception Architecture for Computer Vision” in 2015 [47]. The Inception V3 model suggests a variety of methods for network optimization, including deconstructed convolutions, regularization, dimensionality reduction, and parallel computation. To accelerate the calculation further, the  $(3 \times 3)$  Convolution is decomposed into 2 one-dimensional convolution concatenations ( $1 \times 3$  and  $3 \times 1$ ). It ameliorates the network depths, hence raising the nonlinearity of the network (the ReLU activation function is essential for each additional layer). The considerable improvement in Inception V3 is the Factorization, which separates the  $(7 \times 7)$  convolution into two one-dimensional concatenations form. We used the transfer learning strategy in the IV3 model, which retains the previous layer’s parameters while removing the last layer from the IV3 model by setting “`include_top=False`”. In the experiment, in the IV3 model, we set the last 15 layers as “`layer.trainable = True`”. Then a new flattened sub-layer is designed, which comprises seven layers (four dense layers with ReLu activation, three dropout layers, and one softmax layer) illustrated in Fig. 4.2. The IV3 model is then integrated with the new sub-model to train the trainable layers with the prepared data and predict the output from the final layer. From the output layer, the number of output nodes denotes the category numbers in the dataset. As we are operating on binary classification, the output nodes will be two in this case. To compare the IV3 TFL model with VGG-16 TFL, here we also used the popularly preferred dataset “ImageNet” dataset as the model’s prior knowledge. In the model, from a total of 59,880,610 parameters, trainable parameters were 38,472,322, and the non-trainable parameters were 21,408,288.

During the training phase, the proposed IV3 TFL model used a “sparse categorical cross-entropy” loss function and “`RMSprop(lr=2e-5)`” optimizer to compute the volume that the IV3 model should aim to minimize.

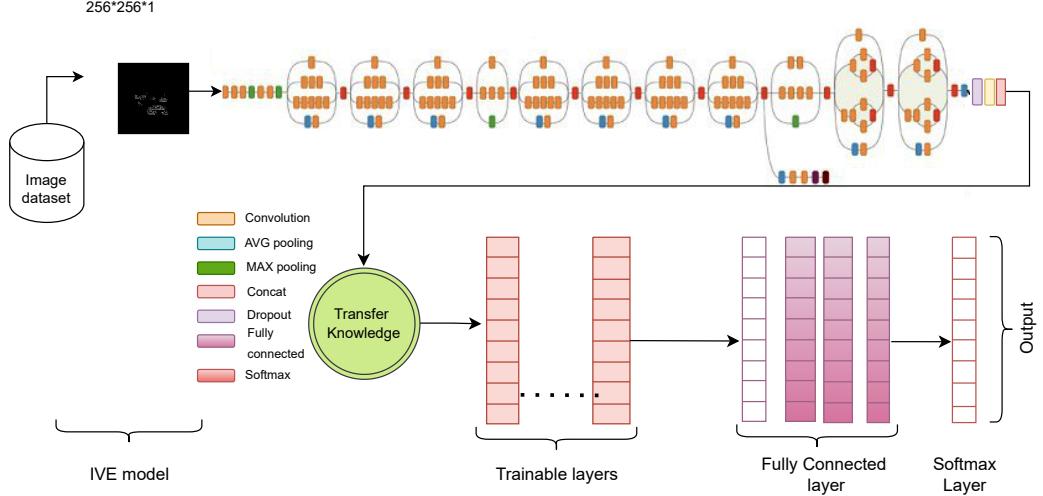


Figure 4.2: Architecture of the Proposed IVEwIV3 Transfer Learning Model

## 4.3 Performance Evaluation

For experimental analysis, the traditional feature extraction approach (CED) and the proposed IVE-based feature extraction method are integrated with the various transfer learning models that were described previously. Finally, extensive experiments are conducted to evaluate all algorithms to classify Melanoma skin cancer and Nevus mole.

### 4.3.1 IVE with VGG-16 based Transfer Learning Algorithm

In the experiment, we utilized CED and IVE methods with a VGG-16 transfer learning model to analyze the effectiveness of our proposed IVE model. Table 4.1 shows 5-fold CEDwVGG-16 TFL results with average performances. For the CED model, VGG-16 TFL performs best at the Fold-3 for five metrics sensitivity (94.35%), NPV (96.03%), accuracy (94.07%), F1-score (92.85%), and ROC-AUC (94.12%). However, the Fold-5 contains the least performed result of four metrics, i.e., sensitivity (80.58%), NPV (85.79%), ROC-AUC (89.68%), and for the accuracy metric, both the Fold-1 and Fold-5 performed worst (90.46%). The Fold-1 performed least for specificity (92.70%) and PPV (89.43%) metrics, whereas Fold-5 achieved the best result of 98.79% and 98.25%, respectively. However, the average outcome shows their performances above 90% except for sensitivity (88.16%).

Our proposed model in Table 4.2 shows the best-performed results of sensitivity (96.03%),

Table 4.1: 5-Fold cross-validation experimental result(%) analysis using using CEDwVGG-16 transfer learning model (best accuracy shown in **bold**)

Metrics	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Average
Sensitivity	87.30	88.03	<b>94.35</b>	90.52	80.58	88.16
Specificity	92.70	96.91	93.89	95.74	<b>98.79</b>	95.61
PPV	89.43	96.15	91.41	92.92	<b>98.25</b>	93.63
NPV	91.16	90.23	<b>96.03</b>	94.24	85.79	91.49
Accuracy	90.46	92.76	<b>94.07</b>	93.75	90.46	92.30
F1-score	88.35	91.91	<b>92.85</b>	91.70	88.54	90.67
ROC-AUC	89.99	92.47	<b>94.12</b>	93.13	89.68	91.88

Table 4.2: 5-Fold cross-validation experimental result(%) analysis using using IVEwVGG-16 transfer learning model (best accuracy shown in **bold**)

Metrics	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Average
Sensitivity	94.02	92.56	91.26	<b>96.03</b>	95.69	93.91
Specificity	94.65	97.27	<b>98.31</b>	96.07	92.02	95.66
PPV	91.67	95.73	<b>97.46</b>	94.53	88.10	93.50
NPV	96.19	95.19	93.09	97.15	<b>97.19</b>	95.76
Accuracy	94.41	95.40	95.39	<b>96.05</b>	93.42	94.93
F1-score	92.83	94.18	94.26	<b>95.28</b>	91.74	93.66
ROC-AUC	94.33	94.91	94.79	<b>96.04</b>	93.86	94.79

accuracy (96.05%), F1-score (95.28%), and ROC-AUC (96.04%) at the Fold-4. Along with the two best-performed metrics, specificity (98.31%) and PPV (97.46%), the Fold-3 also has two worst-performed metrics, sensitivity (91.26%) and NPV (93.09%). Though Fold-5 contains the best-performed metric NPV (97.19) however, the five least-performed metrics specificity (92.02%), PPV (88.10%), accuracy (93.42%), F1-score (91.74%), and ROC-AUC (93.86%) are also found for IVEwVG-16 TFL. Overall, for every metric, the VGG-16 transfer learning model achieved more than 93% result for the IVE model.

To sum up the overall performance of VGG-16 TFL for efficient Melanoma detection, we summarize the average performance in Fig. 4.3 Though two metrics specificity and PPV

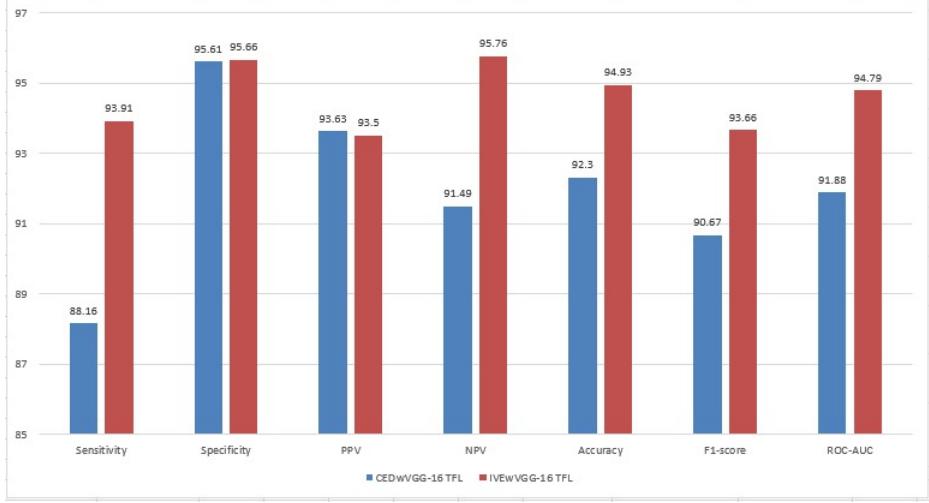


Figure 4.3: Average performance analysis of the proposed IVE and CED model embedded with VGG-16 transfer learning for Melanoma detection and classification

performed very closely for both the models (i.e. CED and IVE). However, the figure states that every other metric in the IVE model outperformed the CED model by a large margin.

### 4.3.2 IVE with Inception V3 Transfer Learning Algorithm

The inception V3-based transfer learning (IV3 TFL) approach is evaluated with the IVE and CED models. In Table 4.3, least performances are shown in Fold-1 and Fold-4. The best sensitivity (92.75%) achieves at Fold-5 and falls to 87.90% in Fold-4. Fold-2, Fold-3 and Fold-5 contain the best results of NPV (94.08%), ROC-AUC (92.58%) and F1-score (91.76%) respectively. However, the Fold-1 shows significantly lesser performances of specificity (91.52%), NPV (90.41%), accuracy (90.13%), F1-score (89.13%), and ROC-AUC (90.00%). The least PPV value (88.98%) was found at the Fold-2. Moreover, the average performances of the CEDwIV3 TFL model are found more than 90.0%.

The performances of the suggested IVEwIV3 TFL model are delineated in Table 4.4. In the Fold-5 we found four metrics PPV (94.96%), accuracy (94.74%), F1-score (94.29%), and ROC-AUC (94.70%) achieved the best results. The least performance for specificity (93.75%) and PPV (91.60%) are shown in Fold-3; however, the fold also performed best for sensitivity (93.75%) and NPV (95.38%). The notable specificity (96.91%) and accuracy (94.74%) are found in Fold-4. Other than the sensitivity metric (91.08%) all metrics achieved more than 92.0% average result.

Table 4.3: 5-Fold cross-validation experimental result(%) analysis using using CEDwIV3 transfer learning model (best accuracy shown in **bold**)

Metrics	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Average
Sensitivity	88.49	90.52	92.02	87.90	<b>92.75</b>	90.34
Specificity	91.52	93.09	92.17	<b>96.67</b>	92.17	93.12
PPV	89.78	88.98	90.71	<b>94.78</b>	90.78	91.01
NPV	90.41	<b>94.08</b>	93.29	92.06	93.87	92.74
Accuracy	90.13	92.10	92.12	<b>93.09</b>	92.43	91.97
F1-score	89.13	89.74	91.37	91.21	<b>91.76</b>	90.64
ROC-AUC	90.00	91.80	<b>92.58</b>	92.28	92.50	91.83

Table 4.4: 5-Fold cross-validation experimental result(%) analysis using using IVEwIV3 transfer learning model (best accuracy shown in **bold**)

Metrics	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Average
Sensitivity	90.58	86.55	<b>93.75</b>	90.90	93.62	91.08
Specificity	94.58	96.76	93.75	<b>96.91</b>	95.71	95.54
PPV	93.28	94.50	91.60	94.34	<b>94.96</b>	93.74
NPV	92.35	91.79	<b>95.38</b>	94.95	94.55	93.80
Accuracy	92.76	92.76	93.75	<b>94.74</b>	<b>94.74</b>	93.75
F1-score	91.12	90.35	92.66	92.59	<b>94.29</b>	92.20
ROC-AUC	92.60	91.70	93.75	93.90	<b>94.70</b>	93.33

Overall, the least performances of specificity (91.52%), NPV (90.41%), Accuracy (90.13%), F1-score (89.13%), and ROC-AUC (90.0%) found in Fold-1 by using CED method. In contrast, the top performance of the F1-score (94.29%) and ROC-AUC (94.70%) is achieved by the IVE model in Fold-5. The best Accuracy (94.74%) is achieved both in the Fold-4 and Fold-5 using our developed IVE model. Moreover, in every fold, the Accuracy is significantly higher using the IVE model. The top sensitivity (93.75%) and NPV (95.38%) were found in the Fold-3 using the IVE model. The best PPV (94.96%) and specificity (96.91%) were achieved using the IVE model for Fold-5 and Fold-4, respectively. However, both metrics performed worst using the CED model. Fig. 4.4 demonstrates that the average

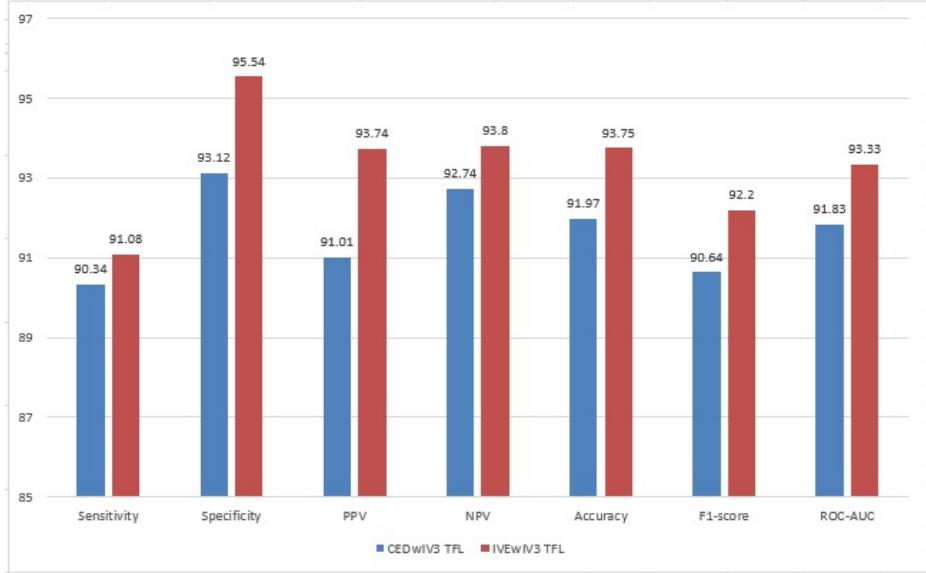


Figure 4.4: Average performance analysis of the proposed IVE and CED model embedded with Inception V3 transfer learning for Melanoma detection and classification

performance of the proposed IVEwIV3 TFL method outperformed the conventional CED-wIVE3 TFL model.

### 4.3.3 Discussion

The proposed IVE model estimates the high-intensity pixel values from the training dataset, which are incorporated into our developed DNN models (CNN, VGG-16 TFL, and IV3 TFL) to detect and classify Melanoma skin cancer and Nevus mole. Table 3.3, 4.4 and 4.2 show the 5-fold cross-validation experimental results of the proposed IVEwCNN, IVEwIV3 TFL, and IVEwVGG-16 TFL model respectively. We showed a comprehensive performance analysis between the popular choice in Melanoma identification i.e., Canny Edge detection (CED), and our proposed intensity value estimation (IVE) technique to extract more relevant features in DNN convolutional layers.

## 4.4 Conclusion

Here we discussed the experimental results by incorporating the proposed IVE model and the popularly used CED with a DNN model (i.e., VGG-16 TFL and IV3 TFL) considering seven popularly used metrics. The results of two models, IVE and CED, are compared to

evaluate the significance of the proposed IVE model. To conclude, we plotted the average, which summarized that our proposed model IVEwVGG-16 TFL outperforms. Finally, to conclude the overall performance, we delineate the average results of all algorithms that are described in the next Chapter.

# **Chapter 5**

## **Experimental Analysis**

---

### **5.1 Introduction**

This chapter includes the experimental result discussion of our proposed model. We aim to quantify the proposed model's significance by comparing state-of-the-art methods considering the sensitivity, specificity, PPV, NPV, accuracy, F-1 score, and ROC-AUC value.

### **5.2 Analysis**

In this thesis, instead of conventional visual observations, an efficient expert system is developed with the collaboration of a DNN model to assist expert physicians in the early-stage detection and classification of Melanoma skin cancer. The proposed IVEwDNN models (CNN, VGG-16 TFL, and IV3 TFL) adopt the new intensity value estimation (IVE) technique in which the high-intense pixel values are estimated from each segmented lesion image after rescaling the image. Most of the study has been conducted by performing lesion segmentation and detecting the edges of the segmented lesion, which is then fed into a deep neural network. The contrast enhancement and texture analysis in the pre-processing step increase the discrimination between the intensity of values of an image that improves the overall performance [56].

Our proposed method performs this more precisely and accurately using three stages. Firstly,

in pre-processing step, a new image is generated consisting of a shape of  $(N_r \times N_c)$  where the lower dimension of segmented image pixels are mapped to retain the original lesion shape. In order to keep the lesion information, such as the lesion area, size, diameter, and border, we kept the lesion shape as the original. Then we performed the proposed intensity value estimation (IVE) model on the final pre-processed images to aid the convolutional layer in identifying the most significant features. Finally, we used three DNN models (i.e., CNN, VGG-16 TFL, and IV3 TFL) to train the models and efficiently classify Melanoma from Nevus mole. Then the IVE model is employed as shown in Fig. 5.1 and Fig. 5.2 to compute the high-intensity pixels that store descriptive features on object or region of interest identification [11].

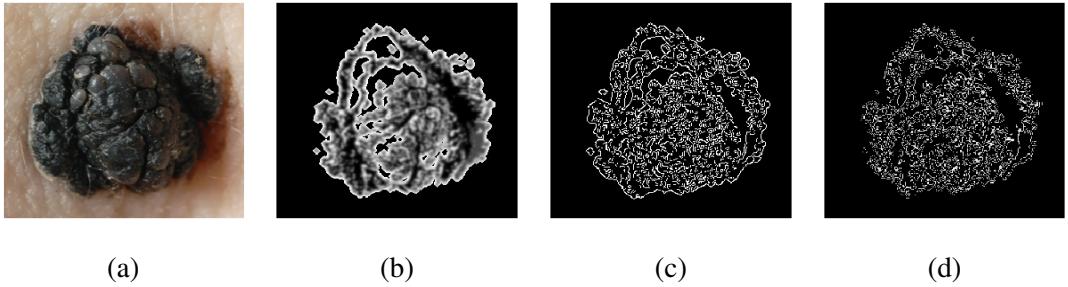


Figure 5.1: Melanoma Skin-Cancer images (a) Melanoma skin lesion (b) Resized segmented lesion (c) Canny Edge detection (d) IVE model output

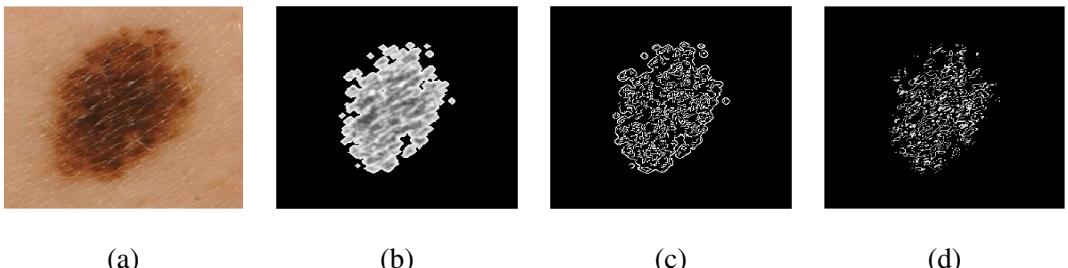


Figure 5.2: Nevus Mole Images (a) Nevus skin lesion (b) Resized segmented lesion (c) Canny Edge detection (d) IVE model output

From Table 5.2 and Fig. 5.3, it is shown that the average performance of our proposed IVEwDNN model is significantly better than the Canny Edge detection (CED) with DNN models. In the proposed IVE model, we used the same threshold value for this experiment's performance evaluation metrics. For experimental evaluation of the algorithm, the proposed model IVEwVGG16 TFL is compared with some existing works described in

Table 5.1: Experimental result(%) evaluation of the Proposed Methodology with the State-of-the-art Methods

Methods	Metrics				
	Sensitivity (Recall)	Specificity	PPV (Precision)	NPV	Accuracy
Illumination Correction [54]	81.0	80.0	75.0	86.0	81.0
S. R. S. Jianu et al. [82]	72.0	89.0	87.0	76.0	81.0
IVEwCNN Proposed Methodology (IVEwVGG-16 TFL)	88.28 <b>93.91</b>	92.22 <b>95.66</b>	89.25 <b>93.50</b>	91.98 <b>95.76</b>	90.66 <b>94.93</b>

[49, 54, 83, 82, 51] as they all worked with the identical or smaller dataset. However, for more accurate comparison, we incorporated the models [54, 82] into our experimental setup and the results are presented in Table 5.1. The values in Table 5.1 are updated to two decimal points to compare with other state-of-the-art methods. In Table 5.1 the results in bold format show the experimental results. For performance measurement, we operated on five commonly utilized metrics (sensitivity, specificity, PPV, NPV, and accuracy) that are widely used in the study of Melanoma identification and classification problems.

In an overall method evaluation, our proposed model achieved 93.91% sensitivity, which gets 6.91% higher than the second-best sensitivity (87.0%) attained by [51]. The specificity (95.66%) we achieved shows a considerably 6.66% higher outcome than the mostly nearer specificity (89.0%) found in [82]. Moreover, for the other metrics, PPV, NPV, and accuracy, our model achieved 6.5%, 9.76%, and 7.93%, rising performance consecutively; to the mostly closer values next to them. Our proposed methodology (IVEwVGG-16 TFL) shows overall better performance evaluation metrics. Hence, we can conclude that the effective use of the DNN model along with a well-processed image generates a superior result when we use highly intense pixel values (IVE) from a segmented lesion skin image.

### 5.3 Scope and data sensitivity

As we found, the proposed IVEwVGG-16 TFL model outperforms the other DNN models in previous sections; so here we investigate the scope of data variations on the model. At

Table 5.2: Average experimental result(%) comparison between CNN, VGG-16 TFL and Inception V3 TFL model (best results are shown in **bold**)

Metrics	CNN		IV3 TFL		VGG-16 TFL	
	CED	IVE	CED	IVE	CED	IVE
Sensitivity	86.85	88.28	90.34	91.08	88.16	<b>93.91</b>
Specificity	90.36	92.22	93.12	95.54	95.61	<b>95.66</b>
PPV	87.21	89.25	91.01	<b>93.74</b>	93.63	93.50
NPV	90.14	91.98	92.74	93.80	91.49	<b>95.76</b>
Accuracy	88.63	90.66	91.97	93.75	92.30	<b>94.93</b>
F1-score	87.52	88.63	90.64	92.20	90.67	<b>93.66</b>
ROC-AUC	88.71	90.25	91.83	93.33	91.88	<b>94.79</b>

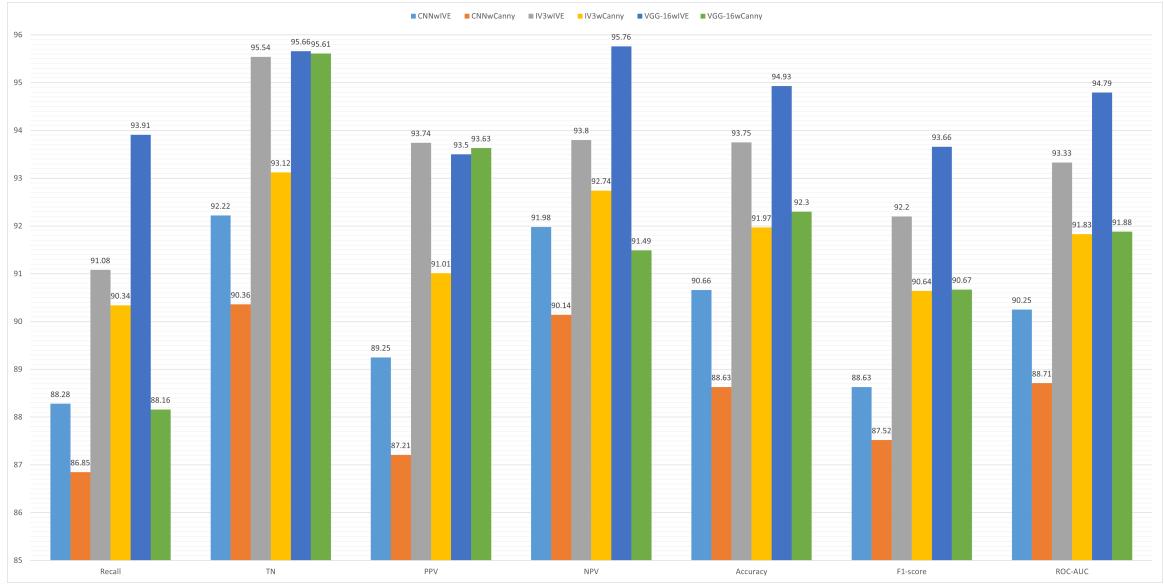


Figure 5.3: Average performance analysis for CNN, IV3 transfer learning and VGG-16 transfer learning for Melanoma detection and classification

first, the generated augmented dataset is split into three sets of data (i.e., (i) training, (ii) validation, and (iii) test set). From Table 5.3 we can see that the outcomes are highly influenced by data variation. Accuracy increases by 8.56% for 1014 data from 507 data. While using the 4056 data size achieves the highest accuracy of 99.18% which is 2.8% better than the result of using 2028 data. Perhaps, on the 4056 data, all the metrics (i.e., recall, TNR, PPV, NPV, F1-score, and accuracy) show significant results compared to the other data variations. The AUC-ROC curve denotes the model's classification performance

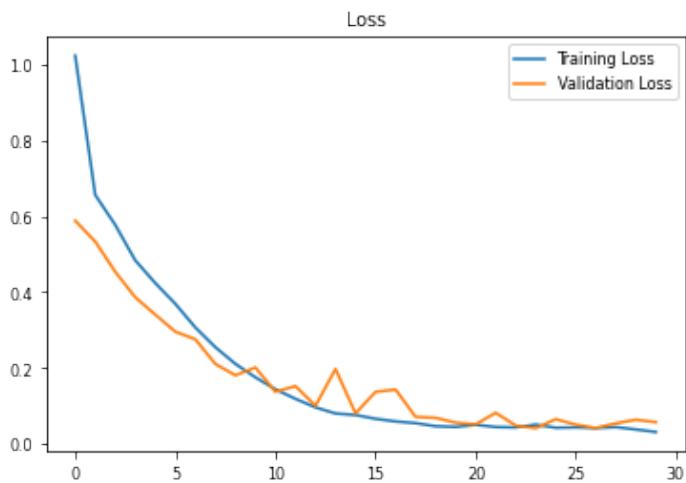
in Fig. 5.4(c), and the value (0.978) is close to the perfect value of 1.

Table 5.3: Result (%) sensitivity through the amount of data variation for the best-proposed model IVEwVGG-16 TFL (best results are indicated as bold)

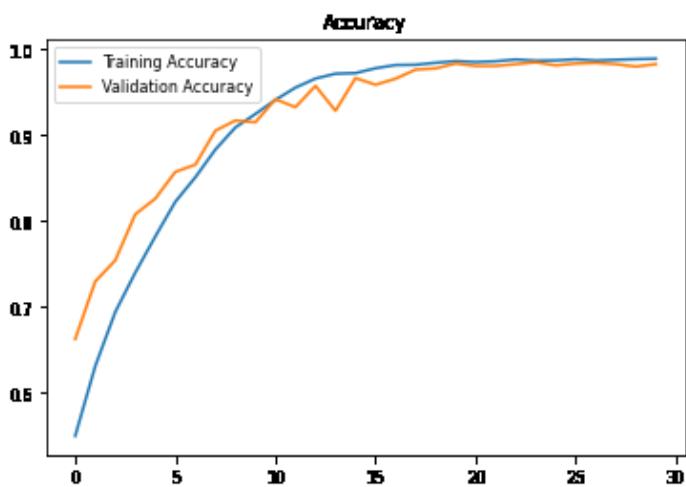
Data size	Sensitivity	Specificity	PPV	NPV	F1-score	Accuracy
507	73.68	92.11	90.32	77.78	81.16	82.89
1014	91.07	91.67	86.44	94.62	88.70	91.45
1521	94.79	95.45	93.81	96.18	94.30	95.18
2028	92.86	98.44	97.20	95.94	94.98	96.38
4056	<b>98.80</b>	<b>99.44</b>	<b>99.20</b>	<b>99.16</b>	<b>99.00</b>	<b>99.18</b>

The loss curves of a model show us a better insight into how the learning performance varies over the number of epochs and aid us in diagnosing any learning issues that can lead the model to an underfit or overfit. As Fig. 5.4(a) illustrates, with the training loss reducing, the validation loss also decreases to the epoch, which denotes that the model did not overfit. The training data is already known to the model, so the training accuracy is usually at least a minimum higher than the validation accuracy shown in Fig. 5.4(b); however, the validation data act as the group of new data points to the model. So the validation sets are used to evaluate the model's performance during training. Fig. 5.4(b) shows better training and validation accuracy.

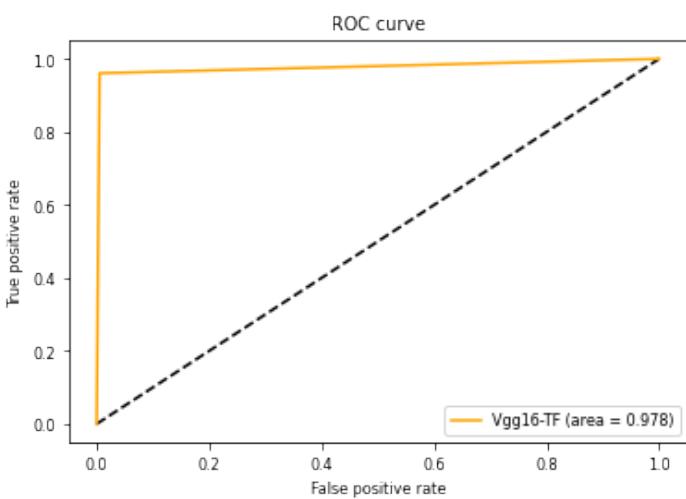
Though the plots show better model acquisition; the unbalanced image data could act as a highly influential element while training and testing a DNN model with a comparatively smaller dataset.



(a)



(b)



(c)

Figure 5.4: Proposed IVEwVGG-16 TFL model's (a) training-validation loss (b) training-validation accuracy (c) AUC-ROC curve for 4056 data

## **5.4 Conclusion**

In this chapter, we briefly discussed how the suggested model achieves the best performances in all the comparable metrics then we compare the outcomes with the state-of-the-art standards. Finally, data sensitivity on the DNN model is discussed.

# **Chapter 6**

## **Conclusion and Future works**

---

### **6.1 Conclusion**

This thesis presented intensity value estimation with a deep neural network (IVEwVGG-16 TFL) based algorithm for detection and classification. The method led the system to achieve high accuracy, sensitivity, specificity, precision, NPV, F1-score, and AUC-ROC to identify and classify Melanoma skin cancer and Nevus mole. As well, pre-processed images increase the learnability of any system. Here, we chose to take the pixels with higher intensity than a threshold value from the segmented lesion image instead of edge detection. The technique preserves more features than only edge detection. So, this aids our proposed system to escalate the overall learnability, which further predicts Melanoma skin cancer more accurately. Our proposed system (IVEwVGG-16 TFL) takes 39 (average) seconds to detect and predict Melanoma skin cancer and Nevus mole. To evaluate the proposed system performance for IVEwDNN, we considered the seven most popular performance evaluation metrics and compared them with the CED model to find the best deep neural network for detecting Melanoma skin cancer. We conclude IVEwVGG-16 transfer learning model outperforms other DNN models. Finally, we compare the IVEwVGG-16 transfer learning outcomes with some notable existing works on the same or smaller dataset. In

this experiment, large datasets were not considered to evaluate the proposed IVE model due to machine limitations. Hence, the proposed model's performance is only evaluated by comparing it to the state-of-the-art models that used the same or smaller dataset for their model. The experimental comparison demonstrates that the proposed algorithm has more satisfactory results than all the state-of-the-art models. We also conclude how the number of data variations influences achieving better results. The proposed deep neural network-based automatic can be implemented for dermatological diagnosis. This intelligent system can be significantly effective in aiding doctors in early Melanoma detection. Moreover, the proposed system needs to be trained on a large clinical dataset before going to a clinical application.

## 6.2 Limitations

- In the experiment, we used only three deep neural networks that are relevant to medical image analysis.
- Due to machine limitation we were able to test the model for limited times.
- Before deploying the model in the clinical environment it needs to train and test with more complex and real images.

## 6.3 Future Works

In the future, we will look forward to using other pre-trained deep neural network models with the proposed intensity value estimation model for more than 4056 data. Data balancing could be another work. Moreover, an optimized and efficient mobile familiar application could be developed to aid all people in investigating and monitoring their condition.

# References

- [1] M. R. Peram, S. Jalalpure, V. Kumbar, S. Patil, S. Joshi, K. Bhat, and P. Diwan, “Factorial design based curcumin ethosomal nanocarriers for the skin cancer delivery: in vitro evaluation,” *Journal of liposome research*, vol. 29, no. 3, pp. 291–311, 2019.
- [2] American Cancer Society, “Key statistics for melanoma skin cancer.” <https://www.cancer.org/cancer/melanoma-skin-cancer/about>
- [3] L. Agilandeswari, M. Sagar, and N. Keerthana, “Skin lesion detection using texture based segmentation and classification by convolutional neural networks (cnn),” *Art Int J Innov Technol Explor Eng (IJITEE)*, vol. 9, no. 2, 2019.
- [4] A. Adegun and S. Viriri, “An enhanced deep learning framework for skin lesions segmentation,” in *International conference on computational collective intelligence*, pp. 414–425, Springer, 2019.
- [5] American Cancer Society, “Melanoma warning signs.” <https://www.skincancer.org/skin-cancer-information/melanoma/melanoma-warning-signs-and>
- [6] G. Argenziano, G. Fabbrocini, P. Carli, V. De Giorgi, E. Sammarco, and M. Delfino, “Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions: comparison of the abcd rule of dermatoscopy and a new 7-point checklist based on pattern analysis,” *Archives of dermatology*, vol. 134, no. 12, pp. 1563–1570, 1998.
- [7] M. A. Kadampur and S. Al Riyae, “Skin cancer detection: Applying a deep learning based model driven architecture in the cloud for classifying dermal cell images,” *Informatics in Medicine Unlocked*, vol. 18, p. 100282, 2020.

- [8] J. Pauline, S. Abraham, and J. Bethanney Janney, “Detection of skin cancer by image processing techniques,” *Journal of chemical and Pharmaceutical research*, vol. 7, no. 2, pp. 148–153, 2015.
- [9] H. Praveena, K. Sudha, and P. Geetha, “Support vector machine based melanoma skin cancer detection,” *J. Univ. Shanghai Sci. Technol.*, vol. 22, pp. 1075–1081, 2020.
- [10] S. Saravanan, B. Heshma, A. A. Shanofer, and R. Vanithamani, “Skin cancer detection using dermoscope images,” *Materials Today: Proceedings*, vol. 33, pp. 4823–4827, 2020.
- [11] M. Q. Khan, A. Hussain, S. U. Rehman, U. Khan, M. Maqsood, K. Mehmood, and M. A. Khan, “Classification of melanoma and nevus in digital images for diagnosis of skin cancer,” *IEEE Access*, vol. 7, pp. 90132–90144, 2019.
- [12] B. P. Hibler, Q. Qi, and A. M. Rossi, “Current state of imaging in dermatology,” *Semin Cutan Med Surg*, vol. 35, no. 1, pp. 2–8, 2016.
- [13] N. R. Abbasi, H. M. Shaw, D. S. Rigel, R. J. Friedman, W. H. McCarthy, I. Osman, A. W. Kopf, and D. Polsky, “Early diagnosis of cutaneous melanoma: revisiting the abcd criteria,” *Jama*, vol. 292, no. 22, pp. 2771–2776, 2004.
- [14] P. Accetta, J. Accetta, and J. Kostecki, “The use of digital cameras by us dermatologists,” *Journal of the American Academy of Dermatology*, vol. 69, no. 5, pp. 837–838, 2013.
- [15] J. P. Banky, J. W. Kelly, D. R. English, J. M. Yeatman, and J. P. Dowling, “Incidence of new and changed nevi and melanomas detected using baseline images and dermoscopy in patients at high risk for melanoma,” *Archives of dermatology*, vol. 141, no. 8, pp. 998–1006, 2005.
- [16] S. Q. Wang, A. W. Kopf, K. Koenig, D. Polsky, K. Nudel, and R. S. Bart, “Detection of melanomas in patients followed up with total cutaneous examinations, total cutaneous photography, and dermoscopy,” *Journal of the American Academy of Dermatology*, vol. 50, no. 1, pp. 15–20, 2004.

- [17] L. T. Dengel, G. R. Petroni, J. Judge, D. Chen, S. T. Acton, A. T. Schroen, and C. L. Slingluff Jr, “Total body photography for skin cancer screening,” *International journal of dermatology*, vol. 54, no. 11, pp. 1250–1254, 2015.
- [18] S. A. Oliveria, D. Chau, P. J. Christos, C. A. Charles, A. I. Mushlin, and A. C. Halpern, “Diagnostic accuracy of patients in performing skin self-examination and the impact of photography,” *Archives of dermatology*, vol. 140, no. 1, pp. 57–62, 2004.
- [19] A. A. Marghoob, R. Usatine, and N. Jaimes, “Dermoscopy for the family physician,” *American family physician*, vol. 88, no. 7, pp. 441–450, 2013.
- [20] M. Vestergaard, P. Macaskill, P. Holt, and S. Menzies, “Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting,” *British Journal of Dermatology*, vol. 159, no. 3, pp. 669–676, 2008.
- [21] H. Kittler, H. Pehamberger, K. Wolff, and M. Binder, “Diagnostic accuracy of dermoscopy,” *The lancet oncology*, vol. 3, no. 3, pp. 159–165, 2002.
- [22] M.-L. Bafounta, A. Beauchet, P. Aegerter, and P. Saiag, “Is dermoscopy (epiluminescence microscopy) useful for the diagnosis of melanoma?: Results of a meta-analysis using techniques adapted to the evaluation of diagnostic tests,” *Archives of dermatology*, vol. 137, no. 10, pp. 1343–1350, 2001.
- [23] N. Zhang, Y.-X. Cai, Y.-Y. Wang, Y.-T. Tian, X.-L. Wang, and B. Badami, “Skin cancer diagnosis based on optimized convolutional neural network,” *Artificial Intelligence in Medicine*, vol. 102, p. 101756, 2020.
- [24] P. Carli, V. De Giorgi, E. Crocetti, F. Mannone, D. Massi, A. Chiarugi, and B. Gannotti, “Improvement of malignant/benign ratio in excised melanocytic lesions in the ‘dermoscopy era’: a retrospective study 1997–2001,” *British Journal of Dermatology*, vol. 150, no. 4, pp. 687–692, 2004.
- [25] C. Rosendahl, P. Tschandl, A. Cameron, and H. Kittler, “Diagnostic accuracy of dermatoscopy for melanocytic and nonmelanocytic pigmented lesions,” *Journal of the American academy of dermatology*, vol. 64, no. 6, pp. 1068–1073, 2011.

- [26] S. K. T. Que, N. Fraga-Braghioli, J. M. Grant-Kels, H. S. Rabinovitz, M. Oliviero, and A. Scope, “Through the looking glass: basics and principles of reflectance confocal microscopy,” *Journal of the American Academy of Dermatology*, vol. 73, no. 2, pp. 276–284, 2015.
- [27] G. Pellacani, P. Pepe, A. Casari, and C. Longo, “Reflectance confocal microscopy as a second-level examination in skin oncology improves diagnostic accuracy and saves unnecessary excisions: a longitudinal prospective study,” *British Journal of Dermatology*, vol. 171, no. 5, pp. 1044–1051, 2014.
- [28] J. Champin, J.-L. Perrot, E. Cinotti, B. Labeille, C. Douchet, G. Parrau, F. Cambazard, P. Seguin, and T. Alix, “In vivo reflectance confocal microscopy to optimize the spaghetti technique for defining surgical margins of lentigo maligna,” *Dermatologic Surgery*, vol. 40, no. 3, pp. 247–256, 2014.
- [29] H. Nadiminti, A. Scope, A. A. Marghoob, K. Busam, and K. S. Nehal, “Use of reflectance confocal microscopy to monitor response of lentigo maligna to nonsurgical treatment,” *Dermatologic surgery*, vol. 36, no. 2, pp. 177–184, 2010.
- [30] E. Cinotti, J. L. Perrot, N. Campolmi, B. Labeille, M. Espinasse, D. Grivet, G. Thuret, P. Gain, C. Douchet, F. Forest, *et al.*, “The role of in vivo confocal microscopy in the diagnosis of eyelid margin tumors: 47 cases,” *Journal of the American Academy of Dermatology*, vol. 71, no. 5, pp. 912–918, 2014.
- [31] S. Debarbieux, J. Perrot, N. Erfan, S. Ronger-Savlé, B. Labeille, E. Cinotti, L. Depaepe, N. Cardot-Leccia, J. Lacour, L. Thomas, *et al.*, “Reflectance confocal microscopy of mucosal pigmented macules: a review of 56 cases including 10 macular melanomas,” *British Journal of Dermatology*, vol. 170, no. 6, pp. 1276–1284, 2014.
- [32] J. Welzel, “Optical coherence tomography in dermatology: a review,” *Skin Research and Technology: Review article*, vol. 7, no. 1, pp. 1–9, 2001.
- [33] T. Gambichler, V. Jaedicke, and S. Terras, “Optical coherence tomography in dermatology: technical and clinical aspects,” *Archives of dermatological research*, vol. 303, no. 7, pp. 457–473, 2011.

- [34] A. Masood and A. Ali Al-Jumaily, “Computer aided diagnostic support system for skin cancer: a review of techniques and algorithms,” *International journal of biomedical imaging*, vol. 2013, 2013.
- [35] D. D. Patil and S. G. Deore, “Medical image segmentation: a review,” *International Journal of Computer Science and Mobile Computing*, vol. 2, no. 1, pp. 22–27, 2013.
- [36] M. Vijayalakshmi, “Melanoma skin cancer detection using image processing and machine learning,” *International Journal of Trend in Scientific Research and Development (IJTSRD)*, vol. 3, no. 4, pp. 780–784, 2019.
- [37] M. H. ROBERT, K. SHANMUGAM, and D. ITS’HAK, “Textural features for image classification,” 1973.
- [38] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, “Speeded-up robust features (surf),” *Computer vision and image understanding*, vol. 110, no. 3, pp. 346–359, 2008.
- [39] L. Cai, J. Gao, and D. Zhao, “A review of the application of deep learning in medical image classification and segmentation,” *Annals of translational medicine*, vol. 8, no. 11, 2020.
- [40] K. Zaman and S. S. Maghdid, “Medical images classification for skin cancer using convolutional neural network algorithms,” *Advances in Mechanics*, vol. 9, no. 3, pp. 526–541, 2021.
- [41] K. Zaman, J. I. Bangash, S. S. Maghdid, S. Hassan, H. Afridi, and M. Zohaib, “Analysis and classification of skin cancer images using convolutional neural network approach,” in *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, pp. 1–8, IEEE, 2020.
- [42] A. R. Lakshminarayanan, R. Bhuvaneshwari, S. Bhuvaneshwari, S. Parthasarathy, S. Jeganathan, and K. M. Sagayam, “Skin cancer prediction using machine learning algorithms,” in *Artificial Intelligence and Technologies*, pp. 303–310, Springer, 2022.
- [43] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, “A survey on deep transfer learning,” in *International conference on artificial neural networks*, pp. 270–279, Springer, 2018.

- [44] S. S. Yadav and S. M. Jadhav, “Deep convolutional neural network based medical image classification for disease diagnosis,” *Journal of Big Data*, vol. 6, no. 1, pp. 1–18, 2019.
- [45] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708, 2017.
- [46] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
- [47] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826, 2016.
- [48] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.
- [49] I. Giotis, N. Molders, S. Land, M. Biehl, M. F. Jonkman, and N. Petkov, “Med-node: a computer-assisted melanoma diagnosis system using non-dermoscopic images,” *Expert systems with applications*, vol. 42, no. 19, pp. 6578–6585, 2015.
- [50] A. Maiti and B. Chatterjee, “Improving detection of melanoma and naevus with deep neural networks,” *Multimedia Tools and Applications*, pp. 1–20, 2019.
- [51] S. Mukherjee, A. Adhikari, and M. Roy, “Malignant melanoma detection using multi layer preceptron with visually imperceptible features and pca components from med-node dataset,” *International Journal of Medical Engineering and Informatics*, vol. 12, no. 2, pp. 151–168, 2020.
- [52] R. Garg, S. Maheshwari, and A. Shukla, “Decision support system for detection and classification of skin cancer using cnn,” *arXiv preprint arXiv:1912.03798*, 2019.
- [53] S. Baldi, I. Michailidis, V. Ntampasi, E. B. Kosmatopoulos, I. Papamichail, and M. Papeorgiou, “Simulation-based synthesis for approximately optimal urban traffic light management,” in *2015 American Control Conference (ACC)*, pp. 868–873, IEEE, 2015.

- [54] E. Nasr-Esfahani, S. Samavi, N. Karimi, S. M. R. Soroushmehr, M. H. Jafari, K. Ward, and K. Najarian, “Melanoma detection by analysis of clinical images using convolutional neural network,” in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 1373–1376, IEEE, 2016.
- [55] R. Fatima, M. Z. A. Khan, A. Govardhan, and K. D. Dhruve, “Computer aided multi-parameter extraction system to aid early detection of skin cancer melanoma,” *Int. J. Comput. Sci. Netw. Secur.*, vol. 12, no. 10, pp. 74–86, 2012.
- [56] C. Zheng and D.-W. Sun, “Image segmentation techniques,” in *Computer Vision Technology for Food Quality Evaluation*, pp. 37–56, Elsevier, 2008.
- [57] R. Refianti, A. B. Mutiara, and R. P. Priyandini, “Classification of melanoma skin cancer using convolutional neural network,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 3, pp. 409–417, 2019.
- [58] K. Zaman and S. S. Maghdid, “Medical images classification for skin cancer using convolutional neural network algorithms,” *Advances in Mechanics*, vol. 9, no. 3, pp. 526–541, 2021.
- [59] A. R. Lakshminarayanan, R. Bhuvaneshwari, S. Bhuvaneshwari, S. Parthasarathy, S. Jeganathan, and K. M. Sagayam, “Skin cancer prediction using machine learning algorithms,” in *Artificial Intelligence and Technologies*, pp. 303–310, Springer, 2022.
- [60] 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.
- [61] L. Torrey and J. Shavlik, “Transfer learning,” in *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*, pp. 242–264, IGI global, 2010.
- [62] F. Gerges and F. Y. Shih, “A convolutional deep neural network approach for skin cancer detection using skin lesion images,” *International Journal of Electrical and Computer Engineering*, vol. 15, no. 8, pp. 475–478, 2021.

- [63] R. Indraswari, R. Rokhana, and W. Herulambang, “Melanoma image classification based on mobilenetv2 network,” *Procedia Computer Science*, vol. 197, pp. 198–207, 2022.
- [64] D. N. Le, H. X. Le, L. T. Ngo, and H. T. Ngo, “Transfer learning with class-weighted and focal loss function for automatic skin cancer classification,” *arXiv preprint arXiv:2009.05977*, 2020.
- [65] M. Hasan, S. D. Barman, S. Islam, and A. W. Reza, “Skin cancer detection using convolutional neural network,” in *Proceedings of the 2019 5th International Conference on Computing and Artificial Intelligence*, pp. 254–258, 2019.
- [66] D. Koundal and B. Sharma, “Advanced neutrosophic set-based ultrasound image analysis,” in *Neutrosophic set in medical image analysis*, pp. 51–73, Elsevier, 2019.
- [67] H. M. Ünver and E. Ayan, “Skin lesion segmentation in dermoscopic images with combination of yolo and grabcut algorithm,” *Diagnostics*, vol. 9, no. 3, p. 72, 2019.
- [68] V. Jaya and R. Gopikakumari, “Iem: a new image enhancement metric for contrast and sharpness measurements,” *International Journal of Computer Applications*, vol. 79, no. 9, 2013.
- [69] D. J. Bora, “Importance of image enhancement techniques in color image segmentation: A comprehensive and comparative study,” *arXiv preprint arXiv:1708.05081*, 2017.
- [70] C. Poynton, “Frequently asked questions about color,” *Retrieved June*, vol. 19, no. 449, p. 2004, 1997.
- [71] A. K. Yadav, R. Roy, A. P. Kumar, *et al.*, “Survey on content-based image retrieval and texture analysis with applications,” *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 7, no. 6, pp. 41–50, 2014.
- [72] C. Kanan and G. W. Cottrell, “Color-to-grayscale: does the method matter in image recognition?,” *PloS one*, vol. 7, no. 1, p. e29740, 2012.
- [73] M. Čadík, “Perceptual evaluation of color-to-grayscale image conversions,” in *Computer graphics forum*, vol. 27, pp. 1745–1754, Wiley Online Library, 2008.

- [74] J. Zhang and J. Hu, “Image segmentation based on 2d otsu method with histogram analysis,” in *2008 international conference on computer science and software engineering*, vol. 6, pp. 105–108, IEEE, 2008.
- [75] O. Salih and S. Viriri, “Skin lesion segmentation using local binary convolution-deconvolution architecture,” *Image Analysis & Stereology*, vol. 39, no. 3, pp. 169–185, 2020.
- [76] S. H. Kassani and P. H. Kassani, “A comparative study of deep learning architectures on melanoma detection,” *Tissue and Cell*, vol. 58, pp. 76–83, 2019.
- [77] E. A. Sekehravani, E. Babulak, and M. Masoodi, “Implementing canny edge detection algorithm for noisy image,” *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 4, pp. 1404–1410, 2020.
- [78] I. Gogul and V. S. Kumar, “Flower species recognition system using convolution neural networks and transfer learning,” in *2017 fourth international conference on signal processing, communication and networking (ICSCN)*, pp. 1–6, IEEE, 2017.
- [79] R. Islam, S. Imran, M. Ashikuzzaman, and M. M. A. Khan, “Detection and classification of brain tumor based on multilevel segmentation with convolutional neural network,” *Journal of Biomedical Science and Engineering*, vol. 13, no. 4, pp. 45–53, 2020.
- [80] C. Shorten and T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” *Journal of Big Data*, vol. 6, no. 1, pp. 1–48, 2019.
- [81] I. Guyon *et al.*, “A scaling law for the validation-set training-set size ratio,” *AT&T Bell Laboratories*, vol. 1, no. 11, 1997.
- [82] S. R. S. Jianu, L. Ichim, and D. Popescu, “Automatic diagnosis of skin cancer using neural networks,” in *2019 11th International Symposium on Advanced Topics in Electrical Engineering (ATEE)*, pp. 1–4, IEEE, 2019.
- [83] S. Mukherjee, A. Adhikari, and M. Roy, “Malignant melanoma detection using multi layer perceptron with optimized network parameter selection by pso,” in *Contemporary Advances in Innovative and Applicable Information Technology*, pp. 101–109, Springer, 2019.

- [84] A. Sagar and D. Jacob, “Convolutional neural networks for classifying melanoma images,” *bioRxiv*, pp. 2020–05, 2021.
- [85] V. Pillay, D. Hirasen, S. Viriri, and M. Gwetu, “Melanoma skin cancer classification using transfer learning,” in *International Conference on Computational Collective Intelligence*, pp. 287–297, Springer, 2020.