T.C.

ISTANBUL AYDIN UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

MELANOMA SKIN CANCER DETECTION USING MACHINE LEARNING TECHNIQUES

MASTER'S THESIS

Muhammad Ali ABBASI

Department of Software Engineering

Artificial Intelligence and Data Science Program

JANUARY, 2024

T.C.

ISTANBUL AYDIN UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

MELANOMA SKIN CANCER DETECTION USING MACHINE LEARNING TECHNIQUES

MASTER'S THESIS

Muhammad Ali ABBASI

(Y2113.140006)

Department of Software Engineering
Artificial Intelligence and Data Science Program

Thesis Advisor: Prof. Dr. Ali OKATAN

JANUARY, 2024

APPROVAL PAGE

DECLARATION

I hereby state respectfully that the study "Melanoma Skin Cancer Detection Using Machine Techniques," which I submitted as a master's thesis, was written without any assistance in violation of scientific ethics and traditions in all steps from the project phase to the thesis' conclusion, and that the works I used as a resource were those listed in the references. (8/01/2024)

Muhammad Ali ABBASI

FOREWORD

I'd want to say thank you to Dr. Ali Okatan, who served as my dissertation adviser, for giving me all the required guidance. He has helped me so much at this period, and I appreciate it everything. He made it feasible for me to develop my thesis and finish it on time thanks to his contributions of suggestions and encouragement. Every time he felt I was lost, he tried his best to point me in the proper route.

I also want to use this chance to express my gratitude to my family and friends for their continuous support and inspiration as I worked to finish my thesis.Without the help of the people I've mentioned, this success would not have been possible. I'm grateful.

January, 2024 Muhammad Ali ABBASI

MELANOMA SKIN CANCER DETECTION USING MACHINE TECHNIQUES

ABSTRACT

public health problems worldwide. So far, the application of machine learning algorithms has shown that earlier and more accurate diagnosis is better for patient health. This article describes the use of his three new CNNs to detect skin cancers, including melanoma. The dataset used for the study includes dermoscopy photographs from different datasets ensuring diversity of melanoma and non-melanoma cases. An extensive training and validation process improved the CNN that differentiates between benign and malignant diseases. The fact that the accuracy for this model was amazingly high at 89% percent clearly sets it apart from all of the others.This research is of great importance to the prediction of the progress in skin cancer diagnosis in the near future. Machine learning model, such as ResNet50v2 can be used in the healthcare sector for the early detection and diagnosis of melanoma which will result into changed healthcare. The high rate of precision in the ResNet50v2 model will aid in early detection and ultimately improve patient results. Going forward, there are high hopes that other better screening techniques for early melanoma would become available especially those involving minimal invasiveness and thus better prognosis and lesser melanoma-related deaths.

Keywords: CNN; Machine Learning; Melanoma skin Cancer;

MAKİNE TEKNİKLERİ KULLANILARAK MELANOM CİLT KANSERİNİN TESPİTİ

ÖZET

Dünya çapında halk sağlığı sorunları. Şimdiye kadar makine öğrenimi algoritmalarının uygulanması, daha erken ve daha doğru teşhisin hasta sağlığı açısından daha iyi olduğunu göstermiştir. Bu makale, melanom da dahil olmak üzere cilt kanserlerini tespit etmek için üç yeni CNN'nin kullanımını açıklamaktadır. Çalışma için kullanılan veri seti, melanom ve melanom dışı vakaların çeşitliliğini sağlayan farklı veri setlerinden dermoskopi fotoğraflarını içermektedir. Kapsamlı bir eğitim ve doğrulama süreci, iyi huylu ve kötü huylu hastalıkları birbirinden ayıran CNN'yi geliştirdi. Bu modelin doğruluğunun yüzde 94 gibi inanılmaz derecede yüksek olması, onu diğerlerinden açıkça ayırıyor. Bu araştırma, yakın gelecekte cilt kanseri teşhisinde ilerlemenin öngörülebilmesi açısından büyük önem taşıyor. Xception gibi makine öğrenimi modeli, sağlık sektöründe melanomun erken tespiti ve teşhisi için kullanılabilir ve bu da sağlık hizmetlerinde değişikliğe yol açabilir. Xception modelinin yüksek hassasiyet oranı, erken tespite yardımcı olacak ve sonuçta hasta sonuçlarını iyileştirecektir. İleriye dönük olarak, erken melanom için diğer daha iyi tarama tekniklerinin, özellikle minimal invazivliği içerenlerin ve dolayısıyla daha iyi prognoz ve daha az melanomla ilişkili ölümlerin mevcut olabileceğine dair büyük umutlar var.

Anahtar Kelimeler: CNN ; Makine öğrenme ; Melanom cilt kanseri;

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

As a leading public health problem in modern medicine, melanoma—which can lead to skin cancer—is well known. Early identification of ailments helps promote quick responses and positive results in patients. The traditional way to detect melanoma relied upon the examination by an ophthalmologist, thereby increasing the risk of bias and misjudgment resulting from human mistakes. The emergence of CNNs and deep learning in skin cancer diagnosis marks a new era of development. The use of machine learning has brought a major strides forward in the arena of medical diagnostics. Our study aims at improving accuracy and efficiency in prognosis of melanomas using computers. This change aims at streamlining the screening procedure so as to realize immediate production of results that matter much in melanoma. Moreover, the development seeks to improve diagnostics precision through multiple machine learning solutions. Thus, CNN leads this field. Due to excellent picture identification and classification accuracy. In this regard, neural networks are strongly aligned to the field of dermatology because of their reliability while they provide fast learning, as opposed to manual procedures. These device can identify fine details and intricate patterns in dermography image of the skin lesion which human eye cannot distinguish alone. Researchers will be able to carry out exhaustive assessments on large databases with skin lesion records courtesy of this technology. This will enable the generation of exceptionally accurate and dependable assessments regarding these lesions, all while maintaining a rapid processing speed.Our research is centered on the use of artificial intelligence in the field of melanoma diagnostics, including many methodologies that have demonstrated more favorable results.The use of algorithmic-pushed algorithms can help inside the detection of outstanding symptoms of cancer, such as the advent of aberrant paperwork, coloration adjustments, and structural anomalies. We are looking for to store lives through detecting this lethal situation early and lowering the burden it places on human beings and healthcare systems.

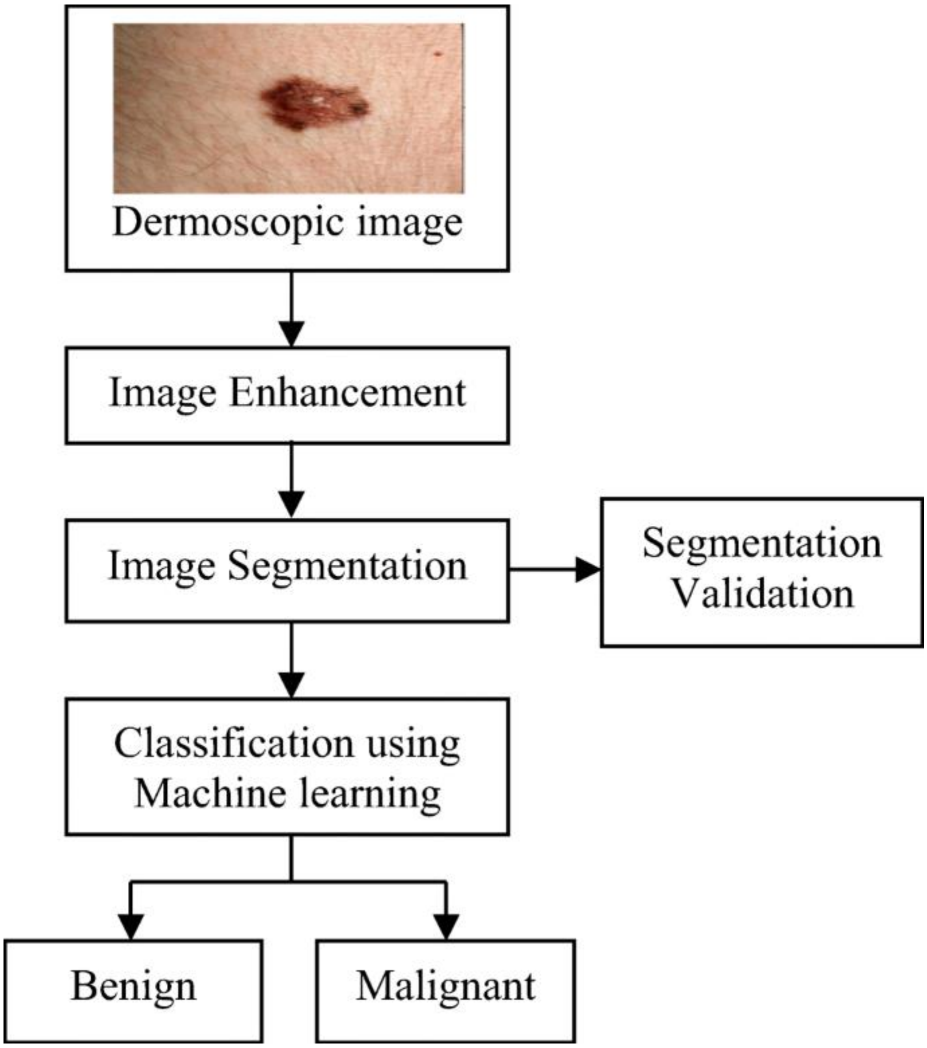

Figure 1 Diagram showing system flow

A. Motivations

Melanoma algorithms for gadget mastering in dermatological healthcare.The significance of using these algorithms stems from the capability to cause early and precise discovery.When it involves melanoma, one of the most severe forms of pores and skin cancer, early treatment improves the affected person's possibilities of getting better.Algorithms based totally on system getting to know were taught in current years to test extensive quantities of pores and skin photos and distinguish benign from malignant cancers.It is notion that this may useful resource in early detection and likely shop lives.Another persuasive reason is the scalability and effectiveness of system mastering in figuring out melanoma.The number of skin cancer cases continues to upward thrust, putting stress on dermatologists' ability to adopt well timed examinations.Multiple photos may be analyzed rapid using gadget gaining knowledge of strategies, reducing the strain for fitness group of workers.Furthermore, they may be applied for far flung prognosis, allowing those in underserved areas and people who are not able to reach professional health care establishments to have their instances checked through experts.

Machine Learning can also enhance the objectivity and consistency of cancer detection. Visual inspections are prone to human mistake and subjectivity, while algorithms follow targeted styles and standards, assuring a consistent approach. This uniformity is beneficial not simply to dermatologists but also to patients, who can have more trust in the outcomes. Machine gaining knowledge of techniques assist significantly to improve the overall excellent of cancer prognosis and, in the long run, affected person effects by way of lowering variability and increasing accuracy.

1. Enhancing Early Detection

The primary driving force behind this research is the urgent need to improve melanoma early detection. Due to melanoma's risk for death, early detection greatly enhances patient outcomes. Detecting melanoma more quickly and accurately holds the potential to save lives by catching the illness in its early, most curable stages.

2. Reducing Human Error

The evaluation of skin lesions by human vision is inherently fallible and liable to inaccuracy. Machine learning approaches, powered by neural networks and algorithms, offer a reliable and impartial method of melanoma diagnosis. The goal is to lessen the possibility of incorrect or delayed diagnoses brought on by human subjectivity.

3. Handling the Growing Melanoma Incidence

Melanoma cases are continuously rising over the world. More effective and scalable diagnostic techniques are urgently needed as the impact of this disease spreads. Given its ability to analyse enormous volumes of data, machine learning is an effective tool for handling the growing number of situations

4. Reducing Healthcare Costs

Early diagnosis and treatment not only lead to better patient outcomes but also can lower the overall cost of melanoma therapy. We may be able to lessen the cost load on healthcare systems and patients by spotting illnesses at an earlier stage.

5. Technology Advancement

The field of artificial intelligence and machine learning is constantly developing. Researchers are driven by the chance to use cutting-edge technical developments to develop more precise and effective melanoma detection techniques. Our capacity to successfully treat this condition grows as technology does as well.

B. Objectives

Study uses machine learning techniques based on pictures to detect melanoma skin cancer. The following are the key goals:

- To present an in-depth analysis illustrating the capability of melanoma skin cancer detection and ensemble strategies for learning in recognizing melanoma skin cancer detection Dataset photos.
- To improve the predicted accuracy of melanoma skin cancer detection by employing multiple ensemble learning algorithms.
- Improving ensemble melanoma skin cancer models' accuracy and recall by reducing false positives and negatives.
- An assortment of performance indicators, including F1-score, were employed to assess the robustness and generalizability of the suggested ensemble different CNN technique. Accuracy .
- To look into the potential benefits of changing the issue from multi-class to binary classification, which might improve the overall performance of ensemble CNN learning models in melanoma skin cancer diagnosis.

C. Thesis Outline

The thesis is structured as follows:

- Chapter 1 is a review of the writers' previous work.
- Chapter 2 discusses study and methods, as well as some key terms.
- Chapter 3 describes the implementation details in detail, as well as each basic

model used.

- Chapter 4 presents results and debate around research. It also contains information about the analysis of experimental data.
- Chapter 5 In the part under "Conclusions," we discuss the study's implications, possible improvement areas, future work, and the benefits of employing an ensemble.

II. BACKGROUND STUDIES AND RELATED WORKS

A variety of techniques and methodologies have been applied to the investigation of gender classification; the following are some of the most recent:

The medical sector's implementation of machine learning (ML) algorithms to potentially improve diagnostics and prognostic clinical outcomes is referred to as artificial intelligence (AI) in healthcare (Jiang et al., 2017). Advances in computational power and large amounts of data curation inside health systems have resulted in the creation of algorithms that can assist healthcare clinicians as clinical decision-support (CDS) technologies. AI has found extensive application in the healthcare sector. For instance, electronic health record data has been utilized to develop risk predictors (Juhn and Liu, 2020; Lauritsen et al., 2020). Additionally, wearable devices have been implemented to facilitate continuous disease monitoring and early detection of diseases such as sepsis (Goh et al., 2021; Komorowski et al., 2018). Subsequent to the National Institutes of Health Chest X-Ray Dataset (Wang et al., 20171) and DeepLesion, an extensive collection of 32,000 computed tomography images intended for scientific research (Yan et al., 2018), ingenious endeavors have been undertaken to acquire substantial volumes of medical image datasets, either for public or institutional use.

The field of Vision for computers is a branch of artificial intelligence which a machine learns to understand visual images. It has made medical picture evaluation more accurate and efficient (Voulodimos et al., 2018). Convolutional neural networks (CNN) are a sort of artificial neural network that has revolutionized image analysis by eliminating the requirement for traditional handmade elements such as colors, intensity value, topological structure, and texture information (Carin and Pencina, 2018). Deep learning models that have been trained on millions of photographs for tasks such as image categorization, object detection, and image recognition have been developed by researchers. Training and testing on millions of images is used to construct models for computer vision tasks such as image categorization and objection recognition. These models were developed which were inspired primarily by the ImageNet (Deng et al., 2009), CIFAR (Krizhevsky and Hinton, unpublished data), Modified National Institute of Standards and Technology (MNIST) (Deng, 2012), COCO (Common Objects in Context) (Lin et al., 2014), Open Images (Kuznetsova et al., 20202), and SUN (Xiao et al.

Therefore, a typical model, generally trained on a common dataset (such as ImageNet), can be selected and fine-tuned to match a given situation. Contrary to popular belief, ImageNet-pretrained models have produced human-level accuracy in pathology (Ehteshami Bejnordi et al., 2016; Gown et al., 2008; Qaiser and Mukherjee, 2018) and dermatology (Cho et al., 2020; Haenssle et al., 2018; Maron et al., 2019). In dermatology, AI systems that use transfer learning outperform dermatologists in diagnosing skin diseases (Esteva et al., 2017; Haenssle et al., 2020). The rapidly increasing worldwide incidence of skin cancer, the development of teledermatology during the COVID-19 epidemic, and the supplydemand imbalance for dermatologists all indicate to an urgent need for effective triaging systems backed by AI for dermatological disease detection and diagnosis. The purpose of this research is to give a complete assessment of published applications of pretrained models on dermatological images, as well as their associated datasets, restrictions, and results.The growing globally incidence of skin cancer, the development of teledermatology during the COVID-19 epidemic, and the supply-demand imbalance for dermatologists all point to an increasing need for effective triaging systems backed by AI for dermatological disease detection and diagnosis. This paper intends to give a complete assessment of documented applications of pretrained models on dermatological images, as well as their associated datasets, restrictions, and outcomes.

Table 1 A summary of melanoma skin cancer detection researches

Study	Classifier	Result/Accuracy
(Bi et al., 2017)	ResNet	79.4%
(Shahin et al., 2018)	ResNet-50 and Inception V3	89.9%
(Yap et al., 2018)	Two ResNet50 fusion	86%
(Brinker et al., 2019b)	ResNet50	81%
(Salamaa and Aly, 2021)	VGG16 and ResNet50 with SVM ResNet50 performed the best.	99%
(Jojoa Acosta et al., 2021)	ResNet152	90%
(Li and Li, 2018)	Segmentation: ResNet Classification: ResNet, DenseNet, Inception	97%
(Mahbod et al., 2019)	AlexNet, VGG16, and ResNet18 for feature extraction SVM for classification	97%
(Pomponiu et al., 2016)	AlexNet	93%
(Shorfuzzaman, 2022)	Ensemble of EffcientNetB0, DenseNet121, and Xception	95%
(Yang et al., 2021)	EfficientNet-b4	92%
(Cassidy et al., 2021)	YOLO, EfficientDet, FRCNN (resnet, inception resnet)	69%

III. RESEARCH AND METHODOLOGY

The following section discusses several aspects of melanoma skin cancer detection. This article provides an overview of melanoma, a type of skin cancer, focuses on the training and evaluation datasets used, investigates how advanced machine learning methodologies, specifically CNN, have been used to improve detection, discusses the use of transfer learning, and touches on the integration of different machine learning model approaches. These topics contribute to a complete understanding of melanoma skin cancer detection. encompassing dataset selection, melanoma skin cancer detection performance, the effectiveness of Convolutional Neural Networks (CNNs), the leveraging of pre-trained models using transfer learning, and the potential benefits of Different machine learning model approaches for improved detection performance.

A. Melanoma Skin Cancer

Melanoma, a type of skin cancer, manifests itself in two ways: benign and malignant. Benign melanocytic nevi, sometimes known as moles, are non-cancerous skin growths. These moles are often small, well-defined, and uniform in color, usually brown or black. Most moles are innocuous and remain so throughout a person's life. However, even though most moles are benign, there is a tiny possibility that they will change and eventually turn into malignant melanoma.

Malignant cancer of the skin, on the other hand, is a deadly and aggressive cancer that arises from melanocytes, the cells in the skin. Malignant melanomas, as opposed to benign moles, may exhibit irregular characteristics. These characteristics include an irregular shape, uneven pigmentation, and a proclivity to develop rapidly. Malignant melanoma can infect neighboring tissues and metastasis, spreading to other parts of the body, making it far more dangerous if left untreated. Early detection and rapid medical intervention are crucial in the case of malignant melanoma, since they considerably improve the odds of effective treatment and long-term survival.

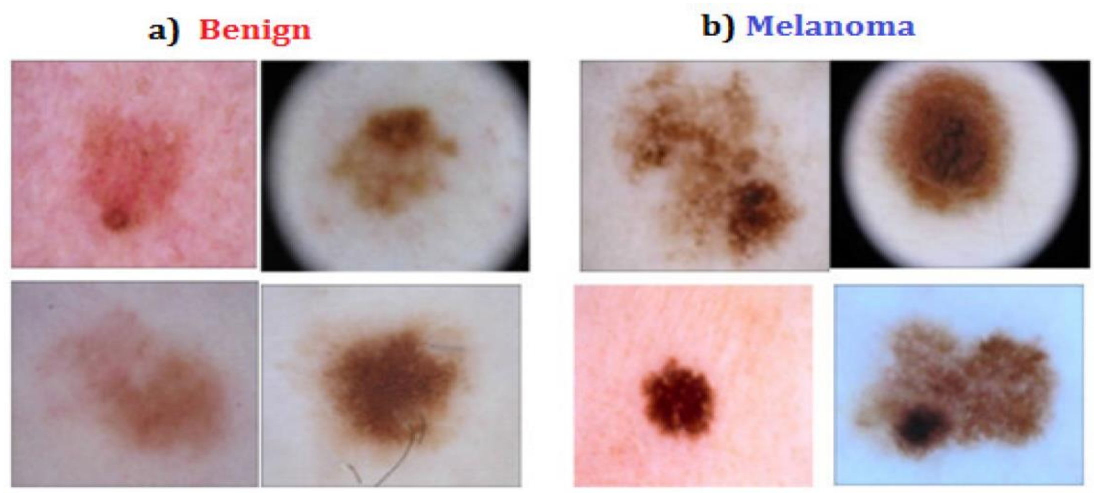

Figure 2 Melanoma skin cancer problems visualized.

Medical practitioners use a variety of diagnostic methods to distinguish benign from malignant melanocytic lesions, including clinical examination, dermoscopy (a procedure for magnified skin lesion assessment), and, where necessary, biopsies. Regular skin self-examinations and annual dermatologist checkups are critical for early diagnosis of melanoma, since any worrisome changes or new growths on the skin should be evaluated immediately. identifying these two types of melanocytic skin lesions are critical steps in the early detection, treatment, and management of melanoma, which can have a substantial impact on a patient's prognosis and quality of life.

B. Dataset Analyzation

The Kaggle melanoma skin cancer dataset contains a large collection of 10,000 retinal images that were methodically acquired for the purpose of automating the identification of melanoma skin cancer. It is regarded as the third-largest dataset of its kind and serves as a critical resource for training and assessing diagnostic models.

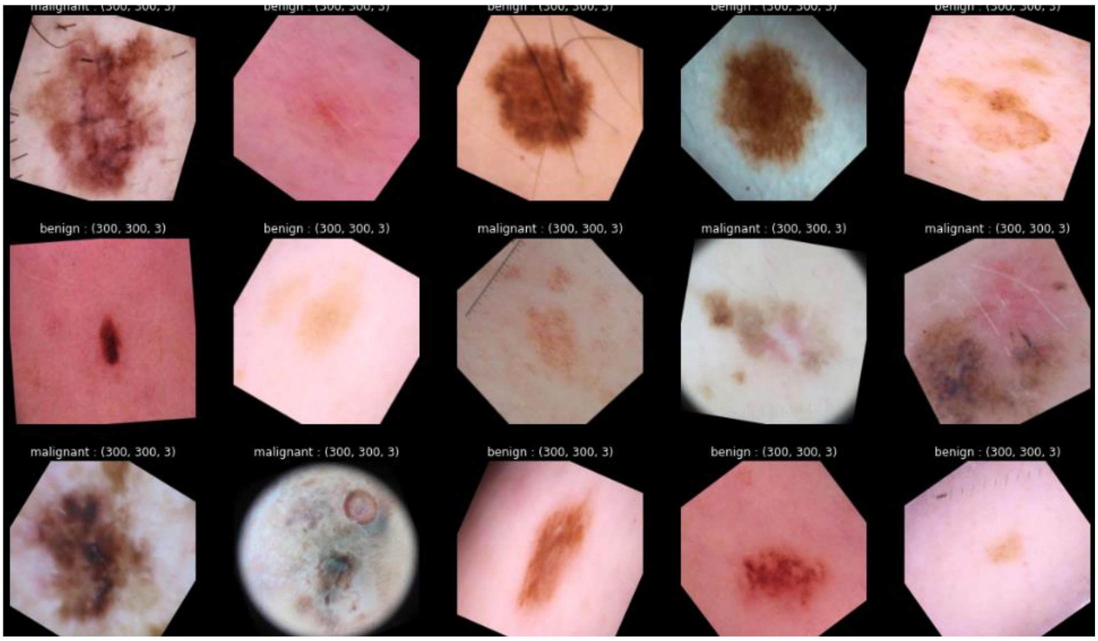

Figure 3 large collection of 10,000 retinal images

The labeling in the dataset is consistent with widely accepted guidelines for identifying melanoma skin cancer, as seen in Table 2. The photographs have been meticulously classified into two groups.

Benign (Class 1)

Malignant (Class 2)

Particularly, Class 2, which represents malignant instances, has fewer samples than Class 1, which represents benign cases. Both classes have data that is less than half the size of Class 1, indicating that the dataset has a class imbalance. The training set contains 5,500 photos, while the validation set has 700 images for Class 1 (Benign). The training set for Class 2 (Malignant) contains 4,500 photos, while the validation set has 600 images.

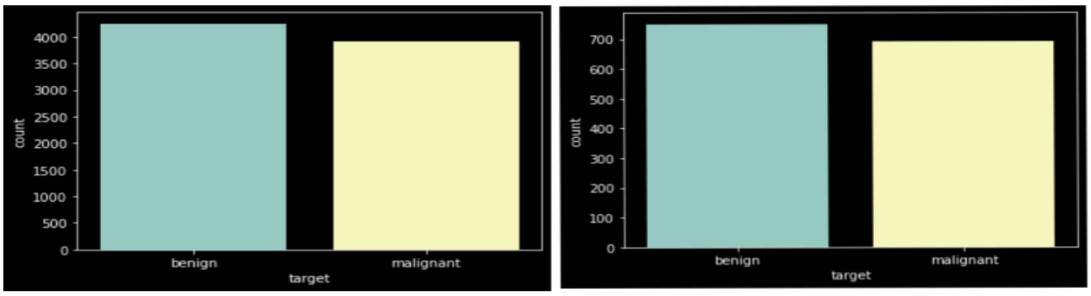

Figure 4 Inbalanced dataset label class distribution

C. Terminologies

Several core AI and ML terminology, such as CNN and Deep Learning, play critical roles in extending the capabilities of models and algorithms. Each term relates to a specific concept or strategy within the broader topic of machine learning.

1. Deep Learning

Deep learning, a subclass of artificial intelligence, is a cutting-edge technique to machine learning that has attracted substantial attention and success in recent years. Deep learning relies on artificial neural networks inspired by the form and function of the human brain. These networks are made up of numerous layers of interconnected neurons that allow them to independently learn and represent complicated patterns in data. The name "deep" alludes to the several layers involved, which allow the model to acquire hierarchical elements, making it particularly adept at spotting intricate patterns.

Deep learning training frequently includes backpropagation, in which the model adjusts its internal parameters to reduce the difference between its predictions and the actual target values in the training data. The efficiency of deep learning is directly related to the availability of large datasets, which allow models to learn and generalize from various examples. This technology has had a impact on several domains, including computer vision, natural language processing, and speech recognition, resulting in notable breakthroughs in tasks such as picture classification, language translation, and voice assistants. As a result, deep learning has been essential in generating advancements across industries, from healthcare to autonomous Car, and it continues to be a driving force in AI research and development.

2. Convolutional Neural Network

A CNN is a form of artificial neural network that is specifically developed for processing and interpreting visual input such as photos and videos . CNNs excel in computer vision problems due to their capacity to automatically learn hierarchical characteristics from images. CNNs' main innovation is the use of convolutional layers, which apply filters or kernels to small portions of an input image to detect various properties such as edges, textures, and patterns. As the network advances through its layers, it learns to detect increasingly complicated and abstract properties.

CNNs have demonstrated great effectiveness in image classification tasks, where they can accurately classify objects inside photos. This technology has numerous applications, including autonomous vehicles, medical imaging, and facial recognition. Furthermore, when paired with recurrent neural networks (RNNs) in models like Convolutional Neural Network-Long Short-Term Memory for sequence data analysis, CNNs can be used for tasks other than image analysis, such as natural language processing. CNNs have played a critical role in pushing the boundaries of many machine learning and artificial intelligence applications, making them a core tool in modern deep learning.

Convolutional layers, pooling layers, completely linked layers, and an output layer are common components of CNN architecture. Convolutional layers filter the incoming data, producing feature maps that emphasize key spatial information. Pooling layers subsample the data to minimize its dimensionality, maintaining only the most significant information from each feature. The Danse layers connect the convolutional and pooling layers' output to the network's final output layer, which generates predictions. CNNs may learn to recognize patterns and objects in images by mixing these layers and training on big datasets, allowing them to excel at many computer vision applications.

The information provided here is CNN model can carry out specific four operations:

- Convolution operation
- Max pooling
- Flattening
- Dense Layer

a. Convolution Operation

Convolution is a key concept in signal processing and deep learning that is used to perform a wide range of tasks such as image and speech recognition. The resultant function is a third function that shows how one function influences the other. Convolution facilitates in the extraction of features from input data in the context of neural networks. A kernel, also known as a filter or a convolutional filter, is a common method for carrying out the convolution process. To construct the output feature map, the kernel is slid along the input data, multiplying its corresponding values by those of the input at each place and adding them together. The network may extract local scales of feature hierarchies from the data using this method.

In the context of image processing, for example, a convolution process could be viewed as a specialized filter that moves around an image revealing various elements such as edges and texture. It is vital in lowering the dimensional size of the input while preserving critical material. In order to perform well in visual or pattern recognition tasks, CNN employs this approach to extract the hierarchical feature representation from raw input data.

b. Max pooling

When faced with intricate images that have a naturally elongated shape, it is critical to effectively decrease their dimensions while preserving essential portions or structures. This can be seen in the graphic next to it, where the complexity and size of the image necessitate an effective downsampling method. In this circumstance, max pooling is a preferred strategy, with the purpose of retaining essential features by selecting the highest normalized values within certain regions. During the the max pooling process, picture is divided into unique and not overlapping in pieces, and the highest value is retained for each sector. This technique enables the extraction of the most significant features while decreasing size of data without jeopardizing the fundamental information given in the image.

Max pooling is a downsampling technique that prioritizes delivering the most important features within small regions. It effectively selects the most relevant values from specific areas. compresses the representation of the image. This facilitates in subsequent processing steps while maintaining the essential structural and distinguishing elements of the complex shot.

c. Flattening

Displaying the features of an image in a distinct area, as illustrated by the well-resolved value, is an important step in learning systems and deep learning processes. The concept includes image attribute mapping in dependent form, which enables for quick analysis and processing. In this case, Pulling down is a key phase in the feature extraction process. The technique involves flattening for reducing a two-dimensional attribute matrix to a one-dimensional vector by converting dimension spatial linkages into a linear sequence of appropriate values and applying it to downstream responsibilities.

By flattening the feature matrix, the intrinsic spatial structure of the image is condensed into a linear form, which corresponds to the criteria of many ML algorithms. This method enables smooth integration of image features into a broader model architecture, allowing for successful learning and extraction of complex patterns from the flattened feature vector.

d. Dense Layer

A thorough connection mechanism (seen below) facilitates interaction between the flattened components and the neural network. Following the completion and submission of the order, a four-day time frame is made out for the delivery of on-demand responses and the correction of errors. Connecting each element in the flat feature vector to neurons in the network's following layers is the second essential stage in neural network building. A fully linked layer ensures that every component of the flattened feature vector contributes to the neural network's comprehension and final conclusion. The flattened structure of neural networks helps with tasks like classification, regression, and any other sort of learning.

The complete connection method facilitates in the synthesis of information derived from flat features for simple connections at various points within the neural network. This allows the network to recognize small connections between the flattened pattern and allows it to make informed decisions about the incoming data's reduced yet rich initial features.

3. Machine learning

Machine learning is a branch of artificial intelligence that focuses on creating algorithms and models that allow computer systems to learn and make predictions or judgments without being explicitly programmed. It is predicated on the notion that machines can analyze and comprehend data patterns, allowing them to improve their performance on a certain activity over time. Machine learning comprises a wide range of techniques and approaches, such as supervised learning, unsupervised learning, and reinforcement learning.

One of the most common kinds of machine learning is supervised learning, in which a model is trained on a labeled dataset and learns to make predictions and categorize information based on input data and associated target values. Unsupervised learning, on the other hand, is training a model using unlabeled data to uncover patterns and structure within the data. Reinforcement learning is a sort of machine learning that is used to train agents to make sequences of decisions in an environment using rewards or penalties as feedback.

Machine learning has applications ranging from natural language processing and computer vision to recommendation systems and autonomous robots. Its potential for automation, data-driven insights, and predictive skills has led to its acceptance in a variety of industries, altering how organizations and researchers approach problem-solving and decision-making. The capacity of machine learning models to continuously develop and adapt to new data makes them a strong tool for addressing complex and dynamic difficulties in the modern world.

4. Transfer learning

In machine learning, transfer learning is a technique in which a model learned on one job is used to do another related activity. Transfer learning allows you to use a pre-trained model's expertise and adjust it for the specific challenge at hand rather than building a model from start for a new task. This method is particularly beneficial when there is a scarcity of data for the new job or when training a model from scratch would be computationally costly.

Transfer learning is becoming more prominent in the area of natural language processing (NLP). Models like as BERT and GPT (have been pre-trained on enormous text corpora, providing them with a deep knowledge of language. These trained models may then be fine-tuned for specific NLP tasks including text categorization, sentiment analysis, language translation, and others. The knowledge of language structure and semantics from the pre-trained model is transferred to the new task, considerably enhancing performance.

The use of transfer learning in NLP has produced developments in a variety of applications, which include chatbots, language translation, and summary generation, because it allows developers to build on the wealth of linguistic knowledge encapsulated in pre trained models without having to start from start, saving time and resources.

IV. IMPLEMENTATION

The details of melanoma skin cancer detection and classification were divided into different steps in this study. The initial step was devoted to configuring the base models, which included ResNet50V2, Xception, and VGG16. These models were chosen for their ability to extract significant characteristics from retinal images and their performance in image classification tasks. Each model was independently trained for acquire precise structures and traits linked with melanoma skin cancer. The predictions of numerous models were combined in ensemble learning, whereas the predictions were used as input features.A comprehensive evaluation and training split was performed to ensure accurate evalution and minimize overfitting. The research aims to achieve reliable detection and classification of melanoma skin cancer through these procedures, allowing for timely diagnosis and treatment for those affected.

A. Models

To teach ensemble learning, we must first train and prepare the fundamental models. After the dataset has been trained on the model. We only used three models for this investigation.

1. VGG16

VGG16, short for Visual Geometry Group 16, is a well-known convolutional neural network design in the field of computer vision. It was created by the University of Oxford's Visual Geometry Group and is noted for its ease of use and efficacy in image categorization tasks. VGG16 is a deep neural network consisting of 16 weight layers, 13 convolutional layers and 3 fully connected layers. The network architecture follows a simple pattern of employing 3x3 convolutional filters with a stride of 1 and max-pooling layers with a 2x2 window. VGG16 is popular among scholars and practitioners in the field due to its ease of understanding and implementation.

One distinguishing feature of VGG16 is its uniform architecture, which means that it utilizes the same filter size and stride across the network. This design choice has made it simple to modify and fine-tune for a variety of image categorization jobs. VGG16 is frequently used as a pre-trained model for transfer learning, in which learned characteristics from a big dataset (such as ImageNet) are used to improve the performance of a neural network on a separate task. While VGG16 was revolutionary when it was first released, more contemporary designs, including as ResNet and Inception, have outperformed it in terms of effectiveness and functionality. However, VGG16 is an important landmark in the history of deep learning and computer vision, illustrating the usefulness of deep convolutional networks in image recognition.

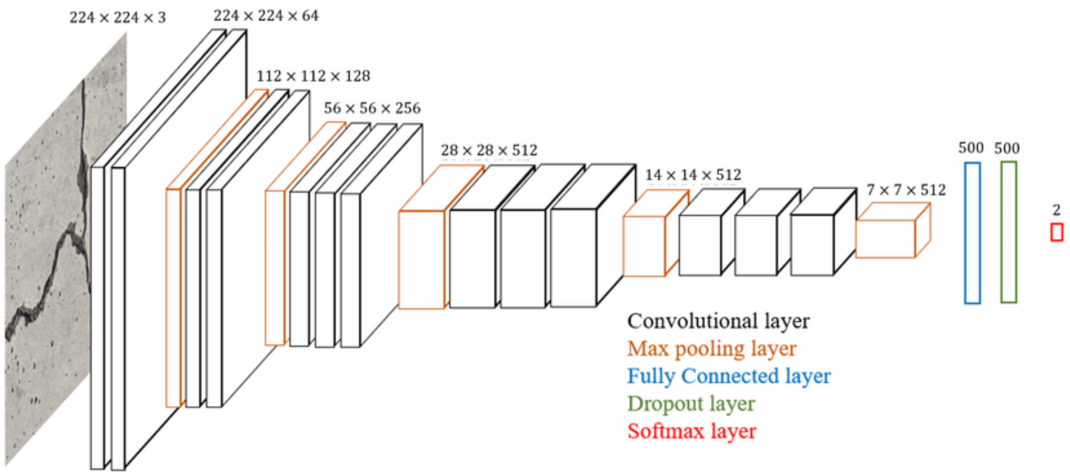

Figure 6 VGG16 Model

2. ResNet50V2

ResNet50V2 is a sophisticated convolutional neural network design that has significantly advanced computer vision. This network is a subset of the ResNet (Residual Network) family, which was developed to solve the vanishing gradient problem in deep neural networks. The "50" in ResNet50V2 refers to the number of weight layers in the network, while the "V2" refers to the second version, which contains improvements to the original architecture.

One of ResNet50V2's notable characteristics is the usage of residual connections, which allow information to travel directly across the network's levels by skipping certain tiers. Shortcut connections have proven to be extremely useful in training very deep networks, resulting in better performance in image recognition applications. The number "50" in the name refers to the network's depth, which comprises of 50 layers. These layers are grouped into bottleneck blocks to reduce computing complexity while retaining excellent accuracy. Additionally, ResNet50V2 incorporates batch normalization and rectified linear unit (ReLU) activations to improve training stability and convergence.

ResNet50V2 has become a typical benchmark for image classification and object recognition applications, because to its 50-layer deep architecture and residual connections. ResNet50 "V2" includes enhancements that improve overall training efficiency and performance. It is extensively used for transfer learning, which involves fine-tuning pre-trained models on huge datasets for specific image recognition applications, making it a versatile and valuable tool for deep learning and computer vision researchers and practitioners.

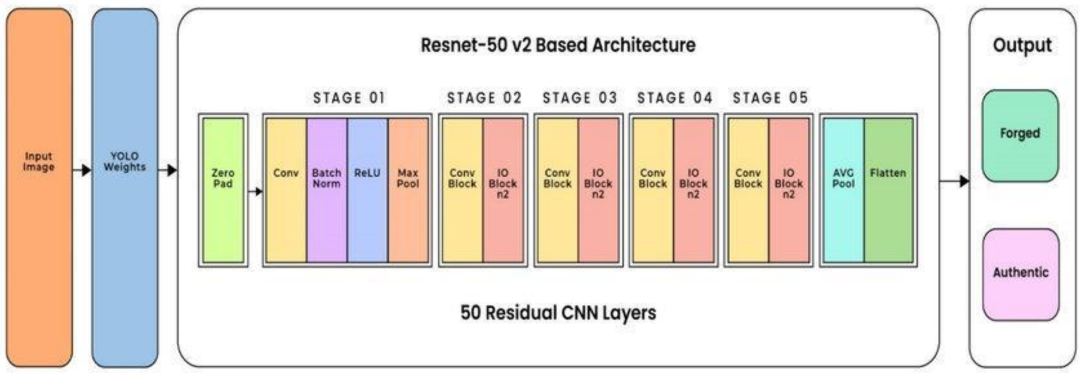

Figure 7 ResNet50V2 Model

3. Xception

Xception, an abbreviation for "Extreme Inception," is a novel convolutional neural network design inspired by the Inception architecture. Xception, which was developed by Google researchers, is intended to push the limits of convolutional neural networks (CNNs) and achieve even greater performance in image classification and computer vision applications.

The unique approach to convolutional layers that distinguishes Xception from other architectures. Xception uses depthwise separable convolutions rather than the typical 3x3 or 5x5 filters used in CNNs. The spatial filtering and channel-wise filtering are decoupled in this approach, lowering the number of parameters and computing complexity. It enables the network to capture finer-grained features while remaining efficient.

The architecture of Xception has shown outstanding performance on difficult picture recognition tasks, and it has received praise for its efficiency in terms of both processing resources and parameter count. It has proven very useful in circumstances with limited computational resources, making it an excellent alternative for a variety of real-world applications. Overall, Xception marks an important advancement in the evolution of convolutional neural network architectures, combining enhanced performance with a more efficient design.

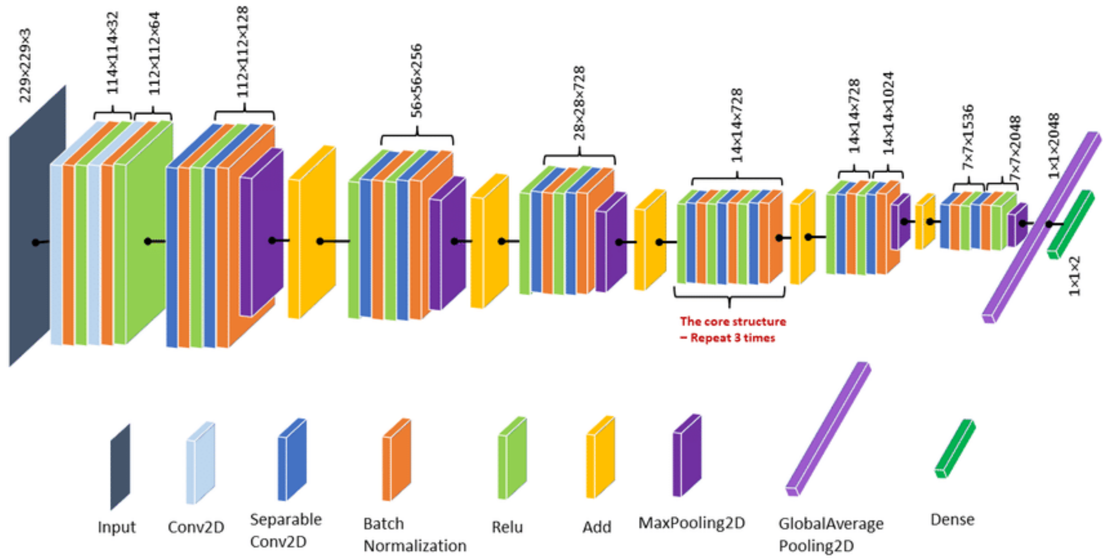

Figure 8 Xception Model

a. Dense Connectivity

Dense Connectivity, also known as Densely Connected Convolutional Networks, is a deep neural network architecture that promotes rich feature reuse by densely connecting each layer in a feedforward fashion. Unlike standard CNNs, which flow information sequentially from one layer to the next, Dense Connectivity allows each layer to accept direct inputs from all previous layers and share its feature maps with all subsequent layers. This interconnectedness improves gradient flow, training efficiency, and accuracy, making Dense Connectivity, as represented by DenseNet, a powerful and widely-used architecture in deep learning and computer vision, particularly for image classification problems.

b. Dense Blocks

Dense blocks comprise an essential part of DenseNet (Densely Connected Convolutional Networks), a deep neural network design. A dense block is made up of a series of tightly connected convolutional layers, where each layer receives feature maps from all preceding layers and passes its own feature maps to all following layers.

This dense interconnection supports considerable feature reuse and encourages the network to record a diverse variety of features across its depth, resulting in very efficient and accurate feature extraction. Dense blocks are a major innovation in DenseNet, allowing deep neural networks to maintain high gradient flow and perform remarkably well in many computer vision applications such as picture classification and object recognition.

c. Growth Rate

"Growth rate" is a hyperparameter in neural network topologies, particularly in DenseNet (Densely Connected Convolutional Networks). The amount of feature maps or channels added to each layer within a dense block is specified by the growth rate. Each layer within a dense block in a DenseNet contributes a growing amount of feature mappings to the subsequent levels. This idea is critical to comprehending the architecture's design since it has a direct impact on the model's capacity and parameter count. A faster growth rate means that more feature maps are added to each layer, resulting in a more expressive model with more parameters, which may offer better performance but necessitates more processing resources.

d. Max Polling

Max pooling is a down-sampling technique often employed in convolutional neural networks (CNNs) for feature extraction in computer vision tasks. It includes dividing the input data into non-overlapping sections (usually 2x2 or 3x3) and selecting the greatest value from each region, which becomes part of the down-sampled output. This approach helps to decrease the spatial dimensions of the data, maintaining the most important information while minimizing computing complexity. Max pooling is extremely good in preserving and enhancing significant features, making it an important component in CNN architectures for tasks such as image classification, object identification, and segmentation, where spatial hierarchies and scale-invariance are required.

e. Convolution layer

A convolutional layer, which is frequently a core component of convolutional neural networks (CNNs), is crucial in extracting and learning features from input data, which is typically images. Convolution procedures are used, which slide small learnable filters (kernels) across the input data to produce feature maps by computing element-wise dot products. These feature maps record many features such as edges, textures, and more complicated patterns, allowing for hierarchical feature representation. By sharing weights inside each kernel, the convolution layer helps to preserve the spatial structure of the data while minimizing the number of parameters. It is a key component of deep learning in computer vision and other fields, facilitating tasks such as picture identification, object detection, and others.

f. BatchNormalization

Batch normalization, a critical approach in deep learning, normalizes neural network layer activations. It computes the mean and variance of activations across a set of training data and then standardizes these values, which aids in the stabilization and acceleration of the training process. Batch normalization, by addressing the vanishing/exploding gradient problem, enables deeper and more stable neural network training. Furthermore, it functions as a type of regularization, lowering the risk of overfitting. Batch normalization is commonly used in neural network designs and has considerably improved the training efficiency and overall performance of deep learning models in a variety of applications.

g. Separate Convolution layer

A separable convolution layer, which is commonly employed in deep neural network designs, is a variation of classic convolution layers aimed to reduce computational complexity while keeping the ability to learn significant features from input data. Instead of using a typical convolution with a big kernel size, separable convolutions split the operation into two steps: depthwise convolution and pointwise (or 1x1) convolution. The depthwise convolution performs a distinct convolution operation on each input channel, whereas the pointwise convolution combines these channel-wise features via 1x1 convolutions to generate the output feature maps. This separation of spatial and channel-wise information considerably decreases the amount of parameters and calculations, making separable convolutions appropriate for resource-constrained applications such as mobile and embedded devices. Despite the computational savings, they frequently retain comparable or even better performance in a variety of tasks such as image classification, object recognition, and semantic segmentation.

B. Pre processing

A modification pipeline was performed to the retinal pictures during the preparation step of the Melanoma Skin Cancer Dataset, which had 10,000 images, to prepare them for analysis. Initially, an outlier check was performed to identify and address any abnormalities, but no duplicate or irrelevant images were identified. The photos were then resized to ensure they all had consistent dimensions for subsequent processing. This standardization of picture sizes is a frequent preprocessing step in machine learning and computer vision applications, helping to ease model training and image comparison while eliminating potential differences in image dimensions.

The resizing phases serves as essential for ensuring that all photographs in the dataset have the same dimensions, which facilitates processing and analysis of the dataset. Following that, the pictures were transformed into tensors, which are numerical representations that are well-suited for use in melanoma skin cancer detection algorithms. This translation allows for fast computing as well as smooth integration with deep learning structures. Furthermore, the picture pixel values were normalized, which is a procedure that aligns these values to a defined scale with a mean and a standard deviation . Normalization is used to normalize data and reduce the impact of different picture features. These preparation processes prepare the Melanoma Skin Cancer Dataset in a consistent and systematic manner.These preparation methods work together to create the Melanoma Skin Cancer Dataset in a compatible and uniform manner, which is necessary to training and assessing deep learning models on Skin Cancer dataset.

C. Training and Validation Split

To facilitate appropriate training and assessment of models in this study, the dataset was painstakingly divided into two subsets: a training subset and a validation subset. This separation allows the models to be trained on a portion of the data (the training set) while their performance is rigorously evaluated on a distinct fraction of the data (validation set) things they did not come across throughout their training period. This technique simulates real-world circumstances and aids in assessing how well-trained models will adapt to new situations, unknown data. The Sample' function of the pandas package was used to achieve this dataset split. The value was set to 0.8, which indicated that 80% of the data was allocated for the training set.

To ensure reproducibility, the 'random_state' option was assigned to a specified value, resulting in consistent data selection across code runs. The validation set was created after the training set was established by removing its indicators from the whole dataset, so deleting the 20% of data that was not part of the training set. This strategy to data partitioning is a widely used practice in machine learning, ensuring a A sensible strategy to training and testing methods. It was used for this study as well as other initiatives, such as the detection of diabetic retinopathy and melanoma skin cancer.

Table 2 Numerical illustration of the dataset's training and test pictures.

Images	Training	Validation
Selected Images	80%	20%
Whole Images 10000	9300	700

V. RESULTS

A. Training Approach

In this work, During the training phase for melanoma skin cancer detection, the initial step involves employing data augmentation techniques to enhance the model's ability to generalize across diverse picture samples.

The random occlusion procedure encompasses several transformations applied to input pictures, including rotations by 90, 180, or 270 degrees, center cropping, adjustments to brightness, and mirroring. These processes simulate alterations that occur in real-world scenarios, enabling the model to effectively capture valuable traits and patterns.

The ultimate phase involves the training and validation of a model that utilizes data obtained from a specifically curated dataset focused on Melanoma skin cancer, subsequent to the use of data augmentation techniques.

Typically, the model acquires knowledge from diverse datasets for the purposes of training and validation, in order to evaluate its efficacy when presented with previously unknown samples. During the training process, the hyperparameters of the model undergo adjustments using the back-propagation algorithm, therefore enhancing its ability to accurately identify cases of melanoma.In this scenario, the model will be subjected to testing using a distinct collection of images depicting melanoma skin cancer. In order to evaluate the performance of the model, many assessment measures are employed, including accuracy, recall, F1 score, and area under the curve (AUC). The model's performance metrics encompass specificity, recall, F1 score, and AUROC.This study employs three distinct convolutional neural network (CNN) models, namely VGG16, Xception, and ResNet50V2. The performance parameters associated with each design provide valuable insights into the efficacy of melanoma detection. The process of comparing several models enables the identification of the most suitable technique that exhibits accuracy, robustness, and generalizability.

B. Model Performance

In the evaluation of model performance for the detection of melanoma skin cancer, three notable CNN architectures, namely VGG16, Xception, and ResNet50V2, were utilized. The findings demonstrate that the VGG16 model attained an accuracy rate of 87%, a recall rate of 84%, an F1 score of 87%, and an area under the curve (AUC) value of 87%. This implies that the VGG16 model has a praiseworthy performance overall, effectively achieving a balance between accuracy and sensitivity for detecting cases of melanoma.

Regarding Xception, the model demonstrated an accuracy of 87%, a recall rate of 82%, an F1 score of 86%, and an AUC of 86%. Although Xception has a commendable level of accuracy, its recall rate is somewhat lower, indicating a modest decrease in its capacity to accurately detect genuine positive situations when compared to VGG16. Nevertheless, the F1 score and AUC metrics demonstrate a commendable equilibrium between accuracy and recall, establishing it as a promising contender for the diagnosis of melanoma.

The ResNet50V2 model demonstrated superior performance in this evaluation, with an accuracy of 89%, a recall rate of 92%, an F1 score of 89%, and an AUC of 88%. The findings demonstrate that ResNet50V2 exhibits a remarkable capacity to effectively classify instances of melanoma, with particularly high levels of sensitivity and overall accuracy. The elevated recall rate shown in the results indicates that ResNet50V2 has superior performance in accurately detecting a greater percentage of genuine positive cases. Consequently, based on this comparative research, ResNet50V2 emerges as the preferred model for melanoma detection.

Table 3 Results from the three main models before collaborative learning

VGG 16	Xception	ResNet50V2
87%	87%	89%

1. Accuracy

The primary statistic utilized in the field of machine learning is accuracy, which quantifies the overall correctness of a model's predictions. Termed as accuracy, this metric pertains to the proportion of accurate predictions divided by the total number of data points in the dataset. While accuracy is a straightforward and comprehensible measure for measuring a model's performance, it may not always be the sole indicator, since there might be instances where one class significantly outweighs others in imbalanced datasets. Nevertheless, in such instances, it is important to note that a high accuracy score might be misleading due to potential bias towards the dominant class, resulting in inadequate performance on the minority class.

Machine learning experts acknowledge the limitations of accuracy as a sole metric and instead employ additional measures like precision, recall, F1-score, or AUC to more comprehensively evaluate the effectiveness of the model. The additional metrics provide insight into the model's capacity to handle false positive and false negative errors, as well as its performance in balancing accuracy and recall. Specifically, in conditions of importance of attaining well-balanced performances across various categories, the precision and recall allow identifying weaknesses or deviations present within the predictions of the model that could negatively impact classification accuracy.

In due course, however, accuracy emerges as the sole criterion for determining the proper amount of precision in a model's predictions. However, it is crucial to consider that the problem must be well characterized, since the effectiveness of any proposed solution will be directly contingent upon the accuracy of this definition. Furthermore, it is important to consider the distribution of classes within the dataset and the potential impact of the error rate on the ultimate implementation. Practitioners have the option to integrate an accuracy measure with additional metrics, enhancing their comprehension of the suitability and dependability of a machine learning model.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

2. Recall/Sensitivity

Another important measure used in machine learning is recall, which is sometimes called sensitivity or true positive rate. It provides an estimate of how well a model recognizes every case of a given class. This is derived by dividing the number of true positives predicted by the total of true positives plus false negatives. Basically, the recall rate expresses how many of the true-positives were caught by the model. A good recall means that the model has low chances of getting false negative and thus identifying most positive cases. This may be very vital for those areas where if such positive cases are missed then it can result into catastrophic outcomes.

Recollection has particular value, particularly if costs are very high for negative results (e.g., in medical diagnosis or fraud detection). For example, in medical settings, having a good recall implies that the model recognizes people affected by a specific disease to minimize possible oversights of crucial cases. Although, if such a model is designed to achieve high recall, then there must be an increased risk for generating false positives due to the model's performance.

Recall is important, as it gives an indication if the model has indeed included all positive cases. This is a central measure in circumstances wherein it is crucial identifying as several positive cases as conceivable, and one must take its comprehension into account when other measures like precision or F1 score are being employed for an impartial evaluation of a model's performance.

$$Sensitivity = \frac{TP}{TP+FN} \qquad (9)$$

3. Specificity

In addition to sensitivity, specificity of performance is another crucial criterion. This metric is sometimes referred to as specificity, which may be defined as the ratio of true negatives to the sum of true negatives and false positives. Essentially, the model's efficacy in accurately identifying true negatives is quantified by its calculation. The issue of specificity has significant importance in classification models, since the occurrence of even a single false positive might result in substantial harm.

Specialization encompasses the concept of specificity, which is particularly prevalent in scenarios where the consequences of false alarms carry significant costs. This is evident in contexts such as pharmaceutical testing and security screening. For example, a high level of specificity in medical diagnostics signifies that the model effectively identifies individuals who do not possess the specified ailment, hence minimizing the likelihood of unnecessary interventions or treatments. It is important to acknowledge that there is typically a trade-off between sensitivity and specificity, since enhancing one may have a detrimental effect on the other. It is crucial to strike a balance between the two in terms of the job's requirements or consequences.

Therefore, specificity provides crucial information on the accuracy of a model in correctly detecting cases classified as negative. In these instances, the prioritization of avoiding false positives necessitates the consideration of sensitivity and accuracy in order to fully evaluate the model's effectiveness.

$$Specificity = \frac{TN}{TN+FP} \qquad (10)$$

4. Precision

Precision is a fundamental metric within the field of machine learning, serving as a quantitative measure of the correctness of positive predictions generated by a given model. The calculation involves determining the proportion of true positive instances in relation to the combined total of true positive and false positive instances. In alternative terms, accuracy measures the ratio of accurately identified positive cases to the total expected positive instances. A high precision score signifies the model's efficacy in limiting false positives, hence instilling trust in the accuracy of its positive result predictions. Precision is of utmost importance in domains where the consequences or expenses associated with false positives are substantial, such as in the fields of medical diagnostics or financial fraud detection.

In situations when the emphasis is on obtaining accuracy, it is imperative to strike a harmonious equilibrium with other measures such as recall. There is typically an inverse relationship between precision and recall, whereby enhancing one metric may result in a decline in the other. Achieving an optimal equilibrium is contingent upon the particular demands of the undertaking. The F1 score is a valuable statistic for evaluating the overall performance of a model in situations when there is a need to balance accuracy and recall. It does this by combining both precision and recall into a single measure.

In brief, the metric of precision offers vital insights into the efficacy of a model in generating accurate positive predictions. The measure has significant importance in scenarios where the reduction of false positives is crucial, and its analysis should be taken into account in conjunction with other metrics to attain a comprehensive assessment of a model's performance.

$$Precision = \frac{TP}{TP + FP} \qquad (11)$$

5. F1-Score

In scenarios when such a trade-off is present, the F1 score serves as an evaluative metric in the field of machine learning. It effectively combines accuracy and recall, offering a comprehensive assessment of a model's performance. The F1-score is a mathematical measure that combines accuracy and memory through the harmonic mean. It is calculated as 2PQ/(P+Q), where P and Q represent precision and recall, respectively. The F1 formula is capable of including both false positives and false negatives, rendering it valuable in situations where a balance between these two factors needs to be attained.

Specifically, the F1 score is more suitable than accuracy when dealing with imbalanced classes. This is especially true in scenarios when the distribution of classes is uneven. For example, the F1 score provides a more accurate assessment of the model's performance in scenarios when the negative class (such as fraudulent transactions or healthy individuals) is predominant, as observed in applications like fraud detection or uncommon illness diagnosis. The F-measure is a metric that integrates the precision in generating positive predictions with the ability to minimize unnecessary false alarms into a single measure.

In conclusion, the F1 score is a significant metric that integrates accuracy and recall to provide a thorough evaluation of the model's output. This issue becomes particularly relevant in situations when the trade-off between false positives and false negatives is crucial, and its comprehension extends beyond mere accuracy, particularly when dealing with disparate datasets.

$$F1 = 2\ x\ \frac{precision\ x\ recall}{precision + recall} \qquad (12)$$

6. Area Under The Curve (AUC)

The Area Under the Curve (AUC) is a commonly employed measure in the field of machine learning, specifically in the domain of binary classification tasks and ROC curves. ROC curve is a visual depiction of a model's performance at different threshold levels, illustrating the relationship between the true positive rate (sensitivity) and the false positive rate. The AUC is a metric used to evaluate the overall performance of model by computing the integral of the curve. A greater area under the receiver operating characteristic curve (AUC) value, which approaches 1, signifies improved discrimination between positive and negative examples, irrespective of the selected threshold. Consequently, AUC serves as a reliable statistic for assessing the model's capacity to rank and classify instances.

The area under the receiver operating characteristic curve (AUC) is especially advantageous in the context of unbalanced datasets, characterized by a substantial disparity in the representation of one class compared to the other. In instances of this nature, the exclusive consideration of accuracy may not yield a comprehensive assessment of a model's performance. The area under the receiver operating characteristic curve (AUC) is a metric that takes into account the balance between sensitivity and specificity. It assesses the model's ability to correctly prioritize positive examples over negative examples by condensing this information into a single numerical number. This is particularly advantageous in situations such as medical diagnostics or credit scoring, when accurately selecting good cases is of utmost importance.

In brief, the area under the receiver operating characteristic curve (AUC) is a robust measure that effectively quantifies the overall discriminatory capability of a model across various decision thresholds. The usefulness of this approach may be applied to situations where there are uneven class distributions or variable classification thresholds, offering a complete and simply understandable evaluation of a model's binary classification performance assignments.

Table 4 Results from the three base models Accuracy,Recall,F1 score,AUC

Model	Accuracy	Recall	F1 score	AUC
ResNet50V2	89%	92%	89%	88%
VGG16	87%	84%	87%	87%
Xception	87%	82%	86%	86%

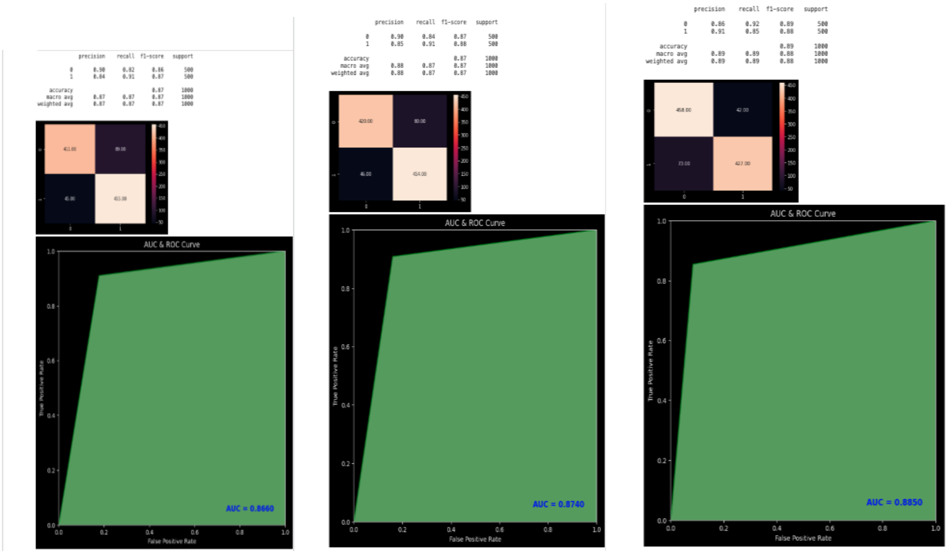

Figure 9 Performance of model ROC area in accuracy, recall, F1 score, and AUC.

C. Summary of Findings

The findings investigation carried out on the Kaggle dataset centered on the identification of melanoma skin cancer through the use of three separate CNN architects, known as VGG16, Xception, and ResNet50V2. The dataset was subjected to preprocessing, which involved performing picture transformations and dividing it into separate training and test sets. The assessment of the models yielded significant performance indicators. The VGG16 model exhibited an accuracy rate of 87%, a recall rate of 84%, an F1 score of 87%, and an area under the curve (AUC) value of 87%. The Xception convolutional neural network (CNN) architecture demonstrated notable performance metrics, including an accuracy of 87%, a recall rate of 82%, an F1 score of 86%, and an AUC of 86%. ResNet50V2 demonstrated superior performance, with notable results including an accuracy rate of 89%, a recall rate of 92%, an F1 score of 89%, and an AUC of 88%.

The findings underscore the efficacy of deep learning models in the identification of melanoma, with ResNet50V2 exhibiting improved performance in several important measures. The combined high recall, accuracy, F1 score, and AUC metrics highlight the effectiveness of the model in reliably detecting cases of melanoma. This is particularly important in the context of early diagnosis, as it can significantly contribute to better patient outcomes. The aforementioned discoveries provide significant contributions to the continuous endeavors in utilizing sophisticated machine learning methods for the identification of skin cancer. These findings underscore the need of carefully selecting appropriate models to improve the accuracy of diagnoses.

n summary, the examination of VGG16, Xception, and ResNet50V2 in the context of melanoma detection using the Kaggle dataset highlights the importance of model design in attaining reliable and effective results. The findings underscore the promise of ResNet50V2 as a viable model for the identification of melanoma, highlighting its better performance in terms of accuracy and recall, which are crucial criteria in the field of medical image classification. The potential for expanding the area of skin cancer diagnosis lies in the further research and refining of these models.

D. Potential Improvements and Future Work

In the domain of enhancing the identification of melanoma skin cancer by the utilization of machine learning techniques, there are several prospective avenues for future research. One potential approach is investigating ensemble approaches that combine the capabilities of many models, such as VGG16, Xception, and ResNet50V2. The utilization of ensemble approaches has the potential to alleviate the limitations of individual models, resulting in a prediction system that is more resilient and precise. Moreover, the utilization of transfer learning techniques, which include leveraging pre-trained models on extensive datasets, has the potential to enhance the ability to generalize to various features of skin lesions. The flexibility of pre-trained models to the complexity of the target domain might potentially be improved by fine-tuning them to effectively capture the unique nuances present in melanoma photos.

An additional domain warranting future investigation pertains to the enhancement of the model's durability and efficacy. The aforementioned objective may be accomplished by integrating advanced data augmentation techniques into the training dataset in order to enhance its diversity and mitigate any existing imbalances. Furthermore, it is important to emphasize the significance of enhancing machine learning models in the field of medical diagnostics by assuring their interpretability and explainability. The prioritization of research endeavors aimed at enhancing the transparency and comprehensibility of these models is crucial in cultivating trust within therapeutic environments. Ultimately, it is of utmost importance to establish collaborative collaborations with dermatologists and healthcare experts. By incorporating clinical insights and domain expertise into the model building process, it is possible to refine the algorithms in order to more effectively address the practical obstacles and complexities that arise in real-world situations. This eventually leads to an improvement in the clinical usefulness and overall impact of the algorithms.

E. Implications of the Study

The work on melanoma skin cancer detection using machine learning has several implications that are of great importance to the field of medical diagnostics and have broader applications in healthcare. The results of the study highlight the promise of advanced deep learning models, specifically ResNet50V2, in improving the accuracy and effectiveness of melanoma detection. The aforementioned has significant ramifications for the timely identification of medical conditions, which is a crucial element in enhancing patient results and potentially preserving lives. This research adds to the expanding corpus of data that endorses the incorporation of artificial intelligence (AI) into clinical processes. It demonstrates the potential of AI to enhance the diagnostic abilities of healthcare workers while dealing with intricate medical situations.

Furthermore, the efficacy of CNNs such as VGG16, ResNet50V2 and Xception, in the examination of medical imaging data underscores the adaptability of these models. The ongoing development of machine learning algorithms establishes a significant precedent for the integration of these technologies within the field of dermatology and the diagnosis of skin cancer. The ramifications of this study have a wider scope than only melanoma, indicating that comparable approaches may be utilized to identify and categorize other dermatological disorders. This has the potential to significantly enhance dermatological treatment on a greater scale.

Nevertheless, it is essential to acknowledge that the incorporation of artificial intelligence (AI) in the healthcare sector entails ethical problems, such as privacy concerns, potential biases, and the need for interpretability. The consequences of the study encompass the need to address ethical problems in order to guarantee responsible and fair implementation of machine learning models in healthcare environments. In brief, the consequences of the study encompass enhanced clinical practices in the identification of melanoma, the wider use of Convolutional Neural Networks (CNNs) in the field of dermatology, and the necessity of ethical considerations to govern the appropriate integration of Artificial Intelligence (AI) in the healthcare sector.

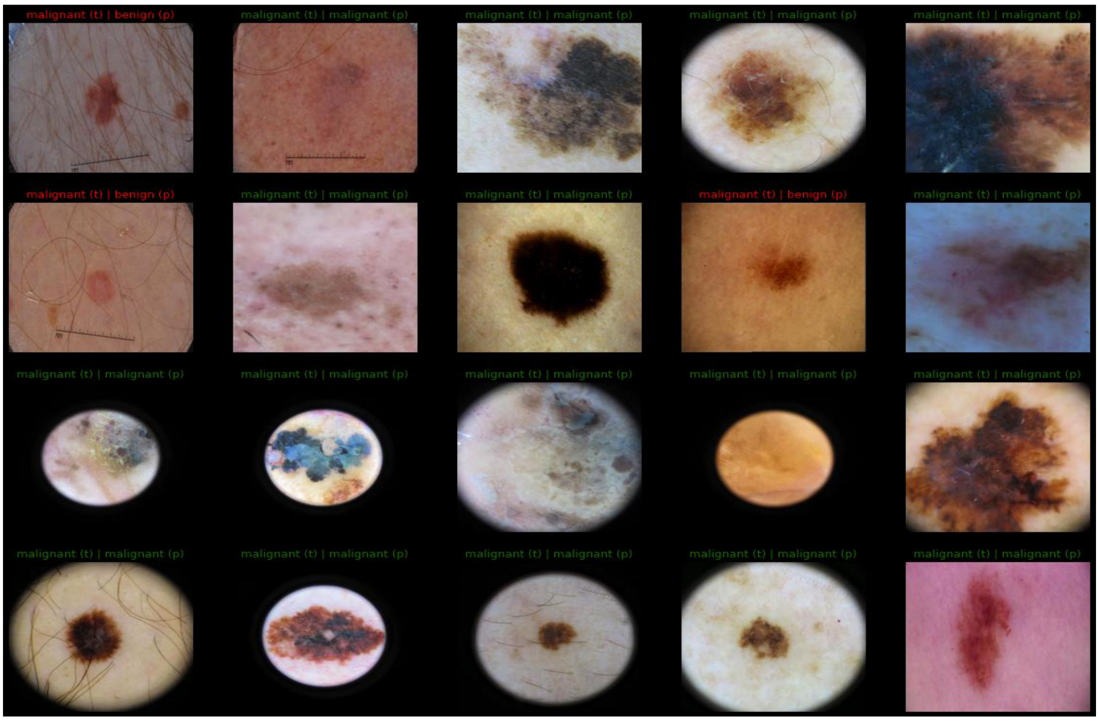

Figure 10 **VGG** Melanoma Skin Cancer Detection

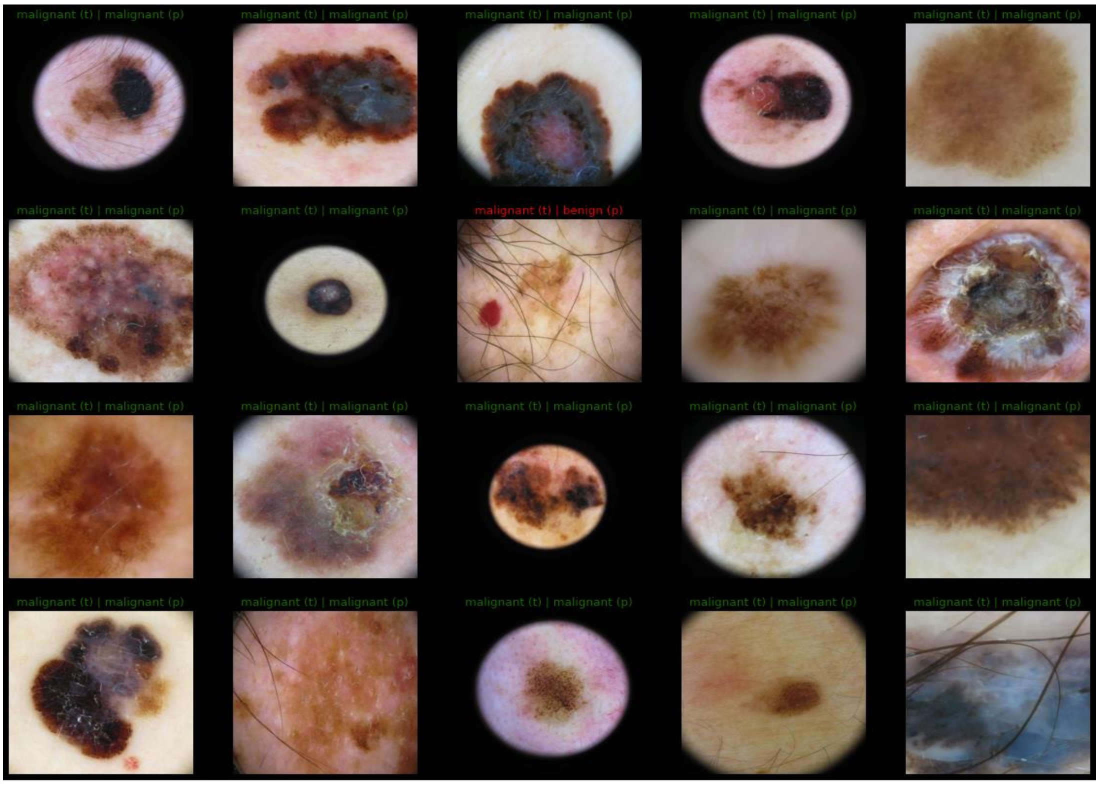

Figure 11 **Xception** Melanoma Skin Cancer Detection

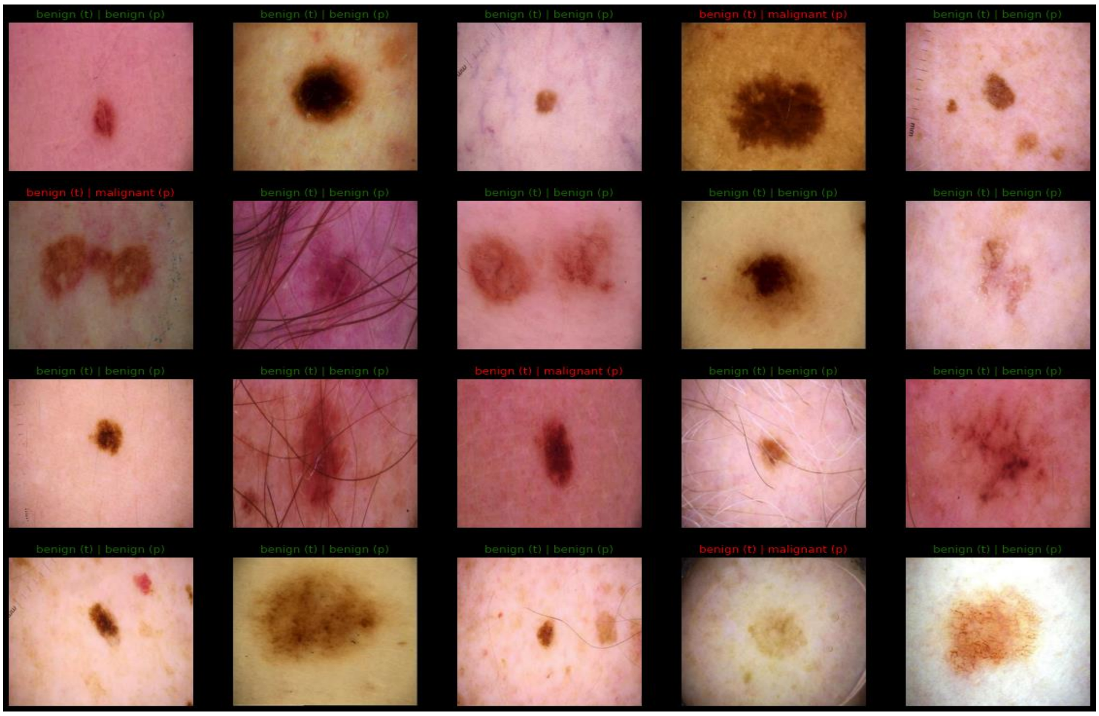

Figure 11 **ResNet50V2** Melanoma Skin Cancer Detection

VI. CONCLUSION

Ultimately, utilizing machine learning models such as VGG16, Xception, and ResNet50V2 for the detection of skin cancer shows great potential for the development of sophisticated diagnostic instruments. After doing thorough research and evaluation, it has been concluded that ResNet50V2 is the most effective model among the options, with an excellent accuracy rate of 89%. This indicates that ResNet50V2 surpasses other models in accurately distinguishing skin cancer from photos.

The results of this study emphasize the importance of choosing the appropriate model structure for tasks involving the classification of dermatological images. With an accuracy rate of 89%, ResNet50V2 demonstrates a strong performance, making it a dependable choice for future development and practical use in real-world situations. This achievement not only showcases the capacity of machine learning in identifying skin cancer but also underscores the significance of ongoing investigation and improvement of models to achieve higher levels of precision and dependability.

In order to make progress, it is important for future efforts in this area to concentrate on investigating supplementary models and architectures to consistently enhance the effectiveness of skin cancer detection. By broadening our research to include a wider range of machine learning models, we can enhance our comprehension of their capabilities and constraints, so aiding the construction of more advanced and precise diagnostic tools for detecting skin cancer.

VII. REFERENCES

BOOKS

CHRISTOPHER M. BISHOP (2016) **Pattern Recognition and Machine Learning**

AURÉLIEN GÉRON (2017) **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow**

TREVOR HASTIE, ROBERT TIBSHIRANI, AND JEROME FRIEDMAN(2001) **The Elements of Statistical Learning**

ARTICLES

ALQURAN H, QASMIEH IA, ALQUDAH AM, ALHAMMOURI S, ALAWNEH E, ABUGHAZALEH A, HASAYEN F (2017) “The melanoma skin cancer detection and classification using support vector machine”. In **IEEE Jordan conference on applied electrical engineering and computing technologies** (AEECT), 1-5

BASLY H, OUARDA W, SAYADI FE, OUNI B, ALIMI AM (2020) “ CNN-SVM learning approach based human activity recognition”. **Proceedings of international conference on image and signal processing. Springer, Cham,** pp 271–281

BRINKER TJ, HEKLER A, ENK AH, BERKING C, HAFERKAMP S, HAUSCHILD A, WEICHENTHAL M, KLODE J, SCHADENDORF D, HOLLAND-LETZ T, KALLE CV, FRÖHLING S, SCHILLING B, UTIKAL JS (2019) Deep neural networks are superior to dermatologists in melanoma image classification. **Eur J Cancer** 119:11–17

BRINKER TJ, HEKLER A, ENK AH, VON KALLE C (2019) “Enhanced classifier training to improve precision of a convolutional neural network to identify images of skin lesions”. **PLoS One** 14(6):e0218713

CHATURVEDI SS, TEMBHURNE JV, DIWAN T (2020) A multi-class skin Cancer classification using deep convolutional neural networks. **Multimed Tools Appl** 79(39):28477–28498

CHEN Y, HE F, WU Y, HOU N (2017) A local start search algorithm to compute exact Hausdorff distance for arbitrary point sets. **Pattern Recogn** 67:139–148

CHEN C, WANG G, PENG C, FANG Y, ZHANG D, QIN H (2021) “Exploring rich and efficient spatial temporal interactions for real-time video salient object detection”. **IEEE Trans Image Process** 30:3995–4007

CHOLLET F (2017) Xception: “deep learning with depthwise separable convolutions ”. **IEEE conference on computer vision and pattern recognition,** 1251-1258

CHOWDHARY CL, ALAZAB M, CHAUDHARY A, HAKAK S, GADEKALLU TR (2021) “ Computer vision and recognition systems using machine and deep learning approaches: fundamentals, Technologies and Applications”. **Inst Eng Technol**

CHOWDHARY CL, REDDY GT, PARAMESHACHARI BD (2022) “ Computer vision and recognition systems” , **research innovations and trends. CRC Press**

CODELLA NC, GUTMAN D, CELEBI ME, HELBA B, MARCHETTI MA, DUSZA SW, HALPERN A (2018) “ Skin lesion analysis toward melanoma detection: a challenge at the 2017 international symposium on biomedical imaging (ISBI) ”, hosted by the international skin imaging collaboration (ISIC). In **IEEE 15th international symposium on biomedical imaging (ISBI 2018),** 168-172

DO MN, VETTERLI M (2005) “ The contourlet transform: an efficient directional multiresolution image Representation ”. **IEEE Trans Image Process 14(12):2091–2106**

ELGAMAL M (2013) “ Automatic skin Cancer images classification. ” **Int J Adv Comput Sci Appl 4(3):287–294**

FAROOQ MA, KHATOON A, VARKARAKIS V, CORCORAN P (2020) “Advanced deep learning methodologies for skin Cancer classification in prodromal stages ”. **arXiv preprint arXiv:2003.06356**

GANSTER H, PINZ P, ROHRER R, WILDLING E, BINDER M, KITTLER H (2001) “ Automated melanoma recognition ” **IEEE Trans Med Imaging** 20(3):233–239

GU J, WANG Z, KUEN J, MA L, SHAHROUDY A, SHUAI B, LIU T, WANG X, WANG L, WANG G, CAI J, CHEN, T (2015) “ Recent advances in convolutional neural networks ”. **CoRR, arXiv preprint arXiv**

HE K, ZHANG X, REN S, SU, J (2016) “ Deep residual learning for image recognition”. In Proceedings of the **IEEE conference on computer vision and pattern recognition,** 770–778. https://doi.org/10.1109/CVPR.

2016.90

HEBBAR N, PATIL HY, AGARWAL K (2020) “ Web powered CT scan diagnosis for brain hemorrhage using deep Learning ”. In **2020 IEEE 4th conference on Information & Communication Technology (CICT),** 1–5

HOSNY KM, KASSEM MA, FOAUD MM (2018) “ Skin cancer classification using deep learning and transfer Learning” . **In 9th Cairo international biomedical engineering conference (CIBEC),** 90-93. https://doi.org/10.1109/CIBEC.2018.8641762

HUANG D, SHAN C, ARDABILIAN M, WANG Y, CHEN L(2011) “ Local binary patterns and its application to facial image analysis: a survey ”. **IEEE Trans Syst Man Cybern Part C Appl Rev** 41(6):76

HUANG G, LIU Z, VAN DER MAATEN L, WEINBERGER KQ (2017) “Densely connected convolutional networks” . In proceedings of the **IEEE conference on computer vision and pattern recognition,** 4700-4708

ISMAIL MA, HAMEED N, CLOS J (2021) " Deep learning-based algorithm for skin Cancer classification" . In proceedings of the **international conference on trends in computational and cognitive engineering,** 709–719

JANA E, SUBBAN R, SARASWATHI S (2017) " Research on skin Cancer cell detection using image processing ". In **IEEE international conference on computational intelligence and computing research (ICCIC),** 1–8. https://doi.org/10.1109/iccic.2017.8524554

KAWAHARA J, HAMARNEH G (2016) " Multi-resolution-tract CNN with hybrid pretrained and skin-lesion trained Layers ". In: **International workshop on machine learning in medical imaging. Springer, Cham,** pp 164–171

KINGMA DP, BA J (2014) " a method for stochastic optimization ". **arXiv preprint arXiv:**1412.6980

LABANI S, ASTHANA S, RATHORE K, SARDANA K (2020) "Incidence of melanoma and nonmelanoma skin cancers " in **Indian and the global regions.** 17:906–911. https://doi.org/10.4103/jcrt.JCRT_785_19

LIANG H, SUN X, SUN Y, GAO Y (2017) " Text feature extraction based on deep learning ". **EURASIP J Wirel Commun Netw 2017**(1):1–12

LIANG Y, HE F, ZENG X (2020) " 3D mesh simplification with feature preservation based on whale optimization algorithm and differential evolution ". **Integr Comput-Aided Eng** 27(4):417–435

LOPEZ AR, GIRO-I-NIETO X, BURDICK J, MARQUES O(2017) " Skin lesion classification from dermoscopic images using deep learning techniques ". In **13th IASTED international conference on biomedical engineering (BioMed),** 49-54

MISHRA SP, SARKAR U, TARAPHDER S, DATTA S, SWAIN DP, SAIKHOM R, LAISHRAM M (2017) Multivariate statistical data analysis-principal component analysis (PCA). **Int J Livestock Res** 7(5):60–78. https://doi.org/10.5455/ijlr.20170415115235

OJALA T, PIETIKÄINEN M, HARWOOD DA (1996) "Comparative study of texture measures with classification based on featured distributions" . **Pattern Recogn** 29(1):51–59

PATIL HY, KOTHARI AG, BHURCHANDI KM (2016) "Expression invariant face recognition using local binary patterns and contourlet transform". **Optik 127(5):2670–2678**

RIKER AI, ZEA N, TRINH T (2010) "The epidemiology, prevention, and detection of melanoma". **Ochsner J** 10(2):56–65

SÁEZ A, SÁNCHEZ-MONEDERO J, GUTIÉRREZ PA, HERVÁS-MARTÍNEZ C (2015) " Machine learning methods forbinary and multiclass classification of melanoma thickness from dermoscopic images ". **IEEE Trans Med Imaging** 35(4):1036–1045

SAHLOL AT, YOUSRI D, EWEES AA, AL-QANESS MA, DAMASEVICIUS R, ELAZIZ MA (2020) " COVID-19 image classification using deep features and fractional-order marine predators algorithm" . **Sci Rep** 10(1):1–15

SHARMA N, JAIN V, MISHRA A (2018) "An analysis of convolutional neural networks for image classification".**Proced Comput Sci** 132:377–384

SHORTEN C, KHOSHGOFTAAR TM (2019) "A survey on image data augmentation for deep learning". **J Big Data** 6(1):1–48. https://doi.org/10.1186/s40537-019-0197-0

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

SZEGEDY C, VANHOUCKE V, IOFFE S, SHLENS J, WOJNA Z (2016) "Rethinking the inception architecture for computer vision". In proceedings of the **IEEE conference on computer vision and pattern recognition,** 2818- 2826. https://doi.org/10.1109/CVPR.2016.308

WEESE J, LORENZ C (2016) Four challenges in medical image analysis from an industrial perspective. **Med Image Anal** 33:44–49**.** https://doi.org/10.1016/j.media.2016.06.023

WEISS K, KHOSHGOFTAAR TM, WANG D (2016) "A survey of transfer learning" . **J Big Data** 3(1):1–40. https://doi.org/10.1186/s40537-016-

WELCH HG, WOLOSHIN S, SCHWARTZ LM (2005) "Skin biopsy rates and incidence of melanoma: **population based ecological study" . BMJ** 331(7515):481

WU Y, HE F, ZHANG D, LI X (2015) “Service-oriented feature-based data exchange for cloud-based designand manufacturing.” **IEEE Trans Serv Comput 11(2):341–353**

WU J, HU W, WEN Y, TU W, LIU X (2020) “Skin lesion classification using densely connected convolutional networks with attention residual learning”. **Sensors** 20(24):7080

WU Z, LI S, CHEN C, HAO A, QIN H (2020) “A deeper look at salient object detection: bi-stream network with a small training dataset” . **arXiv preprint arXiv**:2008.02938

. WU Z, LI S, CHEN C, HAO A, QIN H (2022) “ Recursive multi-model complementary deep fusion for robust salient object detection via parallel sub-networks ”. **Pattern Recogn** 121:108212

XIE F, FAN H, LI Y, JIANG Z, MENG R, BOVIK A (2016) “ Melanoma classification on dermoscopy images using a neural network ensemble model ”. **IEEE Trans Med Imaging** 36(3):849–858

RESUME

NAME: Muhammad Ali Abbasi

EDUCATION:

- **Bachelor**:Computer Science 2012 - 2018 Bachelor in computer science
- **M.S**: 2023, Istanbul Aydin University, Faculty of Computer Sciences, Department of Software Engineering, Artificial Intelligence and data science (English)(Thesis) Program

PROFESSIONAL EXPERIENCE:

- **Flutter developer**

Amnen Tech Dec, 2O21 – Present

- **Android Developer and Flutter developer**

Fiverr Freelancer Jul, 2O21 – Present

- **Android Developer And Flutter Developer (Remote Job)**

Atitsolutionz | UAE Jan 2019 – May 2020

- **Android Developer (Remote Job)**

Apricot Computers | USA Jun 2018 – Jan 2020

Projects:

- **Travel Affiliate App**

People use travel apps to book flights and hotels and travel across the world.

- **Dog Out**

Dog Out is a Flutter Mobile app that allows individuals to find restaurants and other locations where they may take their pets.

- **Amnen Call**

Amnen call is a Flutter Mobile app messenger via which users may communicate with one another.

PUBLICATIONS FROM DISSERTATION, PRESENTATIONS AND PATENTS:

- Muhammad Alı Abbasi, Melanoma Skın Cancer Detectıon in Eurasıan journal of educatıon research

OTHER PUBLICATIONS, PRESENTATIONS AND PATENTS:

-