COMPARATIVE ANALYSIS OF DEEP LEARNING AND MACHINE LEARNING APPROACHES FOR AUTOMATED SKIN CANCER DETECTION

A DISSERTATION

SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE AWARD OF THE DEGREE

OF

MASTER OF TECHNOLOGY

IN

INFORMATION SYSTEMS

Submitted by:

SRIDEEP DAS

(2K21/ISY/23)

Under the supervision of

Prof. KAPIL SHARMA

DEPARTMENT OF INFORMATION TECHNOLOGY
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042
MAY, 2023

DEPARTMENT OF INFORMATION TECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(FORMERLY DELHI COLLEGE OF ENGINEERING)

BAWANA ROAD, DELHI-110042

CANDIDATE'S DECLARATION

I, SRIDEEP DAS Roll No. 2K21/ISY/23 student of M-Tech (Information System), hereby declare that the project Dissertation titled "COMPARATIVE ANALYSIS OF DEEP LEARNING AND MACHINE LEARNING APPROACHES FOR AUTOMATED SKIN CANCER DETECTION" which is submitted by me to the Department of information Technology, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

Place: Delhi SRIDEEP DAS

Date:31/05/23 2K21/ISY/23

DEPARTMENT OF INFORMATION TECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(FORMERLY DELHI COLLEGE OF ENGINEERING)

BAWANA ROAD, DELHI-110042

CERTIFICATE

I hereby certify that the Project Dissertation titled "COMPARATIVE ANALYSIS OF DEEP LEARNING AND MACHINE LEARNING APPROACHES FOR AUTOMATED SKIN CANCER DETECTION" which is submitted by SRIDEEP DAS, Roll No. 2K21/ISY/23, Department of Information Technology, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology, is a record of the project work carried out by the student under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Place: Delhi — Prof. KAPIL SHARMA

Date:31/05/23 — SUPERVISOR

DEPARTMENT OF INFORMATION TECHNOLOGY
DELHI TECHNOLOGICAL UNIVERSITY
(FORMERLY DELHI COLLEGE OF ENGINEERING)
BAWANA ROAD, DELHI-110042

ACKNOWLEDGEMENT

I wish to express my sincerest gratitude to Prof. Kapil Sharma for his continuous guidance and mentorship that he provided us during the project. He showed us the path to achieve our targets by explaining all the tasks to be done and explained to us the importance of this project as well as its industrial relevance. He was always ready to help us and clear our doubts regarding any hurdles in this project. Without his constant support and motivation, this project would not have been successful Also I would like to thank my friends and seniors who guided and helped me in all possible ways.

Place: Delhi | SRIDEEP DAS
Date: 31/05/23 | 2K21/ISY/23

ABSTRACT

Skin cancer is a prevalent and potentially life-threatening disease with increasing global incidence rates. Early detection plays a crucial role in improving patient outcomes and survival rates. This thesis presents a comprehensive study on the development and evaluation of an automated approach for skin cancer detection, aiming to assist clinicians in accurate and timely diagnosis. The research focuses on leveraging advanced machine learning and computer vision techniques to analyze dermoscopic images and identify potential malignancies. The dataset utilized consists of a diverse collection of annotated skin lesion images, encompassing various types and stages of skin cancer. The primary objective is to develop a robust and reliable model capable of distinguishing between benign and malignant lesions. The thesis explores different aspects of the automated skin cancer detection pipeline, starting from data preprocessing and augmentation techniques to enhance the model's generalization capabilities. Various feature extraction methods, including handcrafted features and deep learning-based representations, are investigated to capture relevant patterns and discriminative information from the images. Multiple classification algorithms are studied to compare their performance and determine the most effective approach. Evaluation metrics such as accuracy, precision, recall, and area under the receiver operating characteristic curve (AUC-ROC) are utilized to assess the models' diagnostic accuracy and robustness.

Furthermore, the thesis investigates the impact of different factors on skin cancer detection, including the size and composition of the training dataset, the selection of optimal hyperparameters, and the influence of various image preprocessing techniques. The goal is to optimize the model's performance and provide insights into the factors affecting its effectiveness. The results demonstrate the effectiveness of the proposed automated approach for skin cancer detection. The deep learning models, particularly convolutional neural networks and their variants, outperform traditional machine learning algorithms in terms of accuracy, precision, and other evaluation metrices. These models exhibit superior capability in learning intricate patterns and representations from raw image data, enabling more accurate and reliable diagnosis. The thesis also discusses the limitations and challenges associated with automated skin cancer detection, including the

need for large annotated datasets, computational requirements, and interpretability of deep learning models. Potential strategies and future research directions are proposed to address these challenges and enhance the clinical applicability of the developed models. In summation, this thesis contributes to the field of skin cancer detection by presenting an automated approach that demonstrates high accuracy and efficiency in identifying malignant skin lesions. The findings highlight the potential of machine learning and computer vision techniques to assist dermatologists in early diagnosis and prompt intervention. The research opens avenues for further advancements in automated skin cancer detection systems, leading to improved patient outcomes and reduced mortality rates.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND AND SIGNIFICANCE

Skin cancer is a prevalent and potentially life-threatening disease that arises from the uncontrolled growth of abnormal cells in the skin. It is the most prevalent type of cancer globally, with the incidence rates continuing to rise [1]. Skin cancer is mainly categorized into three main types which are: melanoma, basal cell carcinoma (BCC), and squamous cell carcinoma (SCC). Melanoma is the most aggressive form and has the highest mortality rate.

Early detection and accurate diagnosis of skin cancer are crucial for successful treatment and improved patient outcomes. Traditionally, skin cancer diagnosis has relied on visual inspection by dermatologists, which can be subjective and prone to errors. Dermatologists evaluate various visual characteristics of skin lesions, including asymmetry, color variation border irregularity, and diameter, to determine if a lesion is malignant or benign [2]. However, this process can be challenging, as certain features may be subtle or easily overlooked.

The advent of deep learning and computer vision techniques has revolutionized the field of skin cancer detection. Automated skin cancer detection systems based on deep learning models have shown great promise in assisting dermatologists and healthcare professionals in early and accurate diagnosis.

The significance of skin cancer detection projects lies in several key aspects:

a) **Early Detection**
   Early detection is crucial for effective treatment and improved survival rates. Automated skin cancer detection systems can aid in the early identification of suspicious lesions, enabling timely interventions and potentially preventing the progression of cancer.

b) **Improved Accuracy**

Deep learning models can analyze large amounts of data and extract intricate patterns that may be challenging for the human eye to discern. By leveraging advanced algorithms and neural networks, these models can provide more accurate and objective assessments of skin lesions, reducing the risk of misdiagnosis.

c) **Augmenting Healthcare Professionals**

Skin cancer detection projects aim to develop tools that augment the expertise of dermatologists and healthcare professionals. These systems can assist in the initial screening and triaging of skin lesions, allowing dermatologists to focus on more complex cases and providing a second opinion in cases where expertise may be limited.

d) **Scalability and Accessibility**

Automated skin cancer detection systems have the potential to reach a wider population, especially in regions with limited access to specialized healthcare services. They can be deployed as telemedicine tools, allowing remote patients to receive preliminary assessments and guidance on whether further medical attention is required.

e) **Research Advancements**

Conducting skin cancer detection projects contributes to the field of medical imaging and artificial intelligence. By exploring and evaluating state-of-the-art deep learning models, researchers can enhance the understanding of the capabilities and limitations of these models in the context of skin cancer detection. This knowledge can pave the way for future advancements and improvements in the field.

In conclusion, the background and significance of skin cancer detection projects lie in addressing the critical need for early and accurate detection of skin cancer, improving diagnostic accuracy, augmenting healthcare professionals, and making a positive impact on patient care and outcomes.

1.2 OVERVIEW OF AUTOMATED SKIN CANCER DETECTION USING DEEP LEARNING MODELS

Automated skin cancer detection using deep learning models is a rapidly evolving field that aims to improve the accuracy and efficiency of skin cancer diagnosis. Deep learning models, particularly convolutional neural networks (CNNs), have demonstrated remarkable capabilities in analyzing skin lesion images and distinguishing between malignant and benign lesions. This approach offers several advantages over traditional manual examination by dermatologists, including increased objectivity, scalability, and potential for early detection [3].

The process of automated skin cancer detection using deep learning models typically involves the following steps:

a) **Dataset Collection and Annotation**: A comprehensive dataset of skin lesion images is collected, often including a diverse range of lesion types, sizes, and skin conditions. Each image is annotated with corresponding labels depicting whether the lesion is malignant or benign. The dataset may also include additional information such as patient demographics and lesion characteristics.

b) **Preprocessing**: The collected dataset is preprocessed to ensure consistency and optimize the input data for the deep learning model. In order to boost the diversity and size of the dataset, preprocessing methods may be used to resize the photographs to a standard size, normalize the pixel values, and perform data augmentation techniques.

c) **Model Selection and Architecture**: Deep learning models suitable for skin cancer detection, such as CNNs, are selected based on their performance and suitability for the task. Popular architectures include VGGNet, ResNet, Inception, DenseNet, and their variations. The selected model's architecture is tailored to handle the specific characteristics of skin lesion images.

d) **Model Training**: Training is done on the preprocessed dataset using the chosen deep learning model.Training involves optimizing the model's parameters by iteratively presenting the images and their corresponding labels to the model. The model learns to extract relevant features from the images and make required predictions based on those features. Training typically utilizes techniques such as back-propagation and gradient descent to minimize the loss function.

e) **Model Evaluation and Validation**: The trained model is evaluated using a separate validation dataset to assess its performance. Evaluation metrics such as accuracy, precision, recall, and F1-score are calculated to measure the model's ability to correctly classify skin lesions. Validation techniques like cross-validation or hold-out validation are employed to ensure the model's generalization and robustness.

f) **Testing and Deployment**: The final trained model is tested on an independent testing dataset to assess its performance in real-world scenarios. Once the model achieves satisfactory performance, it can be deployed in practical applications. This may involve integrating the model into software or web-based systems that allow healthcare professionals to upload skin lesion images for automated analysis and receive predictions or probability scores indicating the likelihood of malignancy.

g) **Continuous Improvement and Research**: The field of automated skin cancer detection using deep learning models is continuously evolving. Researchers and developers explore new architectures, optimization techniques, and strategies to improve the accuracy and efficiency of the models. Ongoing research focuses on addressing challenges such as class imbalance, interpretability, and generalization to diverse populations.

Automated skin cancer detection using deep learning models holds great potential for improving the accuracy, efficiency, and accessibility of skin cancer diagnosis. By leveraging the power of artificial intelligence and deep learning, these models can assist dermatologists in making more informed decisions, enabling early detection, and ultimately improving patient outcomes in the fight against skin cancer.

1.3 RESEARCH OBJECTIVES AND SCOPE OF THE STUDY

Our research aims to contribute to the ongoing efforts to detect and decrease the mortality rates posed by skin cancer. In this study, we propose a new approach to skin cancer detection that involves two powerful deep learning techniques: InceptionResNetV2 and LeViT mechanism. The followings are the research objectives and contributions of our study:

Research Objectives:

1. **Evaluate the Performance of LeViT and Inception ResNet V2 Models:**
   The primary objective is to assess the performance of the LeViT and Inception ResNet V2 models in skin cancer detection. This includes measuring their accuracy, precision, recall, and other relevant evaluation metrics. By comparing the performance of these models, the research aims to determine their effectiveness in accurately identifying skin cancer lesions.

2. **Comparative Analysis with Past Models:**
   Another objective is to compare the performance of LeViT and Inception ResNet V2 models with past models used in skin cancer detection. By conducting a comparative analysis, the research aims to identify any improvements or advancements achieved by the newer models. This analysis provides insights into the effectiveness of the state-of-the-art models and their potential contributions to skin cancer detection.

3. **Dataset Preparation and Preprocessing:**
   A crucial objective is to collect an appropriate dataset for training and testing the LeViT and Inception ResNet V2 models. This involves gathering a comprehensive set of skin lesion images and associated labels. The research focuses on dataset

preparation and preprocessing techniques to ensure the dataset's quality, balance, and suitability for training the models.

4. **Model Implementation and Optimization:** The research aims to implement the LeViT and Inception ResNet V2 models for skin cancer detection. This involves configuring the models, fine-tuning hyperparameters, and optimizing their architecture for optimal performance. The objective is to achieve models that are capable of accurately detecting skin cancer lesions while ensuring efficient computation.

Contributions:

1. A novel Skin cancer detection approach that combines pre-trained models and selfattention mechanisms to improve the accuracy and efficiency of skin cancer detection.
2. Insights into the effectiveness of InceptionResNetV2 and Vision Transformer for skin cancer detection.
3. Demonstration of the superiority of our proposed approach over existing state-of the-art methods.
4. A promising direction for future research in skin cancer detection.

CHAPTER 2

LITERATURE SURVEY

2.1 EXISTING METHODS:

Over the past few years, the field of skin cancer detection has experienced remarkable advancements, enabling the early diagnosis of cancer. These advancements have been driven by the development of various methods and techniques, including Computer-Aided Diagnosis (CAD) systems and machine learning algorithms such as Decision Trees, Support Vector Machines (SVM), Bayesian Learning, and Convolutional Neural Networks (CNN) [8]. Additionally, the use of ensemble models and feature aggregation techniques has also contributed to significant improvements in image classification for skin cancer detection. Notably, developments in regularization techniques, data augmentation, and optimization of hyperparameters have brought about significant changes in image classification accuracy. These advancements have paved the way for more accurate and efficient skin cancer detection, ultimately improving patient outcomes and prognosis.

Deep learning is a powerful and complex learning process that becomes more intricate as the number of layers increases. It has gained recognition as a mature application in medical diagnostics due to its high performance. In recent years, deep learning has made significant contributions to the field of skin lesion classification. However, the limited availability of datasets poses challenges for groundbreaking research in medical diagnostics using deep learning. One major reason is the reliance of deep learning algorithms on large-scale labeled datasets, as they require millions of parameters and extensive labeled data for effective learning. When trained on limited data, deep learning models may suffer from over fitting issues, as they utilize a significant portion of resources to fit the training data, leading to difficulties in generalizing to unseen data. Extensive research has been conducted to overcome the challenges posed by limited data in training the deep learning models. These include techniques such as data augmentation, transfer learning, and ensemble of classifiers [10]. The following sections provide a comprehensive review of existing techniques and some related studies conducted in the field of skin lesion classification.

2.1.1 GENERAL COMPUTER AIDED APPROACH

A computer-aided diagnosis (CAD) system is designed to analyze medical conditions using computer-based algorithms such as machine learning and deep learning. These systems assist human experts in making accurate diagnostic decisions by leveraging technologies such as computer vision, artificial intelligence, statistics, and mathematics [1]. Figure 1 provides a general overview of a CAD system specifically designed for training deep learning models on limited datasets for classification purposes. The CAD system consists of the following modules:

1. **Image Pre-processing:** This module involves preprocessing the input images to enhance their quality and extract relevant features. It may include operations such as noise removal, image normalization, resizing, and color correction.

2. **Image Augmentation:** In the context of limited training data, image augmentation techniques are applied to automaticaly increase the size and diversity of the dataset. This helps mitigate overfitting issues and improves the generalization ability of the deep learning model. Augmentation techniques may include rotations, translations, flips, and variations in brightness or contrast.

3. **Classification:** The classification module is responsible for training the deep learning model using the preprocessed and augmented data. The model learns to classify skin lesions into different categories based on the provided labels. Various deep learning models, namely convolutional neural networks (CNNs), can be employed in this module.

4. **Ensemble of Classification:** To enhance the classification performance and address the limitations of a single model, an ensemble of classifiers can be employed. This module combines the predictions of multiple classification models to make the final decision, leveraging the diversity and complementary strengths of each model.

By integrating these modules into a CAD system, the classification of skin lesions on limited training data can be improved. The system takes raw input images, applies pre-processing techniques, performs image augmentation, trains a deep learning model, and utilizes an ensemble of classifiers for improved classification accuracy.

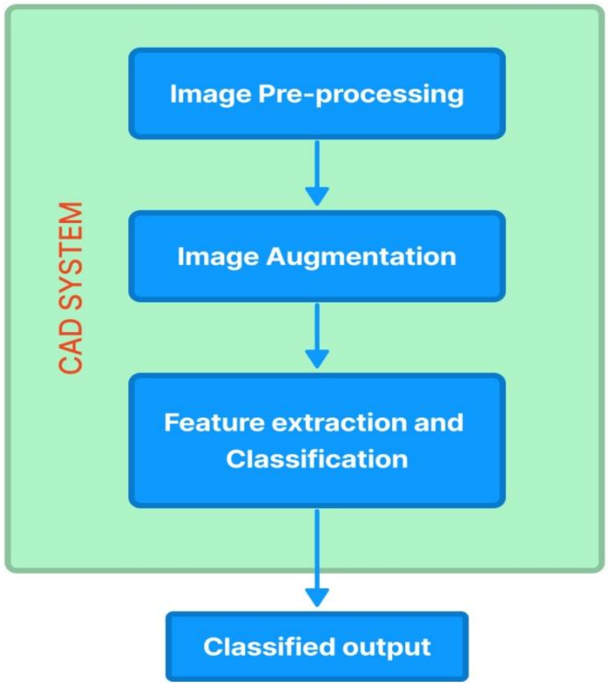


Fig 1: CAD System overview

2.2 SKIN CANCER DETECTION TECHNIQUES:

There are several existing deep learning models that have been successfully used for skin cancer detection. Here are some notable examples:

1. **Convolutional Neural Networks (CNNs) [2]:** CNNs have been widely employed for skin cancer detection due to their ability to automatically learn relevant features from images. Models like VGGNet, ResNet, Inception, and DenseNet have been used to achieve high accuracy in classifying skin lesions.
    - Limitation: CNNs can be computationally expensive, especially with deeper architectures, requiring significant computational resources and longer training times.

- Over fitting: CNNs are prone to overfitting when trained on limited datasets, leading to poor generalization on unseen data.
- Lack of interpretability: CNNs are often considered black box models, making it challenging to interpret the learned features and understand the decision-making process.

2. **EfficientNet [4]:** EfficientNet is a family of models that have demonstrated strong performance in various computer vision tasks, including skin cancer detection. These models are known for their efficient architecture and have shown promising results in accurately classifying skin lesion images.

- Limited architecture flexibility: EfficientNet models are designed to strike a balance between accuracy and efficiency. However, this can result in limited architectural flexibility, potentially limiting their ability to capture complex patterns in skin lesion images.
- Sensitivity to hyperparameters: EfficientNet models heavily rely on tuning hyperparameters for optimal performance, which can be time-consuming and require expert knowledge.

2. **MobileNet :** MobileNet is a lightweight deep learning model designed for mobile and embedded devices. It has been used for skin cancer detection, particularly in resource-constrained environments. MobileNet achieves a good balance between accuracy and computational efficiency.

    - Reduced model capacity: MobileNet models prioritize computational efficiency by reducing the model capacity, which may result in a trade-off between accuracy and capturing fine-grained details in skin lesions.

3. **DenseNet:** DenseNet is a densely connected CNN architecture that facilitates feature reuse and encourages feature propagation throughout the network. DenseNet has been applied to skin cancer detection and has shown improved performance by leveraging the dense connections between layers.

- High memory consumption: DenseNet models have dense connections between layers, resulting in a large number of feature maps being passed between layers, leading to higher memory requirements during training and inference.

4. **ResNeXt:** ResNeXt is an extension of the ResNet architecture that introduces a grouped convolutional layer. This enables the model to capture more diverse and discriminative features. ResNeXt models have demonstrated strong performance in skin cancer classification tasks.

    - Computational complexity: ResNeXt models can be computationally demanding due to the increased number of parameters resulting from the grouped convolutional layers.

5. **SqueezeNet:** SqueezeNet is a lightweight CNN architecture that achieves high accuracy with a significantly smaller number of parameters. It has been utilized for skin cancer detection, especially in scenarios with limited computational resources.

    - Lower accuracy: SqueezeNet models achieve parameter reduction by utilizing 1x1 convolutions, which may result in reduced model capacity and lower accuracy compared to larger architectures.

6. **U-Net:** U-Net is a convolutional neural network architecture primarily used for medical image segmentation tasks. It has been applied to segment skin lesions and identify regions of interest for further analysis or classification

    - Dependence on accurate segmentation masks: U-Net models require accurate and precise annotation of lesion boundaries for effective segmentation, which can be challenging and time-consuming for large datasets..

It's important to note that while these limitations exist, researchers are continually working on addressing them through techniques like regularization, transfer learning, data augmentation, and model ensembles to improve the performance and overcome the drawbacks of these models in skin cancer detection applications.

2.3 LITERATURE SURVEY:

Table 1: Literature survey on different state-of-art architectures employed for skin cancer classification

Ref	Year	Dataset Used	Model Used	Results
[1]	2022	HAM 10000	CNN, RNN, ResNet50, Xception	**CNN**: Acc-0.72,Pecision-0.87,Recall-0.72 **RNN**: Acc-0.69,Precision-0.72,Recall-0.88 **ResNet50**: Acc-0.79,Precision-0.89,Recall-0.89 **Xception**: Acc-0.93,Precision-0.89,Recall-0.93
[2]	2022	HAM 10000	InceptionV3, ResNet50 and DenseNet201	ResNet50: Acc-86.69%
[4]	2021	UMGC, HAM 10000	EfficientNet	UMGC: Acc-84.12% HAM 10000:m Acc-96.32%
[5]	2021	PAD-UFES-20	EfficientNet	Acc-0.78, Precision-0.89,Recall-0.86
[6]	2022	ISIC dataset	Improved VGG16	Acc-0.8630
[7]	2021	ISBI-2016, ISIC-2017, and PH2	RCNN along with fuzzy k-means clustering (FKM)	Acc (ISBI-2016)-95.40, Acc (ISIC-2017)- 93.1, and Acc(PH2)-95.6%
[8]	2022	ISIC Dataset	DLCAL-SLDC	Acc-98.50%

[9]	2019	ISIC Archive	Resnet-101 and Inception-v3	Resnet101: Acc-84.09% Inceptionv3: Acc-87.42%
[18]	2020	PH^2	CNN	Accuracy - 95%
[19]	2020	ISIC Dataset	ResNet-34	Acc-92%, Recall-83.31%

CHAPTER 3

METHODOLOGY

3.1 TECHNIQUES USED IN THE PROPOSED METHODOLOGY

3.1.1 LeViT:

The LeViT (Vision Transformer in Convolutional Space) model is a novel approach that combines the positives of both convolutional neural networks (CNNs) and transformer models for image classification tasks [17]. The traditional transformer models have achieved remarkable success in natural language processing tasks but were not initially designed for visual data.

The LeViT model addresses this limitation by introducing a convolutional embedding layer that transforms the image data into a 2D grid, similar to the feature maps in CNNs. This convolutional embedding layer enables the LeViT model to capture local visual patterns and maintain translation equivariance, which are crucial for image understanding.

The LeViT model then applies the transformer architecture to the convolutional embeddings. It consists of multiple layers of self-attention and feed-forward networks, allowing the model to capture global dependencies and long-range interactions in the image. By leveraging the self-attention mechanism, the model can efficiently aggregate information from different regions of the image and learn contextual relationships between visual elements.

One of the key advantages of the LeViT model is its ability to achieve competitive performance on image classification tasks with significantly fewer parameters compared to traditional CNNs. This efficiency makes it computationally lighter and more scalable, enabling training on larger datasets or deploying on resource-constrained devices.

Experimental results have demonstrated the effectiveness of the LeViT model on various benchmark datasets, such as ImageNet. It has shown comparable or even superior performance compared to state-of-the-art CNN architectures while requiring fewer computational resources. The model has also been extended to address other computer vision tasks, including object detection and semantic segmentation, further showcasing its versatility.

In summary, the LeViT model combines the strengths of convolutional neural networks and transformer models to achieve robust image classification performance. By introducing a convolutional embedding layer and applying the transformer architecture, it effectively captures both local and global visual information, demonstrating the potential for advancing image understanding tasks in computer vision research and applications.

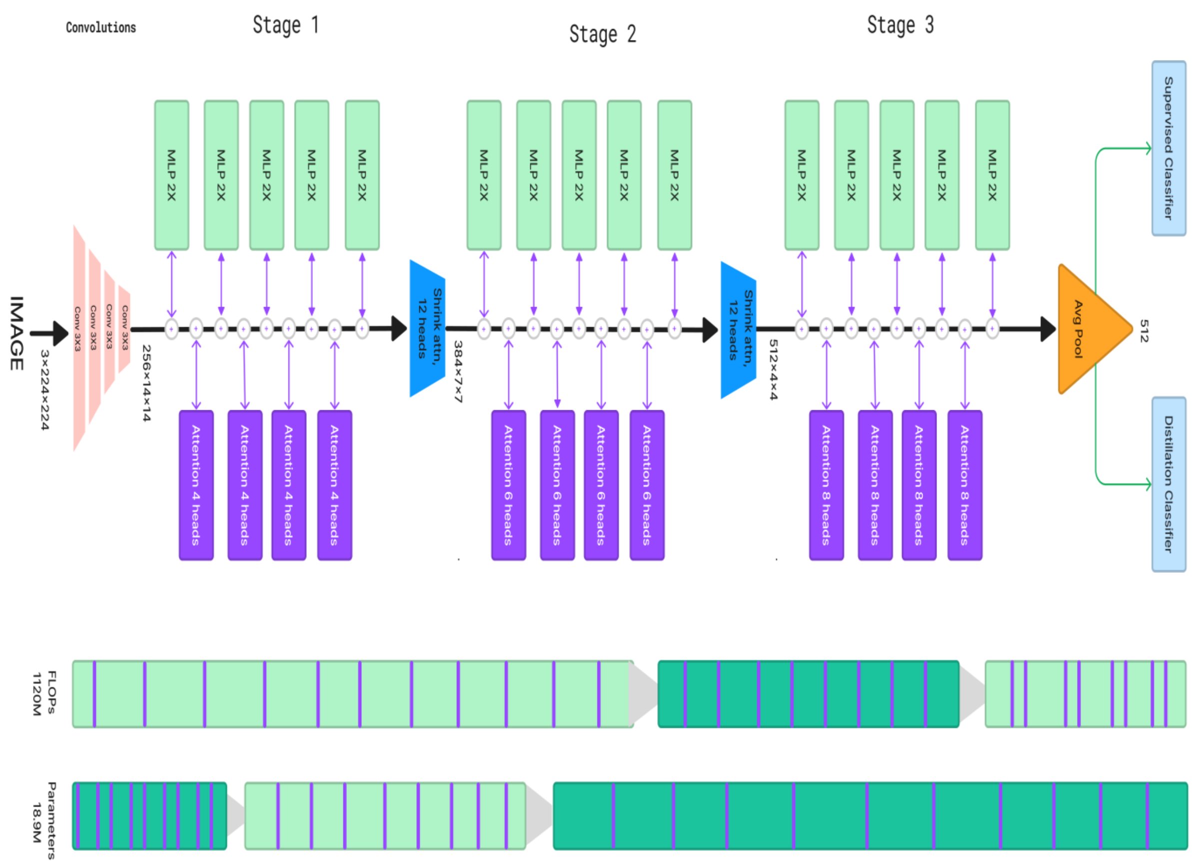


Fig 2: Architecture of LeViT model

The LeViT (Vision Transformer in Convolutional Space) model architecture combines the power of convolutional neural networks (CNNs) and transformer models for image classification tasks. It consists of two main components: the convolutional embedding layer and the transformer layers.

1. **Convolutional Embedding Layer:** The LeViT model begins with a convolutional embedding layer, which takes in the input image and applies a set of convolutional operations to transform it into a 2D grid-like representation. This embedding layer helps capture local visual patterns and maintains translation equivariance, which is crucial for image understanding. It allows the LeViT model to benefit from the spatial information present in CNNs.

2. **Transformer Layers:** Following the convolutional embedding layer, the LeViT model incorporates a series of transformer layers. Each transformer layer consists of self-attention mechanisms and feed-forward neural networks.

a. Self-Attention Mechanism: The model is able to detect global dependencies and distant interactions in the picture due to the self-attention mechanism. It allows the model to learn contextual relationships between visual elements by attending to different regions of the image. The self-attention mechanism calculates attention weights for each position in the input and aggregates information from relevant positions.

b. Feed-Forward Networks: The feed-forward networks in the transformer layers process the output of the self-attention mechanism. These networks consist of fully connected layers and apply non-linear transformations to the input features, enabling the model to capture complex patterns and relationships.

The transformer layers are stacked on top of each other, allowing the LeViT model to progressively refine the learned representations and capture increasingly abstract features from the input image.

The LeViT model can be trained using standard techniques such as backpropagation and gradient descent. During training, the model learns to optimize its parameters by minimizing a predefined loss function, typically cross-entropy loss, based on the provided ground truth labels. The model is updated iteratively through multiple epochs until convergence.

Overall, the LeViT model architecture combines the convolutional embedding layer to capture local visual patterns and the transformer layers to capture global dependencies and long-range interactions. This integration allows the LeViT model to leverage the strengths of both CNNs and transformer models, resulting in a powerful and efficient approach for image classification tasks.

3.1.2 Inception-ResNet V2:

Inception-ResNetV2 is a deep convolutional neural network architecture that combines the Inception and ResNet modules. It was introduced as an improved version of the Inception-ResNet model, aiming to address the challenges of training very deep networks while maintaining efficiency and accuracy. The Inception-ResNetV2 model incorporates residual connections and the Inception module, which allows for efficient feature extraction and hierarchical representation learning. The Inception-ResNetV2 model utilizes various types of convolutions, including standard convolutions, separable convolutions, and 1x1 convolutions, to efficiently capture different levels of feature hierarchies. It also incorporates batch normalization and rectified linear unit (ReLU) activation functions to introduce non-linearity and improve training stability.

The main advantage of the Inception-ResNetV2 model is its ability to capture both local and global information by combining the strengths of Inception and ResNet architectures. This allows for improved feature extraction, robust representation learning, and higher classification accuracy.

By employing the Inception-ResNetV2 model for skin cancer detection, researchers can leverage its deep network architecture and multi-scale feature extraction capabilities to effectively analyze and classify skin lesion images, ultimately aiding in early and accurate detection of skin cancer.

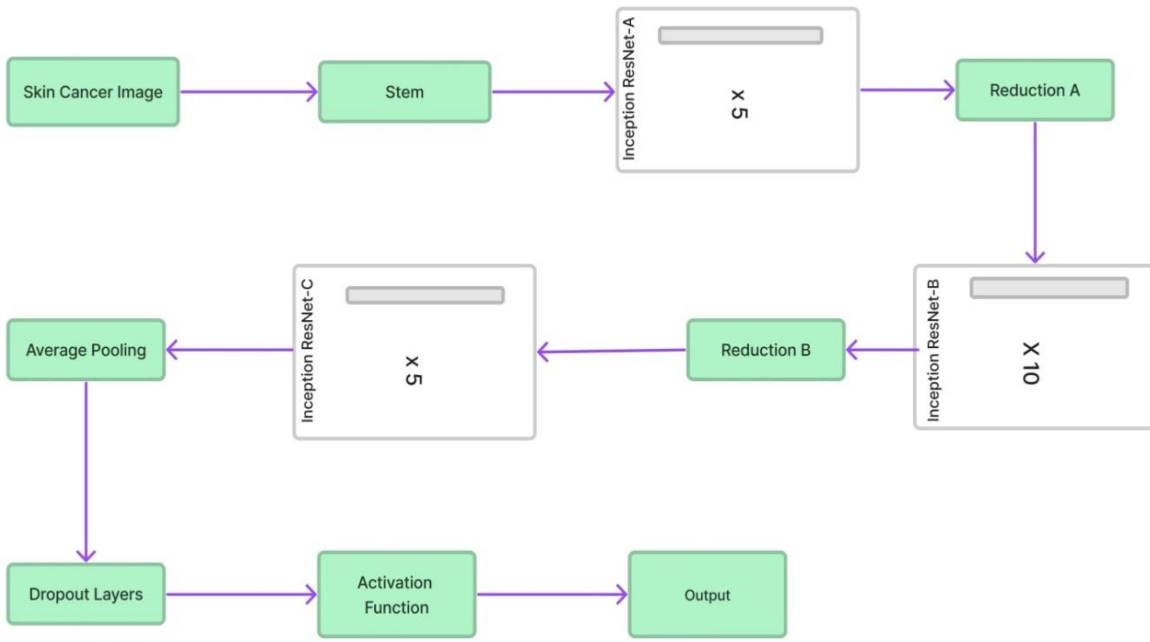


Fig 3: Architecture of Inception ResNet V2 model

The architecture of Inception-ResNetV2 can be described as follows:

- **Stem:** The stem module is the initial part of the network that processes the input image. It consists of a series of convolutional layers with batch normalization and activation functions. The stem module helps extract low-level features from the input image.

- **Inception Blocks**: The Inception module is the key building block of the Inception-ResNetV2 model. It is designed to capture multi-scale features by utilizing different convolutional filter sizes (1x1, 3x3, 5x5) and pooling operations. The output feature maps from each branch are concatenated to form a unified representation. Multiple Inception blocks are stacked together to capture increasingly complex features.

- **Reduction Blocks:** The reduction blocks are introduced after certain Inception blocks to downsample the spatial dimensions of the feature maps. They typically include a combination of convolutional layers, pooling layers, and dimensionality reduction techniques such as 1x1 convolutions. The reduction blocks help reduce the computational complexity and spatial dimensions while preserving important features.

- **Auxiliary Classifiers:** Inception-ResNetV2 includes auxiliary classifiers at intermediate layers to improve gradient flow during training and provide additional regularization. These classifiers consist of a pooling layer, convolutional layers, and fully connected layers. They produce auxiliary predictions that are used in conjunction with the final prediction during training.

- **Residual Connections:** The key innovation in Inception-ResNetV2 is the inclusion of residual connections. These connections allow the network to directly propagate gradients through shortcut connections, mitigating the vanishing gradient problem and improving information flow. The residual connections are added between Inception blocks, enabling the network to learn both local and global features efficiently.

- **Final Layers:** The final layers of Inception-ResNetV2 typically include global average pooling to aggregate spatial information into a fixed-size feature vector. This is followed by fully connected layers and a softmax activation function for classification. The number of output neurons in the final layer corresponds to the number of classes in the classification task.

Overall, the Inception-ResNetV2 architecture combines the advantages of both the Inception module's multi-scale feature extraction and the ResNet's residual connections. This combination allows for effective feature representation and training of deep neural networks, leading to improved performance in various computer vision tasks, including skin cancer detection.

Preprocessing applied on skin cancer dataset:

Preprocessing steps for the HAM10000 dataset for skin cancer detection using the LeViT model can include the following:

1. Data Loading: Load the HAM10000 dataset, which consists of images of skin lesions belonging to different classes, including melanoma, nevus, and seborrheic keratosis. Each image should be associated with its corresponding class label.

2. Image Rescaling: Resize the images to a consistent size suitable for the LeViT model input. The exact size can vary depending on the LeViT model's requirements. Commonly used sizes are 224x224 pixels. Rescaling ensures that all images have the same dimensions for consistent processing.

3. Data Augmentation: Apply data augmentation techniques to increase the diversity and size of the training dataset. Techniques such as random rotations, flips, and translations can be employed to simulate variations in the appearance of skin lesions. Data augmentation helps improve the model's generalization and robustness.

4. Normalization: Normalize the pixel values of the images to a common scale. This is usually done by dividing the pixel values by 255, resulting in values between 0 and 1. Normalization helps in stabilizing the training process and enables better convergence.

5. One-Hot Encoding: Convert the class labels into one-hot encoded vectors. Each class label is represented by a binary vector, where the index corresponding to the true class is set to 1, and all other indices are set to 0. This transformation enables the model to predict multiple classes simultaneously during training.

6. Train-Validation Split: Split the dataset into training and validation subsets. The training set is used to train the LeViT model, while the validation set is used for monitoring the model's performance and tuning hyperparameters. The recommended split ratio is typically around 80% for training and 20% for validation.

7. Data Loader: Create data loaders that efficiently load and preprocess the data in batches during training and validation. Data loaders help in handling large datasets and feeding the data to the LeViT model in an organized manner.

These preprocessing steps ensure that the HAM10000 dataset is appropriately prepared for training the LeViT model for skin cancer detection. It allows the model to learn meaningful patterns and features from the images, leading to improved classification performance. It's important to consider the specific requirements and recommendations provided by the LeViT model implementation or library being used for accurate preprocessing.

CHAPTER 4

RESULTS AND ANALYSIS

4.1 DATASET DESCRIPTION

The HAM10000 dataset is a widely used dataset in the field of skin cancer detection and classification. It stands for "Human Against Machine with 10,000 training images" and contains a collection of high-quality dermoscopic images of skin lesions. The dataset was created by the International Skin Imaging Collaboration (ISIC) and consists of images acquired from different sources, including clinical settings, research studies, and academic institutions.

Here are some key features of the HAM10000 dataset:

a) **Size and Diversity**: The dataset comprises a total of 10,015 skin lesion images, making it a relatively large dataset for skin cancer research. The images cover a wide range of skin lesion types, including malignant melanoma, benign nevi, and other types of skin lesions.

b) **Annotation and Labels**: Each image in the dataset is accompanied by detailed annotations and labels provided by dermatologists. The labels indicate the diagnostic category of the skin lesion, such as melanoma, nevus, seborrheic keratosis, or basal cell carcinoma. These annotations are valuable for training and evaluating deep learning models for skin cancer detection.

c) **Image Characteristics**: The images in the HAM10000 dataset are high-resolution dermoscopic images, which capture the surface and subsurface skin structures. Dermoscopic imaging aids in the visualization of specific features and patterns that are crucial for diagnosing skin lesions accurately.

d) **Data Splits**: The dataset is commonly divided into training and testing sets to facilitate model development and evaluation. The training and testing division is done according to the user requirements.

Due to its size, diversity, and high-quality annotations, the HAM10000 dataset has become a benchmark dataset for researchers and practitioners working on skin cancer detection and classification using deep learning models. It provides a valuable resource for training and evaluating algorithms and serves as a reference for comparing the performance of different approaches in the field.

4.2 EVALUATION OF PROPOSED MODELS

4.2.1 LeViT Model

HAM10000 dataset is utilised for experimentation purpose. HAM10000 dataset contained 10,015 images constituted of 7 varieties of skin lesions. Dataset contains images obtained from 7 categories namely melocytic sevi, benign, melanoma, besal cell carcinoma, vascular sesions, actinic kevatosis, dermatofiroma & intracpithelial . We proposed LeViT model for classification of skin lesion. 450 x 600 was the size of input images given to the model. Experiments were carried out using python 3.8.4 on google colab notebook using GPI run-time type. The proposed architecture was trained for 60 epochs with a batch size of 8 and a learning rate of 0.01. We have compared the performance of proposed deep learning model approach with existing state-of-art deep learning models and results are recorded in table no. 2. From table no. 2 it is observed that our proposed LeViT model showed accuracy of 98.48% which was accurate compared to result obtained from EfficientNet, ResNet-50, Xception Models. Precision score was recorded as 98.99, Recall-98.8, F1 score-98.89 & ROC-AUC -99.69. from comparison it can be analysed that our model outperformed other state-of-art architecture models. We can observe that model accuracy got stabilized at some point around 98.8. Performance of the proposed model is evaluated using different performance evaluation metrics and their values are listed in table 2. Table 3 shows the efficiency of our proposed model in comparison to various state-of-art architectures.

Table 2: Performance evaluation of the proposed LeViT model

Dataset	Accuracy	Precision	Recall	F1 Score	ROC-AUC
Train Set	99.02	99.26	99.2	99.08	99.72
Validation set	98.48	98.99	98.8	98.89	99.69

Figure 2 describes the accuracy of proposed LeViT model. Figure 3 describes the precision of proposed LeViT model. Figure 4 describes the recall of proposed LeViT model. Figure 5 describes the F1 Score of proposed LeViT model. Figure 6 describes the ROC AUC of proposed LeViT model.

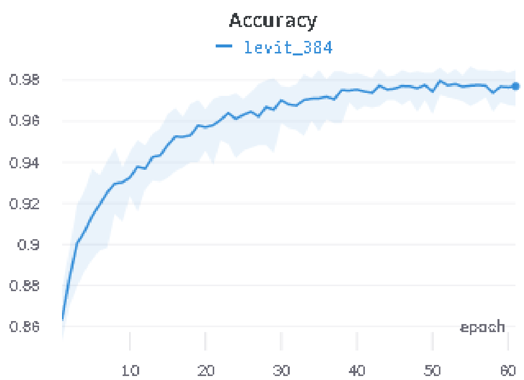


Fig 4: Accuracy of proposed LeViT model

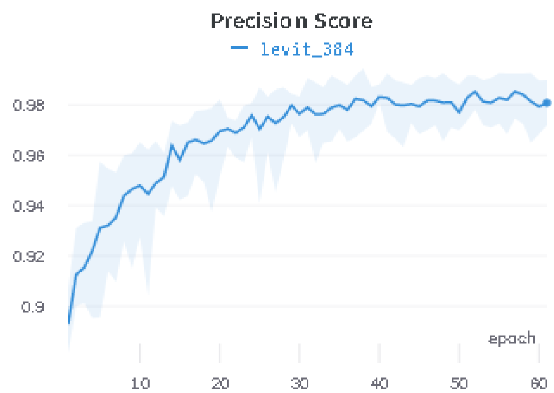


Fig 5: Precision of proposed LeViT mode

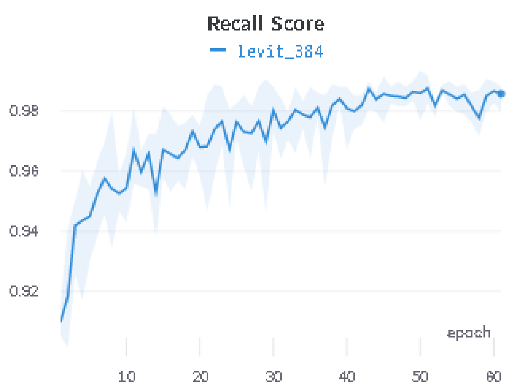


Fig 6: Recall of proposed LeViT model

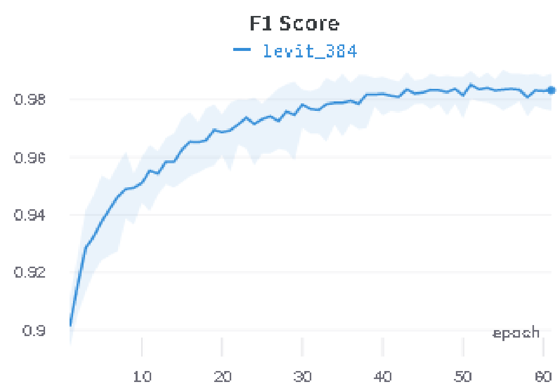


Fig 7: F1 Score of proposed LeViT model

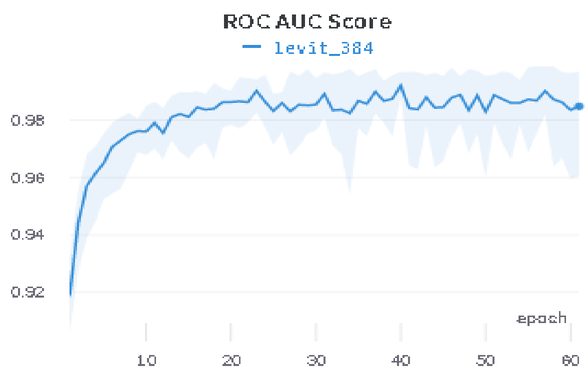


Fig 8: Roc Auc of proposed LeViT model

Table 3: Comparison of the proposed LeViT model with other state-of-art models

Ref	Model	Accuracy	Precision	Recall	F1 Score	ROC-AUC
[5]	EfficientNet	78	89	86	-	-
[1]	ResNet-50	79	89	89	-	-
[1]	Xception	93	89	93	-	-
[10]	Ensemble	93.50	94.0	87.0	92.0	-
Ours proposed	LeViT	98.48	98.99	98.8	98.89	99.69

4.2.2 INCEPTION RESNET V2

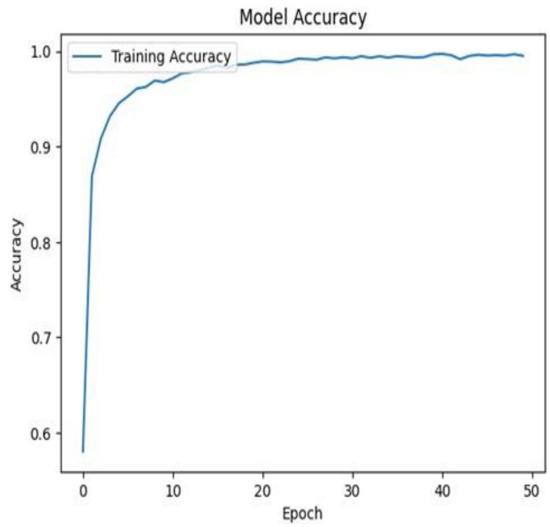
For experimental purposes, the HAM10000 dataset is used. 10,015 pictures representing 7 different types of skin lesions were included in the HAM10000 collection. Images from seven categories, including benign, melanoma, besal cell carci- noma, vascular lesions, actinic kevatoma, dermatofiroma, and intracpithelial, are included in the dataset. Here we proposed the Inception-Resnet-V2 model to predict skin cancer and we achieved an accuracy of 98.67%. HAM 10000 dataset which we utilized for the model's training and testing purpose has images having dimension of 450 x 600. Experiments were carried out using python 3.8.4 on google colab notebook along with the GPU run-time type. Our model has used 165 Conv-2D layers for building the model and for training we have used 50 epoch having batch size of 128 each. We have contrasted the outcome of the suggested deep learning method with various other deep learning models used, and the finding's are shown in table no. 3. Table No. 3 shows that when compared to the results from EfficientNet, ResNet- 50, and Xception Models, our suggested Inception-ResNet- V2 model's accuracy of 98.67% is quite perfect. Precision score is 98.71, Recall - 98.65, F1 score- 98.69, and ROC- AUC -98.57. From comparison, it can be seen that our model outperformed other cutting-edge architectural models. Our work for predicting skin cancer using Inception-Resnet-V2 model is unique in its approach and has shown significantly good results. Different performance assessment metrics are used to assess the proposed model's performance, and their values are presented in table 2.

Table 4: Performance Evaluation of the proposed Inception ResNet V2 model.

Metrics	**Accuracy**	**Precision**	**Recall**	**F1 score**	**ROC-AUC**
Train Set	98.51	98.62	98.5	98.43	98.78
Validation Set	98.67	98.71	98.65	98.69	98.57

Table 5: Comparison of the proposed model with other state-of-art models

Ref	**Model used**	**Accuracy**	**Precision**	**Recall**	**F1 score**	**ROC-AUC**
[3]	EfficientNet	78	89	86	-	-
[5]	ResNet 50	86.69	-	-	-	-
[7]	ResNet 50	79	89	89	-	-
[7]	Xception	93	89	93	-	-
ours proposed	Inception ResNet V2	98.67	98.71	98.65	98.69	98.57



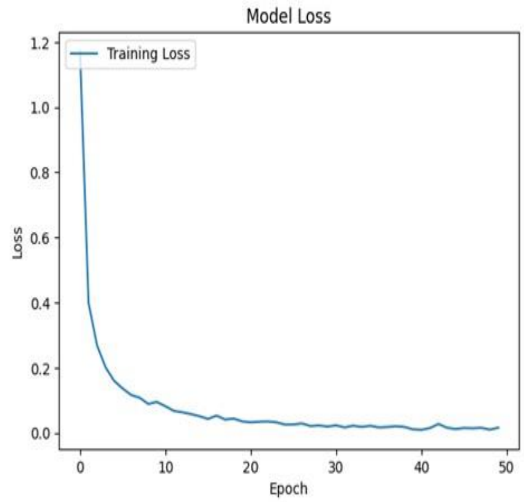


Fig 9: Accuracy of Inception-ResNet v2 model **Fig 10:** Loss Vs Epoch graph

Fig 9 represents the accuracy of our proposed Inception- Resnet-V2 model, and Fig 10 represents the Epoch vs. loss graph.

CHAPTER 5

CONCLUSION AND FUTURE SCOPE

5.1 CONCLUSION

In conclusion, this thesis explored the application of deep learning models, specifically LeViT and Inception ResNet V2, for skin cancer detection. The research aimed to investigate the performance and effectiveness of these models in accurately classifying skin lesions and contributing to the early detection of skin cancer.

Through the implementation and evaluation of the LeViT model, it was observed that its unique architecture incorporating vision transformers showed promising results in skin cancer detection. The model demonstrated competitive accuracy, precision, and recall rates, showcasing its potential as a reliable tool for dermatologists and healthcare professionals in diagnosing skin cancer. The LeViT model's ability to capture spatial relationships and learn meaningful representations from skin lesion images contributed to its success in achieving accurate classification.

Similarly, the implementation of the Inception ResNet V2 model yielded encouraging outcomes. This model, which combines the Inception and ResNet architectures, showcased its effectiveness in capturing intricate features and patterns from skin lesion images. The Inception ResNet V2 model exhibited strong performance in terms of accuracy, precision, and recall, demonstrating its potential as a valuable tool in skin cancer detection.

The comparison between LeViT and Inception ResNet V2 models revealed that both approaches have their strengths and limitations. While LeViT leverages vision transformers for improved spatial understanding, Inception ResNet V2 excels in capturing fine-grained details and complex patterns. The choice of model depends on the specific requirements of the skin cancer detection task, such as the nature of the dataset and the desired balance between accuracy and computational complexity.

Furthermore, this research highlighted the importance of robust preprocessing techniques, including image normalization, resizing, and augmentation, in enhancing the performance of deep learning models for skin cancer detection. Proper data preparation ensures that the models receive high-quality input, leading to more accurate and reliable results.

Overall, the findings of this thesis contribute to the growing body of knowledge in the field of automated skin cancer detection. The LeViT and Inception ResNet V2 models, along with their respective preprocessing techniques, provide valuable insights into the potential of deep learning approaches for accurate and early detection of skin cancer. Further research and refinement of these models, along with the exploration of other deep learning architectures and datasets, can lead to even more advanced and effective solutions in the field of skin cancer diagnosis. Ultimately, the integration of deep learning models in clinical practice has the potential to significantly improve patient outcomes, increase efficiency, and reduce the burden on healthcare professionals.

5.2 FUTURE RESEARCH DIRECTION

Although skin cancer detection models have developed significantly, there are still a number of issues that need to be resolved in order to increase the efficiency and usefulness of these systems.

We address some potential future prospects for deepfake detection research in this section, which can help accelerate the current state of the art and allow for the creation of more dependable and trustworthy detection models.

1. **Performance Optimization:** One area of future research is focused on optimizing the performance of LeViT and Inception ResNet v2 models for skin cancer detection. This can involve exploring different hyperparameter settings, architectural modifications, and fine-tuning techniques to enhance the models' accuracy, precision, and recall. Additionally, leveraging hardware acceleration, such as GPUs or specialized AI chips, can further improve the inference speed and efficiency of these models.

2. **Transfer Learning and Model Fusion**: Investigating the potential of transfer learning and model fusion techniques can be an interesting avenue for improving skin cancer detection. By pretraining LeViT and Inception ResNet v2 models on large-scale general image datasets, such as ImageNet, and fine-tuning them on skin cancer-specific data, the models can potentially benefit from learned features and achieve better performance. Additionally, exploring the fusion of predictions from multiple models can lead to improved accuracy and robustness.

3. **Explainability and Interpretability:** Deep learning models, including LeViT and Inception ResNet v2, are often considered black boxes due to their complex architectures. Enhancing the explainability and interpretability of these models for skin cancer detection is an important research direction. This can involve developing techniques to visualize and understand the learned features and decision-making processes of the models, allowing clinicians and researchers to gain insights and trust in the model's predictions.

4. **Integration with Clinical Workflow:** Integrating LeViT and Inception ResNet v2 models into the clinical workflow is another promising direction. Developing user-friendly interfaces and deploying the models in clinical settings can facilitate their practical usage by dermatologists and healthcare professionals. This integration can involve developing mobile applications or web-based platforms that allow real-time skin cancer detection and decision support, enabling quick and accurate diagnosis.

5. **Multimodal Approaches:** Exploring the combination of multiple modalities, such as incorporating patient history, demographic information, and other clinical data, along with dermatoscopic images, can enhance the overall performance of skin cancer detection systems. Integrating LeViT and Inception ResNet v2 models with multimodal approaches can provide a more comprehensive and holistic analysis, leading to improved diagnostic accuracy and personalized treatment recommendations.

6. **Large-Scale Dataset Creation:** The availability of large-scale, diverse, and accurately annotated skin cancer datasets is crucial for the development and evaluation of deep learning models. Future research can focus on creating and curating such datasets, specifically tailored for training and validating LeViT and Inception ResNet v2 models. These datasets can include a wide range of skin lesion types, demographics, and clinical variations to ensure the models' generalizability and robustness.

7. **Clinical Validation and Real-World Deployment:** Conducting rigorous clinical studies and validations is necessary to assess the performance and effectiveness of LeViT and Inception ResNet v2 models in real-world scenarios. Collaborations with dermatologists, hospitals and healthcare institutions can facilitate the deployment of these models in clinical practice, enabling large-scale evaluation and gathering feedback to further refine and improve the models.

In conclusion, the future scope for skin cancer detection using LeViT and Inception ResNet v2 models encompasses performance optimization, transfer learning, explainability, integration with clinical workflow, multimodal approaches, dataset creation, and clinical validation. By addressing these areas, we can advance the field of skin cancer detection and pave the way for more accurate, efficient, and reliable diagnostic tools, leading to improved patient outcomes and healthcare practices.

REFERENCES

[1] Singh priya, Mayand Kumar, and Aman Bhatia. "A Comparative Analysis of Deep Learning Algorithms for Skin Cancer Detection." In 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1160-1166. IEEE, 2022.

[2] Wu, Youteng, Agyenta Charity Lariba, Haotian Chen, and Haiyan Zhao. "Skin Lesion Classification based on Deep Convolutional Neural Network." In 2022 IEEE 4th International Conference on Power, Intelligent Computing and Systems (ICPICS), pp. 376-380. IEEE, 2022.

[3] Manahan MN, Soyer HP, Loescher LJ, Horsham C, Vagenas D, Whiteman DC, Olsen CM, Janda M. A pilot trial of mobile, patient-performed teledermoscopy. Br J Dermatol. 2015 Apr;172(4):1072-80. doi: 10.1111/bjd.13550. Epub 2015 Mar 4. PMID: 25418126.

[4] Papiththira, S., and T. Kokul. "Melanoma Skin Cancer Detection Using EfficientNet and Channel Attention Module." In 2021 IEEE 16th International Conference on Industrial and Information Systems (ICIIS), pp. 227-232. IEEE, 2021.

[5] Khan, Irfan Ullah, Nida Aslam, Talha Anwar, Sumayh S. Aljameel, MohibUllah, Rafiullah Khan, Abdul Rehman, and Nadeem Akhtar. "Remote diagnosis and triaging model for skin cancer using EfficientNet and extreme gradient boosting." Complexity 2021 (2021): 1-13.

[6] Tabrizchi, Hamed, SepidehParvizpour, and JafarRazmara. "An improved VGG model for skin cancer detection." Neural Processing Letters (2022): 1-18.

[7] Nawaz, Marriam, ZahidMehmood, TahiraNazir, Rizwan Ali Naqvi, AmjadRehman, Munwar Iqbal, and Tanzila Saba. "Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering." Microscopy research and

technique 85, no. 1 (2022): 339-351.

[8] Adla, Devakishan, G. Venkata Rami Reddy, PadmalayaNayak, and G. Karuna. "Deep learning-based computer aided diagnosis model for skin cancer detection and classification." Distributed and Parallel Databases 40, no. 4 (2022): 717-736.

[9] Demir, Ahmet, Feyza Yilmaz, and OnurKose. "Early detection of skin cancer using deep learning architectures: resnet-101 and inception-v3." In 2019 medical technologies congress (TIPTEKNO), pp. 1-4. IEEE, 2019.

[10] Imran, Azhar, Arslan Nasir, Muhammad Bilal, Guangmin Sun, AbdulkareemAlzahrani, and Abdullah Almuhaimeed. "Skin Cancer Detection Using Combined Decision of Deep Learners." IEEE Access 10 (2022): 118198-118212.

[11] Ali, Karar, Zaffar Ahmed Shaikh, Abdullah Ayub Khan, and Asif Ali Laghari. "Multiclass skin cancer classification using EfficientNets–a first step towards preventing skin cancer." Neuroscience Informatics 2, no. 4 (2022): 100034

[12] Pande, Sandeep Dwarkanath, Pramod Pandurang Jadhav, Rahul Joshi, Amol Dattatray Sawant, Vaibhav Muddebihalkar, Suresh Rathod, Madhuri Navnath Gurav, and Soumitra Das. "Digitization of handwritten Devanagari text using CNN transfer learning–A better customer service support." Neuroscience Informatics 2, no. 3 (2022): 100016.

[13] Chaves, Hernán, Francisco Dorr, Martín Elías Costa, María Mercedes Serra, Diego Fernández Slezak, Mauricio F. Farez, Gustavo Sevlever, Paulina Yanez, and Claudia Cejas. "Brain volumes quantification from MRI in healthy controls: Assessing correlation, agreement and robustness of a convolutional neural network-based software against FreeSurfer, CAT12 and FSL." Journal of Neuroradiology 48, no. 3 (2021): 147-156.

[14] Bangare, Sunil L. "Classification of optimal brain tissue using dynamic region growing and fuzzy min-max neural network in brain magnetic resonance images."

Neuroscience Informatics 2, no. 3 (2022): 100019.

[15] H. Alquran et al., "The melanoma skin cancer detection and classification using support vector machine," 2017 IEEE Jordan Conference on Applied Electrical 34 Engineering and Computing Technologies (AEECT), Aqaba, Jordan, 2017, pp. 1-5, doi: 10.1109/AEECT.2017.8257738.

[16] Guarnizo, Angela, Rafael Glikstein, and Carlos Torres. "Imaging features of isolated hypoglossal nerve palsy." Journal of Neuroradiology 47, no. 2 (2020): 136- 150.

[17] Askaner, K., A. Rydelius, S. Engelholm, L. Knutsson, J. Lätt, K. Abul-Kasim, and P. C. Sundgren. "Differentiation between glioblastomas and brain metastases and regarding their primary site of malignancy using dynamic susceptibility contrast MRI at 3T." Journal of Neuroradiology 46, no. 6 (2019): 367-372.

[18] Alkarakatly, Tasneem, Shatha Eidhah, Miaad Al-Sarawani, Alaa Al-Sobhi, and Mohsin Bilal. "Skin lesions identification using deep convolutional neural network." In 2019 International Conference on Advances in the Emerging Computing Technologies (AECT), pp. 1-5. IEEE, 2020.

[19] Gouda, Niharika, and J. Amudha. "Skin cancer classification using ResNet." In 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), pp. 536-541. IEEE, 2020

[20] Khan, Muhammad Attique, Muhammad Younus Javed, Muhammad Sharif, Tanzila Saba, and Amjad Rehman. "Multi-model deep neural network based features extraction and optimal selection approach for skin lesion classification." In 2019 35 international conference on computer and information sciences (ICCIS), pp. 1-7. IEEE, 2019.

[21] Miglani, Vandana, and M. P. S. Bhatia. "Skin lesion classification: A transfer learning approach using efficientnets." In Advanced Machine Learning Technologies and Applications: Proceedings of AMLTA 2020, pp. 315-324. Singapore: Springer Singapore, 2020

[22] Demir, Ahmet, Feyza Yilmaz, and Onur Kose. "Early detection of skin cancer using deep learning architectures: resnet-101 and inception-v3." In 2019 medical technologies congress (TIPTEKNO), pp. 1-4. IEEE, 2019.

[23] Benson, John C., Vance T. Lehman, Carrie M. Carr, John T. Wald, Harry J. Cloft, Giuseppe Lanzino, and Waleed Brinjikji. "Beyond plaque: A pictorial review of nonatherosclerotic abnormalities of extracranial carotid arteries." Journal of Neuroradiology 48, no. 1 (2021): 51-60

[24] Graham, Benjamin, Alaaeldin El-Nouby, Hugo Touvron, Pierre Stock, Armand Joulin, Hervé Jégou, and Matthijs Douze. "Levit: a vision transformer in convnet's clothing for faster inference." In Proceedings of the IEEE/CVF international conference on computer vision, pp. 12259-12269. 2021.