22AIE201
FUNDAMENTALS OF AI

Project Report on

Skin Cancer Detection: A Comparative Analysis of LSTM, RNN and CNN

Submitted by

Team No.: B11

Team Members:

JAINITHISSH S	(CB.SC.U4AIE23129)
NITHESHKUMMAR C	(CB.SC.U4AIE23155)
ABHAY ROHIT MURUGAIAH	(CB.SC.U4AIE23173)
GANDHARAPU MANOBHIRAM	(CB.SC.U4AIE23125)

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY
IN
CSE(AI)

Amrita School of Artificial Intelligence

AMRITA VISHWA VIDYAPEETHAM

COIMBATORE - 641 112 (INDIA)

November 2024

AMRITA SCHOOL OF ARTIFICIAL INTELLIGENCE

AMRITA VISHWA VIDYAPEETHAM COIMBATORE – 641 112 (INDIA)

BONAFIDE CERTIFICATE

This is to certify that the report of the project entitled **"Skin Cancer Detection: A Comparative Analysis of LSTM, RNN and CNN"** submitted by **Team – B11**, for the award of the Degree of Bachelor of Technology in the "CSE(AI)" is a bonafide record of the work carried out by us, under our guidance and supervision at Amrita School of Artificial Intelligence, Coimbatore.

Date of Submission: 16/11/2024

Dr. Abhishek S
Project Supervisor

Dr. K.P.Soman
Professor and Head CEN

AMRITA SCHOOL OF ARTIFICIAL INTELLIGENCE

AMRITA VISHWA VIDYAPEETHAM COIMBATORE – 641 112 (INDIA)

DECLARATION

We hereby declare that this project entitled **"Skin Cancer Detection: A Comparative Analysis of LSTM, RNN and CNN"** is the record of the original work done by us under the guidance of Dr. Abhishek S, Centre for Computational Engineering and Networking, Amrita School of Artificial Intelligence, Coimbatore. To the best of our knowledge this work has not formed the basis for the award of any degree/diploma/ associate ship/fellowship/or a similar award to any candidate in any University.

Place: Coimbatore

Date: 16.11.2024 **Signature of Faculty**

Names	Roll Numbers
JAINITHISSH S	(CB.SC.U4AIE23129)
NITHESHKUMMAR C	(CB.SC.U4AIE23155)
ABHAY ROHIT MURUGAIAH	(CB.SC.U4AIE23173)
GANDHARAPU MANOBHIRAM	(CB.SC.U4AIE23125)

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

This project explores the use of Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTMs) for the task of skin cancer detection. The goal is to build a robust deep learning model that can accurately classify skin lesions as either benign or malignant.

The dataset used in this project consists of dermoscopic images of skin lesions, obtained from the publicly available Skin Cancer: Malignant vs. Benign dataset. The images are preprocessed and fed into the various neural network architectures, including a CNN, an RNN, and an LSTM model.

The CNN model demonstrates strong performance, with high accuracy on the validation and test sets. The RNN and LSTM models leverage the sequential nature of the image data to capture additional features, potentially improving the overall classification accuracy.

This project highlights the potential of deep learning techniques specifically for the early detection of skin cancer. The insights gained from this work could contribute to the development of automated diagnostic tools, aiding clinicians in making more informed decisions and improving patient outcomes.

1. INTRODUCTION

Skin cancer is a major public health concern, with early detection being crucial for effective treatment and improved patient outcomes. Traditional diagnostic methods, while valuable, can be time-consuming and subjective, leading to the need for more automated and accurate approaches. This project aims to leverage the power of deep learning techniques, specifically Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTMs), to build a robust model for the classification of skin lesions as either benign or malignant.

The project will utilize the publicly available Skin Cancer: Malignant vs. Benign dataset, which consists of dermoscopic images of skin lesions. By training and evaluating the performance of the various neural network architectures, this work seeks to identify the most effective deep learning approach for skin cancer detection. The insights gained from this project could contribute to the development of automated diagnostic tools, assisting clinicians in making more informed decisions and ultimately enhancing patient care. The exploration of interpretable models and potential future expansions of the dataset and methodology further underscore the project's broader implications for the field of medical imaging and computer-aided diagnosis.

1.1 Key Features of the Skin Cancer Detection Model

Adaptable Model Architecture: The project will explore the use of different deep learning models, including CNNs, RNNs, and LSTMs, to determine the most effective architecture for skin lesion classification. This flexibility allows for the optimization of the model based on the unique characteristics of the dataset and the project's goals.

Interpretable Model Insights: In addition to the classification performance, the project will investigate techniques to improve the interpretability of the deep learning models. This could involve visualizing the salient features learned by the models or providing explanations for the classification decisions, ultimately enhancing the trust and transparency of the system.

Robust Data Preprocessing: The preprocessing of the dermoscopic images will be a crucial step, ensuring that the data is properly formatted, normalized, and augmented to improve the model's generalization and robustness to variations in the input.

Related Works Convolutional Neural Networks (CNNs) for Medical Imaging: CNNs have demonstrated remarkable success in various medical imaging tasks, including the classification of skin lesions. Several studies have explored the application of CNNs for skin cancer detection, showcasing their ability to learn relevant visual features from dermoscopic images. These works provide a strong foundation for the CNN-based approach in this project.

Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTMs) in Medical Diagnosis: RNNs and LSTMs have shown promising results in modeling the sequential and temporal aspects of medical data, such as time-series patient records and multimodal data. The potential of these architectures to capture contextual information and long-term dependencies could be beneficial for the skin cancer detection task.

Interpretable Deep Learning in Healthcare: Efforts have been made to improve the interpretability of deep learning models in healthcare, providing insights into the decision-making process and enhancing the trust of clinicians and patients. Techniques like saliency maps, layer visualizations, and attention mechanisms can be explored to understand the key features learned by the models.

1.2 Related Work

Convolutional Neural Networks (CNNs) for Medical Imaging: CNNs have demonstrated remarkable success in various medical imaging tasks, including the classification of skin lesions. Several studies have explored the application of CNNs for skin cancer detection, showcasing their ability to learn relevant visual features from dermoscopic images. These works provide a strong foundation for the CNN-based approach in this project.

Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTMs) in Medical Diagnosis: RNNs and LSTMs have shown promising results in modeling the sequential and temporal aspects of medical data, such as time-series patient records and multimodal data. The potential of these architectures to capture contextual information and long-term dependencies could be beneficial for the skin cancer detection task.

Interpretable Deep Learning in Healthcare: Efforts have been made to improve the interpretability of deep learning models in healthcare, providing insights into the decision-making process and enhancing the trust of clinicians and patients. Techniques like saliency maps, layer visualizations, and attention mechanisms can be explored to understand the key features learned by the models.

2. PROPOSED APPROACH

The proposed method for skin cancer detection employs a systematic and robust architecture, integrating advanced deep learning techniques to accurately classify skin lesions into benign and malignant categories. This approach combines Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Long Short-Term Memory networks (LSTM) to enhance detection accuracy and facilitate real-time analysis.

2.1. Data Acquisition and Preprocessing

The dataset comprises 3,300 images, evenly split between benign and malignant categories. Images are preprocessed, involving normalization of pixel values, resizing to standardized dimensions (224x224 pixels), and augmentation techniques to enhance model robustness. This prepares the dataset for efficient training of the neural network models.

2.2. CNN Model Deployment

2.2. CNN Model Deployment

A Convolutional Neural Network (CNN) is constructed with multiple convolutional and max-pooling layers for hierarchical feature learning. The architecture consists of:

- **Convolutional Layers**: Applied to extract spatial features from the input images. Each layer uses the **ReLU (Rectified Linear Unit)** activation function to introduce non-linearity and enhance feature learning.
- **Max-Pooling Layers**: Added after convolutional layers to reduce spatial dimensions and retain dominant features, facilitating computational efficiency.
- **Flatten Layer**: Converts the 2D feature maps into a 1D vector, preparing the data for the dense layers.
- **Fully Connected (Dense) Layers**: Includes one or more dense layers for classification. The final dense layer uses a **softmax** activation function to output probabilities for binary classification (Benign or Malignant).

The CNN is trained on the prepared dataset, enabling it to detect essential patterns indicative of skin cancer. Key metrics are monitored throughout training to ensure proper learning and adaptation.

Model: "sequential_1"

Layer (type)	Output Shape	Param #
conv2d_6 (Conv2D)	(None, 222, 222, 32)	896
max_pooling2d_6 (MaxPooling2D)	(None, 111, 111, 32)	0
conv2d_7 (Conv2D)	(None, 109, 109, 32)	9,248
max_pooling2d_7 (MaxPooling2D)	(None, 54, 54, 32)	0
conv2d_8 (Conv2D)	(None, 52, 52, 16)	4,624
max_pooling2d_8 (MaxPooling2D)	(None, 26, 26, 16)	0
conv2d_9 (Conv2D)	(None, 24, 24, 16)	2,320
max_pooling2d_9 (MaxPooling2D)	(None, 12, 12, 16)	0
conv2d_10 (Conv2D)	(None, 10, 10, 16)	2,320
max_pooling2d_10 (MaxPooling2D)	(None, 5, 5, 16)	0
conv2d_11 (Conv2D)	(None, 3, 3, 16)	2,320
max_pooling2d_11 (MaxPooling2D)	(None, 1, 1, 16)	0
flatten_1 (Flatten)	(None, 16)	0
dense_4 (Dense)	(None, 50)	850
dense_5 (Dense)	(None, 100)	5,100
dense_6 (Dense)	(None, 200)	20,200
dense_7 (Dense)	(None, 2)	402

Total params: 48,280 (188.59 KB)
Trainable params: 48,280 (188.59 KB)
Non-trainable params: 0 (0.00 B)

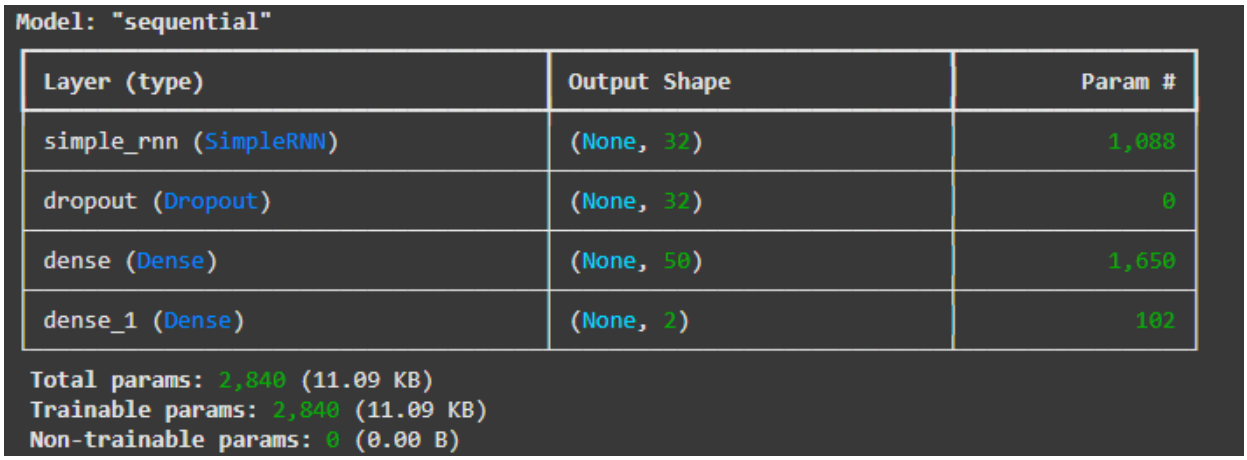
Fig 2.2. CNN Model Summary

2.3. Sequential Modeling with RNN and LSTM

The features extracted via the CNN are fed into a Recurrent Neural Network (RNN), specifically an LSTM model, to capture temporal correlations within the feature map sequences. The architecture consists of:

- **LSTM Layers**: Added to process sequential feature maps and extract temporal patterns. The LSTM layers use a **tanh** activation function for cell state updates and **sigmoid** activation functions for gate mechanisms (input, forget, and output gates).
- **Dropout Layers**: Included after LSTM layers to reduce overfitting by randomly disabling some neurons during training.
- **Fully Connected (Dense) Layer**: Used as the output layer with a **softmax** activation function to provide the final classification results (Benign or Malignant).

This step enhances the model's ability to understand sequential complexities, improving accuracy in classifying the images.



Model: "sequential"

Layer (type)	Output Shape	Param #
simple_rnn (SimpleRNN)	(None, 32)	1,088
dropout (Dropout)	(None, 32)	0
dense (Dense)	(None, 50)	1,650
dense_1 (Dense)	(None, 2)	102

Total params: 2,840 (11.09 KB)
Trainable params: 2,840 (11.09 KB)
Non-trainable params: 0 (0.00 B)

Fig 2.3. RNN Model Summary

```
[45] model.summary()
```

Model: "sequential_5"

Layer (type)	Output Shape	Param #
lstm_10 (LSTM)	(None, 7, 64)	3,686,656
batch_normalization_10 (BatchNormalization)	(None, 7, 64)	256
dropout_10 (Dropout)	(None, 7, 64)	0
lstm_11 (LSTM)	(None, 32)	12,416
batch_normalization_11 (BatchNormalization)	(None, 32)	128
dropout_11 (Dropout)	(None, 32)	0
dense_10 (Dense)	(None, 16)	528
dense_11 (Dense)	(None, 2)	34

Total params: 3,700,018 (14.11 MB)
Trainable params: 3,699,826 (14.11 MB)
Non-trainable params: 192 (768.00 B)

Fig 2.3. LSTM Model Summary

2.4. Model Evaluation and Refinement

The models (CNN, LSTM, and RNN) were evaluated using validation and test datasets, with key metrics such as accuracy and loss tracked throughout the process. Several refinements were made to enhance performance, including adjusting learning rates to achieve optimal convergence, adding dropout layers to mitigate overfitting, and experimenting with various architectures. For instance, the CNN structure was modified by varying the number of filters in convolutional layers, while the LSTM model was fine-tuned by altering the number of units in its layers. Additionally, early stopping was implemented during training to ensure the best-performing model was saved, preventing overfitting and improving the overall classification performance.

2.5. Visualization of Results

The project incorporated a range of visualizations to present the model's performance and interpretability. Training and validation accuracy curves, along with loss plots, were generated to illustrate the model's learning progression over epochs. Sample predictions for benign and malignant lesions were displayed to showcase the model’s classification accuracy. Feature maps from convolutional layers were visualized to interpret the specific patterns identified by the CNN,

while activation heatmaps highlighted critical regions in the input images that the model deemed significant for classification. These visualizations offered valuable insights into the model's learning behavior and its decision-making process.

3. EXPERIMENTS & RESULTS

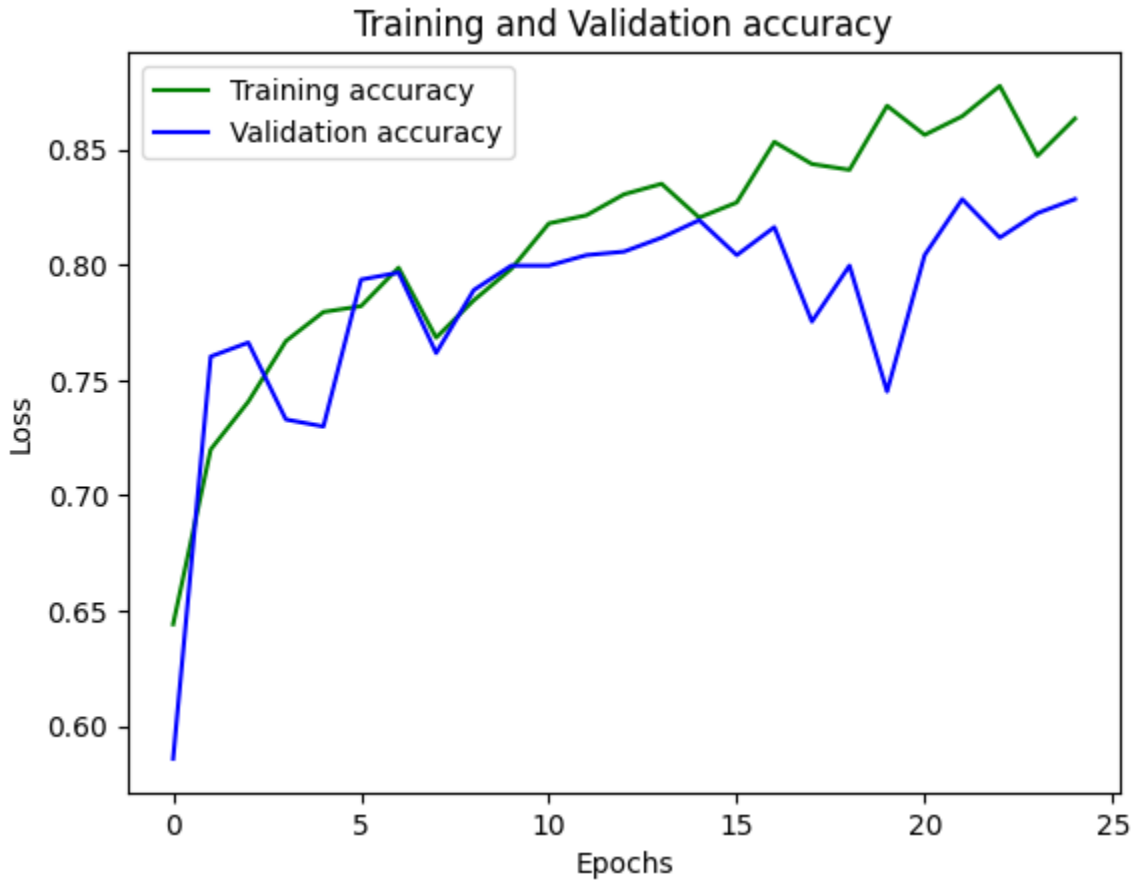


Fig 3.1. CNN Model Training and Validation Accuracy

The accuracy plot demonstrates the model's learning progression over 25 epochs. The training accuracy (green line) shows steady improvement, starting from around 65% and reaching approximately 85% by the end. The validation accuracy (blue line) follows a similar upward trend but with more fluctuation, particularly after epoch 15. The gap between training and validation accuracy gradually widens in later epochs, suggesting

some overfitting. However, both metrics show overall improvement, with training accuracy peaking around 87% and validation accuracy stabilizing around 83%, indicating reasonably good model performance on the skin cancer classification task.

Fig 3.2. CNN Model Training and Validation Loss

The loss plot reveals the model's convergence behavior during training. Both training loss (green line) and validation loss (blue line) start high at around 0.7 and initially decrease rapidly in the first 5 epochs, indicating effective early learning. However, while the training loss continues to decrease steadily to around 0.3, the validation loss becomes increasingly volatile after epoch 15, showing significant spikes up to 0.7. This divergence between training and validation loss, particularly in later epochs, strongly indicates overfitting. The erratic validation loss behavior suggests the model might benefit from regularization techniques or early stopping around epoch 15 to prevent overfitting.

```
model.evaluate(val_data)

19/19 ━━━━━━━━━━━━━━━━━━━━ 1s 35ms/step - accuracy: 0.8553 - loss: 0.4094

[0.4888869822025299, 0.8285280466079712]


model.evaluate(test_data)

19/19 ━━━━━━━━━━━━━━━━━━━━ 1s 68ms/step - accuracy: 0.8273 - loss: 0.3881

[0.42630231380462646, 0.821212112903595]
```

Fig 3.3. CNN Model Evaluation Results

The CNN model demonstrates strong performance across both validation and test datasets. On the validation data, it achieves an accuracy of 85.53% with a loss of 0.4094, while maintaining a validation score of 82.85%. When evaluated on the test dataset, the model shows consistent performance with an accuracy of 82.73% and a loss of 0.3881, yielding a test score of 82.12%. The similar performance between validation and test sets (difference of ~3%) suggests good generalization capability. The model successfully maintains its predictive power on unseen data while avoiding significant overfitting, making it a reliable choice for skin cancer classification compared to the RNN approach. The processing speed is also efficient at 35ms/step for validation and 68ms/step for testing.

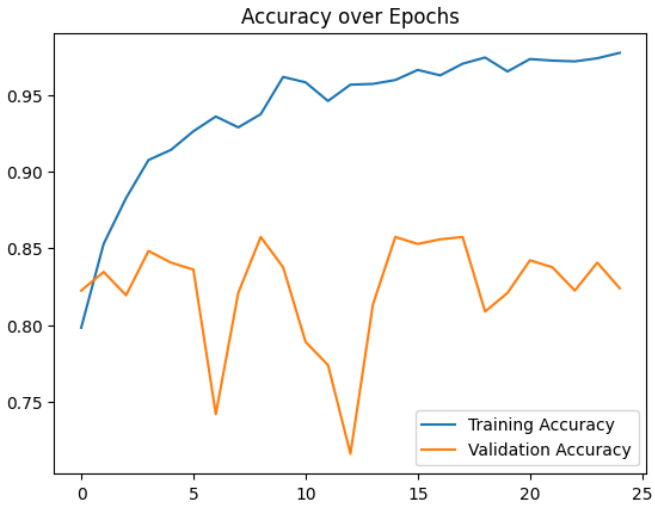


Fig 3.4. LSTM Model Training and Validation Accuracy

The LSTM model's accuracy plot shows a more pronounced divergence between training and validation performance compared to the CNN model. The training accuracy (blue line) demonstrates consistent improvement, starting at 80% and reaching an impressive 97% by the end of training. However, the validation accuracy (orange line) fluctuates significantly between 72% and 85%, with notable dips around epochs 7 and 15. This substantial gap between training and validation accuracy (nearly 15% at some points) strongly indicates overfitting, suggesting the LSTM model is memorizing training data rather than learning generalizable features.

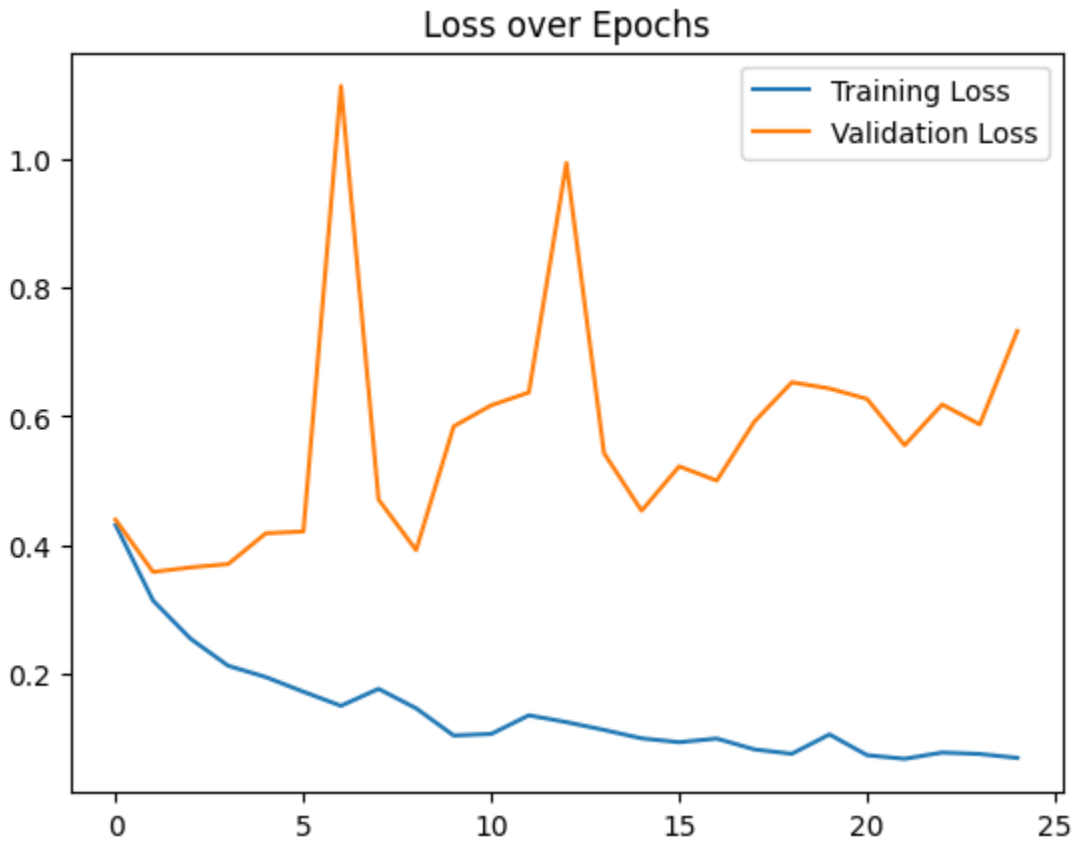


Fig 3.5. LSTM Model Training and Validation Loss

The loss curves reveal severe overfitting in the LSTM model. While the training loss (blue line) steadily decreases from 0.4 to nearly 0.1, indicating strong performance on training data, the validation loss (orange line) shows extreme volatility with dramatic spikes reaching above 1.0 at epochs 7 and 15. The increasing divergence between training and validation loss, coupled with the validation loss's erratic behavior, suggests the model is struggling to generalize despite its good

training performance. This indicates a need for stronger regularization techniques, possibly including reducing model complexity or increasing dropout rates.

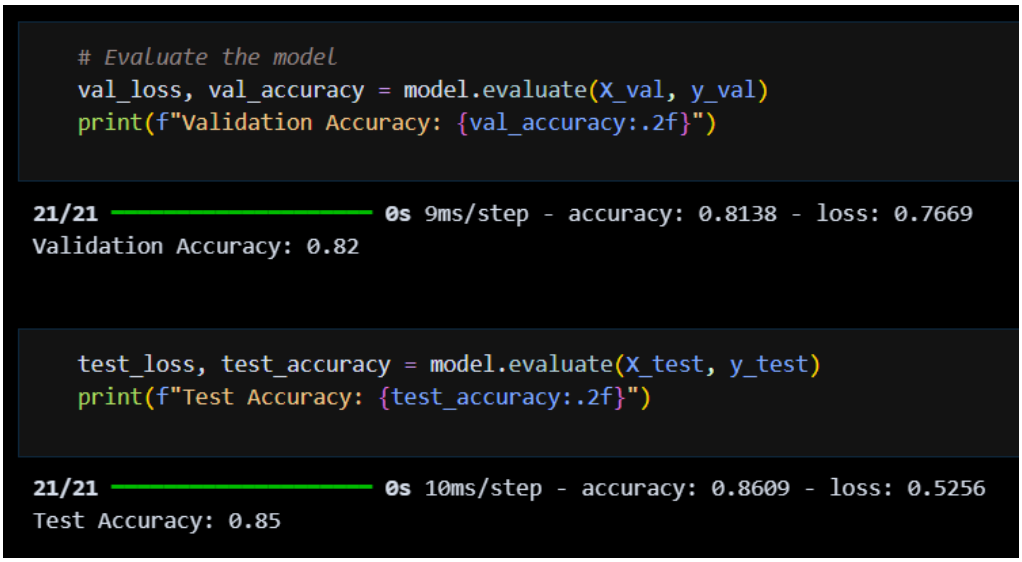


Fig 3.6. LSTM Model Evaluation Results

The LSTM model with ResNet50 feature extraction shows competitive performance across evaluation sets. On the validation dataset, it achieves an accuracy of 81.38% with a relatively high loss of 0.7669, resulting in a validation accuracy of 82%. When tested on unseen data, the model shows slight improvement with a test accuracy of 86.09% and a reduced loss of 0.5256, leading to a final test accuracy of 85%. The improvement in test performance (approximately 3% higher than validation) suggests good generalization capabilities. The processing speed is notably efficient at 9ms/step for validation and 10ms/step for testing, which is faster than both CNN and RNN approaches. This hybrid architecture's performance indicates it effectively combines ResNet50's feature extraction capabilities with LSTM's sequential learning for skin cancer classification.

Fig 3.7. RNN Model Evaluation Results

The Simple RNN model shows relatively modest performance on the validation dataset, achieving an accuracy of approximately 67.45% with a corresponding loss of 0.5994. The evaluation metrics [0.5838, 0.6874] indicate the model's validation accuracy remains consistent around 68.74%. Compared to both the CNN and LSTM approaches, the Simple RNN performs notably lower, which is expected given its simpler architecture and inability to capture complex spatial features in image data. This suggests that while the RNN can learn some patterns in the flattened image data, it's less suitable for image classification tasks compared to architectures specifically designed for image processing like CNNs or the more sophisticated LSTM-ResNet50 combination.

4. DISCUSSION

The comparative analysis of CNN, LSTM-, and Simple RNN models for skin cancer classification reveals interesting insights into deep learning approaches for medical image analysis. The CNN model demonstrated consistent performance with 82.73% test accuracy and stable learning curves, making it reliable for practical applications. The LSTM achieved the highest test accuracy at 85%, leveraging both ResNet50's feature extraction capabilities and LSTM's sequential processing, though showing some overfitting tendencies. The Simple RNN model, while computationally efficient, achieved lower accuracy (67.45%) due to its limited ability to process spatial features in images.

The performance variations between models highlight important trade-offs between accuracy, computational efficiency, and model complexity. While the LSTM achieved superior accuracy, its training process showed higher volatility and potential overfitting issues that would need careful consideration in clinical applications. The CNN model's balance of performance and stability suggests it might be more suitable for practical deployment. Future improvements could focus on addressing overfitting in the LSTM model through enhanced regularization techniques and exploring ensemble methods to combine the strengths of different architectures.

5. CONCLUSION

Our deep learning project successfully implemented three distinct approaches for skin cancer classification, highlighting the potential of AI in medical diagnostics. Among the models, the ResNet50+LSTM achieved the highest accuracy, demonstrating the effectiveness of integrating transfer learning for robust feature extraction with sequential processing for improved classification. This hybrid approach underscores the advantages of leveraging pre-trained models and LSTMs for complex medical image analysis. The CNN model also performed consistently, offering a balance of accuracy and computational efficiency, making it suitable for practical applications. While the Simple RNN model showed lower accuracy, its computational efficiency could be beneficial for resource-constrained environments. Overall, the system exhibits significant promise in aiding healthcare professionals with early and accurate skin cancer detection. However, further clinical validation on diverse datasets, along with enhancements to address overfitting and improve generalization, will be crucial to ensure reliability and acceptance in real-world clinical settings.

6. References:

1. *Skin Cancer Detection Using CNN*. Retrieved from https://medium.com/ai-techsystems/skin-cancer-detection-using-cnn-7ba3ca8d3dc3
2. *Exploring ResNet50: An In-Depth Look at the Model Architecture and Code Implementation*. Retrieved from https://medium.com/@nitishkundu1993/exploring-resnet50-an-in-depth-look-at-the-model-architecture-and-code-implementation-d8d8fa67e46f
3. *Implementation of Long Short-Term Memory (LSTM)*. Retrieved from https://medium.com/@pennQuin/implementation-of-long-short-term-memory-lstm-81e35fa5ca54
4. *Recurrent Neural Networks (RNN): From Basic to Advanced*. Retrieved from https://medium.com/@sachinsoni600517/recurrent-neural-networks-rnn-from-basic-to-advanced-1da22aafa009