

Efficient and Accurate Skin Lesion Analysis using CNN–ResNet and LSTM Hybrid Model

Y. MALLIKHARJUNA REDDY

Department of E.C.E
Tirumala Engineering College
Andhra Pradesh, India
Email: malli4b4@gmail.com

NANDIGAM ANUSHA

Department of E.C.E
Tirumala Engineering College
Andhra Pradesh, India
Email: anushanandigam99@gmail.com

MAMIDI DEEPTHI

Department of E.C.E
Tirumala Engineering College
Andhra Pradesh, India
Email: mamidideepthi123@mail.com

PATTEPOGU NIREEKSHANA

Department of E.C.E
Tirumala Engineering College
Andhra Pradesh, India
Email: neerikshanapattepogu@gmail.com

NIDAMANURI MOULIKA

Department of E.C.E
Tirumala Engineering College
Andhra Pradesh, India
Email: moulikanidamanuri@gmail.com

Abstract—Skin cancer is one of the most dangerous and rapidly growing diseases worldwide, particularly melanoma which has a high mortality rate if not detected early. Traditional diagnostic methods rely on manual examination of dermoscopic images by dermatologists, which can be time-consuming, subjective, and prone to errors. In this research, a hybrid deep learning model combining Convolutional Neural Networks (CNN), Residual Networks (ResNet), and Long Short-Term Memory (LSTM) is proposed for accurate skin lesion classification. The system automatically extracts spatial and deep features from images and captures relationships among features to improve classification performance. The model is trained and evaluated using dermoscopic image datasets, achieving high accuracy and reliability. The proposed system provides an efficient and scalable solution for early detection of skin cancer.

Index Terms—Skin Cancer, CNN, ResNet, LSTM, Deep Learning, Medical Imaging

I. INTRODUCTION

Skin cancer has become one of the most critical health concerns worldwide due to its increasing incidence and high mortality rate. Among different types of skin cancer, melanoma is considered the most aggressive and life-threatening form. Early detection and diagnosis play a crucial role in reducing mortality rates and improving patient survival. How-

ever, traditional diagnostic methods rely heavily on dermatologists' expertise, visual inspection, and biopsy procedures, which are often time-consuming, subjective, and prone to misclassification.

With the advancement of artificial intelligence, deep learning techniques have emerged as powerful tools in medical image analysis. Convolutional Neural Networks (CNNs) are widely used for extracting spatial features from images, while deep architectures such as ResNet improve learning capability by overcoming vanishing gradient problems. Additionally, Long Short-Term Memory (LSTM) networks are capable of capturing relationships between features, enhancing classification accuracy.

In this work, a hybrid CNN–ResNet–LSTM model is proposed to improve the accuracy and efficiency of skin lesion classification. The system aims to provide an automated, reliable, and scalable solution for early detection of skin cancer.

II. MOTIVATION

The increasing number of skin cancer cases and the limitations of traditional diagnostic methods highlight the need for automated and intelligent systems. Manual diagnosis is not only time-consuming

but also depends on the experience of dermatologists, leading to inconsistencies in results.

Another major challenge is the variability in skin lesion images, including differences in color, texture, size, and shape. These variations make it difficult to accurately classify lesions using conventional methods. Moreover, the availability of large medical image datasets requires efficient computational techniques for analysis.

Therefore, this project is motivated by the need to develop a deep learning-based system that can automatically analyze dermoscopic images and accurately classify skin lesions.

III. PROBLEM STATEMENT

Skin lesion classification is a challenging task due to the complexity and variability of dermoscopic images. Traditional methods fail to capture intricate patterns and often result in inaccurate classification. Additionally, the presence of noise, low-quality images, and class imbalance further complicates the problem.

Existing deep learning models such as CNN and ResNet provide good performance but may not fully capture relationships between features. Therefore, there is a need for a hybrid approach that combines spatial feature extraction, deep learning, and sequential modeling to improve classification accuracy.

IV. OBJECTIVES

The objectives of this work are to develop a hybrid CNN-ResNet-LSTM model for skin lesion classification, to accurately classify lesions as benign or malignant, to improve feature extraction using deep learning techniques, to reduce overfitting and improve generalization, to achieve high accuracy, precision, recall, and F1-score, and to build a scalable and efficient system for medical applications.

V. SYSTEM ARCHITECTURE

The system architecture of the proposed skin lesion classification model is designed as a structured pipeline that processes dermoscopic images through

multiple stages, enabling efficient feature extraction and accurate classification. The architecture integrates Convolutional Neural Networks (CNN), Residual Networks (ResNet), and Long Short-Term Memory (LSTM) into a unified hybrid framework.

The overall architecture of the system is illustrated in Fig. 1, which represents the flow of data from input image acquisition to final classification output. Each stage of the architecture performs a specific function, contributing to improved performance and reliability.

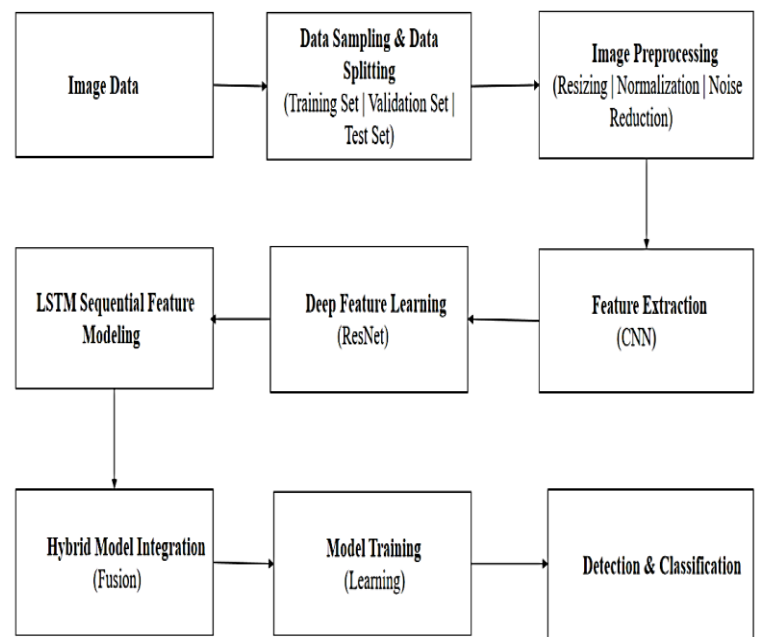


Fig. 1. System Architecture of Proposed CNN-ResNet-LSTM Model

The architecture consists of the following main components:

1. Input Image Acquisition: The system takes dermoscopic images of skin lesions as input. These images are collected from medical datasets and may include different classes such as benign and malignant lesions.

2. Image Preprocessing: The input images are preprocessed to improve quality and consistency. This includes resizing images to a fixed dimension, normalization of pixel values, and noise reduction. These steps ensure that the data is standardized before being passed to the model.

3. CNN-Based Feature Extraction: The preprocessed images are fed into a Convolutional Neural

Network (CNN), which extracts spatial features such as edges, textures, and shapes. The convolution and pooling layers help in reducing dimensionality while preserving essential information.

4. ResNet-Based Deep Feature Learning: The features extracted by CNN are further processed using a ResNet architecture. The residual connections in ResNet enable deeper network training and help in capturing complex and fine-grained patterns in skin lesion images.

5. LSTM-Based Sequential Modeling: The deep features obtained from ResNet are passed to an LSTM network. LSTM captures relationships among extracted features and learns dependencies, improving the model's ability to distinguish between similar-looking lesions.

6. Classification Output: Finally, the processed features are passed through fully connected layers, and a SoftMax classifier is used to categorize the input image into classes such as benign or malignant.

The proposed architecture effectively combines spatial feature extraction, deep learning, and sequential modeling, resulting in improved classification accuracy and robustness in skin lesion analysis.

VI. WORKING METHODOLOGY

The proposed system follows a structured and systematic approach for the classification of skin lesion images using a hybrid deep learning model. The methodology is designed to ensure efficient data processing, robust feature extraction, and accurate classification by integrating CNN, ResNet, and LSTM architectures.

A. Data Acquisition

The initial step involves collecting dermoscopic images from publicly available medical datasets. The dataset consists of labeled images categorized into different classes such as benign and malignant skin lesions. These images may vary in size, resolution, and lighting conditions, which necessitates preprocessing before model training.

B. Data Preprocessing

To ensure uniformity and improve model performance, preprocessing is applied to all input images. The preprocessing steps include resizing images to

a fixed dimension suitable for the model input, normalization of pixel values to a standard range, and removal of noise. Additionally, data augmentation techniques such as rotation, flipping, and zooming are applied to increase dataset diversity and reduce overfitting.

C. Feature Extraction using CNN

After preprocessing, the images are passed through a Convolutional Neural Network (CNN) for feature extraction. The CNN consists of multiple convolutional layers that capture spatial features such as edges, textures, and color variations. Pooling layers are used to reduce the dimensionality of feature maps while retaining important information. This stage transforms raw image data into meaningful feature representations.

D. Deep Feature Learning using ResNet

The features extracted by the CNN are further processed using a Residual Network (ResNet). ResNet introduces skip connections that allow the network to learn deeper representations without suffering from vanishing gradient problems. This stage enhances the feature extraction process by capturing complex patterns and subtle variations in skin lesion images.

E. Sequential Feature Modeling using LSTM

The deep features obtained from the ResNet model are passed to a Long Short-Term Memory (LSTM) network. The LSTM layer is responsible for capturing relationships and dependencies among features. This improves the model's ability to distinguish between visually similar lesions by analyzing feature sequences and patterns.

F. Classification Layer

The output from the LSTM network is fed into fully connected layers for final classification. A SoftMax activation function is used in the output layer to classify the input image into predefined categories such as benign or malignant. The model predicts the probability of each class and assigns the label with the highest probability.

G. Model Training and Evaluation

The model is trained using labeled data with an appropriate optimization algorithm such as Adam optimizer and a loss function such as categorical cross-entropy. The dataset is divided into training and testing sets to evaluate the performance of the model. Performance metrics such as accuracy, precision, recall, and F1-score are used to assess the effectiveness of the proposed system.

The overall methodology ensures efficient processing of input data, improved feature learning, and high classification accuracy by combining multiple deep learning techniques in a unified framework.

VII. IMPLEMENTATION

The implementation of the proposed skin lesion classification system is carried out using a deep learning framework that integrates CNN, ResNet, and LSTM models. The system is developed using Python programming language with the support of TensorFlow and Keras libraries, providing an efficient environment for building and training deep learning models.

A. Software and Hardware Requirements

The implementation is performed using Python in platforms such as Google Colab and Visual Studio Code. These platforms provide the necessary computational resources and libraries required for deep learning. The system requires a machine with sufficient memory and processing capability to handle image data and model training. GPU support is utilized to accelerate the training process.

B. Dataset Preparation

The dataset used in this project consists of dermoscopic skin lesion images categorized into different classes such as benign and malignant. The dataset is organized into separate folders for training and testing purposes. Before feeding the data into the model, all images are resized to a fixed dimension to maintain uniformity.

Data augmentation techniques such as rotation, horizontal flipping, zooming, and scaling are applied to increase the size of the dataset and improve the generalization capability of the model. This step helps in reducing overfitting and enhances the robustness of the system.

C. Model Development

The model is developed using a hybrid architecture that combines CNN, ResNet, and LSTM. Initially, a pre-trained ResNet model is used for feature extraction. The top layers of the ResNet model are removed, and custom layers are added to adapt it to the specific classification task.

The CNN layers extract spatial features from input images, which are then passed through the ResNet model to learn deeper representations. The output features are reshaped and provided as input to the LSTM layer, which captures relationships between features and improves classification accuracy.

Fully connected layers are added after the LSTM layer, followed by a SoftMax activation function for final classification.

D. Model Compilation

The model is compiled using the Adam optimizer, which provides efficient gradient-based optimization. The loss function used is categorical cross-entropy, which is suitable for multi-class classification problems. Accuracy is used as the primary evaluation metric.

E. Training Process

The model is trained using the prepared dataset over multiple epochs. During training, the model learns to identify patterns and features associated with different skin lesion classes. The dataset is divided into training and validation sets to monitor the model's performance and avoid overfitting.

Batch processing is used to efficiently handle large datasets. The training process involves forward propagation, loss calculation, backpropagation, and weight updates. Validation accuracy and loss are monitored at each epoch to evaluate the learning progress.

F. Testing and Evaluation

After training, the model is tested using unseen data to evaluate its performance. The testing phase measures how well the model generalizes to new images. Performance metrics such as accuracy, precision, recall, and F1-score are calculated to assess the effectiveness of the model.

G. Prediction

The trained model is used to predict the class of new input images. A test image is preprocessed and passed through the trained model, which outputs the probability of each class. The class with the highest probability is selected as the final prediction.

The implementation demonstrates that the hybrid CNN-ResNet-LSTM model is capable of accurately classifying skin lesion images, making it suitable for real-world medical applications.

VIII. RESULTS AND ANALYSIS

The performance of the proposed CNN-ResNet-LSTM hybrid model is evaluated using standard performance metrics and visual analysis techniques. The model is trained and tested on dermoscopic skin lesion images, and the results demonstrate its effectiveness in accurate classification.

A. Performance Metrics

The evaluation of the model is carried out using the following metrics:

- **Accuracy:** Measures the overall correctness of the model.
- **Precision:** Indicates how many predicted positive cases are actually correct.
- **Recall:** Measures the model's ability to detect all actual positive cases.
- **F1-Score:** Provides a balance between precision and recall.

The proposed model achieves high performance across all these metrics, indicating reliable classification of skin lesions.

B. Accuracy Analysis

The training and validation accuracy of the model over multiple epochs is shown in Fig. 2. The graph shows a steady increase in accuracy, indicating that the model is learning effectively.

It can be observed that the validation accuracy closely follows the training accuracy, which indicates that the model generalizes well and does not suffer from overfitting.

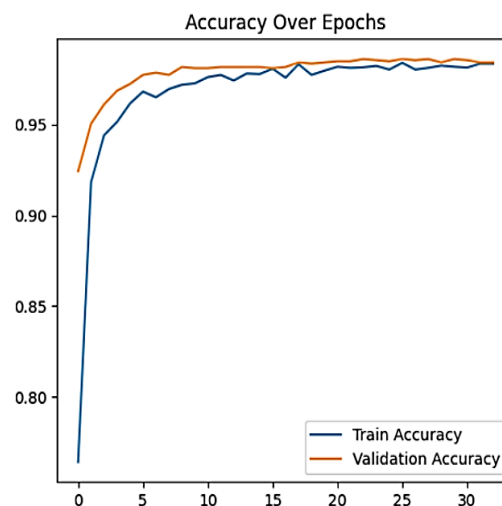


Fig. 2. Training and Validation Accuracy

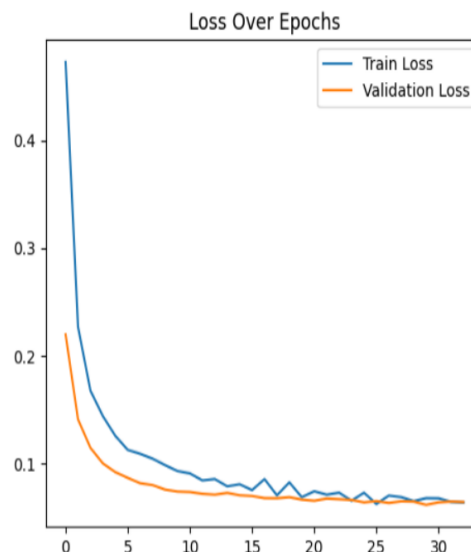


Fig. 3. Training and Validation Loss

C. Loss Analysis

The training and validation loss curves are shown in Fig. 3. The decreasing trend of loss values indicates that the model is optimizing effectively during training.

The validation loss stabilizes after a certain number of epochs, indicating that the model has reached an optimal learning stage.

D. Confusion Matrix Analysis

The confusion matrix of the model is presented in Fig. 4. It provides a detailed view of classification performance by showing correctly and incorrectly classified instances.

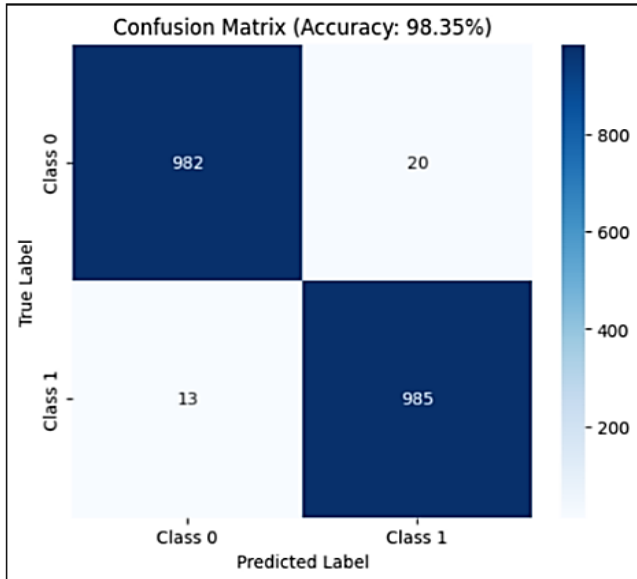


Fig. 4. Confusion Matrix

The matrix shows a high number of true positives and true negatives, indicating that the model performs well in distinguishing between benign and malignant lesions.

E. Comparative Analysis

The proposed hybrid model is compared with traditional models such as CNN, ResNet, and LSTM. The results indicate that the hybrid CNN–ResNet–LSTM model achieves superior performance in terms of accuracy, precision, recall, and F1-score. The improvement is due to the integration of spatial feature extraction, deep learning, and sequential modeling techniques.

The comparison of different models is shown in Fig. 5.

From Fig. 5, it is observed that the proposed model achieves higher accuracy.

F. Discussion

The experimental results demonstrate that the proposed model provides accurate and reliable classification of skin lesion images. The combination of

Model	Accuracy	Precision	Recall	F1-score
RESNET	94.12	93.60	94.00	93.80
CNN	95.20	94.80	95.10	94.85
LSTM	98.00	97.70	98.20	97.95

Fig. 5. Comparison of CNN, ResNet, and LSTM Models Based on Performance Metrics

CNN, ResNet, and LSTM enhances feature learning and improves performance. The model shows strong generalization ability and performs well on unseen data. However, minor misclassifications may occur due to poor image quality or similarity between lesion types.

IX. APPLICATIONS

The proposed CNN–ResNet–LSTM based skin lesion classification system has a wide range of applications in the healthcare and medical imaging domain. The ability of the model to accurately classify skin lesions enables its use in several real-world scenarios.

1. Early Detection of Skin Cancer: The system can be used for early detection of skin cancer, particularly melanoma. Early diagnosis significantly increases the chances of successful treatment and reduces mortality rates.

2. Clinical Decision Support Systems: The proposed model can assist dermatologists by providing a second opinion during diagnosis. It helps in reducing human error and improving diagnostic accuracy in clinical settings.

3. Telemedicine Applications: The system can be integrated into telemedicine platforms, allowing

patients to upload images of skin lesions remotely. Doctors can analyze the results without requiring physical presence, making healthcare more accessible.

4. Automated Medical Image Analysis: The model can be used in hospitals and diagnostic centers for automated analysis of dermoscopic images. This reduces the workload on medical professionals and speeds up the diagnostic process.

5. Mobile Health Applications: The system can be deployed as a mobile application where users can capture images using smartphones and receive preliminary analysis. This enables quick screening and awareness among the general public.

6. Research and Medical Studies: The proposed model can be used by researchers for analyzing large-scale medical image datasets and studying patterns in skin diseases. It can also aid in the development of improved diagnostic tools.

7. Healthcare Monitoring Systems: The system can be integrated into healthcare monitoring platforms to track changes in skin lesions over time, helping in continuous patient monitoring and follow-up.

Overall, the proposed system provides a scalable, efficient, and reliable solution for medical image analysis and has the potential to significantly improve the quality of healthcare services.

X. ADVANTAGES

The proposed CNN–ResNet–LSTM based skin lesion classification system offers several advantages over traditional diagnostic methods and standalone deep learning models. The integration of multiple architectures enhances the overall performance, efficiency, and reliability of the system.

1. High Classification Accuracy: The hybrid model combines CNN, ResNet, and LSTM, resulting in improved accuracy compared to individual models. It effectively captures spatial, deep, and sequential features, leading to better classification of skin lesions.

2. Automated Feature Extraction: The system automatically extracts relevant features from dermoscopic images without the need for manual intervention. This reduces dependency on expert knowledge and minimizes human error.

3. Improved Generalization: The use of data preprocessing and augmentation techniques helps the model generalize well to unseen data. This ensures consistent performance across different types of skin lesion images.

4. Reduction in Diagnostic Time: The automated nature of the system significantly reduces the time required for diagnosis. This enables faster decision-making in clinical environments.

5. Robust Performance: The model performs well under varying conditions such as changes in image quality, lighting, and lesion characteristics. The use of ResNet and LSTM enhances robustness and stability.

6. Scalability: The system can be scaled to handle large datasets and can be extended to classify multiple types of skin diseases beyond binary classification.

7. Support for Medical Professionals: The system acts as a decision support tool for dermatologists by providing reliable predictions, thereby assisting in accurate diagnosis and reducing workload.

8. Integration Capability: The proposed model can be easily integrated into healthcare systems, telemedicine platforms, and mobile applications, making it suitable for real-world deployment.

Overall, the proposed system provides a reliable, efficient, and scalable solution for automated skin lesion classification, improving the quality and accessibility of healthcare services.

XI. CONCLUSION

In this paper, a hybrid deep learning approach for skin lesion classification using CNN, ResNet, and LSTM has been presented. The proposed system is designed to improve the accuracy and reliability of skin cancer detection by effectively combining spatial feature extraction, deep feature learning, and sequential modeling techniques.

The model processes dermoscopic images through multiple stages, including preprocessing, feature extraction, and classification, ensuring that important visual patterns such as texture, color, and shape are captured efficiently. The integration of ResNet helps in learning deeper representations without degradation, while the LSTM layer enhances the model's ability to understand relationships among extracted features.

Experimental results demonstrate that the proposed model achieves high performance in terms of accuracy, precision, recall, and F1-score. The training and validation analysis indicates that the model generalizes well to unseen data, with minimal overfitting. The confusion matrix further confirms that the system is capable of correctly distinguishing between benign and malignant lesions with a high degree of reliability.

Compared to traditional methods and individual deep learning models, the proposed hybrid approach shows improved classification performance. The system also reduces manual effort and provides faster and more consistent results, making it suitable for real-world medical applications.

Overall, the proposed CNN–ResNet–LSTM based system provides an efficient, scalable, and reliable solution for automated skin lesion analysis. It has strong potential to assist healthcare professionals in early detection and diagnosis of skin cancer, thereby contributing to improved patient outcomes.

XII. FUTURE SCOPE

The proposed CNN–ResNet–LSTM model can be further enhanced by using larger and more diverse datasets to improve generalization and robustness. Future work can focus on extending the system from binary classification to multi-class classification for identifying different types of skin diseases.

Additionally, advanced preprocessing and segmentation techniques can be incorporated to improve feature extraction. The model can also be optimized for faster execution and deployed in real-time applications such as mobile or web-based platforms.

Furthermore, integrating explainable AI techniques can help in understanding model decisions, making the system more reliable and suitable for practical healthcare applications.

REFERENCES

- [1] A. Esteva et al., “Dermatologist-level classification of skin cancer with deep neural networks,” *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.
- [2] M. Codella et al., “Skin lesion analysis toward melanoma detection: A challenge at ISBI 2017,” in *Proc. IEEE Int. Symp. Biomedical Imaging*, 2017, pp. 168–172.

- [3] N. Codella et al., “Deep learning ensembles for melanoma recognition in dermoscopy images,” *IBM Journal of Research and Development*, vol. 61, no. 4/5, pp. 5:1–5:15, 2017.
- [4] P. Tschandl, C. Rosendahl, and H. Kittler, “The HAM10000 dataset: A large collection of multi-source dermatoscopic images,” *Scientific Data*, vol. 5, 2018.
- [5] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE CVPR*, 2016, pp. 770–778.
- [6] A. Krizhevsky, I. Sutskever, and G. Hinton, “ImageNet classification with deep convolutional neural networks,” in *Proc. NIPS*, 2012, pp. 1097–1105.
- [7] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [8] G. Litjens et al., “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
- [9] J. Kawahara et al., “Deep features to classify skin lesions,” *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 2, pp. 362–373, 2017.
- [10] D. Shen, G. Wu, and H. Suk, “Deep learning in medical image analysis,” *Annual Review of Biomedical Engineering*, vol. 19, pp. 221–248, 2017.
- [11] O. Russakovsky et al., “ImageNet large scale visual recognition challenge,” *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.
- [12] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.