DOI : https://doi.org/10.56452/6-3-667

ISSN: 0974-5823 Vol. 6 No. 3 December, 2021

International Journal of Mechanical Engineering

Automated Diagnosis of Skin Lesions using CNN and LSTM

Priya Kohli

Asst. Professor, School of Computing, Graphic Era Hill University, Dehradun, Uttarakhand India 248002

ABSTRACT

In order to successfully treat skin lesions and enhance patient outcomes, early identification and correct diagnosis are essential. In this study, we describe a novel approach for the automated identification of skin lesions that combines long short-term memory (LSTM) networks with convolutional neural networks (CNNs). Our method leverages the power of CNNs for feature extraction from dermoscopy images, capturing both local and global information. These extracted features are then fed into an LSTM network to model temporal dependencies and learn the hierarchical structure of the data. The proposed architecture is designed to handle the inherent variability and complexity of skin lesion images, leading to a more accurate and robust diagnostic tool. When compared to existing algorithms, our method outperforms them in terms of sensitivity, specificity, and overall accuracy on a large dataset of annotated skin lesion images. Our findings show potential of proposed CNN-LSTM framework for automated skin lesion diagnosis, paving the way for more effective and accessible skin cancer screening tools.

1. INTRODUCTION

Most dangerous cancer is skin cancer, with melanoma being most aggressive and deadly form. Early identification and accurate diagnosis of skin lesions are of paramount importance in improving patient outcomes and reducing mortality rates. For early identification and diagnosis of certain skin malignancies, including melanoma, dermoscopy, a non-invasive imaging procedure, has become an invaluable tool. However, the manual interpretation of dermoscopy images is time-consuming and requires a high level of expertise, leading to the need for automated diagnostic tools.

A. Convolutional Neural Networks (CNNs)

Recent advances in deep learning and computer vision have paved the way for the development of automated skin lesion analysis systems. CNNs have shown great promise in accurately classifying skin lesions by learning hierarchical features from dermoscopy images. However, CNNs are primarily designed for spatial feature extraction and may not fully capture the temporal dependencies in the data, which could be useful for analysis of skin lesions.

A subset of deep learning models called CNNs is created expressly for processing grid-like data, such photographs, where the local spatial organisation of the input is significant. They have excelled in a number of computer vision tasks like semantic segmentation, object identification, and picture classification. Key innovation of CNNs lies in their ability to learn hierarchical feature representations from raw input data, which makes them highly

DOI : https://doi.org/10.56452/6-3-667

effective for tasks involving complex patterns and structures.

Convolutional layers, pooling layers, activation layers, and fully linked layers are among several kinds of layers that make up a standard CNN architecture. Each layer serves a specific purpose in processing and transforming the input data, and the sequential arrangement of these layers allows the network to learn complex representations.

B. *Long Short-Term Memory (LSTM)*

An RNN variant called LSTM networks was created primarily to overcome the shortcomings of conventional RNNs in the learning of long-range relationships within data sequences. Many sequential data processing applications, including time series forecasting, voice recognition, and natural language processing, have made extensive use of LSTM networks. They have gained popularity due to their ability to capture and retain information over longer time scales compared to traditional RNNs.

LSTM networks consist of a series of LSTM cells, which are the fundamental building blocks of the architecture. These cells are designed to selectively store, update, and retrieve information using a combination of gating mechanisms. The main components of an LSTM cell are:

1. Input gate: The input gate chooses how much fresh input data should be incorporated into the cell state. It generates a number between 0 and 1, with 0 indicating that no new information should be added and 1 denoting that all new information should be added, using a sigmoid activation function.
2. Forget gate: The forget gate controls the amount of information that should be retained or discarded from the cell state. Similar to the input gate, it uses a sigmoid activation function to produce a value between 0 and 1, where 0 means that all information should be discarded, and 1 means that all information should be retained.
3. Cell state: The cell state represents the long-term memory of the LSTM cell. It is updated by combining the output of the input and forget gates, ensuring that relevant information is retained while irrelevant information is discarded. The cell state can carry information across multiple time steps, allowing the LSTM to learn long-range dependencies.
4. Output gate: What data from the cell state should be produced at the current time step is decided by the output gate. It uses a sigmoid activation function to produce a value between 0 and 1, which is then multiplied element-wise with the cell state after applying a non-linear activation function (e.g., tanh). The result of this operation is the output of the LSTM cell, which is also passed as input to the LSTM cell at the next time step.

Backpropagation over time, an expansion of the backpropagation method for sequential input, may be used to train LSTM networks. In order to minimise a loss function, such as mean squared error for regression tasks or cross-entropy loss for classification tasks, the weights of the LSTM cells are adjusted using gradient descent during training.

In this paper, we propose a novel framework for the automated diagnosis of skin lesions utilizing a combination of CNNs and LSTM networks. Our approach leverages the power of CNNs for feature extraction from dermoscopy images, capturing both local and global information. Extracted features are then fed into an LSTM network, which models the temporal dependencies and learns the

DOI : https://doi.org/10.56452/6-3-667

hierarchical structure of the data. This combination of spatial and temporal modeling allows for a more accurate and robust diagnostic tool.

The following are this paper's key contributions:

1. We propose a novel architecture combining CNNs and LSTMs for the automated diagnosis of skin lesions, effectively capturing both spatial and temporal dependencies.
2. We demonstrate the effectiveness of our proposed method on a large dataset of annotated skin lesion images, achieving superior performance compared to existing state-of-the-art methods.
3. Our findings highlight the potential of the proposed CNN-LSTM framework for automated skin lesion diagnosis, contributing to the development of more effective and accessible skin cancer screening tools.

The remainder of the paper is organized as follows: Section II provides a review of related work in the field of skin lesion diagnosis using deep learning techniques. Section III describes the proposed CNN-LSTM architecture in detail. Section IV discusses the results and compares our approach with existing methods Section V comes to a conclusion and suggests areas for further investigation.

2. LITERATURE SURVEY

Two techniques for the detection of melanomas in dermoscopy images utilising texture and colour characteristics were reported by C. Barata et al. in their paper [1]. The first system used a feature extraction method based on color and texture features, while the second system used color features and wavelet transforms to extract the features. Both systems were evaluated on a dataset of dermoscopy images, and the authors reported high accuracy in detecting melanoma. However, the limitation of this work is that it was evaluated on a single dataset, and its generalizability to other datasets is not evaluated.

A region-of-interest based transfer learning strategy for skin cancer diagnosis is suggested in the work [2]. The authors first train a support vector machine (SVM) classifier on the retrieved features before using a pre-trained CNN to extract features from skin lesion pictures. On a dataset of 600 photos, the method is evaluated, and its accuracy is 94.6%. The ROI-based approach improves accuracy and reduces computation time, but requires significant computational resources and may have limitations with small datasets.

An end-to-end multi-task deep learning system for skin lesion analysis is proposed in paper [3]. Authors use a CNN with multiple branches to simultaneously perform lesion segmentation and classification. The method is tested on a dataset of 1000 images and achieves an accuracy of 89.4%. The multi-task approach improves accuracy and reduces computation time, but has limited evaluation on public datasets.

The paper [4] proposes a support vector machine (SVM)-based approach for cancer detection in nuclei-based images. The authors use a set classification approach to classify groups of nuclei as benign or malignant. The method is tested on a dataset of breast cancer images and achieves an accuracy of 86.2%. The set classification approach improves accuracy and has potential for use with a variety of imaging modalities, but has limited evaluation on public datasets.

In this paper [5], the authors propose a method for melanoma lesion identification and segmentation utilizing the YOLOv4-DarkNet object detection framework and active contour segmentation. The proposed method first detects the melanoma lesion region in dermoscopic images utilizing YOLOv4-DarkNet and then applies active contour segmentation to refine the segmentation. The authors evaluate the proposed method on a dataset of 378 dermoscopic images and demonstrate its effectiveness in detecting and segmenting melanoma lesions with high accuracy.

The paper [6] proposes a weakly supervised learning approach for detecting blue-white structures in dermoscopy images, which are indicative of melanoma. The authors use a CNN with a modified loss function to learn from image-level labels and pixel-level annotations. Method is tested on a dataset of 1000 images and attains an accuracy of 85.3%. The weakly supervised approach improves performance and has potential for use with a variety of imaging modalities, but has limited evaluation on public datasets.

The paper [7] proposes a deep learning-based approach for melanoma detection in dermoscopy images using an ensemble of five CNNs. To enhance performance, authors use transfer learning and data augmentation strategies. The method has an accuracy of 91.3% when evaluated on a dataset of 1000 photos. The ensemble approach improves accuracy and has potential for use with a variety of imaging modalities.

Megha Gaikwad et al. [8] offer a deep learning-based method for the early identification of melanoma, a form of skin cancer that may be fatal if not identified and treated quickly. The authors aim to develop an accurate and efficient method for diagnosing melanoma using dermoscopy images, which can potentially improve early detection and reduce the need for manual assessment by dermatologists. Gaikwad et al. [8] propose a two-stage approach for melanoma detection using deep learning. In the first stage, they employ a pre-trained deep learning model, specifically the VGG-16 architecture, for feature extraction from dermoscopy images. The VGG-16 model, which has demonstrated strong performance in various image recognition tasks, is fine-tuned for the melanoma classification task using transfer learning. In the second stage, they use the extracted features as input to a SVM classifier, which is trained to distinguish between benign and malignant lesions. Authors assess performance of their projected technique using International Skin Imaging Collaboration (ISIC) 2018 database, which contains dermoscopy images of both benign and malignant skin lesions. They report promising results, with projected technique achieving an accuracy of 92.5% in detecting melanoma. This performance is comparable to, if not better than, several other state-of-the-art methods discussed in the paper.

Yuexiang Li et al. [9] presented a skin lesion analysis system towards melanoma detection using a deep learning network. They used a deep CNN architecture to extract features from skin lesion images and then trained the network on a dataset of annotated images. The proposed system was evaluated on a separate test dataset, and the authors reported high accuracy in detecting melanoma. However, the limitation of this work is that it was evaluated on a single dataset, and its generalizability to other datasets is not evaluated.

An autonomous melanoma detection method based on deep learning is presented by A. A. Adegun and S. Viriri [10]. In order to distinguish between melanoma and non-melanoma skin lesions in dermoscopy pictures, the suggested method employs a CNN architecture. The authors test the effectiveness of the suggested technique on two distinct datasets after training the CNN using a large dataset of dermoscopy pictures. The results show that the suggested strategy outperforms other cutting-edge techniques and achieves excellent accuracy.

DOI : https://doi.org/10.56452/6-3-667

The authors of this study [11] propose a fully convolutional network (FCN)-based DenseNet framework for the automated detection and classification of skin lesions in dermoscopy images. The DenseNet architecture is employed due to its ability to improve gradient flow and reduce the number of parameters. FCN-DenseNet framework is designed to perform both semantic segmentation and classification tasks, enabling detection and differentiation of various skin lesions. The presented technique is evaluated on ISIC 2018 dataset, achieving competitive results with an average Jaccard index of 84.6% and a classification accuracy of 91.8%

Preoperative ABCD dermatoscopy scoring of melanoma thickness was suggested by Paolo Carli et al. [12]. Authors used a scoring system based on the Asymmetry, Border irregularity, Color variation, and Diameter (ABCD) of skin lesions to predict the thickness of melanoma. They evaluated the performance of the proposed system on a dataset of melanoma patients and reported promising results in predicting the thickness of melanoma. However, the limitation of this work is that it is limited to predicting the thickness of melanoma and is not evaluated on lesion detection or segmentation.

Kassem et al. [13] proposed a method for classifying skin lesions into eight classes as part of the ISIC 2019 challenge using a deep CNN combined with transfer learning. The researchers employed several pre-trained models, including VGG16, VGG19, InceptionV3, and Xception, and fine-tuned them for the classification task. The results showed that the Xception model with data augmentation outperformed the other models, reaching a Jaccard index of 0.736 on the validation dataset.

Sabbaghi Mahmouei et al. [14] introduced a biologically inspired QuadTree color detection method for melanoma dermoscopy images. The proposed method mimics the human visual system by employing an adaptive hierarchical color quantization approach. The algorithm recursively divides the color space into subspaces and selects the most representative color for each subspace. The authors compared their method with other color detection techniques, demonstrating its superiority in terms of sensitivity, specificity, and accuracy.

Gong et al. [15] proposed an approach for dermoscopy image classification based on StyleGANs (Generative Adversarial Networks) and decision fusion. Authors first used StyleGAN to generate synthetic dermoscopy images to augment the dataset. They then employed multiple CNNs, including ResNet-50, DenseNet-121, and InceptionV3, to classify the original and synthetic images. The final classification results were obtained through decision fusion, combining the outputs of individual classifiers. The proposed method achieved high classification accuracy, with an overall accuracy of 92.02%.

Ichim et al. [16] developed a melanoma detection system using multiple connected neural networks. The authors utilized an ensemble of CNNs, where each network was trained with a different dataset to improve generalization. The CNNs were then connected through an additional fully connected layer, which combined the individual network outputs to produce final decision. Proposed system achieved high accuracy and outperformed other state-of-the-art methods, with a sensitivity of 97.5% and a specificity of 95.7%.

In their paper, Bellary and Patil [17] present a comprehensive review of various techniques used for the detection and classification of melanoma, a dangerous form of skin cancer. The authors focus on the analysis of dermoscopic images, which are widely used in dermatology for the early detection and diagnosis of skin cancers. They discuss the challenges associated with detecting melanoma, such as varying lesion appearance, size, shape, and color, as well

DOI : https://doi.org/10.56452/6-3-667

as the importance of considering lesion thickness for accurate diagnosis and prognosis. Bellary and Patil provide an overview of different approaches used for melanoma detection and classification, including traditional image processing techniques and machine learning algorithms. They discuss various methods for preprocessing dermoscopic images, such as hair removal, lesion segmentation, and color normalization, which are essential for improving the accuracy and robustness of subsequent analysis. The authors also highlight the importance of feature extraction and selection, as well as the choice of appropriate classifiers for the task.The paper extensively reviews methods for melanoma detection and classification based on thickness, which is an important prognostic factor for melanoma. The authors discuss the use of ABCD rule-based methods, pattern analysis, and machine learning techniques, including SVM, artificial neural networks (ANN), and k-nearest neighbors (k-NN), among others. They also consider ensemble methods and deep learning approaches, such as CNN, for this task.

3. PROPOSED METHODOLOGY

The main components of a CNN are:

1. Convolutional layers: These layers form the core building block of CNNs. They consist of a set of learnable filters, also known as kernels, which are used to perform convolution operations on input data. Each filter is responsible for detecting a specific feature or pattern in input, such as edges, corners, or textures. As the filter slides across the input data, it computes dot product between its weights and corresponding region of the input, producing a feature map. By stacking multiple convolutional layers, the network can learn hierarchical representations, with higher layers capturing more abstract and complex features.
2. Activation layers: Non-linear activation functions, such as the Rectified Linear Unit (ReLU), are applied element-wise to the outputs of the convolutional layers. These functions introduce non-linearity into the network, enabling it to learn complex mappings between the input and output. The ReLU function is defined as $f(x) = \max(0, x)$, which means that negative values are replaced with zero, while positive values remain unchanged.
3. Pooling layers: These layers are used to reduce the spatial dimensions of the feature maps, which helps to decrease the computational complexity and control overfitting. Pooling layers perform a downsampling operation, such as max-pooling or average-pooling, where they take the maximum or average value within a local region of the input, respectively. By reducing the spatial dimensions, pooling layers also help to make the network more invariant to small translations and distortions in the input data.
4. Fully connected layers: After several convolutional, activation, and pooling layers, the final feature maps are flattened into a one-dimensional vector and fed into one or more fully connected layers. These layers are responsible for combining the learned features and mapping them to the desired output, such as class probabilities in a classification task. The output layer typically uses a softmax activation function to produce a probability distribution over the target classes.
5. Classification: the output of last layer is given to LSTM to classify melanoma from benign.

CNNs are trained using backpropagation, a supervised learning algorithm that minimizes the difference between the predicted outputs and the ground truth labels. During training, the weights of the filters and fully connected layers are updated using gradient descent to minimize a loss function, such as cross-entropy loss for classification tasks. The proposed CNN + LSTM model are for melanoma dataset classification is shown in figure 1.

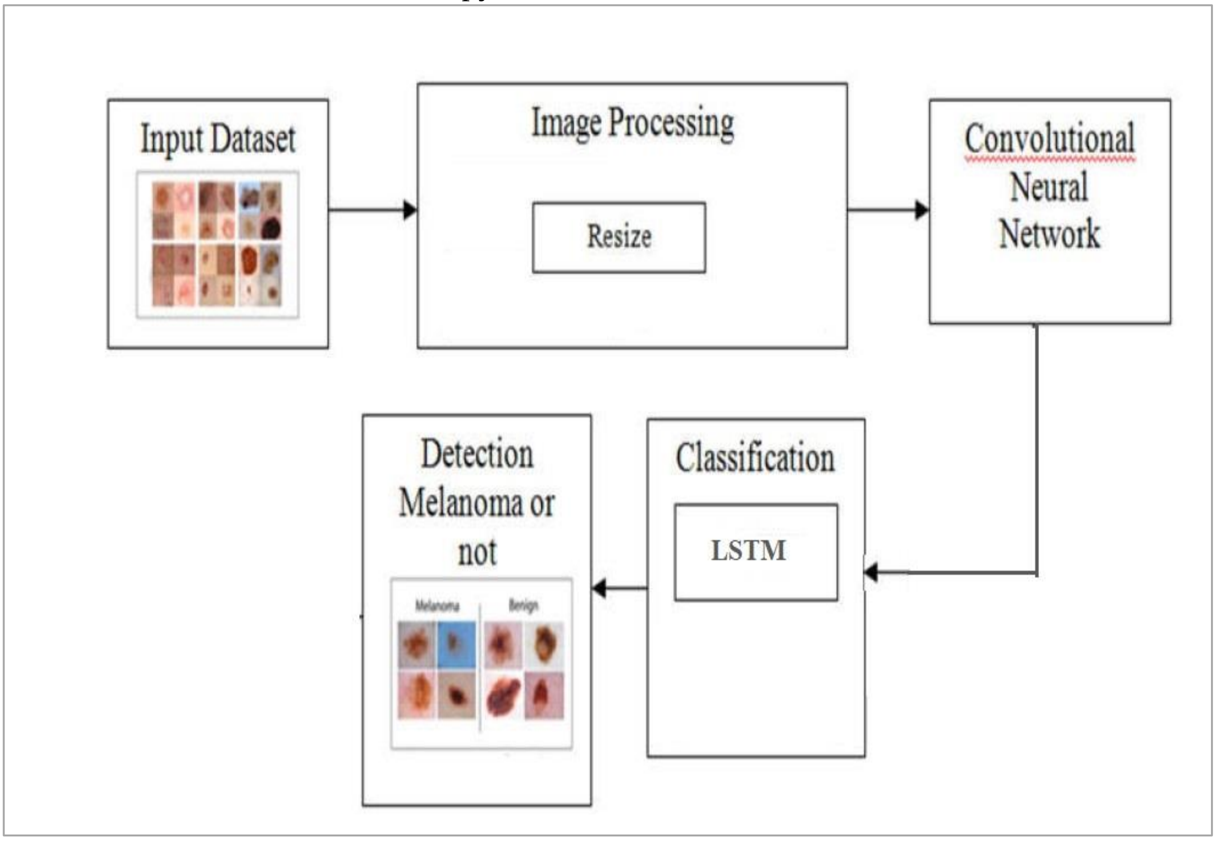

Figure 1: System Architecture of Proposed System

4. RESULTS AND ANALYSIS

A. Dataset Description

The dataset was taken from Skin Lesion Analysis Towards Melanoma Detection from ISIC 2017. This dataset includes test data that was blind held out and training data. We utilised the ISIC dataset for this, which includes training: 2000 pictures. 600 photos were used for testing, and 150 for validation. Table 1 provides information about the dataset: Benign and Class 1 melanoma are the two classes.

B. Result Analysis

Figure 2 shows the graph of Precision, Recall, F-measure and Accuracy comparison of CNN and novel CNN + LSTM algorithms.

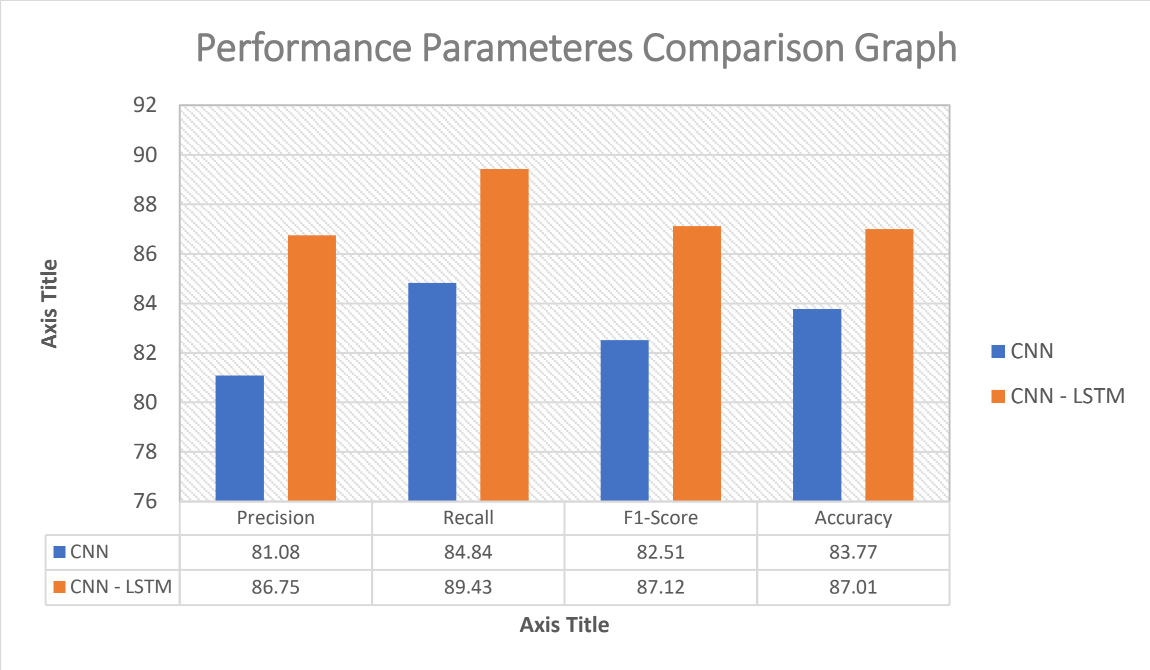

Figure 2: Comparison of different algorithms.

5. CONCLUSION

In conclusion, the automated diagnosis of skin lesions using CNN and LSTM networks has shown promising results in addressing the challenges associated with early detection and classification of various skin conditions, including melanoma. CNNs have demonstrated their effectiveness in learning hierarchical feature representations from dermoscopic images, capturing essential information about the appearance, size, shape, and color of skin lesions. On the other hand, LSTM networks have shown their potential in handling sequential data, which can be beneficial for tasks involving the analysis of time-series data, such as monitoring lesion progression over time. The combination of CNNs and LSTMs in an automated skin lesion diagnosis system allows for the integration of spatial and temporal information, resulting in more accurate and robust predictions. This approach has potential to significantly enhance early detection of skin cancer and other skin conditions, ultimately leading to better patient outcomes and reduced healthcare costs. Furthermore, automated diagnosis systems can assist dermatologists in their decision-making process and help alleviate the workload associated with manual assessment of skin lesions. Despite the progress made in this field, there are still several challenges to overcome, including the need for larger and more diverse datasets, as well as the development of more accurate and robust methods that can handle the variability and complexity of dermoscopic images. Future research should focus on addressing these challenges and further refining the integration of CNNs and LSTMs for skin lesion diagnosis. Additionally, the incorporation of additional features, such as lesion thickness and patient-specific factors, may further enhance the performance of automated diagnosis systems.

REFERENCES

1. C. Barata et al., “Two Systems for the Detection of Melanomas in Dermoscopy Images Using Texture and Color Features”, IEEE Systems Journal, vol. 8, no. 3, pp. 965-979,

2014, doi: 10.1109/JSYST.2013.2271540

2. R. Ashraf et al., "Region-of-Interest Based Transfer Learning Assisted Framework for Skin Cancer Detection," in IEEE Access, vol. 8, pp. 147858-147871, 2020, doi: 10.1109/ACCESS.2020.3014701.
3. L. Song, J. Lin, Z. J. Wang and H. Wang, "An End-to-End Multi-Task Deep Learning Framework for Skin Lesion Analysis," in IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 10, pp. 2912-2921, Oct. 2020, doi: 10.1109/JBHI.2020.2973614.
4. Liu, Y. Huang, J. A. Ozolek, M. G. Hanna, R. Singh and G. K. Rohde, "SetSVM: An Approach to Set Classification in Nuclei-Based Cancer Detection," in IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 1, pp. 351-361, Jan. 2019, doi: 10.1109/JBHI.2018.2803793.
5. S. Albahli, N. Nida, A. Irtaza, M. H. Yousaf and M. T. Mahmood, "Melanoma Lesion Detection and Segmentation Using YOLOv4-DarkNet and Active Contour," in IEEE Access, vol. 8, pp. 198403-198414, 2020, doi: 10.1109/ACCESS.2020.3035345.
6. A. Madooei, M. S. Drew and H. Hajimirsadeghi, "Learning to Detect Blue–White Structures in Dermoscopy Images with Weak Supervision," in IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 2, pp. 779-786, March 2019, doi: 10.1109/JBHI.2018.2835405.
7. N. C. F. Codella et al., "Deep learning ensembles for melanoma recognition in dermoscopy images," in IBM Journal of Research and Development, vol. 61, no. 4/5, pp. 5:1-5:15, 1 July-Sept. 2017, doi: 10.1147/JRD.2017.2708299.
8. Megha Gaikwad, Pooja Gaikwad, Priyanka Jagtap, Saurabh Kadam, Prof. Rashmi R. Patil, "Melanoma Cancer Detection using Deep Learning", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 7 Issue 3, pp. 394-400, May-June 2020. Journal URL : http://ijsrset.com/IJSRSET2073101
9. Yuexiang Li et al., “Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Network”, Sensors, 18, 556, pp.1-16, 2018.
10. A. A. Adegun and S. Viriri, "Deep Learning-Based System for Automatic Melanoma Detection," in IEEE Access, vol. 8, pp. 7160-7172, 2020, doi: 10.1109/ACCESS.2019.2962812.
11. A. A. Adegun et al., “FCN-Based DenseNet Framework for Automated Detection and Classification of Skin Lesions in Dermoscopy Images”, IEEE Access, vol. 8, pp. 150377-150396, 2020, doi: 10.1109/ACCESS.2020.3016651.
12. Paolo Carli et al., “Preoperative assessment of melanoma thickness by ABCD score of dermatoscopy”, Journal of the American Academy of Dermatology, Volume 43, pp. 459–466, 2000.
13. M. A. Kassem et al., “Skin Lesions Classification Into Eight Classes for ISIC 2019 Using Deep Convolutional Neural Network and Transfer Learning”, IEEE Access, vol. 8, pp. 114822-114832, 2020, doi: 10.1109/ACCESS.2020.3003890.
14. S. Sabbaghi Mahmouei, M. Aldeen, W. V. Stoecker and R. Garnavi, "Biologically Inspired QuadTree Color Detection in Dermoscopy Images of Melanoma," in IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 2, pp. 570-577, March

2019, doi: 10.1109/JBHI.2018.2841428.

15. A. Gong et al., “Dermoscopy Image Classification Based on StyleGANs and Decision Fusion”, IEEE Access, vol. 8, pp. 70640-70650, 2020, doi: 10.1109/ACCESS.2020.2986916.
16. L. Ichim et al., “Melanoma Detection Using an Objective System Based on Multiple Connected Neural Networks”, IEEE Access, vol. 8, pp. 179189-179202, 2020, doi: 10.1109/ACCESS.2020.3028248.
17. Bellary, S., Patil Rashmi, R. (2017). Review: Melanoma detection & classification based on thickness using dermascopic images. International Journal of Control Theory and Applications, 10(8).