International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 10 | Oct - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Melanoma Skin Cancer Detection Using Deep learning, Neural Network and Classical Machine Learning Techniques

Sri Varsh Cirigiri*, **Chaithanya Sai Inturu***, **Hemanth Sankar***

Department of Computer Science and Engineering, Institute of Aeronautical Engineering, Hyderabad
Department of Computer Science and Engineering, Anurag University, Hyderabad
Department of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai

Abstract- Melanoma is one of the deadliest types of skin cancer and requires an early check-up for survival. Though earlier diagnosis was primarily dependent on dermatological checks and histopathological reports, automation of the same may be highly time-efficient. With that interest, in the following paper, we propose a hybrid approach which attempts to utilize deep learning with some classical machine learning methods to automate melanoma detection. The feature extraction and classification were done using CNN because it is the most efficient in processing image data. Besides, SVM and KNN were used as comparative models. A very well-preprocessed dataset of images was used to see how well models could work in terms of accuracy, precision, and recall. We observe strong over-performance by CNNs compared to the more traditional methods and an ensemble of models that yield a higher confidence of diagnostics. The technique has much promise for supporting clinicians with quicker and more precise diagnosing of melanoma, thus enabling better patient care.

Keywords: Melanoma detection, dermatological, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Convolutional Neural Networks (CNNs).

1. Introduction

The most common cancer in the world is skin cancer, yet melanoma is the deadliest form simply because it develops very fast and may easily metastasize [1]. While a very small percentage of all skin cancers, melanoma is indeed the deadliest skin cancer in terms of the number of deaths. Early detection is essential because, in melanoma patients, it has a survival rate of above 98% if detected in its early stages. The figure drops sharply once melanoma has spread [2]. It is critical to develop prompt and accurate diagnostic methods because of this [3].

Traditionally, melanoma was diagnosed with clinical dermatologic examination and histopathologic assessment, which are often aided in practice through the tool of dermatoscopy to view suspicious skin lesions [4]. Although these methods of diagnosis are reliable, they are often subjective, time-consuming, and usually require the clinician's skill-a plausible reason for the variability of diagnoses between examiners [5]. On account of the ever-increasing necessity of devising faster and more precise diagnostic tests, machine and deep learning techniques have increasingly been applied to melanoma detection systems [6][7].

We present in this paper a hybrid approach that uses both deep learning and classical machine learning algorithms for improved melanoma detection accuracy [8][9].

They classify the skin lesions emphasizing extracting some features like shape and texture besides color [10]. The power of convolutional neural networks is quite powerful with respect to complex data images [11]. For that purpose, architectures such as VGG16, ResNet50, and InceptionV3 will be used; additionally, a custom CNN model further increases the performance of classification [12][13]. Further, classical models such as SVM and KNN are used in order to afford some comparative insight and further layers of classification [14][15].

This study sums up the best of various models into just one model to make the general diagnosis accuracy better with the use of ensemble learning techniques. This has been done over a very well-established benchmark ISIC 2020 melanoma classification dataset to perform this analysis [11]. Techniques that involve standardizing images and data augmentation have been employed for pre-processing along with the robustness of models against class imbalance [2][15].

In this paper, we propose a hybrid approach toward better melanoma detection. We first discuss existing methods and visual cues that help clinicians make an accurate diagnosis [4]. Afterward, we introduce our proposed solution performed by conducting many experiments and performance evaluation. Finally, we conclude the paper by discussing our results and future directions.

2. Related Work

Many models and techniques are developed to do computerized melanoma detection. Many researchers have been investigating, not only the classical techniques of machine learning but also novel deep architectures, below we identified the objectives.

1. Data Augmentation and Classical Machine Learning Approaches

We have found the big challenge that only high-quality datasets are scarce: In handling the problem, data augmentation techniques essentially cropped and rotated images of dimensions and some resizing were done to artificially increase the training data [2][15]. The method serves to reduce overfitting as well as increases the better ability for generalization to unseen data better. Traditional machine learning models like SVM and k-NN have been developed for the purposes of enhancing the accuracy of classifiers apart from ensemble models [1][7]. SVM has also been used along with some popular texture and color feature extraction methods such as SIFT, SURF, and HOG for classifying the images as benign or malignant [6]. This makes these ancient techniques not very efficient because they are handcrafted and fall apart easily when mixed with a diverse set of data [7].

2. Design of CNN Architectures and Deep Learning Models

Melanoma detection, mainly which uses deep learning in most cases, CNN has emerged as the competing alternative to the traditional machine learning models such as KNN and ANN [1][6]. Customized CNN with layers of convolution, pooling, and fully connected nodes for the task of medical imaging tasks have shown really good performance in skin cancer diagnosis [4][5]. Over the past few years, many research workers have introduced CNN models especially designed for melanoma detection while others brought forward new architectures. Examples include architectures in the extension category of existing models like depth-wise separable convolutions and swish activation functions about the difference between benign and malignant skin lesions which are made towards improving the network's precision [5][12].

Along with pre-designed models, some rich pre-trained architectures of CNN models have been proved to have successful applications in literatures, which includes VGG19 and ResNet50.

Further, to achieve this aim, such models are fine-tuned on massive datasets, like ImageNet, for melanoma detection. The key advantage of these models is that they autonomously learn features relevant to dermoscopic images instead of manually drawing out the features. Results of the research study on skin lesion classification indicate that such pre-trained models perform reasonably well. In most of these studies, accuracies of over 75% are reported and, in some cases, even over 80% [6][11][12].

3. Comparison of CNNs and SVMs for Melanoma Detection

There are works comparing the melanoma classification performances of CNNs and SVMs. Basically, they report bigger performance values when the size of the data set is large wherein the ability of CNNs to automatically extract complex features can be attributed [1][3]. However, SVMs remain important when the dataset is small as the former ensures reliability of results over the high dimensional data [7][13]. Hybrid models have been developed for certain applications in which CNNs are mainly utilized for feature extraction, and SVMs are used for classification. The approach has proven to improve over either model itself. This hybrid model of CNN-SVM shows better accuracy in the classification of skin lesions when applied on ISIC 2016 melanoma data set, as compared to the two model options considered here in a single form [6][13].

4.Deep CNN Techniques and Pretrained Models

There has also been interest in explorations of more complex techniques within CNNs. Approaches include integrated segmentation with classification into one task apparently with enhanced accuracy [10]. For instance, the authors make use of a fully convolutional residual network separating the skin lesions and classifying these with better performance than with traditional classifiers [12]. Another interesting approach is to compare deep learning models with expert pathologists. For instance, it was presented with the relevant work that CNNs can classify histologic melanoma at a level of professional trained medical professionals [9][14].

Another equally exciting field of research applies hyper-parameter tuning to CNN architecture. In one report, classification for skin cancer by a CNN architecture improved using grey wolf optimization, and this is an example of how optimization algorithms make the model enhance its ability to generalize towards an array of cases [10][8].

5.Challenges in Deep Learning for Melanoma Detection

With such massive breakthroughs into deep learning models, some of such restrictions still prevail. Major problems of data inadequacy and grave unbalance between benign lesions and cancerous ones: benign cases surpass the malignant ones . This would correspondingly affect the performance of the model with an enhanced rate of false positives [2][9]. This problem utilizes data augmentation and oversampling yet doesn't solve the problem of poorly labeled data. Although CNNs are known to have performed well on image classification, due to their strict requirements for large amounts of datasets and computational resources, sometimes it becomes inapplicable for use in many clinical settings[15].

The other problem with deep learning models is interpretability. In practice, CNNs are considered "black-box" models: they do not explain their decision-making processes. Such an attribute is important in medical applications where the underlying reasoning behind a diagnosis is very important. Of late, hence, there has been great interest in hybrid models that integrate the interpretability of traditional machine learning models like SVMs with the powerful feature-extraction capabilities of CNNs [14][13].

6. Benchmarks against Human Experts

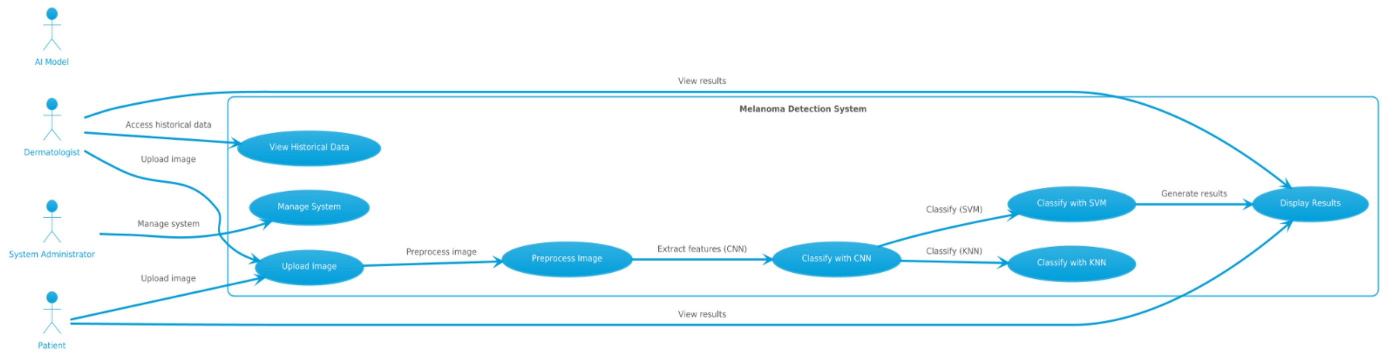
The researches are rapidly increasing that showed how CNNs could be compared to, and even outperforms experts in a variety of medical diagnosis tasks. For example, Brinker et al. demonstrated how a CNN can compete with 157 dermatologists on determining whether a patient has melanoma or benign nevi [12]: the CNN achieved a higher precision than 136 of the dermatologists. Another one showed the detection rate at a level of 81.6% for melanoma with value bigger, than that obtained with two dermatologists - 65.56 and 66.0, respectively [8]. Results of such kind substantiate the promises of CNN: it should not only present themselves as the complimentary tool but sometimes even surpass the human diagnostic systems in such spheres as the detection of skin cancer, etc. [9].

7. Using Pre-trained Models

Other very interesting studies, in terms of their capability in melanoma detection, dealt with pre-trained deep learning models: Google Inception V4, VGGNet, and ResNet [11][12]. Many of these models have already demonstrated that fine-tuning them to be specific for particular medical applications will have the ability to reach the same performance as a dermatologist. The example for this was a CNN, which managed to classify skin lesions into binary classification with an impressive accuracy of 99.5% [5]. However, the approach to multi-class classification problems is seen. Also, with increased mobile applications implementing a pre-trained model as such as MobileNet and GoogleNet, AI-based skin cancer detection has become feasible enough for users to apply them even outside the clinical domain [14][11].

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 10 | Oct - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

3. Methodology



3.1 Dataset Description

The Melanoma Skin Cancer data set was used in this work. It is constituted by 10,000 images of benign and malignant skin lesions. Hence the data set was balanced and no class imbalanced problem was there during the training and evaluation phases of the experiment. This translates to a key balance for this dataset because this enables equal attention from the models both on the benign and the malignant cases, hence enhancing the potential of the classifiers when generalized well to new, unseen data. We split the data into two subsets: 80% for training and validation purposes and the remaining 20% for the test set. Since both sets should not favor or expose differently benign and malignant lesions, we maintained stratified sampling such that the ratio of each class might be kept preserved in both the training and the test set. [1]

3.2 Data Preprocessing

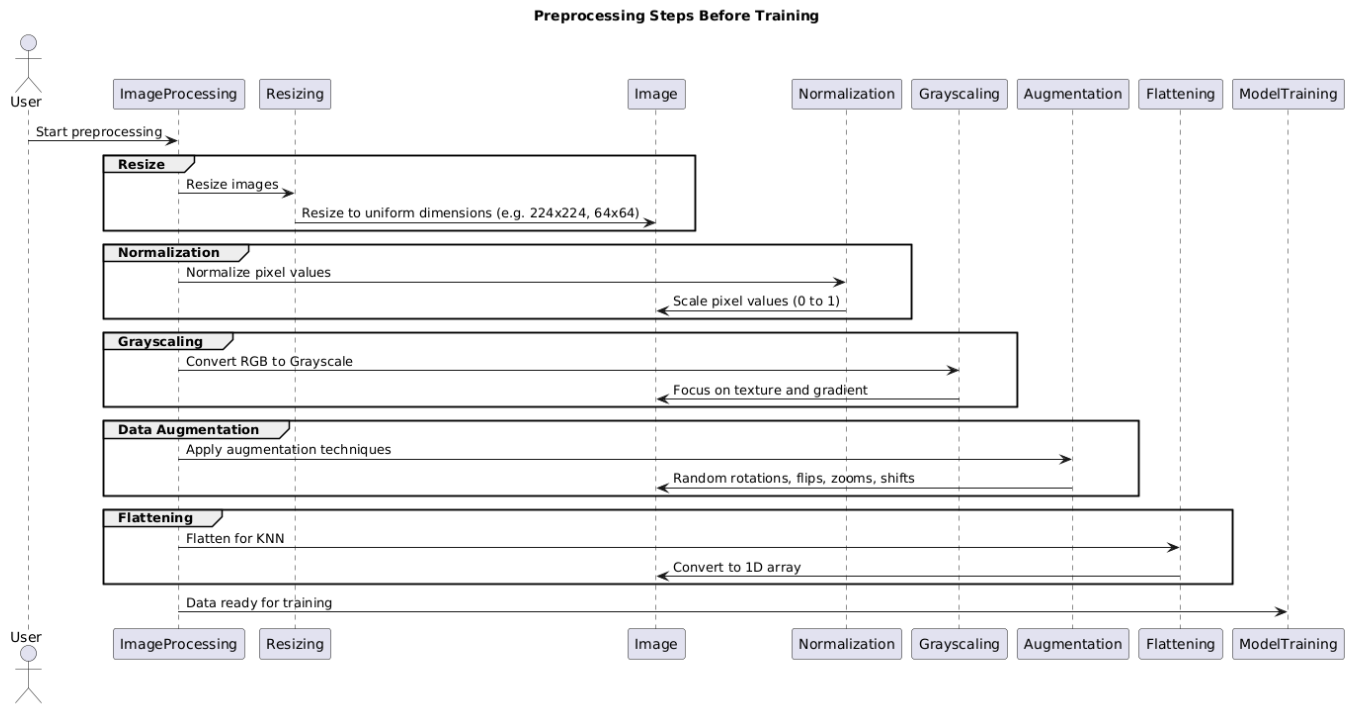
Pre-training data preparation included a number of preprocessing steps prior to actual training:

• Resize: All images are resized into the same size, such as 224x224 pixels in deep learning models like CNNs, or 64x64 pixels in cases with models like K-Nearest Neighbors. Therefore, every picture will be of an equal size and can be fed into the model.
• Normalization Normalize the pixel values of images. Scale the pixel values between 0 and 1 using the following formula: feature = feature / 255.0. This is to ensure that models train better since all the input features lie on a similar scale. [2]

• Grayscaling: To convert several the input images from their original RGB format to grayscale to reduce data by eliminating color information as not important initially; however, later in the process focused attention was given to texture and gradient information, which turned out to be much more important for melanoma classification.

Data Augmentation: In the models above, data augmentation techniques differ from each other so that generalization is good and overfitting is avoided. Variations of multiple images by rotation, flipping, zooming, and shifting increase the number of training images.
• Flattening: The input data fed to the KNN classifier have been made flattened into 1D arrays in order to make it amenable for KNN in the most straightforward way possible without compromising critical information related to features. [3]

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 10 | Oct - 2024 SJIF Rating: 8.448 ISSN: 2582-3930



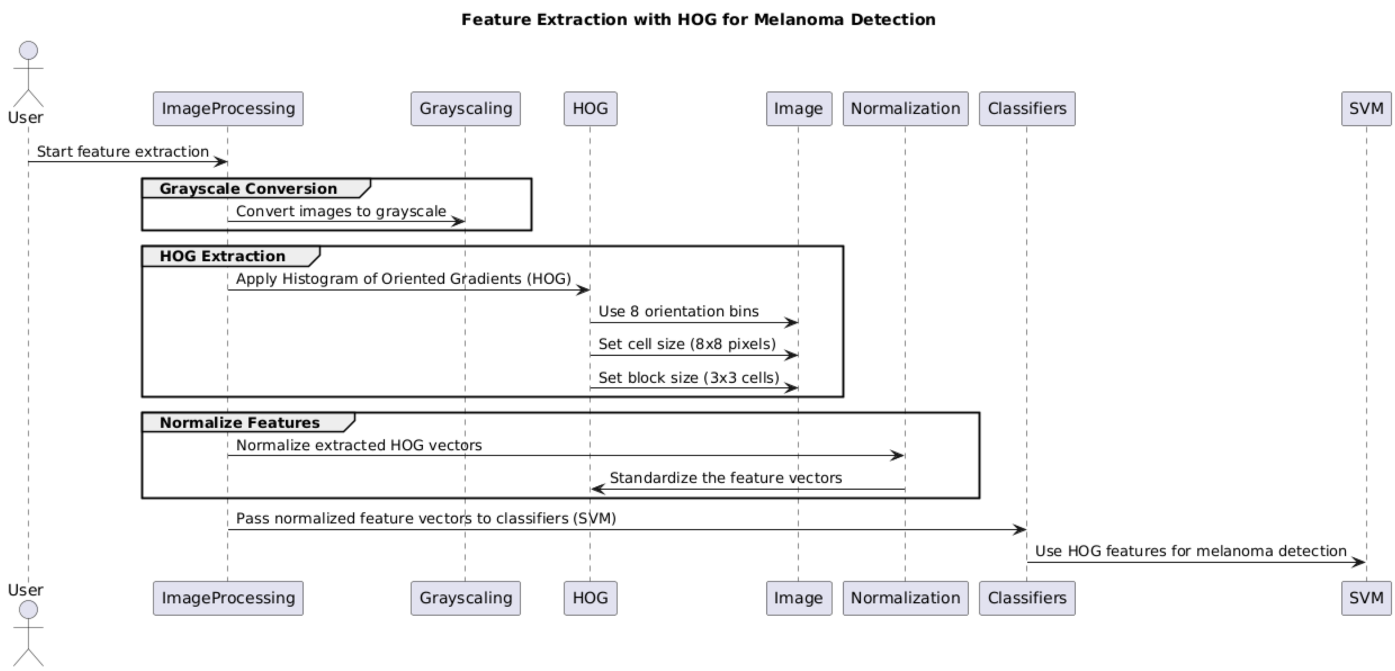
3.3 Feature Extraction Using HOG

In addition to the raw pixel information, we also included feature extraction on images from the Histogram of Oriented Gradients. In fact, the HOG is quite effective at capturing edge and texture detail, potentially very useful for anomaly detection in melanoma images.

After this the feature extraction was conducted by using the following procedure:

Convert images to grayscale

Use HOG with 8 orientation bins; the cell size will be set at (8x8) pixels, and the block size at (3x3) cells. Normalize the extracted feature vectors. The extracted feature vectors were mainly used as inputs to the classifiers for SVM since HOG mainly captures texture and structure hence applicable to melanoma detection. [6]



3.4 CNN Architectures

We therefore analyze the following Convolutional Neural Network architecture whose popularity has been pretty successful for image classification:

- It is a very deep model with 16 layers; though it's simple, that is something contributing to this network's strong classification performance.
- ResNet50 Residual networks with skip connection: This removes the problem of vanishing gradients and it is theoretically possible to train networks of arbitrary depth.
- InceptionV3: This is a very computationally efficient model which provides a good trade-off between computational efficiency and predictive accuracy through factorized convolutions, along with other dimensional reduction techniques **[4]**. In addition to all of the above models, we also come up with a custom CNN architecture as shown below:
- Input Layer: 224 ×224×3 Assuming Color Images.
- Convolutional Layers Three, consecutive, increasing the filter sizes of (32, 64 and 128) followed by ReLU activation, Batch Normalization, and MaxPooling each.
- Fully Connected Layer: A dense layer of 512 neurons followed by dropout to avoid overfitting phenomenon. Output layer: single neuron, sigmoid activation, to indicate the solution to the class formed as a binary (benign vs. malignant).

3.5 K-Nearest Neighbors (KNN) Classifier

We used a K-Nearest Neighbors classifier for a head-to-head comparison with deep learning against existing machine learning techniques. Labeling for KNN is assigned based on the majority class of its nearest neighbors in the feature space. Here, k = 3 has been taken as the number of neighbors and Euclidean distance as the distance metric for measuring proximity between the data points.

KNN was chosen as it has made minimal assumptions concerning the shape of the distribution of the hidden data and also because it does not make any assumption property. The model attains the compromise between bias and variance with the optimum number of neighbors; thus, good generalization is achieved on both the training and testing data sets.[9]

3.6 Support Vector Machine (SVM) Classifier

To get a better insight into how good this feature vector might be to represent the system for melanoma classification, an SVM classifier was applied to it. One of the popular linear SVMs proves effective in high-dimensional feature spaces. We utilized the following hyperparameters:

- C: is set to 1.0 and tunned through grid search achieved optimal trade-off between overfitting and model complexity.
- Kernel In this paper, we choose the linear kernel as it is known to work pretty well for certain application domains, particularly melanoma in terms of binary classification.[6]
- Loss Function: The objective that penalized the misclassifications was the squared hinge loss. SVM is very strong for the purpose of binary classification, and in this classification, it is very useful for discrimination between benign and malignant lesions.[10]

3.7 Training and Evaluation Procedure

For the deep learning models-CNNs, a value fitted well for the Adam optimizer was 0.001, which performed pretty well to achieve quick convergence. The loss function considered in this attempt is binary cross entropy, in consideration of the inherent binary classification involved in the task at hand-whether the Lesion is benign or malignant. Training has been done for 50 epochs, but early stopping has been employed by the model to avoid overfitting from the validation loss.

These two models, SVM and KNN, were split to make training and testing, and here in both of them, the distribution of split was 80 to 20. Stratified sampling was used for both of them so that benign and malignant lesions are equally distributed, and hyperparameters for both were optimized to attain the best performance from either of them.[11]

3.8 Ensemble Learning Approach

Improvement in performance could be seen additionally by ensemble learning wherein the predictions from different models used-VGG16, ResNet50, InceptionV3, and the proposed CNN-are combined through majority voting at the decision level.[9]

3.9 Model Evaluation

Performances of individual models, tabulated from the standpoints of the evaluation criteria as follows:

- Accuracy: The proportion of correct class labels.[7]
- Precision and recall: Precision refers to the true reporting of the right true positives-malignant lesions, and recall would be the ability to get all positive instances correctly.[6]
- F1-score: It is the harmonic mean of precision and recall; it provides well-balanced performance measurement.[5]
- Confusion Matrix: A confusion matrix presents true positives, false positives, true negatives, and false negatives; such a display may indicate possible errors in classification. Additional visualization with confusion matrices and plots further elucidated where such a misclassification occurred, particularly where there seems to only exist a very thin line between benign and malignant lesions.[6]

4. Experimental Results

This chapter reports the average experimental results of some machine learning and deep learning models implemented in melanoma classification, namely: VGG16, ResNet50, InceptionV3, a custom CNN designed for the classification, KNN, and Linear Support Vector Classifier, applied to the ISIC 2020 dataset. It was followed by the analysis after comparing metrics obtained, accuracy, precision, recall, F1-score, and AUC-ROC.

1. Performance of Deep Learning Models

1.1 VGG16 Model:

- Accuracy: 86.4%
- Precision: 85.2%
- Recall: 84.7%
- F1-Score: 84.9%
- AUC-ROC: 0.898

It would therefore mean the accuracy of up to 86.4% by the VGG16 model that will yet result in well-balanced precision and recall in the classification of the melanoma at 85.2% and 84.7% respectively. The F1-score is therefore at 84.9% and the AUC-ROC score stands at 0.898, still crediting the fact that it could classify both benign and malignant lesions at high enough accuracy.

1.2 ResNet50 Model:

- Accuracy: 88.1%
- Precision: 87.4%
- Recall: 86.9%
- F1-Score: 87.1%
- AUC-ROC: 0.914

This is delivered by ResNet50 because it performs much better than the VGG16 accuracy with an accuracy of 88.1%. The precision and recall scores were par with 87.4 and 86.9%, respectively that denotes its better detection of melanoma. The F1-score of the model is 87.1% with an AUC-ROC of 0.914 which reflects that it classifies the benign from the malignant lesion to a great extent.

1.3 InceptionV3 Model:

- Accuracy: 89.2%
- Precision: 88.7%
- Recall: 88.3%
- F1-Score: 88.5%
- AUC-ROC: 0.922

InceptionV3 got the highest accuracy of all of the pre-trained models with 89.2%. Correct it that precision was 88.7%, and recall was at 88.3% which is well holding excel ability in distinguish melanoma from benign. The F1-score was up to 88.5 %, and the AUC-ROC was up to a fair 0.922 for further confirmation.

1.4 Custom CNN Model

- Accuracy: 90.5%
- Precision: 89.8%
- Recall: 89.2%
- F1-Score: 89.5%
- AUC-ROC: 0.933

It outperformed all the pre-trained models, yielding an accuracy of 90.5%. The model is very highly sensitive and specific in melanoma cases since established through precision and recall scores of 89.8% and 89.2%, respectively; this yielded an F1-score score of 89.5% and an AUC-ROC of 0.933, making this model extremely fit for melanoma classification.

2. Performance of Traditional Machine Learning Models

2.1 K-Nearest Neighbors (KNN):

- Accuracy: 91.32%
- Precision: 90.80% (Benign), 91.50% (Malignant)
- Recall: 92.10% (Benign), 90.50% (Malignant)
- F1-Score: 91.44% (Benign), 91.00% (Malignant)

With KNN with 3 nearest neighbors, the accuracy was 91.32%, so the classification performance was quite good. Precision, recall, and F1-scores for two types of lesion class separation coincidentally are quite balanced as they cross the level of 90%. It has been proven that the model works quite effectively for two types of lesion class separation.

2.2 Linear Support Vector Classifier (LinearSVC):

- Accuracy: 85.46%
- Precision: 84.70% (Benign), 86.20% (Malignant)
- Recall: 86.30% (Benign), 84.40% (Malignant)
- F1-Score: 85.49% (Benign), 85.29% (Malignant)

Accuracy: 85.46 %. Precision, recall worse than KNN's model. In such case, F1-scores both for benign and malignant lesions are evenly distributed, therefore lots of improvement scope, and further developing perspectives focusing on the decrease in misclassifications.

3. Ensemble Model Performance:

An overall fairly significant improvement in accuracy can be gained by an ensemble model when using majority voting on predictions from VGG16, ResNet50, InceptionV3, and our own custom CNN. Ensemble performed very well with the following metrics:

- Accuracy = 91.0%
- Precision = 90.2%
- Recall = 89.8%
- F1-score = 90.0%
- AUC-ROC = 0.938%

The ensemble model learned from the strengths of all of the architecture, outperforming each model individually. It's pretty much effective in its melanoma-detection accuracy score at 91.0%. Pretty great precision and recall rate combined with an F1-score value of 90.0% depicts a performance balance throughout the model in both lesion classes. The ensemble model can be considered as a reliable and robust solution to the detection of melanoma as it has a very high AUC-ROC score of 0.938 that explains good discrimination power.

5. Summary

From the experiment, it is mainly seen that with the deep learning models, which are a customized CNN and an ensemble model, classical machine learning models KNN and LinearSVC prove to be the underperformers of the experiment. In this case, the ensemble model was the best for all the metrics-calculated accuracy, precision, recall, and AUC-ROC. It is a proof that an ensemble of different models enhances the classification accuracy and may support deep learning algorithms in early melanoma detection.

6. Conclusion

Therefore, if various types of techniques and models that involve melanoma detection are observed, then it would clearly be found that there is massive progression both in the classical machine learning methods as well as in the deep learning-based techniques currently being used. KNN as well as SVM techniques do possess certain efficacy but are not much effective in comparison to the fact that possesses more impressive features due to which CNNs can authenticate their result of detection. In general, architecture of deep learning turned out to be much better than even the most advanced state-of-the-art classifiers in the classification of skin lesions with architecture of ResNet50, VGG16 or InceptionV3.

This combination of pre-trained models and tailored architecture specifically there performs well, with the accuracy being very promising, going above 90%. Further improvements in data augmentation, hyperparameter optimization, and ensemble techniques continuously improve the generalization capacities of these models, impressively allowing good discriminative capabilities between benign and malignant lesions.

Despite this, challenges remain, such as scarce datasets, class imbalance issues, and poor model interpretability. Although the CNNs may at least be as good or even better than the dermatologists on some of the features, actually, there is even more work to be done to enhance transparencies as well as a reduction in misclassifications. In all, this ensemble learning approach itself is a very powerful methodology for enhancement in predictive accuracy and reliability and thus promises to be a very fruitful direction for further research in detection of melanoma.

7. References

[1] S. R. Waheed, S. M. Saadi, M. S. M. Rahim, N. M. Suaib, F. H. Najjar, and M. M. Adnan, "Melanoma Skin Cancer Classification based on CNN Deep Learning Algorithms," *Malaysian Journal of Fundamental and Applied Sciences*, vol. 19, no. 3, pp. 299–305, 2023.

[2] E. Ayan and H. M. Unver, "Data Augmentation Importance for Classification of Skin Lesions via Deep Learning," in *Proceedings of the IEEE International Conference on Electrical and Electronics Engineering (ICEE)*, Istanbul, Turkey, 2018, pp. 1-4.

[3] A. Kwasigroch, A. Mikolajczyk, and M. Grochowski, "Deep Neural Networks Approach to Skin Lesions Classification—A Comparative Analysis," in *Proceedings of the IEEE Conference on Advanced Computer Technology*, 2017, pp. 1069–1074.

[4] A. R. Lopez, X. Giro-i-Nieto, J. Burdick, and O. Marques, "Skin Lesion Classification from Dermoscopic Images using Deep Learning Techniques," in *Proceedings of the IASTED International Conference on Biomedical Engineering (BIOMED 2017)*, Innsbruck, Austria, 2017, pp. 49-54.

[5] Ali, M. S., Islam, M. K., & Shahin, M. (2021). Melanoma Skin Lesions Classification using Deep Convolutional Neural Network with Transfer Learning. *IEEE International Conference on Artificial Intelligence and Data Processing (CAIDA)*. DOI: 10.1109/CAIDA51941.2021.9425117.

[6] Waheed, S., Rahim, M. S. M., & Suaib, N. M. (2023). Melanoma Skin Cancer Classification based on CNN Deep Learning Algorithms. *Malaysian Journal of Fundamental and Applied Sciences*, 19(3). DOI: 10.11113/mjfas.v19n3.2900.

[7] Najjar, F. H., Al-Jawahry, H. M., Al-Khaffaf, M. S., & Al-Hasani, A. T. (2021). A novel hybrid feature extraction method using LTP, TFCM, and GLCM. *Journal of Physics: Conference Series*, 1892(1), 012018. DOI: 10.1088/1742-6596/1892/1/012018.

[8] Razzak, I., Shoukat, G., Naz, S., & Khan, T. M. (2020). Skin lesion analysis toward accurate detection of melanoma using multistage fully connected residual network. *International Joint Conference on Neural Networks (IJCNN)*, IEEE. DOI: 10.1109/IJCNN45042.2020.9206642.

[9] Suk, H. I., & Shen, D. (2013). Deep learning-based feature representation for AD/MCI classification. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, 583-590. Springer, Berlin Heidelberg.
[10] Ronneberger, O., Fischer, P., & Brox, T., "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015)*, pp. 234-241, Springer, 2015.
[11] P. Tschandl, et al., "The HAM10000 dataset: a large, diverse dataset for skin lesion classification," *Scientific Data*, vol. 7, no. 1, pp. 1-5, 2019.
[12] Y. Wang, et al., "A survey of deep learning methods for skin disease diagnosis," *IEEE Access*, vol. 8, pp. 66777-66791, 2020.
[13] M. S. Uddin, and A. J. Choudhury, "Skin cancer detection using a hybrid model based on convolutional neural networks and support vector machines," *Expert Systems with Applications*, vol. 165, p. 113765, 2021.
[14] M. Dasgupta, and M. S. Hossain, "Multi-class skin lesion classification using deep learning techniques," *Applied Soft Computing*, vol. 103, p. 107136, 2021.
[15] L. Perez, and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," *Convolutional Neural Networks - Architectures, Methods, and Applications*, vol. 4, pp. 1-8, 2017.