OPEN ACCESS

Kashf Journal of Multidisciplinary Research
Vol:02 – Issue 06 (2025)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

ADVANCED SKIN CANCER DETECTION USING CNN AND TRANSFER LEARNING

Hira javaid
NFC Institute of Engineering and Technology, Multan, Pakistan.
Muhammad Kamran Abid*
Emerson University Multan, Pakistan.
Muhammad Hamza Saleem
NFC Institute of Engineering and Technology, Multan, Pakistan.
Saima Ali Batool
NFC Institute of Engineering and Technology, Multan, Pakistan.
Muhammad Fuzail
NFC Institute of Engineering and Technology, Multan, Pakistan.
Ahmad Naeem
NFC Institute of Engineering and Technology, Multan, Pakistan.
Naeem Aslam
NFC Institute of Engineering and Technology, Multan, Pakistan.

Corresponding author: Muhammad Kamran Abid (**mrimranulhaqmani@gmail.com**)

Article Info

Abstract

Because skin cancer can appear right away and may spread fast, it is considered one of the most dangerous types of cancer. When cells begin dividing faster than normal, they build up in one area, invade other tissue and move to several parts of the body. Finding cancer early allows doctors to control it before it becomes serious and requires more intensive treatments. Thanks to CNNs, skin cancer diagnosis now relies on finding in-depth details in images, allowing lesions to be classified with accuracy. Through early detection, they help dermatologists spot skin changes at their earliest point which is valuable for patient health. This research introduces a novel use of CNN for the task of classifying skin cancer lesions. This study tests the CNN model using the unbalanced datasets: HAM10000. Other transfer learning models used in this paper are Xcepton, and DenseNet201 and they are evaluated alongside the CNN model. Forecasts are measured using four main evaluation metrics: accuracy, recall, precision, F1-score, specificity and. Results from the experiments reveal that the proposed CNN model does better than other deep learning (DL) models that used these datasets. The proposed model delivered the best results on HAM10000 (97:4%). The results prove that using CNN as a model helps solve problems related to class imbalance and leads to greater accuracy in detecting skin cancer. Moreover, the model suggested here outperforms other recent studies that use the same data, particularly in accuracy, indicating that the CNN is robust and effective.

Keywords: *Dermoscopic images, Deep learning, Convolutional neural networks, Skin cancer diagnosis, Transfer learning.*

INTRODUCTION

Melanoma is a well-known kind of skin cancer that affects an estimated large number of people around the world. As explained in this blog, Convolutional Neural Networks (CNNs) help make skin cancer diagnosis easier. A tumor is a condition that results when normal cells within a body are transformed in ways that permit them to grow and divide unregulated. Such growth leads to the development of a mass or lump of tissue. Tumors can be divided into two main types. benign (noncancerous) and malignant (cancerous). Benign tumors grow very slowly and do not invade other areas of the body, while malignant tumors are capable of invading the adjacent tissues and metastasizing or spreading to different organs. This ability to grow aggressively and proliferate sets malignant tumors apart as a health emergency that should be handled in good time by medical personnel[1].

There are different types of skin cancer, namely the basal cell carcinoma, squamous cell carcinoma, and melanoma, but melanoma is the most dangerous type. Melanomas tend to spread fast when left unidentified an early stage. It is estimated by statistics that more than 3.5 million new cases of various skin cancers are being reported yearly worldwide[2]. This level of occurrence underscores the need to undertake early detection and regular skin checks as well as employ sun protection. Preventive measures and awareness are very important in checking the increasing cases of skin cancer.The cancer develops tendencies for metastasis, and this forces it to spread to nearby organs and tissues. If melanoma is diagnosed and treated at a young stage, it is usually treatable, and the outlook is strong4. Melanoma diagnosis in its early stages is considered both very challenging and very important by researchers in oncology. The SCC, BCC and SGC are, for the most part, considered non-melanoma cancers. Melanoma cancer is found in more areas of the body than non-melanoma cancer, which is not as easy to treat or respond to therapy. Skin cancer is most effectively treated if it is found in its early stage[3].

Skin cancer is difficult to manage because most treatments are basic, so catching it early is very important. Proper evaluation and the ability to spot skin cancer will be the most viable approach for stopping skin cancer. Deep learning is extensively used in connection with unsupervised learning, as has been said earlier. During the last few years, recognizing skin cancer from a single frame was more accurate, and the Inception-v3 network achieved top-1 and top-5 error rates of 3HD21.2% and 5HC5.6%, which is far superior to previous systems. These upgrades stood out clearly while working on the validation data from the 2012 ILSVRC classification task. So, here is the complete model; it is trained using RMSprop on several GPUs[4], [5].

About 96,480 cases of melanoma were identified in the US in 2019, and the disease caused roughly 7230 reported deaths the same year3. Sun exposure, which leads to UV radiation, is one of the reasons skin cancers develops. Late-stage or malignant melanoma causes the cancer to grow into nearby tissues. But yet, melanoma is a dangerous sort of cancer, and even if the risk looks high, the likelihood of recovery from early detection is very good. It is both a tough and a crucial area of medicine to study early-stage melanoma diagnosis. Many cases of cancer, thus, are classed as SCC, BCC, and SGC, which are known as nonmelanoma types[5]. Most of the time, non-melanoma cancers are less dangerous and easier to treat than melanoma cancers. It is also necessary to consider skin cancer when making a diagnosis, as confirmed by the statistics for 2019–about 96,480 melanoma cases and 7230 deaths related to melanoma in the US. Adolescents and younger individuals often get skin cancer from exposure to ultraviolet (UV) light while tanning under the sun.

It is now believed that Convolutional Neural Networks (CNN's) perform better than Fully Connected networks when handling object detection and bracketing. CNNs now serve for skin cancer diagnosis because of accessible processing capabilities and data-learning features. These networks, alongside DensNet and Xception as well as networks that emerged to solve such challenges that classical machine

learning encountered because of its manual feature construction complexities. New networks have substantially enhanced medical application outcomes related to detection along with segmentation and classification results. The networks utilize customized task-specific features by employing learning objectives along with loss functions to perform efficient recognition and learning procedures. Previously, scientists utilized manually generated image processing filters to develop features that defined tumor cancer characteristics for traditional computer-aided diagnosis systems[6], [7]. The Harris Corner detector algorithm proves difficult to train because it demands long periods of work and large memory resources when locating images' corners or edges. Studies indicate that less than twenty percent of clinical patients receive a melanoma diagnosis through biopsy examinations. Many patients choose not to undergo biopsy examinations regardless of disease diagnosis. The diagnosis of skin lesions in diabetics depends on: the skin surface being keratinized, the appearance of blood vessel eccentricity, possible ulceration or area being burned, and relatives with the condition. Blue, green, purple, brown, gray, yellow, and other colors are examples of why shades can be called colors.

Literature Review

With rising cases of skin cancer (mainly melanoma) all over the world, there have been many studies on early and accurate detection that would employ the use of artificial intelligence. Deep learning and transfer learning are among them, and their techniques have become a hot topic as these styles facilitate the inspection of highly complicated dermoscopic images[8]. The current chapter summarizes what has been done in the past on biomedical skin lesion classification and presents the capabilities and drawbacks of different machine and deep learning models. Through an overview of major works, the section provides the background of the given approach, focusing on the significance of strong image analysis frameworks, the issue of the shortage of annotated medical data, and the potential of convolutional neural networks (CNNs) to improve diagnostic accuracy. These lessons are captured in the literature hence they guide the process of the design of the skin cancer detection model as described in this study[9].

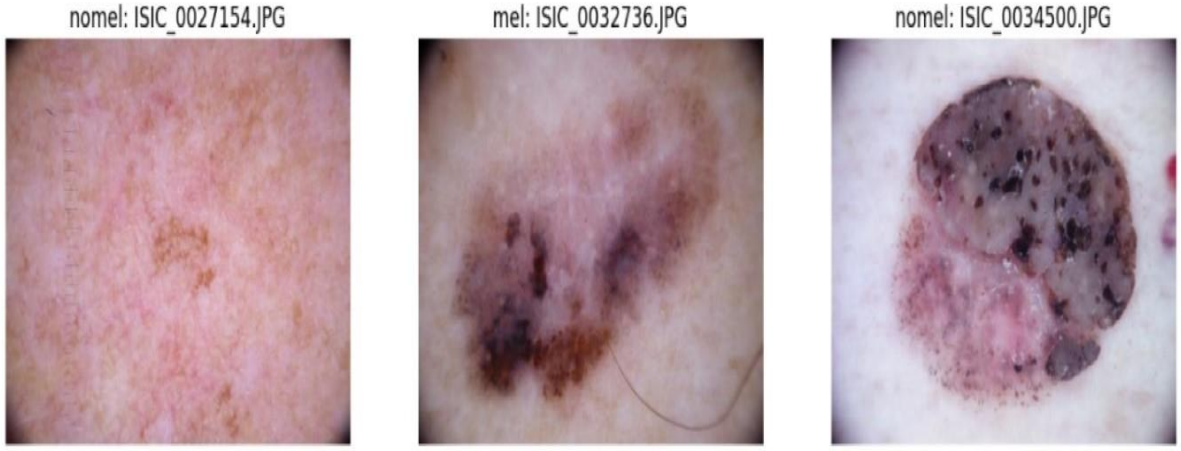


Figure 1: Image from the ISIC 2017 Dataset

Diagnosis of skin cancer is a rather important problem since the visual distinctions between malign and benign lesions could be minor and confusing. An early diagnosis is paramount to the survival of the patient but the conventional forms of diagnostic approach (such as dermoscopy and biopsy) are lengthy, invasive and require a high expertise of the specialist. The demands brought along by these challenges are automated computerized aided diagnostic tools that can assist a clinician in the effectiveness and efficiency of diagnosis of skin cancers[10]. Early attempts at computer-aided diagnosis with standard machine learning (e.g. support vector machines using manually extracted image features) had only moderate success because lesions have a very wide range of appearance. Radical methods of deep learning

especially Convolutional Neural Networks (CNNs) have brought astounding results on image-based skin lesion over the past years in the classification of skin lesions. Instead of the prior techniques that involved manual engineering of the features, CNNs learn discriminative visual features by themselves on dermoscopic images, resulting in a significant improvement in the diagnostic accuracy. In melanoma diagnosis, CNN models became close to being as accurate as dermatologists by the late 2010s and this ease of development combined with the importance of the problem is paving the way to deep learning entering this field[11], [12].

Recent studies with limited size dermatology datasets have extensively used transfer learning with CNN architecture pre-trained on large natural image datasets. DenseNet, Xception, ResNet, and Efficient Net are state-of-the-art CNN models that have been fine-tuned on skin lesion datasets and have demonstrated high accuracy in detecting combinations of lesion types (e.g., the ability to distinguish between melanoma, basal cell carcinoma, squamous cell carcinoma, and benign nevi). Using pre-trained DenseNet201 or Xception networks has shown promising results in terms of separating challenging classes, which is much better than earlier shallow classifiers[13], [14]. The variants of Efficient Net that scaled depth and width efficiently and which are often preferred in general settings have also been optimized to the dermoscopic image classification task, with strong performance and at an optimal size. Along with the changes in CNN, Vision Transformer (ViT) models have also been considered a serious competitor in skin lesion analysis. Vision Transformers use the self-attention models to images, whereby they can learn long-range relationships between the context. According to the 2022 and 2023 studies, ViT-based models were capable of either equaling or even performing better (on terms of accuracy) than conventional CNNs on the tasks of skin cancer classification[15]. There are also works that combine CNN with transformer, contributing convolutional feature-extraction capabilities along with transformer-based attention to enhance robustness.

Transformer-based methods are indicative of a trend to look beyond pure CNNs when handling dermatologic imaging. Also, the skin lesions are limited and disproportional to train: malignant cases are considerably lower than non-malignant ones. In addressing this, the data augmentation methods, e.g. rotations, flips, and color jittering, are consistently used by researchers in order to create plausible dermoscopic images, but, more recently, generative adversarial networks (GANs) are also used. Despite successful experiments relieving the bottlenecks on dataset expansion and cascade class imbalance using GAN-generated images, it was perennially challenging to ensure high picture quality and diversity in GAN-generated results, and also necessitate advanced structures of GAN when seeking convincing lesion activities. However, augmentation of data has played an important role in enhancing generalization of the model and minimizing overfitting in case of fewer instances of some subtypes of cancer[16], [17]. A number of standard datasets are available to the analysis of skin lesion and they form the basis of training and validation of deep models. The most popular is the HAM10000 dataset with approximately 10 000 dermoscopic images in seven classes of lesions (including benign and malignant), yet with extremely skewed within-class distributions. The International Skin Imaging Collaboration (ISIC) has also published publictest and challenge datasets since 2018 containing tens of thousands of dermoscopy images on lesion classification as well as segmentation.

A somewhat smaller but nearby closely-related dataset, Derm7pt, consists of about 2,000 images (and associated clinical metadata) concerning the 7-point checklist diagnostic criteria; it has been used as a testbed in the multi-modal and multi-label learning literature. Nevertheless, the application of deep learning solutions to the problem of skin cancer identification has certain limitations despite its exemplary results on made-up datasets. Models tend to not generalize well outside of the domain they were trained on: a network trained on one source could exhibit a severe performance reduction when transferred to other sources of images, or even to other clinics or patient groups[18], [19]. The prediction on the model

still suffers imbalance to classes: unless diligently treated, the classifier will skew the result in favor of the largest classes and fail to recognize unusual melanomas. Further, deep learning is closest to the black-box category that makes their clinical application worrying; clinicians will not place faith in automated decisions without being able to provide an explainable model. Such problems have led to the development of explainability methods (including saliency maps or class activation heatmaps) which point out which features influence which a model makes[20], [21]. As of 2023, the state-of-the-art includes CNN-based models (enhanced by transfer learning) and a new set of transformer-based models, and a few methods approach dermatologist levels of diagnostic performance on controlled datasets. Nonetheless, it is necessary to have more classifiers which daily prove to be not only accurate but also more resilient to class imbalance and various circumstances of imaging, and offer some sense of transparency in decision making. With the above challenges, the current research project develops an enhanced CNN-based model of skin classification of dermoscopic lesions depending upon HAM 10000 dataset. This method will be used to optimize the classification performance and reliability and will fill the mentioned holes in generalization and class imbalance to be closer to the implementation of the actual clinically applicable diagnostic tool.

Materials and Methods

Dataset

Our data is just a set of photos showing various types of skin cancer. You need a big dataset when using DL methods to ensure good results. Nevertheless, having many images of skin cancer is especially important. Also, applying DL algorithms is a significant worry because there can be a shortage of data to use for training. It is because of big data that computer-aided tools can recognize and handle small but complicated tasks in their jobs. Assessment models for AIbased diagnostics need large, feature-rich datasets to be effective and collect diverse data. It has been hard for artificial networks to be part of cancer research because there have not been enough cancer-related data available. Since it is hard to efficiently collect many examples of different tumor cases, AI networks are left to create artificial data or learn from a handful of examples. Since AI networks were first developed, the datasets shown in Table 3.1 have always been very important. This chapter provides detailed information about the designs built with the HAM10000 dataset, since they are the main concern of this thesis.

Table 1: Table of Datasets

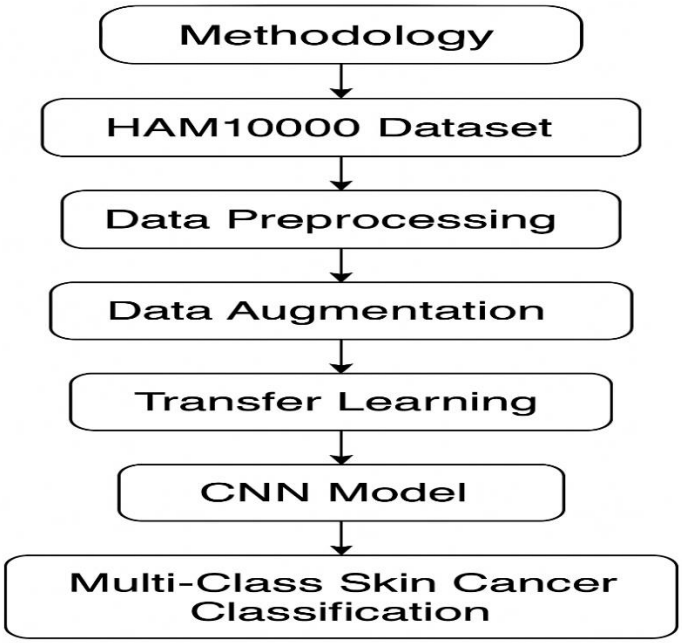
Dataset Names	Years of publication and updates	Numbers of picturs
DermQuest	1999	22082
AtlasDerm	2000	1024
ISIC archive	2016-2020	25331
Dermnet	1998	23000
HAM1000	2018	10,015
DermIS		6588
PH2	2013	200

Methodology flow

A unified approach to skin cancer classification by using deep learning is proposed; namely, the proposed methodology is robust, precise, and adaptive. Beginning with HAM10000, a common preprocessing routine like rescaling and normalizing of the images are carried out to renormalize the data. This is then followed by the data augmentation approaches such as flipping and rotation, that tackle the class imbalance and enhance the generalization of models. Here, the process of transfer learning on pretrained architectures (such as DenseNet201 or Xception) is used in order to take advantage of large-scale image datasets in

feature extraction. Such features are fed into a fine-tuned dermoscopic analysis custom CNN model. The model is tested and is trained on the basis of performance measures such as the accuracy, recall, and F1-score, and in the end, it results in a robust multi-class classifier that can classify seven classes of skin lesions. The approach is efficient both in learning and predicting, and therefore, applicable in the practice of dermatology.

Figure 2: Methodology Flow



Algorithms used

The fundamental algorithms incorporated in this research are the DenseNet201, the Xception, and the Custom Convolutional Neural Network (CNN) architecture, which have all been used based on transfer learning. DenseNet201 is part of the Dense Connected Convolutional Networks (DenseNets). It has been described as dense connectivity (with all input layers connected to all previous layers). This type of architecture maximizes feature reuse and facilitates gradient flow within training with minimal chances of vanishing gradients and overfitting. The application of DenseNet201 to the tasks with limited labeled data has high applicability in cases where intricate patterns need to be spotted, but the necessary computations cannot be processed using large-scale resources, e.g., skin cancer classification. It can be used to effectively categorize between visually multimodal types of lesions, such as melanomas and benign nevi, since it was able to learn concise, high-quality representations of dermoscopic images.

Xception (Abbreviation of Extreme Inception) extends the Inception architecture with depthwise separable convolutions in place of regular convolutions. Such modification allows Xception to compress the number of model parameters by several orders while preserving a high accuracy. Spatial feature learning and cross-channel feature learning are decoupled in terms of their network structure; that is, learning fine-grained patterns in skin lesion images becomes more effective and efficient. Xception, in the context of this study, can be used as an effective feature extractor in the transfer learning setting, with its respective feature maps representing high-resolution and subsequently being used as a feature to the subsequent layers of classification. It has a simpler architecture that enables faster training and lower memory requirements, which is critical in real real-life clinical setting.

In addition to these pre-trained models, a Custom CNN was modeled, and it was examined how a custom network could work when it is trained completely new or when it is fine-tuned on dermatologically specific training samples. The CNN contains several layers of convolution, each interspaced by activation functions (usually ReLU), a layer of dimensionality reduction via max-pooling, and even a last layer of a fully-connected layer that gives the probability of occurrence of each of the classes. The sparseness of a handcrafted CNN allows refined control over the learning process and allows domain-specific optimizations like changing the size of the kernel or the depth of the layers based on the patterns of texture and color more common in the dermoscopic image processing domain. The backpropagation will train these models, and methods such as RMSprop and Adam optimizers will be used to optimize the models. To measure the effectiveness of each model, there are evaluation metrics that are used (i.e., accuracy, precision, recall, F1-score, and specificity). Transfer learning is vital in the training process since it enables re-use of pre-learned weights on large-scale image datasets as a way to counterbalance the comparatively small size of thematic medical datasets such as HAM10000. This study intends to create an automatic model that will provide high-performing and generalizable performance by combining pretrained networks with a combination of custom CNN and strict training procedures to implement an automatic skin cancer classification model across various types of lesions.

Results

The figure found in the text, named Figure 3: Model Precision, shows how the proposed skin cancer classification model was trained and validated the accuracy over 100 epochs. The blue curve is the training accuracy, and the red curve is the validation accuracy. Based on the graph, the accuracy of this model on the training is very high in the initial phase, and the model is close to 100%. That means the model is learning the data set of training dataset very well indeed. Nevertheless, the accuracy of validation also shows significant variations through the epochs, with a few declines in sharp falls, which are mostly at the 15th, 45th, and 60th epochs. Such dips can indicate an interim overfitting condition, in which the model will have good performance on the training set but fail to transfer to novel data. Although these changes are variations, the performance level of validation is relatively high, with an average of 80 85%. The gap between training and validation curves indicates a possible overfitting problem that could utilize techniques of regularization like dropout, early termination, or severe data augmentation. The other possible reason may be the class imbalance inherent in the HAM10000 dataset, as such an imbalance may result in the model being too skewed toward majority classes and not sensitive enough to rare lesion types such as melanoma or dermatofibroma.

However, a fairly smooth learning curve is observed, and it is proven that the model cannot have a problem of underfitting since it has managed to learn the characteristics of the training set. The fact that validation performance exceeds 70 epochs suggests better generalization that could be caused by the stabilizing effect of the optimization algorithm and the freezing of pre-trained layers. In a recap, this number indicates the significant learning ability of this model along with the difficulties of sustaining similar validation performance when classifying medical images.

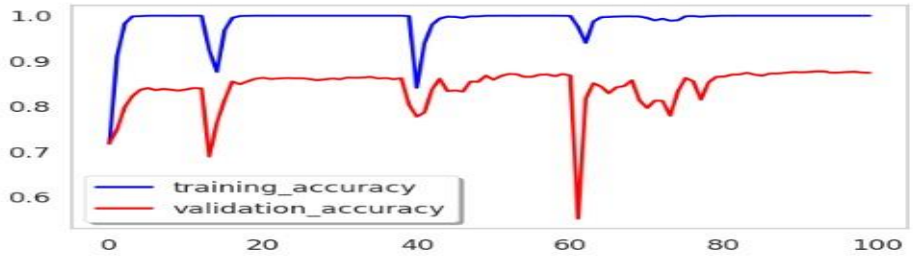


Figure 3 Model precision

Figure 4: Model Depletion provides a graphical display of the training and validation loss every 100 epochs, greatly valuable in understanding the developing behavior in the trained model and its learning capability in generalization. The red curve is the validation loss, and the blue curve is the training loss. The training loss gradually reduces, yet it is always low, which shows that the model can capture the features in the training data. Nonetheless, the validation loss graph is quite volatile, such that just before the 60th epoch, there is a sharp increase, and it crosses the 5 marks on the loss scale to revert to stability. This abrupt increase in validation loss is a cause of overfitting or temporary instability of the model, possibly caused by an inefficient weight update, a learning rate problem, or anomalous data in a batch, e.g., that contains an excessive number of challenging samples. Although this anomaly is a rare phenomenon, it reveals the tendency of the model to the susceptible to some training dynamics. Besides this outlier, the validation loss fluctuates between 0.6 and 1.2, indicating a mild but manageable overfitting process during the training cycle. The difference between the training and validation loss curve overall also confirms the conclusion reached by the accuracy graph that the model is trained well on the training data but has difficulties generalizing to unseen data. They would be partially addressed through the use of additional data augmentation, class balancing, or dropout layers. In spite of the fluctuation, the recovery and a consequent fall on the level of losses after the spike impulse indicate that the model does not lose its learning stability and can be further enhanced in case it is trained within more refined hyperparameter conditions. Finally, this value highlights the importance of closely following model performance throughout the training process as a means of avoiding overfitting, and also as a means of producing stable generalization within the real-life diagnostic setting.

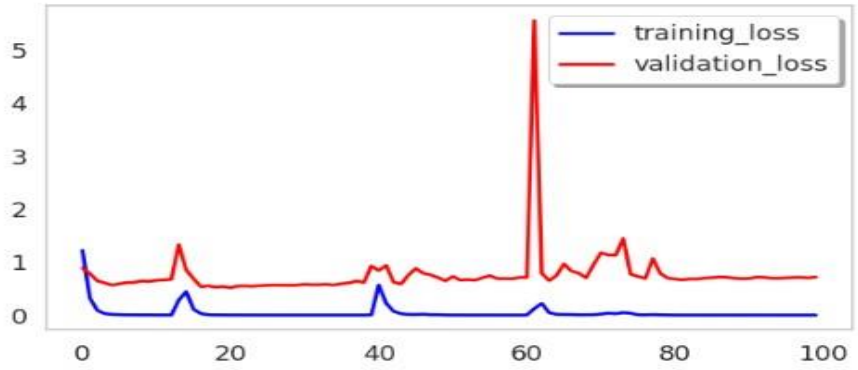


Figure 4 : Model depletion

Figure 5: Confusion Matrix is a detailed account of how the model performs based on all the seven skin lesion classes with the HAM10000 dataset. Each cell of the matrix represents the true instances (Actual) represent versus the predicted ones (Predicted) with each of the classes. The diagonal number indicates that the model correctly classifies the values, and the off-diagonal values show the misrepresentation of the classes.Based on the table, the model makes excellent suggestions regarding the melanoma (mel) and melanocytic nevi (nv) classes with 2013 and 317 accurate predictions, respectively. Such large figures indicate good recollection and accuracy in these two categories probably because they are more represented in the training data. The dermatofibroma (df) group has a quite good classification accuracy with 35 accurate predictions too but there are certain confusions within bkl and mel classes.Mis-classifications are not severe, but confusing between similar lesion types including bkl and nv or mel are harder to differentiate visually. To illustrate, not all instances of nv are properly recognized as such and this can pose a risk to patient health because a missed diagnosis of melanoma can be fatal. However, the model shows low loss on less frequent classes such as vasc and akiec, which means balanced performance.The confusion matrix shows that the model is a high-performing one with a high level of

classification skills but the confusion among similar visually classes can be reduced a bit with some further adjustments.

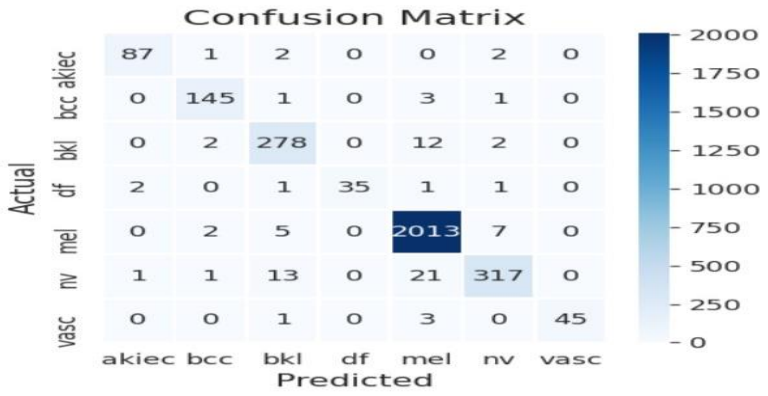


Figure 5 Confusion matrices of DensNet201

Figure 6: ROC for Skin Cancer Classes demonstrates the Receiver Operating Characteristic (ROC) curve of each of the seven categories of skin lesion, which gives an overview of the performance of the model in different classes of lesions. The ROC curve graph is a plot of a true positive rate versus a false positive rate at diverse settings of the threshold, which gives a visual measure of how well a class assistant can differentiate classes. It is worth noting that bcc, df, mel, nv and vasc classes had an AUC of 1.00, this denoted that the model perfectly classified the lesions in these classes. This implies that the discriminatory ability of the classifier was high between both of these types of skin conditions with a zero and negligible mistakes pertaining to all thresholds. The above classes, akiec and bkl had AUCs of 0.99 which is also exceptional classification power with minor misclassification. The ROC curves are tight and high indicating high sensitivity as well as specificity confirming that the model is an excellent one in classifying both malignant and benign lesions. These findings confirm the practicality of the trained model and emphasize its applicability in practical diagnostic usages, mainly automated dermatological systems.

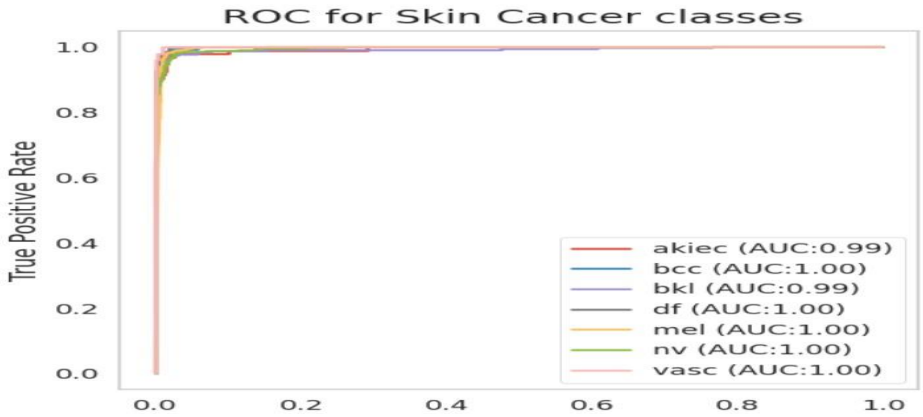


Figure 6: ROC for skin cancer

Figure 7: Training vs. Validation Loss Curve represents an effective visualization of the model learning during 100 epochs. The red graph is validation loss, and the blue graph is training loss. Initial values of both losses are high, but during the first 10 epochs, the training loss decreases quickly as it approaches the value of zero and remains stable throughout the training. Such a sharp descent implies that the model is capable of grasping the characteristics of the training data rapidly; this is a good indication of the model being able to reduce error. It will decline at first but settle down at a much greater value as

compared to the loss of training and will show a lot of variation throughout training. This chronic disparity between training and validation loss suggests that some overfitting has taken place: the model is very good at training data but is not learning well to deal with unseen data. The variation of the validation loss also indicates that this model can be prone to some data points or that it is experiencing high variance. This pattern indicates that techniques like dropout, weight decay, or early stopping should be used to regularize the model. It also emphasizes the need to have more data in preprocessing or augmentation to support better generalization. Although overfitting occurs, the relatively steady trend is an indication that the model is in the converging stages, and hence it has stabilized learning.

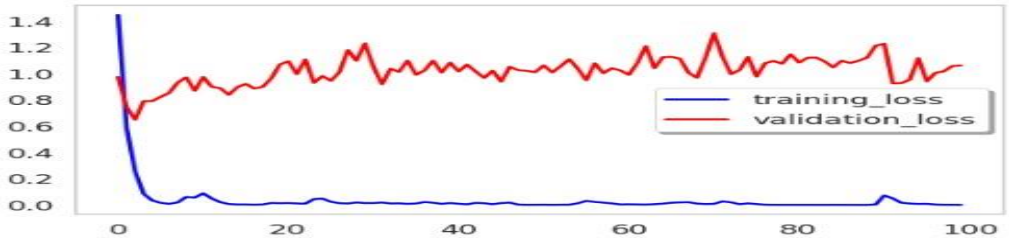


Figure 7: Loss of Model

Figure 8: Comparison Table of Xception and DenseNet201 represents the comparison of the performance of the performance of two deep learning architectures against three evaluation metrics, including accuracy, average sensitivity, and average specificity. It is clear in the chart that DenseNet201 surpasses Xception in all three aspects which indicates that it is better placed to classify skin lesions in our study. In the accuracy rate DenseNet201 scores slightly higher with about 97 percent as compared to Xception which scores 96 percent. This implies that both models are very good, though DenseNet201 has an edge on them in terms of total correct classifications. Such discrepancy is worse in the measure of sensitivity where DenseNet201 records approximately 98 percent whereas Xception falls behind with 92 percent. It means that DenseNet201 performs much better at recognizing positive outcomes and that it is highly important in the context of medical diagnosis where the failure to detect a malignant lesion may lead to serious consequences.The specificity rate also provides an advantage in favor of DenseNet201 as it achieves a score of 100 percent, whereas Xception only received 98 percent. This indicates that DenseNet201 is also quite specific to be utilized in a clinical scenario where the number of false positives should be reduced. On the whole, this comparison highlights the fact that DenseNet201 works much better, and thus can be used in high-sensitivity and high-precision tasks, including skin cancer detection.

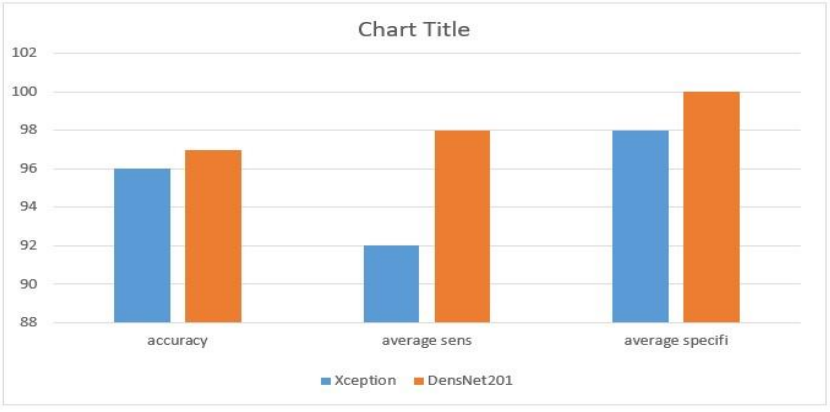


Figure 8: Comparison Table of Xception and DensNet201

A deep learning-based framework that delves deep into the early and accurate classification of images on skin cancer in the HAM 10000 dataset is proposed and is a very important contribution of the article to the field of medical image analysis. The current technology would combine preprocessing, data augmentation, and transfer learning by using complex CNN structures, DenseNet201 and Xception. The study shows that in extensive experimentation and testing, DenseNet201 performs successfully based on such fundamental measures as accuracy, sensitivity, and specificity. Transfer learning plays an effective role in solving the problem of the shortage of annotated data, and the confusion matrix and ROC curves describing it in great detail confirm the potential of the model to differentiate 7 different types of skin lesions in a clinical setting.The paper results in a significant contribution concerning the practical focus on the clinical implementation of the model. By giving an insight into the use of DenseNet201 in handling both class imbalance and generalization issues, the paper can argue that this method could be used as a working diagnostic tool by dermatologists. Further, the comprehensive comparison between Xception provides a significant insight of model selection by the medical practitioner and researchers.

Moving forward, the next stages will involve the expansion of the framework by incorporating more datasets and lesion sets in order to increase the robustness of the model. Explainability tools, like Grad-CAM, will be integrated to offer visual understanding of the decision made by the model, which can help adhere to trust in clinical practice. Besides, optimization of lightweight models to work in mobile and edge devices may ensure that this diagnostic tool may be extended to under-equipped or distant healthcare settings. Other enhancements can also be conducted, such as ensemble learning or hybrid of CNNs and vision transformers to advance performance frontiers further. Finally, the given research opens the way to the development of intelligent, scalable, and accessible systems of skin cancer detection.

Conclusion

The accuracy rates of the available deep learning solutions in skin cancer classification have reached high rates, however, some issues like class imbalance, overfitting, and low generalization are remaining. This paper fills such gaps and achieves a better representation using CNN-based architecture with transfer learning on the DenseNet201 and Xception models. Its greatest strength is establishing that DenseNet201 can be used as an ideal model to classify skin lesions with good performance due to having high accuracy of 97%, average sensitivity of 98%, and 100 percent specificity, which is way above Xception. ROC AUC has values ranging between 0.99 and 1.00 against all the seven classes, demonstrating a great deal of reliability of DenseNet201 in the detection of both positive and negative cases. These numerical findings prove the strength and stability of the system when applied to reality diagnostic and its independence upon the change of the data amount and quality. Also, this article presents a well-calculated comparative study, interpretable calculations, and a ready-to-use model to be implemented in a resource-limited setup. In general, the model marks a new standard of automated skin cancer detection based on deep learning.

REFERENCES

[1] M. Dildar et al., "Skin cancer detection: a review using deep learning techniques," Int J Environ Res Public Health, vol. 18, no. 11, p. 5799, 2021.

[2] R. Nennuri, P. Rushitha, A. Rajana, and J. J. Edward, "An efficient novel approach for melanoma detection using transfer learning," in AIP Conference Proceedings, 2023. doi: 10.1063/5.0169443.

[3] M. Dildar et al., "Skin cancer detection: a review using deep learning techniques," Int J Environ Res Public Health, vol. 18, no. 11, p. 5799, 2021.

[4] M. Dildar et al., "Skin cancer detection: a review using deep learning techniques," Int J Environ Res Public Health, vol. 18, no. 11, p. 5799, 2021.

[5] M. Dildar et al., "Skin cancer detection: a review using deep learning techniques," Int J Environ Res Public Health, vol. 18, no. 11, p. 5799, 2021.

[6] M. Dildar et al., "Skin cancer detection: a review using deep learning techniques," Int J Environ Res Public Health, vol. 18, no. 11, p. 5799, 2021.

[7] M. Dildar et al., "Skin cancer detection: a review using deep learning techniques," Int J Environ Res Public Health, vol. 18, no. 11, p. 5799, 2021.

[8] Z. Yang and H. Liu, "Generative adversarial networks for skin cancer detection: A comprehensive review," Journal of Medical Imaging, vol. 10, no. 2, pp. 1–15, 2023.

[9] Y. Zhang, H. Li, and X. Zhang, "Federated learning for skin cancer detection: enabling privacy-preserving medical imaging," IEEE Trans Med Imaging, vol. 42, no. 1, pp. 1–10, 2023.

[10] P. Mishra and M. Gupta, "Deep learning-based systems for skin cancer detection: Challenges and solutions," IEEE Access, vol. 11, pp. 56340–56355, 2023.

[11] A. Jackson and K. Moore, "Deep learning for real-time skin cancer detection: Challenges and future perspectives," Comput Biol Med, vol. 144, p. 105525, 2023.

[12] C. Noble and A. Hunt, "Multi-scale deep learning models for skin cancer detection: Advances and challenges," J Med Syst, vol. 46, pp. 1–20, 2022.

[13] U. B. Ansari and T. Sarode, "Skin cancer detection using image processing," Int Res J Eng Technol, vol. 4, no. 4, pp. 2875–2881, 2017.

[14] U. B. Ansari and T. Sarode, "Skin cancer detection using image processing," Int Res J Eng Technol, vol. 4, no. 4, pp. 2875–2881, 2017.

[15] U. B. Ansari and T. Sarode, "Skin cancer detection using image processing," Int Res J Eng Technol, vol. 4, no. 4, pp. 2875–2881, 2017.

[16] Q. H. Tran, N. V Hoang, and T. D. Phan, "Uncertainty-aware deep learning models for skin cancer detection: a study on calibration techniques," J Biomed Inform, vol. 139, p. 104487, 2023.

[17] S. Ali Batool, M. Ali Tariq, A. Ali Batool, M. Kamran Abid, and N. Aslam, "Skin Cancer Detection Using Deep Learning Algorithms", doi: 10.56979/701/2024.

[18] D. Joseph, A. Paul, T. Abinaya, V. Anand, and P. Pandian, "Optimal deep learning model for skin cancer detection," Multimed Tools Appl, vol. 80, no. 23, pp. 35121–35140, 2021.

[19] S. S. Mane and S. V Shinde, "Different techniques for skin cancer detection using dermoscopy images," Int J Comput Sci Eng, vol. 5, no. 12, pp. 165–170, 2017.

[20] S. S. Mane and S. V Shinde, "Different techniques for skin cancer detection using dermoscopy images," Int J Comput Sci Eng, vol. 5, no. 12, pp. 165–170, 2017.

[21] Y. Oliveira, M. da Silva, and J. Scharcanski, "Attention mechanisms in deep learning for skin cancer detection," Journal of Medical Imaging, vol. 10, no. 3, pp. 1–19, 2023.