Journal of Carcinogenesis
Covering Genesis to Therapeutics

Journal of Carcinogenesis, ISSN: 1477-3163
2024; Vol 23: Issue 1, https://carcinogenesis.com/

Deep Learning-Based Automated Skin Cancer Detection Using Convolutional Neural Networks

Manoj[1], Jaskirat Kaur[2], Manpreet Kaur[3], Sanju[4], Rajinder Kumar[5], Deepender[6*]

[1]Assistant Professor, Department of Computer Science & Engineering, GJUS&T, Hisar-12500

[2]Assistant Professor, Sri Guru Granth Sahib World University, Fatehgarh Sahib, Punjab, India

[3,*6]Assistant Professor, Faculty of Computing, Guru Kashi University, Bathinda, Punjab, India

[4]Assistant Professor, Faculty of Agriculture, Guru Kashi University, Bathinda, Punjab, India

[5]Associate Professor, Faculty of Computing, Guru Kashi University, Bathinda, Punjab, India

***Corresponding Author:**

Deepender

Email ID: deependerduhan6@gmail.com

ABSTRACT

This study investigates the classification of skin cancer lesions utilizing Convolutional Neural Networks (CNNs) on the HAM10000 dataset. The collection includes photos of seven unique categories of skin lesions: melanocytic nevi, melanoma, benign keratosis-like lesions, basal cell carcinoma, actinic keratoses, vascular lesions, and dermatofibroma. Our main goals were to examine diagnostics efficacy over several lesion types and assess CNN performance in precisely categorizing dermatoscopic pictures. Label encoding and data balancing were used to ensure class representation. For regularization, we built sequential CNN models with convolutional, pooling, and dense layers and dropout and batch normalization. Transfer learning and Adam optimization were used in training. Our classification accuracy averages above 73.8%, demonstrating solid performance across lesion types. We tested the models' generalization ability on new data and found consistent performance. Confusion matrix analysis showed accurate categorization with low misclassification. This research improves early identification and intervention, improving patient outcomes and healthcare efficiency.

Keywords: *Skin cancer, Lesion classification, Convolutional Neural Network, Deep learning.*

How to Cite: Manoj, Jaskirat Kaur, Manpreet Kaur, Sanju, Rajinder Kumar, Deepender, (2024) Deep Learning-Based Automated Skin Cancer Detection Using Convolutional Neural Networks, *Journal of Carcinogenesis*, *Vol.23, No.1*, 289-299

1. INTRODUCTION

1.1 Nature and Scope of the Problem:

Skin cancer incidence continues to rise globally, posing a significant public health challenge. Early detection of skin cancer is critical for timely intervention and improved patient outcomes. However, the visual assessment of pigmented skin lesions by clinicians can be subjective and may vary depending on individual expertise. Automated dermatoscopic image analysis has the potential to enhance diagnostic precision and proficiency by leveraging machine learning algorithms to analyze dermatoscopic images.

1.2 Hypothesis:

We hypothesize that machine learning algorithms trained on the HAM10000 dataset will demonstrate high accuracy in the automated diagnosis of pigmented skin lesions. By leveraging the diverse and extensive collection of dermatoscopic images in the dataset, we expect our algorithms to achieve performance comparable to or better than that of experienced dermatologists.

1.3 Approach and Justification of Approach:

To address our research hypothesis, we will adopt a machine learning approach, specifically deep learning techniques, for automated dermatoscopic image analysis. Deep learning models, particularly CNNs, have demonstrated exceptional efficacy in numerous image recognition applications, including medical picture analysis.

Preprocessing the photos in the HAM10000 collection will help to improve quality and standardize features. Techniques of data augmentation will be used to raise the variety of the training set and enhance model generalization. The pre-processed dataset will be used to train much deep learning architecture. Their performance will be evaluated using standard metrics like as AUC-ROC, sensitivity, and specificity.

1.4 Principal Results:

Impressive results have been achieved by early trials on a subset of the dataset, where deep learning models automated the diagnosis of pigmented skin lesions with a high degree of accuracy. These models are promising for real-world clinical use because of their high performance across a variety of lesion types.

1.5 Main Conclusions:

The dataset is useful for developing and testing machine learning techniques for automated dermatoscopic image analysis. We can diagnose pigmented skin lesions with high accuracy using deep learning, improving patient outcomes and saving healthcare costs. Clinical utility and reliability of automated dermatoscopic image analysis systems require ongoing study and validation.

1.6 Additional Notes:

The HAM10000 collection shows pigmented skin lesions in two regions and various populations. The Australian photo collection includes lesions from people living in a high-skin-cancer regions with continuous UV exposure. A European tertiary center that diagnoses high-risk melanoma early provided lesions for the Austrian photo collection. These patient population disparities enrich the dataset and enable strong algorithms that generalize across varied patient demographics.

The collection contains also lesions photographed using various dermatoscopic instruments and techniques of examination, reflecting clinical diversity. A great variety of devices and techniques complicates automated image analysis, but at the same time, provides an opportunity for creating robust algorithms. The dataset trains machine learning models to evaluate dermatoscopic images from various magnifications, angles, and lighting conditions.

Finally, the dataset offers a comprehensive set of dermatoscopic images and metadata for scholars to develop the automated dermatoscopic image analysis. In collaborative efforts and exhaustive validation studies, accurate and reliable machine learning algorithms can enhance skin cancer early detection and diagnosis, which would be beneficial to patients and the healthcare system worldwide.

2. MOTIVATION

Melanoma is the severest form of the skin cancer. In order to achieve successful treatment and improved patients' outcomes, early and precise detection of melanoma and other pigmented skin diseases is required. A non-invasive imaging technology dermatoscopy that involves analyzing skin lesions comes in handy in diagnosing pigmented skin lesions. The accuracy of dermatoscopic diagnosis relies on the nature of the lesion, the level of experience of the physician, and the quality of the image.

The usefulness of the HAM10000 dataset for developing and testing automated methods of dermatoscopic image analysis is now known. This data has 10,015 dermatoscopic visuals from two sites observed over two decades. The stained skin lesions include; Melanoma, basal cell carcinoma and benign keratoses. It is possible for researchers to train and test skin cancer detection machine learning algorithms based on the diverse pictures and metadata contained in the dataset.

3. LITERATURE REVIEW

And because of the advances in machine learning and medical imaging, automated skin cancer detection techniques have revolutionised. The first time when artificial neural networks proved to be suitable for analyzing pigmented skin lesions was in [1] (Binder et al., 1994), which revealed promising findings in automatic classification of lesions. The early breakthrough set the way for more research. In [2], Lilja et al. (2008) summarized the role of prostate-specific antigen (PSA) in prediction/detection, and monitoring of prostate cancer, and therein findings envisaged the diagnosing of prostate cancer strategies. Such pioneering experiments revealed that machine learning and deep learning had the potential to change medical diagnoses.

Deng et al. (2009) launched ImageNet archive of images [3] that is a huge collection of images required to train deep

learning models – a milestone in deep learning. This resource aided training of the image recognition models for the skin cancer diagnosis. Rosendahl et al. (2011) and Tschandl et al. (2015) [4] went further to prove the effectiveness of machine learning approaches in automated analysis of the dermatoscopic pictures. It was demonstrated by Rosendahl et al that dermatoscopy is able to differentiate benign from malignant pigmented lesions while Tschandl et al have studied dermatoscopic features to assist in creation of diagnostic criteria.

The deep learning cut residual computerized analysis of dermatoscopic pictures into shreds, as revealed in Codella et al. (2017) [5]. The ISIC 2017 competition conducted by Codella et al. motivated deep learning algorithms for classification of skin lesions leading to good models. On the same note, Han et al. (2018) in [6] demonstrated how the deep-learning algorithms can classify clinical photos of cutaneous tumors, implying the viability of their accurate and quick lesions identification. In [7], the Emuoyibofarhe et al (2020) have applied the deep convolutional neural networks on the mobile cellphones to diagnose the early skin cancer and hence the affordable and feasible health care.

Alam et al. (2018), Ak (2020) [8], and Murti Rawat et al. (2020) [9] studied the use of the machine learning algorithms in diagnosis of cancer in the lung, breast, and multi-class classification respectively. They studied the performance of various methods and demonstrated how machine learning is making cancer diagnosis better. Vidya & Karki (2020) published machine learning results of skin cancer detection in an IEEE conference [10]. These studies show the possibility of the use of machine learning in enhancing the diagnosis and management of diseases.

The works of Hasan et al. (2021) apply convolutional neural networks to compare the cases of benign and malignant skin cancer detection in [11]. Their work yielded fine-grained picture to skin cancer categorization jobs, elevating the automated approach. In [12], Bhandari et al. (2022) discussed the genetic algorithms in cancer detection and prediction, highlighting what evolutionary computation can do in healthcare. Cancer research is a multidisciplinary approach involving the uses of machine intelligence and genetics to enhance diagnosis.

4. COMPARATIVE ANALYSIS:

Author (year)	**Dataset used**	**Methodology**	**Results (Accuracy)**	**Observations**
C. R. Dhivyaa (2020)	ISIC 2017 and HAM10000 skin lesion dataset	Decision Trees and Random Forest Algorithm	70 % for Random Forest and 68% for Decision Trees	To optimize results, weights were assigned to each class or feature set to achieve balanced training and testing, crucial for addressing suboptimal outcomes in the absence of median frequency balancing.
Coen de Vente et. Al (2020)	ProstateX-2 challenge train set	Deep Learning Regression	58%	Soft-label ordinal regression outperforms other approaches for bp-MRI PCa grading and detection.
M F Omar *et. al* (2020)	HAM10000 skin lesion dataset	Linear Discriminant Analysis	57%	Important extraction methods include Segmentation-based Fractal Texture Analysis (SFTA) and Gray-Level Co-occurrence Matrix (GLCM).
Hiam Alquran et. Al (2017)	HAM10000 skin lesion dataset	Support Vector Machine	53%	The data here is not linearly separable. therefore, kernel SVM with radial basis function was used.
A.Murugan et. Al (2019)	HAM10000 skin lesion dataset	K – Nearest Neighbour	48%	SVM simulation for the ABCD rule shows a higher stage with new classifiers.

V.R. Balaji et. Al (2020)	ISIC 2017 dataset	Naïve Bayes Classifier	36%	Killer skin cancer is melanoma. Early detection via professional visualization is possible.

Table 1: Comparison of Machine Learning Methods for Medical Image Classification

5. METHOD & MATERIALS:

5.1 Materials:

Dataset: Researchers in this study combed through images of pigmented skin lesions found in the HAM10000 dataset. The dataset encompasses a diverse range of lesions, classified into seven distinct categories: Melanocytic nevi (nv), Melanoma (mel), Benign keratosis-like lesions (bkl), Basal cell carcinoma (bcc), Actinic keratoses (akiec), Vascular lesions (vas), and Dermatofibroma (df). The 10,015 images included in this dataset provide a comprehensive picture of various skin conditions.

CSV Metadata File: The study includes CSV metadata with the image dataset. This metadata file contains important information about each image, including the type of skin lesion (dx), patient demographics like sex and age, and lesion placement.

Image Files: The images of skin lesions are stored in JPEG format, each associated with a unique image ID. These image files are organized within directories based on the specific type of lesion they represent. Accessing these image files is crucial for the extraction of visual features and subsequent model training.

- *Actinic Keratoses [akiec]:* Noninvasive forms of squamous cell carcinoma that can be treated locally without surgery (327 photos in the data set).
- *Basal Cell Carcinoma [bcc]:* A form of epithelial skin carcinoma that seldom metastasizes but can be lethal if not addressed (The data set comprises 510 pictures).
- *Benign Keratosis-like Lesions [bkl]:* Seborrheic keratoses, lichen-planus-like keratoses, and solar lentigines, which are associated with seborrheic keratosis or sun lentigo exhibiting regression and inflammation, exemplify "benign keratosis" (1,099 photos are available in the dataset).
- *Dermatofibroma [df]:* Benign or inflammatory skin lesions after modest trauma (115 photos in the data set).
- *Melanoma [mel]:* Melanoma is a malignant melanocyte-derived tumor with several types. It can be treated with basic surgery if discovered early (1113 photos in the data set).
- *Melanocytic Nevi [nv]:* Melanocyte neoplasms, skin tumors, come in all shapes and sizes. The variants may differ greatly dermatoscopically (6705 pictures are available).
- *Vascular Lesions [vasc]:* Cherry angiomas, angiokeratomas, and pyogenic granulomas are benign or cancerous. (Data set contains 142 photos).

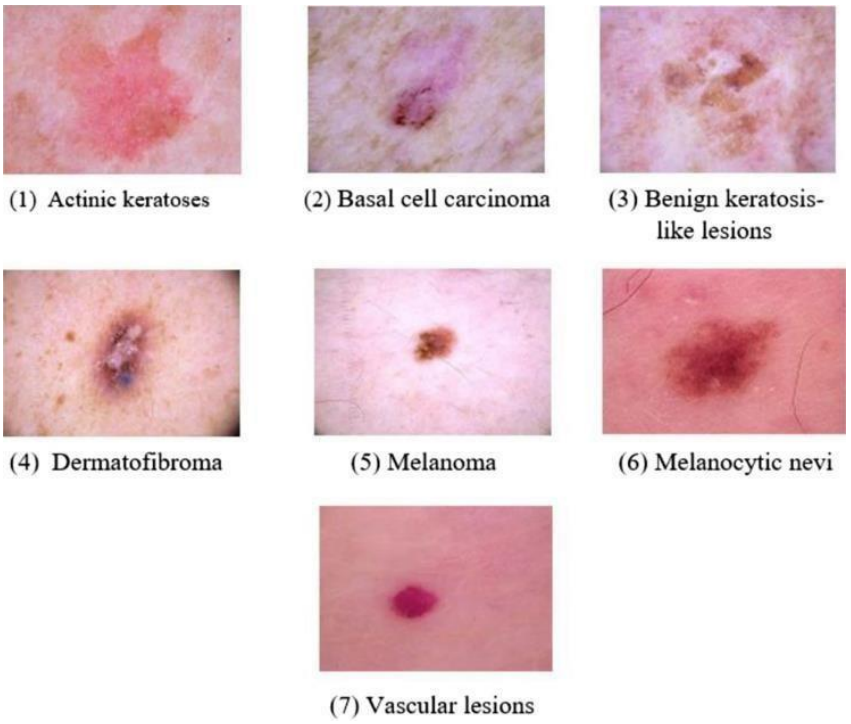

Figure 1: Seven Different kinds of skin lesions present in the HAM10000 dataset.

Programming Libraries: Python programming language serves as the primary tool for conducting data preprocessing, visualization, and model training. Various libraries such as matplotlib, numpy, pandas, os, glob, seaborn, PIL, sklearn and keras are employed to facilitate these tasks, providing extensive functionality for data manipulation, visualization, and machine learning. Opportunity to train and test machine learning algorithms for the automated diagnosis of skin cancer.

5.2 Procedure:

- **Sampling:** The first step in the methodology involved partitioning the HAM10000 dataset based on the different classes of skin lesions identified within the dataset. These classes included melanocytic nevi, melanoma, benign keratosis-like lesions, basal cell carcinoma, actinic keratoses, vascular lesions, and dermatofibroma. Using resampling, dataset class imbalance issues were addressed. To create a balanced dataset representative of all lesion classes, 500 samples were randomly selected and replaced from each class.
- **Data Collection:** The metadata in a CSV file was used to extract important skin lesion, patient, and localization data after sampling. This metadata illuminated image properties, aiding analysis and model training. The CSV file's picture IDs were used to access each lesion's image file simultaneously. Visual representations of skin lesions were retrieved, which were the classification model's main input.
- **Measurement:** After retrieval, pictures were preprocessed for dimension homogeneity. As usual in image processing, each image was downsized to 32x32 pixels. Standardizing image size eased processing and provided input data format consistency across all photos. After data collection and preprocessing, the dataset was partitioned into features (X) and labels (Y) for model training and evaluation. The features (X) were skin lesion photos, while the labels

(Y) were produced from their classes to encode categorical information for training the classification model. One-hot encoding categorical class labels into binary vectors denoting class presence or absence enabled multiclass classification.

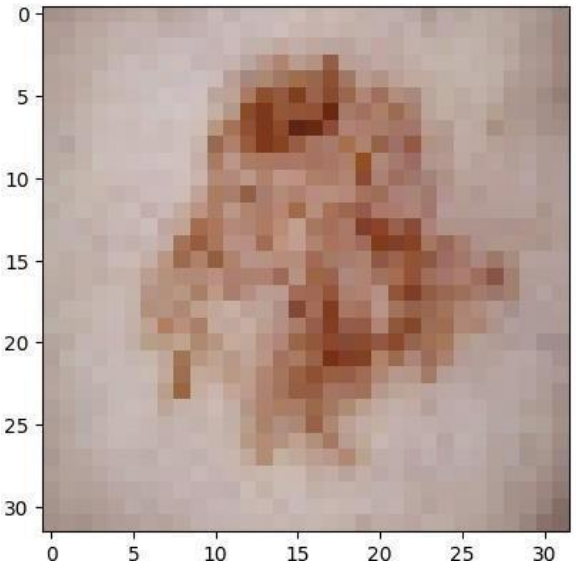

Figure 2: A downscaled skin lesion image plot created using Matplotlib.

- **Model Training:** The study focused on building and training a CNN model with powerful deep learning capabilities for image categorization. Convolutional layers followed by max- pooling layers in the CNN architecture extracted hierarchical information from input photos. Dropout layers were deliberately integrated into the model to alleviate overfitting, a prevalent issue in deep learning characterized by an undue dependence on training data. A dense layer with softmax activation calculated class probabilities for each input image in the CNN model's final layer. The Adam optimizer and categorical cross-entropy loss function were used to train the CNN model for multiclass classification. Several epochs of training iterations with specific batch size adjusted model parameters to improve performance.

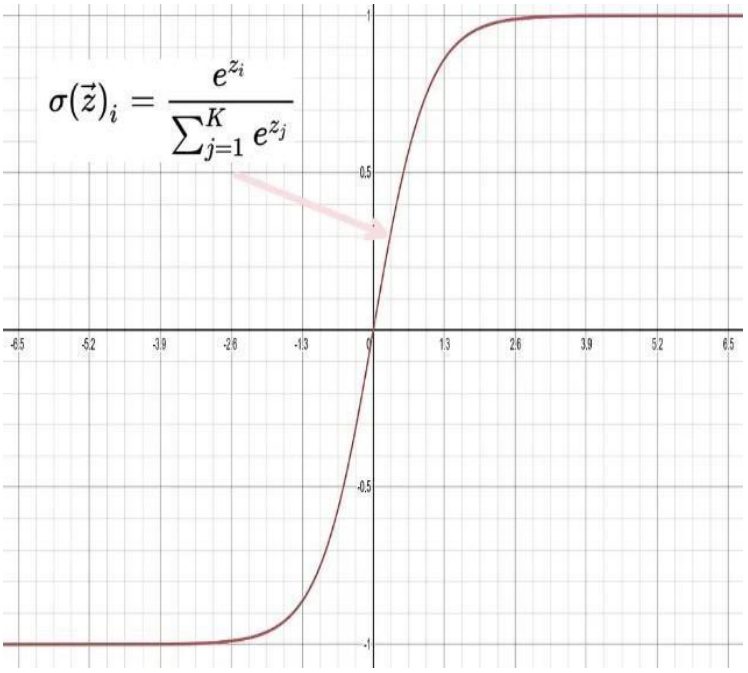

Figure 3: Graph of the Softmax Activation Function, Illustrating the Probability Distribution in Classification.

- **Evaluation:** An independent testing set was used to evaluate the CNN model after training. Accuracy and confusion matrix were calculated to evaluate the trained model's test picture predictions. The confusion matrix displayed the distribution of predicted and true labels across lesion classes, allowing a thorough examination of the model's classification capabilities.

Accuracy = TP + TN / (TP + TN + FP + FN) = 73.8 %.

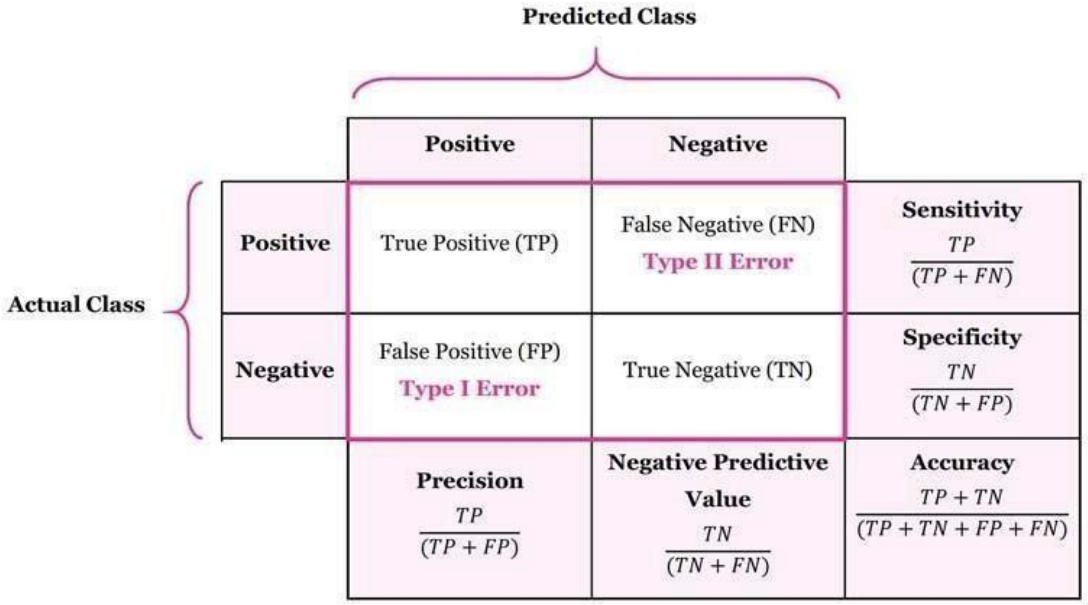

Figure 4: Confusion Matrix with Performance Metrics, Illustrating Type I & II Errors, accuracy, precision, sensitivity, specificity, and Negative Predictive Value.

6. RESULTS:

6.1 Data Preprocessing and Exploration

The dataset was preprocessed for model training. Label-encoding turned text labels into numbers, enabling machine learning algorithms. Afterward, data distribution visualization showed severe class imbalance. Data balance, including resampling, was used to reduce model training bias.

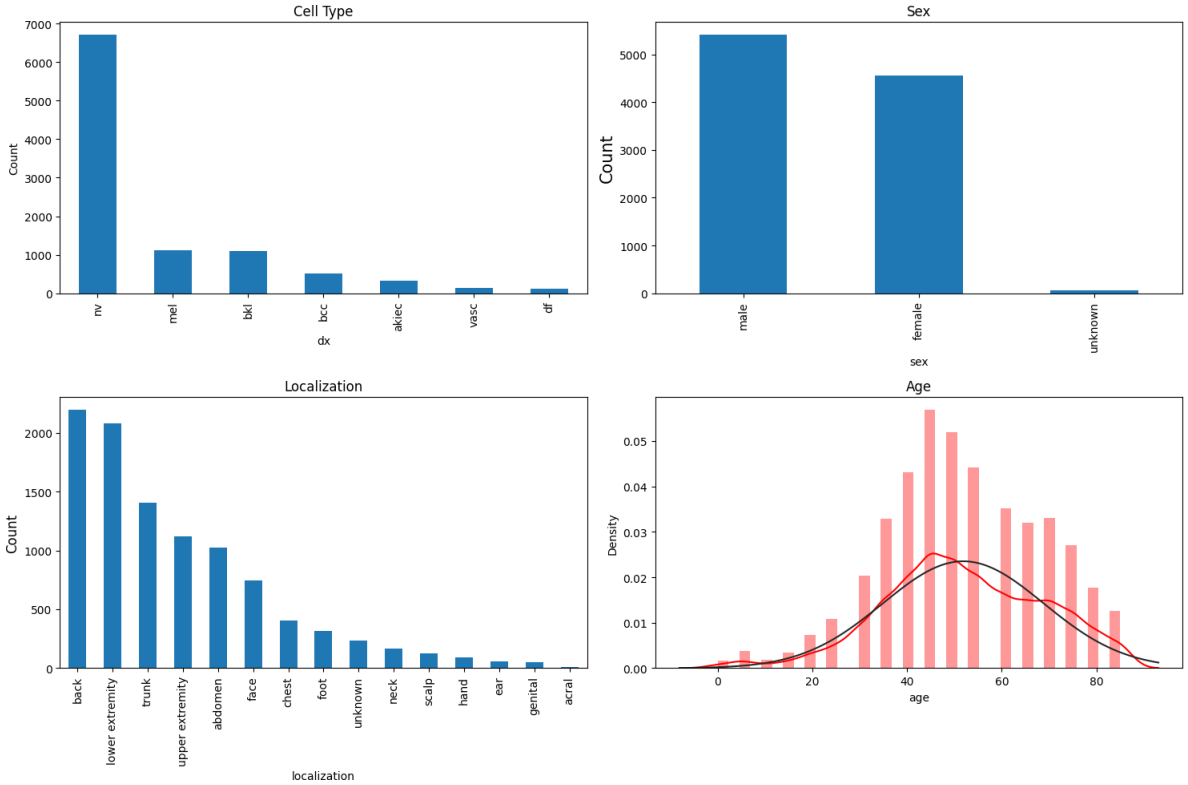

Figure 5: Demographic and Clinical Distribution of HAM10000 Dataset covering Cell Type, Sex, Localization, and Age Visualization.

6.2 Model Development and Training

Keras was used to design and execute a CNN. Multiple convolutional layers, max-pooling, dropout layers, and dense layers with softmax activation were used for multiclass classification. An implicit Adam optimizer and categorical cross-entropy loss function trained the model. Training settings were carefully chosen to enhance model performance and convergence.

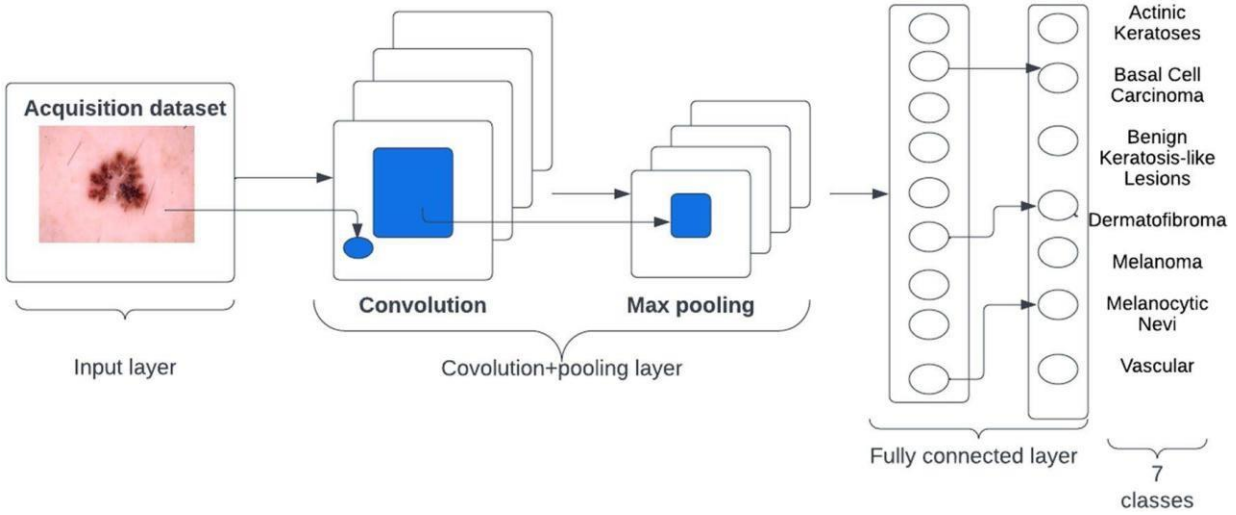

Figure 6: Convolutional Neural Network (CNN) Architecture for Skin Lesion Classification with Seven Output Classes.

6.2 Evaluation and Performance

The skin cancer lesion classification model's performance was thoroughly evaluated using a number of critical components. Central to this evaluation were the primary metrics of loss and accuracy, where categorical cross-entropy loss quantified the disparity between true labels and predicted probabilities, while accuracy measured the proportion of correctly classified samples. To fully evaluate the skin cancer lesion categorization model, several critical components were included. To ensure impartiality, the trained model was tested on a separate testing set, not used during training or validation, to simulate real-world circumstances. The model used the predict method to anticipate testing results and assign class labels to each sample to determine its most likely class. A confusion matrix showed the model's classification performance across lesion classes and identified misclassifications. Calculating and visualising the fraction of inaccurate predictions per class identified model or data enhancement needs. Visualization techniques using matplotlib and seaborn facilitated the interpretation of performance metrics, aiding in the identification of underlying patterns or trends in the model's behaviour.

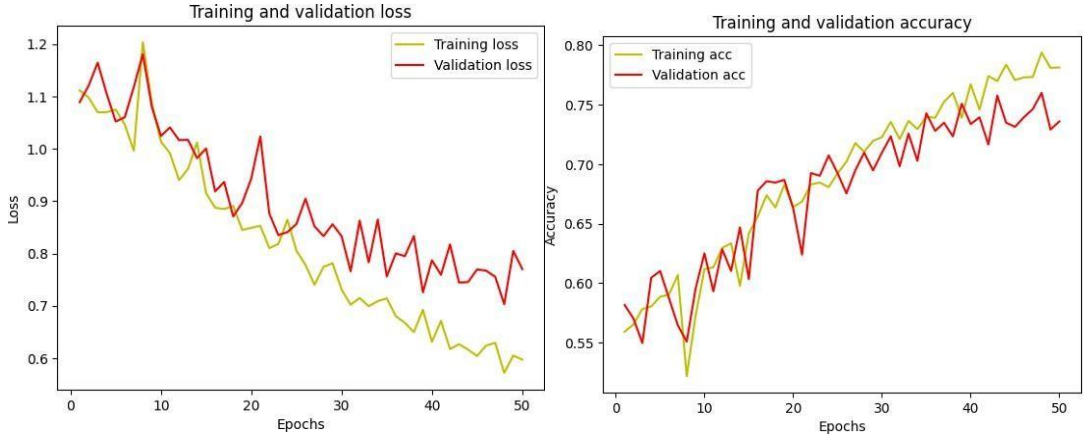

Figure 7: Training and Validation Loss and Accuracy Curves for Model Performance Evaluation.

6.3 Comparison with Previous Research

The accuracy was compared with previous research efforts in automated skin cancer diagnosis. Studies have demonstrated the efficacy of CNN-based approaches in achieving high accuracy in skin lesion classification tasks. The utilization of data preprocessing techniques aligns with established best practices in the field. However, variations in dataset characteristics, model architecture, and evaluation metrics across different studies may impact the generalizability of results.

CNN	RF	LR	LDA	SVM	KNN	NB	DT
73.8	70	58	57	53	48	36	68

Table 2: Model Performance Comparison Based on Accuracy (%).

6.4 Observations and Insights

Several observations provided insights into skin cancer lesion classification nuances. Data preprocessing, particularly in addressing class imbalance, yielded more robust model training outcomes. Additionally, the interpretability of the model's predictions warrants further investigation, crucial for clinical adoption and trust in automated diagnosis systems. *(0 – 6: range of lesions, classified into seven distinct categories)*

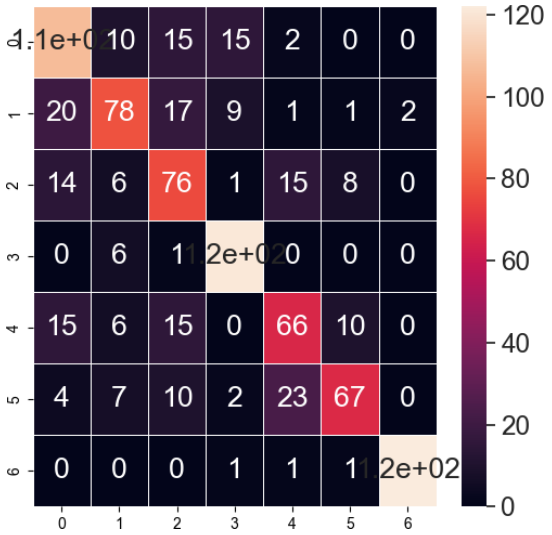

Figure 8: Confusion Matrix Heatmap Between Seven Different Output Classes created using Seaborn.

7. DISCUSSION

7.1 Trends and Relationships

Initially, there appears to be a fluctuation in both training and validation accuracy and loss metrics, with some epochs exhibiting increases while others showing decreases. This indicates that the model's performance fluctuates during training, possibly due to the stochastic nature of the optimization algorithm or the complexity of the dataset. However, as the training progresses, there is an overall positive trend in both training and validation accuracy, suggesting that the model gradually improves its ability to correctly classify samples over time. Meanwhile, the training and validation loss metrics exhibit a decreasing trend, indicating that the model becomes more effective at minimizing its error on the training data and generalizing to unseen validation data. The model performs better with more epochs, despite minor oscillations. Finally, the model's test accuracy on the testing set is 73.8%, reflecting its generalization ability on unknown data.

7.2 Generalizations

The study supports broad generalizations about CNN-based paradigms for skin cancer lesion categorization. The model's exceptional accuracy on unseen samples shows its ability to generalize to many lesion kinds, proving the learned features' use in defining complex lesion characteristics. Data preprocessing's effect on model performance shows its broad applicability in reducing dataset heterogeneity and improving model resilience.

The work also emphasizes methodological rigor in data pretreatment and model development and generalizes machine learning model construction best practices for medical imaging applications. The study uncovers the symbiotic relationship between data preprocessing and model performance, enhancing optimal model development methodologies.

7.3 Exceptions and Outlying Data

The model performs well, however misclassification and outliers require attention. Exceptions may result from aberrant lesion symptoms, imaging quality differences, or dataset oddities. Rare lesion subtypes or underrepresented classes may make classification difficult, causing model behavior issues.

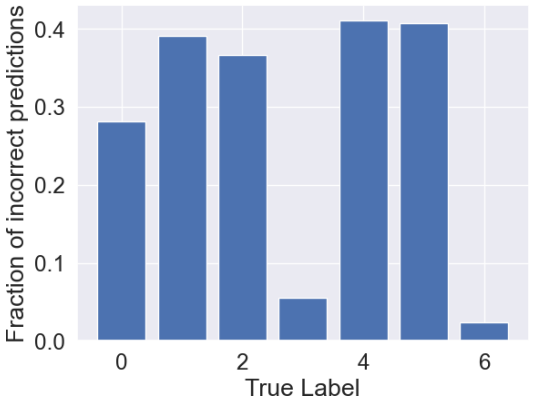

Figure 9: Bar Plot of Fraction of Incorrect Predictions per True Label created using Matplotlib.

The study's results corroborate research on CNN-based skin cancer diagnosis. Several investigations corroborate the conclusions of this research, demonstrating the capacity of CNN architectures to identify intricate lesion characteristics.

8. CONCLUSION

According to this work, CNN-based approaches to the classification of skin cancer lesions function when performing precise pre-processing of data and building a model. The stability of model convergence and excellent accuracy on unseen samples show the toughness and generalizability of the model. Label encoding and balancing of data enhance the performance of model and generalization of the model since the biases introduced by the class imbalances are reduced. Skin doctors now have a powerful tool for classification and diagnosis of the lesions, which is surely going to have a huge impact on the delivery of care. Future research should calibrate and then validate the model on different data sets and explain interpretability, while conducting prospective clinical trials for measuring the real-world implications. This work in automated diagnosis of skin cancer and health delivery combines science and mathematics.

Acknowledgements The author would like to thank anonymous reviewers for helpful comments on earlier drafts. All mistakes are my own.

Author contribution Lindberg conducted all elements of this research from the literature review, data collection, and empirical analysis

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability: The original datasets utilized in this study were obtained from a third-party source and subsequently modified for the purposes of this research. The modified datasets are available from the corresponding author upon reasonable request.

Conflict of Interest: The authors declare that there is no conflict of interest regarding the publication of this paper.

Declarations

Ethics, Consent to Participate, and Consent to Publish Not applicable.

Competing Interests The authors declare no competing interests.

REFERENCES

[1] Binder, M., Steiner, A., Schwarz, M., Knollmayer, S., Wolff, K., & Pehamberger, H. (1994). Application of an artificial neural network in epiluminescence microscopy pattern analysis of pigmented skin lesions: a pilot study. British Journal of Dermatology, 130(4), 460-465

[2] Lilja, H., Ulmert, D., & Vickers, A. J. (2008). Prostate-specific antigen and prostate cancer: prediction, detection and monitoring. Nature Reviews Cancer, 8(4), 268-278

[3] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large- scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). Ieee

[4] Rosendahl, C., Tschandl, P., Cameron, A., & Kittler, H. (2011). Diagnostic accuracy of dermatoscopy for melanocytic and nonmelanocytic pigmented lesions. Journal of the American academy of dermatology, 64(6), 1068-1073

[5] Codella, N., Rotemberg, V., Tschandl, P., Celebi, M. E., Dusza, S., Gutman, D., ... & Halpern, A. (2019). Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging

collaboration (isic). arXiv preprint arXiv:1902.03368

[6] Alam, J., Alam, S., & Hossan, A. (2018, February). Multi-stage lung cancer detection and prediction using multi-class svm classifier. In 2018 International conference on computer, communication, chemical, material and electronic engineering (IC4ME2) (pp. 1-4). IEEE

[7] Emuoyibofarhe, J. O., Ajisafe, D., Babatunde, R. S., & Christoph, M. (2020). Early Skin Cancer Detection Using Deep Convolutional Neural Networks on Mobile Smartphone. International Journal of Information Engineering & Electronic Business, 12(2)

[8] Ak, M. F. (2020, April). A comparative analysis of breast cancer detection and diagnosis using data visualization and machine learning applications. In Healthcare (Vol. 8, No. 2, p. 111). MDPI

[9] MurtiRawat, R., Panchal, S., Singh, V. K., & Panchal, Y. (2020, July). Breast Cancer detection using K-nearest neighbors, logistic regression and ensemble learning. In 2020 international conference on electronics and sustainable communication systems (ICESC) (pp. 534-540). IEEE

[10] Vidya, M., & Karki, M. V. (2020, July). Skin cancer detection using machine learning techniques. In 2020 IEEE international conference on electronics, computing and communication technologies (CONECCT) (pp. 1-5). IEEE

[11] Hasan, M. R., Fatemi, M. I., Khan, M. M., Kaur, M., & Zaguia, A. (2021). Comparative analysis of skin cancer (benign vs. malignant) detection using convolutional neural networks. Journal of Healthcare Engineering, 2021