THE USA JOURNALS
THE AMERICAN JOURNAL OF MEDICAL SCIENCES AND PHARMACEUTICAL RESEARCH
(ISSN – 2689-1026)
VOLUME 06 ISSUE12
PUBLISHED DATE: - 21-12-2024
DOI: - https://doi.org/10.37547/TAJMSPR/Volume06Issue12-10 PAGE NO.: - 92-112

RESEARCH ARTICLE Open Access

OPTIMIZING SKIN CANCER DETECTION IN THE USA HEALTHCARE SYSTEM USING DEEP LEARNING AND CNNS

Md Nasiruddin
Department of Management Science and Quantitative Methods, Gannon University, Erie, PA, USA

Mohammad Abir Hider
Master of Science in Business Analytics, Grand Canyon University, Phoenix, AZ, USA

Rabeya Akter
Master of Science in information technology. Washington University of Science and Technology, Alexandria, VA, USA

Shah Alam
Master of Science in Information Technology, Washington University of Science and Technology, Alexandria, VA, USA.

MD Rashed Mohaimin
MBA in Business Analytics, Gannon University, Erie, PA, USA

MD Tushar Khan
Master of Science in Business Analytics, Trine University, Angola, IN, USA

Abdullah AL Sayeed
Master of Business Administration in Project Management, Central Michigan University, Mt Pleasant, MI, USA

Afrin hoque jui
Management Sciences and Quantitative Methods, Gannon University, Eria, PA, USA

Abstract
Skin cancer is among the most prevalent cancers in the USA, with millions of new cases reported each year. The two main types of skin cancer include aggressive, life-threatening melanoma and less lethal, though potentially very morbid if left unattended, non-melanoma types: basal cell carcinoma and squamous cell carcinoma. The chief aim of this research project is to devise, curate, and propose a deep-learning CNN methodology for skin cancer detection in the USA. The dataset for the current research project was retrieved from the Kaggle website, particularly, The ISIC 2016 Skin Cancer Dataset contained dermoscopic images that were used for skin cancer classification. In this dataset, there were 1271 images of two classes of skin cancer, namely Malignant and Benign. These images were then gathered from the ISIC archive. The dataset was then divided into a training set consisting of 1022 images and a test set consisting of 249 images. The CNN proposed for this work is a deep-learning architecture designed to address skin cancer detection through dermoscopic images. The model follows a sequential architecture with multiple layers dedicated to the extraction of hierarchical features from input images. To assess the performance of the CNN algorithm for skin cancer detection, several proven metrics are utilized, namely, accuracy, precision, recall, and F1-Score. The model obtained a very high precision, recall, and F1-score over all classes, with a general accuracy of 94% for this multi-class problem. This model was very good, both in precision since it correctly identifies the actual positive cases and in recall, where it does not have false positives. The developed proposed CNN model for skin cancer detection has great potential to support human clinical decision-making in all dermatology. This developed model automates the various analyses of dermoscopy images, hence acting as just an adjunct tool for active dermatologists, which shall enable fast and accurate skin lesion assay. Results have shown that this CNN can easily be integrated into diagnosis workflows in normal dermatological practice to offer a second opinion or even a pre-screening tool for dermatologists.

Keywords Skin Cancer Detection; Convolutional Neural Networks; Deep Learning; Dermoscopic Imaging; USA Healthcare System.

INTRODUCTION

Background and Motivation

Skin cancer is the most common cancer in the USA, with millions of new cases reported each year. The two main types of skin cancer include aggressive, life-threatening melanoma and less lethal, though potentially very morbid if left unattended, non-melanoma types: basal cell carcinoma and squamous cell carcinoma. According to the American Cancer Society, that means an over 99% five-year survival rate for early-stage melanoma, against just 27% in the case of late-stage detection (Rahman et al., 2023). This further signifies the all-important question of diagnosis in time and with high accuracy. Despite its importance, early skin cancer detection faces various systemic barriers in the US. These range from a shortage of dermatologists and limited access to care in rural and underserved areas to high costs related to diagnostic procedures. Moreover, diagnostic accuracy significantly varies among health professionals; some depend on subjective visual examination, which is prone to human error. Advanced imaging techniques such as dermoscopy have made the diagnosis more accurate, but their effectiveness greatly relies on the clinician's expertise (Saleh et al., 2023).

Al Amin et al. (2023), states that recent developments in AI, especially deep learning, give promising solutions to these challenges. Deep learning is a subset of artificial intelligence; a class of such algorithms called convolutional neural networks has achieved great success in image recognition tasks, medical imaging included. By learning patterns in large datasets, CNNs can classify skin lesions with accuracy comparable to, or sometimes exceeding, that of dermatologists. These AI-driven solutions, when integrated into the USA health system, might change the face of early diagnosis, reduce disparity in healthcare, and result in improved patient outcomes (Bowmik et al., 2023; Dutta et al., 2024).

Objectives

The principal aim of this research project is to devise, curate, and propose a deep-learning CNN methodology for skin cancer detection in the USA. The specific objectives of this research are: To apply deep learning methodologies for the detection of skin cancer using dermoscopic images, maintaining sensitivity and specificity at high values. To develop a model of a convolutional neural network adapted to classify the different types of skin cancers, especially melanoma, BCC, and SCC. This includes the performance evaluation of the model on standard metrics of accuracy, precision, recall, and F1-score, with further consideration regarding clinical applications within the US healthcare system.

LITERATURE REVIEW

According to Islam et al. (2023), Skin Cancer and its Detection in the USA Skin cancer is the most diagnosed cancer in the United States, with millions of individuals receiving a skin cancer diagnosis every year. The main types of skin cancer are basal cell carcinoma, squamous cell carcinoma, and melanoma. BCC and SCC are usually referred to as non-melanoma skin cancers and are usually highly curable if they are left early. Melanoma, though rare, becomes much more dangerous because it has great tendencies for metastasis. The American Cancer Society estimates that in the USA, melanoma is the cause of more deaths than any other form of skin cancer, with 97,610 new cases and 7,990 deaths projected for 2023(Hossain et al. 2024; Hider et al. 2024). The high prevalence of skin cancers creates a huge burden on the patients, as well as on the healthcare system. Direct costs entail diagnosis, treatments, and follow-up over a long period. Indirect costs are manifested as lost workdays along with a reduction in general quality of life. Accordingly, public health campaigns emphasizing prevention and early detection raise the chances of prognosis being exceedingly better. However, at every turn, disparities in dermatologic care add to the stress-especially in rural versus underserved areas (Ghosh et al., 2024).

Current Diagnostic Techniques of Skin and Their Limitations

1. Visual Inspection by Clinicians

As per Jaber & Akbas (2024), visual inspection remains the first line of defense in diagnosing skin cancer. Clinicians evaluate lesions based on their asymmetry, border irregularity, color variation, diameter, and evolution using the ABCDE criteria. Though easy and cheap, it is highly dependent on a clinician's experience and expertise. Studies demonstrate big variability in diagnostic skills, particularly among non-specialist providers. This subjectivism often may lead to the under-sighting of a suspicious case or unnecessary biopsy.

2. Dermatoscopy

Dermoscopy is a non-invasive method that allows the professional, using polarized light and magnification, to visualize structures subsurface in the skin that may go unrecognized by the naked eye. Thus, this technique is considered one of the finest methods for diagnosing melanoma and distinguishing malignancies from benign lesions (Musthafa et al., 2023). Still, dermoscopy itself does require extensive training and can be highly related to an operator's learning curve concerning diagnostic accuracy. Being only available to a very restricted number of clinicians, in turn, reduces equitable access to care and even further exacerbates cancer outcome disparities in many diverse geographical regions.

3. Histopathology

Histopathology is usually the gold standard in the diagnosis of skin cancers, where microscopic investigations of biopsied tissue are performed. It allows for the final confirmation of malignancy and typing of cancer. However, this technique is invasive, time-consuming, and expensive. There is always a delay for the patients while waiting for biopsy results, and more than 80% of performed biopsies are unnecessary, which indicates the need to improve the pre-biopsy diagnostic tool(Nancy et al., 2023).

4. Reflectance Confocal Microscopy (RCM)

According to Nasiruddin et al. (2024), RCM is a highly advanced imaging technique that enables one to visualize the skin, at a cellular level and in real-time, in high resolution. It allows physicians to diagnose skin lesions without a biopsy. While RCM decreases the need to perform invasive procedures, it is costly and requires special equipment; not only that, it puts a high demand on skilled personnel for image interpretation, so it is not commonly made available in routine clinical practice. 5. Computer-Aided Diagnosis Systems These CAD systems analyze dermoscopic images by using algorithms and then make diagnostic recommendations to clinicians. While the use of such CAD systems enhances diagnostic accuracy and reduces human error, their performance still largely depends on the quality of the training datasets. Most CAD systems are also not integrated into routine clinical workflows and, in practice, remain confined to a few well-equipped health centers (Lilhore et al., 2023).

Deep Learning in Medical Imaging

Deep learning is a subclass of artificial intelligence; it involves neural networks that are multilayered and are trained to find patterns and features in data. As opposed to traditional machine learning, where explicit feature extraction is necessary, deep learning learns the relevant features directly from the training data. It has made it more apt for the analysis of complex datasets such as medical images. Deep learning, has shown phenomenal promise in diagnosis across many medical domains-radiology, pathology, and dermatology(Sha et al., 2023). Such detection ranges from the development of abnormalities within chest X-rays to tumor classification in histopathological slides and assessment of diabetic retinopathy from retinal scans. It has thereby positioned deep learning as probably one of the most transformational tools in modern medicine to process high-dimensional data with unparalleled precision.

Empirical studies conducted by Zareen et al., (2024), deploying CNNs for the detection of skin cancer have presented tremendous improvements in diagnostic accuracy and efficiency. These utilize deep learning techniques in the analysis of dermatoscopic images that allow for the early identification of skin lesions that may be malignant. Among them, one of the notable works proposed by Obayya et al. (2024), an optimized CNN architecture that enhanced skin cancer diagnosis using a very rich dataset called the HAM10000, comprising dermatoscopic images. In the proposed design, they developed a sophisticated model comprising several convolutional, pooling, and dense layers to capture complex visual features in skin lesions. This approach included some interesting data augmentation strategies to handle the class imbalance in the dataset, enabling higher diagnostic precision that could democratize dermatology care, especially where specialist expertise and/or access are limited.

The second major work by Saleh et al. (2023), proposed the "Light-Dermo" model, which can be seen as a light version of the CNN while considering optimization towards real-time applications. The given model used the mechanism of channel-wise attention with integrated Squeeze-and-Excitation blocks for improvements in the classifying accuracy along with computation efficiency. On that account, the given research indicated an improved performance using these state-of-the-art models in previous works, for example, 93.16% training accuracy and 91.93% test accuracy among seven classes for PSLs2. This highlights the potential for accuracy and practical use in the clinical setting that CNNs have. A new deep CNN approach was then proposed to address class imbalance problems inherent in skin cancer datasets. This model enhanced not only the accuracy of classification but also showed resilience in various challenges related to the detection of skin cancer3. Advanced techniques, including transfer learning and data preprocessing, were integrated into CNNs to make them more capable of distinguishing between benign and malignant lesions.

Convolutional Neural Networks

According to Nancy et al., (2023), CNNs are a special form of neural networks designed to operate on grid-like data, such as images. Architecture: Multiple layers, each performing a specialized function:

Convolution Layers: These are the layers through which filters are applied to the input image to extract features that are related to edges, textures, and patterns.

Activation Functions: Non-linear functions such as ReLU introduce non-linearity, thereby helping the network learn even the most complex relationships.

Pooling Layers: Layers that reduce the spatial dimensions of feature maps, hence emphasizing important information and reducing computation.

Fully Connected Layers: This is the last layer; it summarizes the features in the convolutional layers to classify them.

Softmax Layer: The probability over each class is attained from this layer, and the model uses that for labeling an input image.

The hierarchical architecture of CNNs facilitates them to learn low-level features (e.g., edges) in initial layers and higher-level features (e.g., lesion shapes or patterns) in deeper layers. This capability also makes CNNs particularly effective in tasks such as skin cancer detection, where identifying subtle visual differences is key.

Key Achievements and Applications of CNNs in Healthcare

Due to the improvements in model architecture and computation resources, there has been fast development of CNNs in healthcare. The main achievements include:

Transfer Learning: Pre-trained models, such as ResNet, VGGNet, and Efficient-Net, have been fine-tuned for medical imaging tasks, reducing the need for extensive training data.

Attention Mechanisms: These include various techniques that allow models to concentrate on specific parts of the image when necessary; this helps raise the accuracy in specific tasks such as lesion detection.

Explain-ability Tools: Techniques like Grad-CAM give visual explanations for model predictions, hence enhancing interpretability and clinical trust.

Successful modern deployments of CNNs in healthcare include:

Radiology: CNN is used for detecting several critical features, which include lung nodules and other fractures and tumors based on various types from the given input imaging modality obtained normally by X-rays or CT scans.

Pathology: Whatever the case, the CNNs have certainly brought a revolution in the analysis of histopathology images for identifying cancerous cells.

Ophthalmology: The use of automation in detecting diabetic retinopathy and age-related macular degeneration has enhanced the efficiency of its screening.

DATA COLLECTION AND PREPROCESSING

Data Sources

The dataset for the current research project was retrieved from the Kaggle website, particularly, The ISIC 2016 Skin Cancer Dataset contained dermoscopic images that were used for skin cancer classification. In this dataset, there were 1271 images of two classes of skin cancer, namely Malignant and Benign. These images were then gathered from the ISIC archive. The dataset was then divided into a training set consisting of 1022 images and a test set consisting of 249 images (Zihad, 2023). This dataset was used for training and testing machine learning models for skin cancer classification. The dataset is also useful in the development of new image-processing techniques for skin cancer detection and diagnosis. In light of working with skin cancer data sets, the most important ethical concerns refer to compliance with data privacy legislation to protect patient rights and develop trust in AI-driven health solutions. For instance, in the United States, data should be covered under the Health Insurance Portability and Accountability Act, called HIPAA. It demands vigorous protection for Protected Health Information in terms of security. For example, data de-identification requires removing all identifiable information related to patients; a person's name, date of birth, and also medical record numbers all represent sources of personally identifiable information in healthcare records. In this, some of the key ethics that we considered include informed consent regarding the use of a person's data, insight into where this data is going to be used, and anti-bias resulting in equitable performance of their models, irrespective of different skin types or geographical location.

Data Preprocessing

Effective data preprocessing is critical in making any skin cancer dataset suitable for the training of machine learning models. The first step was cleaning, where images were reviewed to eliminate duplicate samples, mislabeled samples, and corrupted files, ensuring the dataset was representative of the target classes. Secondly, **image normalization** is performed for scaling pixel values, usually within a range between 0 and 1. This normalizes the input and speeds up the convergence of the model during training. Thirdly **resizing** of all images to a fixed dimension, such as 224x224 pixels, simply makes them compatible with typical CNN architectures such as ResNet or VGGNet since they require a fixed input size. To increase diversity in the dataset and prevent overfitting, some augmentation techniques are adopted, including rotation, flipping, zooming, and brightness adjustment. These augmentations simulate real-world variations, enabling the model to generalize better across unseen data. Fourth, the dataset was split into training, validation, and test sets, often using stratified sampling to maintain class balance. The described preprocessing pipeline is complete in ensuring that the dataset is clean, standardized, and representative of a phenomenon, which forms a good basis for robust model training.

Exploratory Data Analysis (EDA)

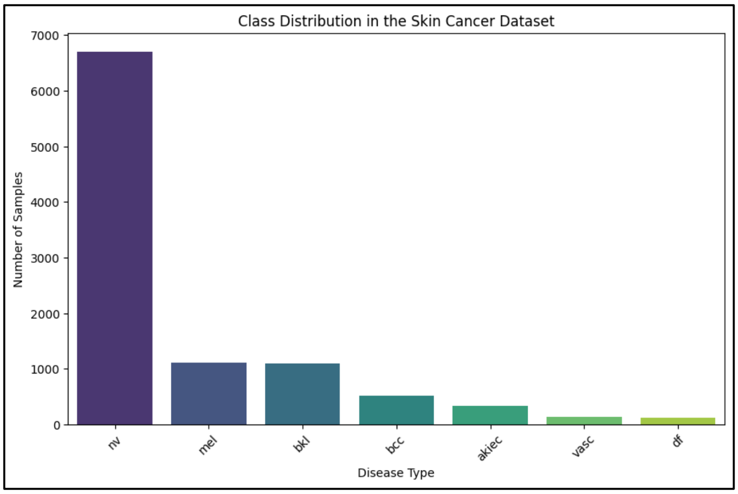


Figure 1: Exhibits Class Distribution in the Skin Cancer Dataset

The histogram above showcases a high-class imbalance in the Skin Cancer dataset, dominated by the class "nv"-most probably with about 7,000 samples, while all other classes have much fewer representatives. The second most represented classes, "mel" for melanoma and "bk" for benign keratosis-like lesions, have about 1,000–1,200 samples each, showing a sharp drop from the dominant class. This is followed by "bcc" standing for basal cell carcinoma, which has 600–800 samples, while "akiec" (Actinic keratosis) has less than 600. The rare classes include "vasc"(vascular lesions) and "df" (dermatofibroma), representing vascular lesions and dermatofibroma respectively, each having less than 200 samples. This is a challenge in the training of machine learning models since the dominance of the class "nv" might result in biased predictions and reduce the capability of the model to detect classes that are less frequent yet clinically critical, such as melanoma. To handle the class imbalance, data augmentation, resampling techniques, or class-weighted loss functions were necessary to make sure the model performance was robust and fair.

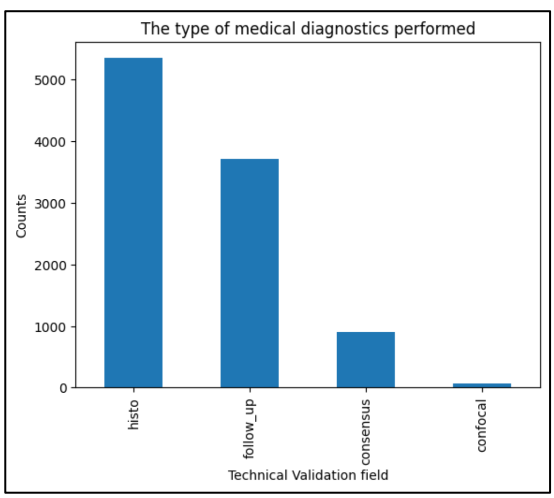


Figure 2: The Type of Medical Diagnostics Performed

This histogram represents the distribution of different medical diagnostic techniques from a dataset. "histo" probably is the abbreviation for histopathology, which was over 5,000 and therefore the most popular diagnostic technique in the dataset. "Follow_up" follows with a count of roughly 3,500, further elaborating that it is also a key technique used mostly for follow-up or tracking certain skin conditions. The less frequently used method, "consensus," ranges at about 1,000 counts, while the least used is "confocal" diagnostics, with less than 100 counts. The above distribution indicates reliance on histopathology as the gold standard for diagnosis, while other methods, such as follow-ups, are supportive. The very minimal utilization of confocal diagnostics perhaps reflects its specialized nature, its cost, or its availability. This distribution underlines the interest in investigating complementary methods to diversify and possibly streamline diagnostic workflows.

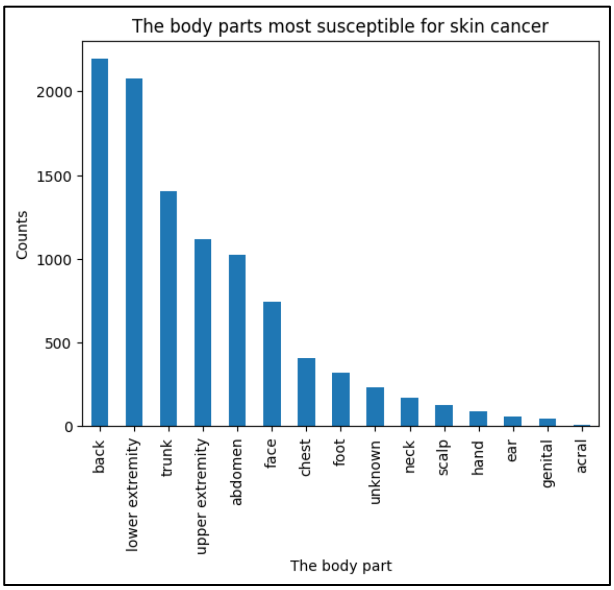


Figure 3: Displays the Body Parts Most Susceptible for Skin Cancer

This chart represents the distribution of skin cancer cases among different body parts and pinpoints those parts that are most prone to this disease. The highest number of cases is recorded on the back, with more than 2,000 cases, followed by the lower extremity and trunk, having counts above 1,500 and 1,000, respectively. The upper extremity and abdomen are also well represented, each having about 800–1,000 cases. There is a moderate count for the face, chest, and foot between 500 and 700; whereas for the neck, scalp, hand, and ear, there are less than 300 cases noted. Rare sites include the genital region and acral sites, with very minimal counts. The "unknown" category indicates a small but notable portion of cases where the location was unspecified. The pattern of distribution underlines that the focused screening has to be performed on high-risk areas, especially the back and lower extremities, considering, however, comprehensive coverage due to less common sites.

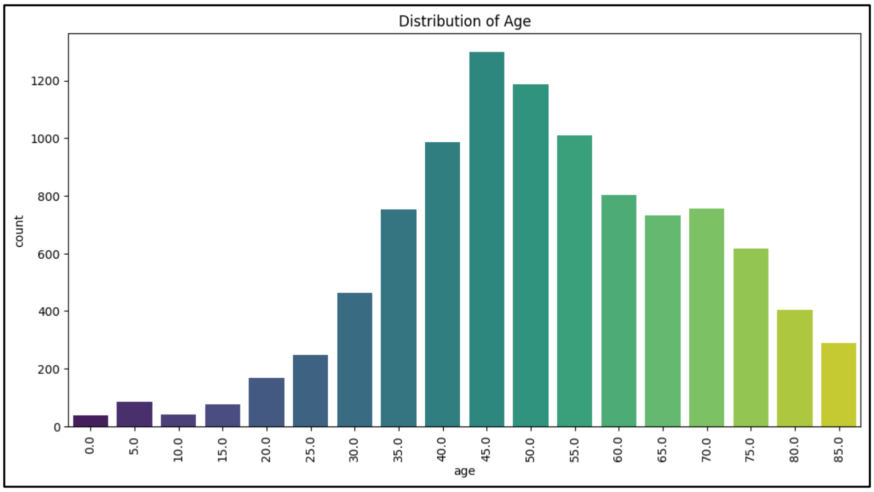


Figure 4: Portrays Distribution of Age

This histogram above reflects the distribution of age in a dataset. The distribution appears to be right-skewed, peaking in the 40-45 age group. This means that the majority of the people who form part of this dataset are middle-aged. With increasing age, the count goes down very gradually to show a smaller proportion of older ones. The long tail to the right is extended more, showing again that this distribution is right-skewed. The overall histogram represents the distribution of age in the dataset, where middle-age brackets are more populated.

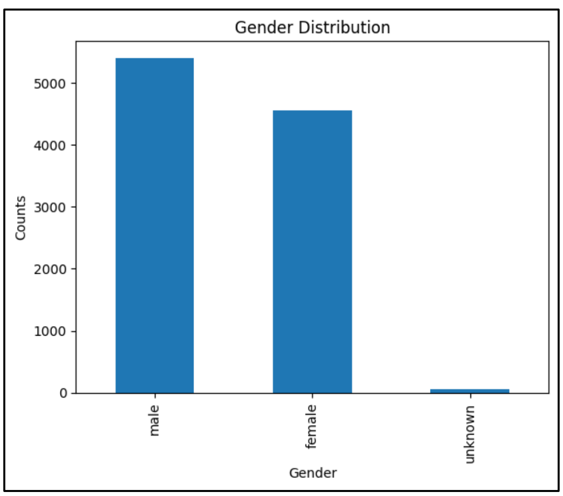


Figure 5: Visualizes Gender Distribution

This histogram portrays the distribution graph of gender in a dataset: the majority of constituent members of the dataset are males, about 5,500. The number is around 4,500 for females, which means there is an imbalance in keeping up the gender level. The third category, "unknown", corresponds to a very small figure compared to the other two sections, meaning that most sexes are known in this lot. Overall, the shown histogram reflects a clear male-gender dominance within the dataset or population.

THE USA JOURNALS
THE AMERICAN JOURNAL OF MEDICAL SCIENCES AND PHARMACEUTICAL RESEARCH
(ISSN – 2689-1026)
VOLUME 06 ISSUE12

Figure 6: Exhibits Dermoscopic Images with Different Kinds of Skin Lesions

Above are dermoscopic images, corresponding to different kinds of skin lesions. Each image is associated with its diagnosis: akiec is the label used for actinic keratoses, bcc means basal cell carcinoma, bkl stands for benign keratosis, mel refers to melanoma, nv stands for melanocytic nevus, and vascular lesions are defined as vasc. Different colors, texture patterns, and shapes of lesions represent their wide variability. While some lesions appear as elevated nodules or plaques, others may be flat or ulcerated. The colors range from light brown to various tones of dark brown and black; some lesions contain shades of red or blue. The great variability in the appearance of skin cancer accentuates the challenge in diagnosis, thus placing considerable demands on methods for distinguishing benign from malignant lesions.

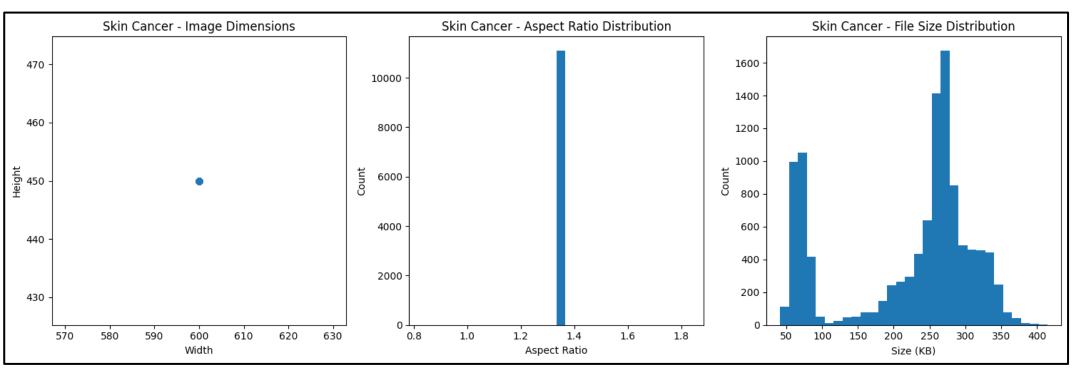


Figure 7: Depicts the Skin Cancer Dimensions, Aspect Ratio, and File Size Distribution

As displayed in the first histogram, it is uniformly distributed, with the majority of the images having a width of about 600 pixels and a height of about 460 pixels. The second histogram shows that this aspect ratio is highly positively skewed, with most images having an aspect ratio of about 1.3. The third histogram represents the file size distribution of the images; it appears to be right-skewed. Thus, most of the images hold a file size between 100 and 200 KB. However, some may get up to over 300 KB. These are broad indications that the size and shape of the majority of images are quite consistent but varying in their file size presumably owing to quality and compression parameters of the camera or otherwise.

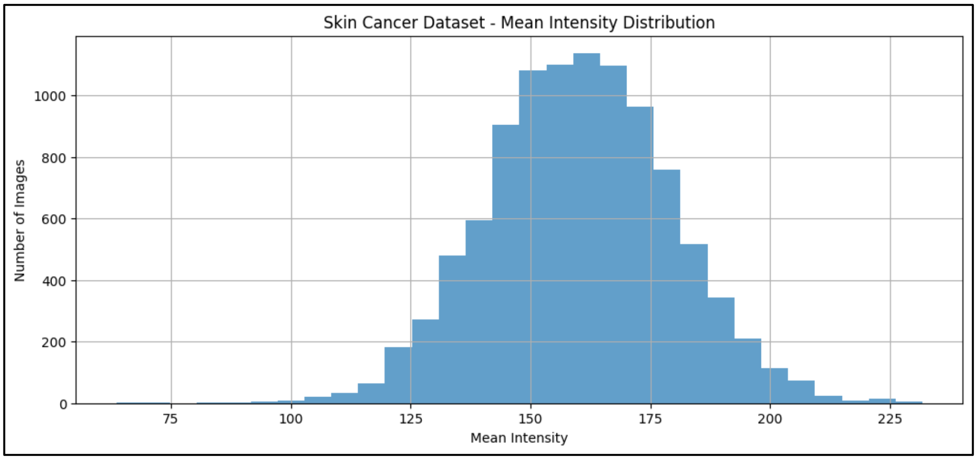


Figure 8: Showcases Skin Cancer Dataset-Mean Intensity Distribution

The histogram below shows the distribution of mean intensity values across the skin cancer dataset. This distribution is approximately normal, peaking at around 160 and spreading from about 75 to 225. This would suggest that most of the images in this collection have a mean intensity in the middle range, with fewer with very low or very high mean intensities. Due to the normal shape, one would conclude that the variation across the dataset is relatively stable in terms of mean intensity value.

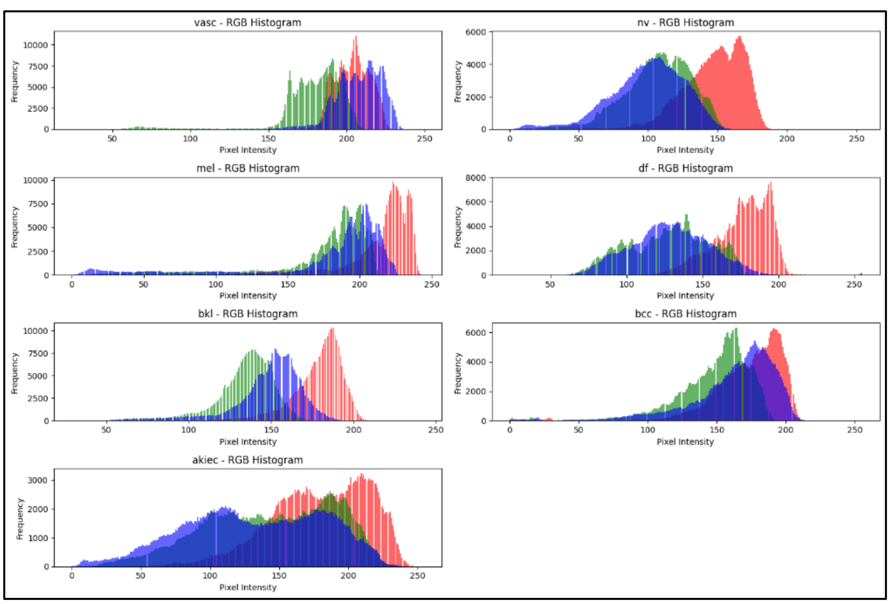


Figure 9: Illustrates the Distribution of Pixel Intensities

The above figure depicts the distributions of pixel intensities in red, green, and blue channels each for vascular (vasc), melanocytic nevus (nv), melanoma (mel), dermato-fibroma (df), benign keratosis (bkl), basal cell carcinoma (bcc), and actinic keratosis (akiec). These enable us to get an idea regarding the color characteristics of different kinds of skin lesions. For example, the melanoma and dermato-fibroma histograms are shifted toward the higher intensity values in the red channel, indicating a reddish color. In contrast, benign keratosis and basal cell carcinoma have more concentration in the lower intensity values in the red channel, indicating a brown or gray color. These variations in color distribution can help distinguish the type of skin lesions and are possibly useful in automated classification systems.

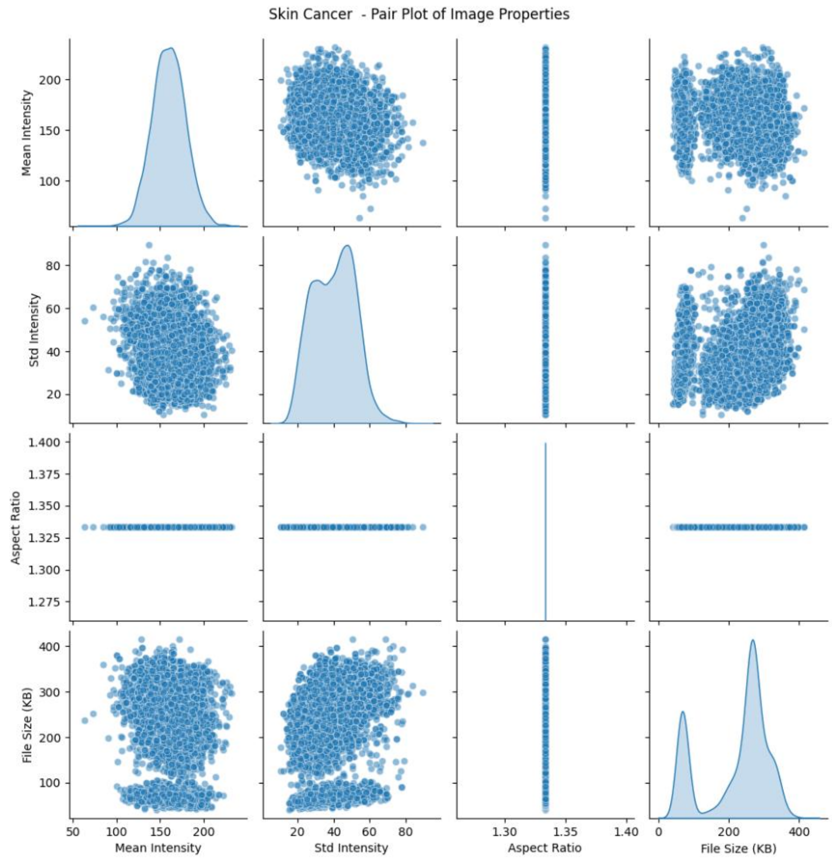


Figure 10: Illustrates the Distribution of Skin Cancer- pair plot of image properties

Above is a pairs plot. A pairs plot is useful in visualizing the relationship between several different properties of images within the Skin Cancer dataset. For the images, the diagonal plots give an idea of the distribution for each separate property; these are: right-skewed for the means of intensity and the size of the file while the highly-aspected ratio is strongly focused around 1.3 while the off-diagonal plots represent a scatter plot of one property versus another. We observed weak positive correlations: average intensity vs. file size as well as average intensity, and standard deviation of the intensities of the file. Weak anticorrelation in aspect ratio about file size. By and large, though it should be that most associations there in fact appear to be of real tenuosity, or else non-existent, as against the assumption that this ensemble of image properties all will be highly interlinked -.

METHODOLOGY

Model Architecture

The CNN proposed for this work is a deep-learning architecture designed to address skin cancer detection through dermoscopic images. The model follows a sequential architecture with multiple layers dedicated to the extraction of hierarchical features from input images. The architecture begins with an input layer where preprocessed images are fed into the network, usually resized to a fixed dimension, such as 224x224 pixels. After that, several convolutional layers are applied to identify the local features automatically, such as edges, textures, and patterns, which form an essential part of identifying unique features of skin lesions. These convolutional layers are then followed by a nonlinear actuation function like ReLU, which would introduce non-linearity, hence guaranteeing the network learns complex patterns. The max-pooling layers serve to down-sample the feature maps by convolution, hence reducing its spatial dimensionality and subsequent computational load, while retaining the most salient features. To avoid overfitting, several dropout layers technique that randomly deactivates part of the neurons during training into place, promoting better generalization capability in the trained model. These feature maps are then fed into fully connected layers that combine all the learned features for final classification. The output layer consists of a soft-max activation function, which gives probability scores for each class, such as benign, malignant, and so on, to facilitate multi-class classification of skin conditions.

This chosen architecture for CNN is based on an appropriate theoretical rationale: it inherently learns and selects spatial patterns from the dermoscopic image in a hierarchical manner, to build a decision boundary for the fine-grained diagnosis of skin lesions. In this sense, convolutional layers are best for extracting the relevant spatial features, ideal in image analysis tasks. The max-pooling layers contribute to reducing the dimensionality such that the network is efficient, and running while retaining the most relevant features. Besides, dropout was implemented to avoid overfitting, which is an important point when medical image datasets are used since much variability in image quality and characteristics of lesions may easily result in overfitting. This last softmax layer is important for multi-class classification, outputting a probability distribution that indicates the likelihood of a given skin lesion belonging to a specific class. The overall architecture strikes a good balance between feature extraction, efficiency, and robustness, hence suitable for real-world clinical skin cancer detection.

Training and Testing Framework

The skin cancer detection dataset was divided, into clear-cut subsets for the fair assessment of the performance of the model, in the training, validation, and testing sets. This division was done in a 70-20-10 split for devoting 70% of the data to training, 20% going to validation, and testing getting 10%. Testing data is used to allow the CNN model to learn the different features across the dermoscopic images. The validation set was applied to monitor the performance of a model iteratively during the training, where changes may be made to avoid over-fitting. The test set was used only for the very final evaluation of the generalization capability of the model, Fairly giving an unbiased estimate of the performance. Finally, to add more robustness, k-fold cross-validation was done. This technique has some variations in which the dataset is divided into k subsets, and then multiple trainings of the model are done, each time using one subset as the validation set while the rest of the data are used for training. Cross-validation helps ensure that a model is not biased toward a particular subset of the data and may provide a more general estimate of its performance across the entire dataset.

Hyperparameter Tuning

The optimization of hyperparameters is very crucial to optimize the performance of the model. The three major hyperparameters of a CNN are learning rate, batch size, and the number of epochs. The learning rate controls the rate at which changes in the weights of the model take place during training, which impacts the speed and stability of convergence. The batch size defines the number of samples that will be processed before updating the weights of the model. This value is a trade-off that influences the efficiency and memory usage during training. The number of epochs is the number of times a model iterates over the whole dataset. Now, for this, there is a set of hyperparameters that one often optimizes using techniques such as grid search and random search. Grid search exhaustively checks all possible combinations of hyperparameters within predefined ranges and assures that the best set is chosen. On the other hand, random search randomly selects hyperparameters in a certain range and is often considerably faster, particularly when large search spaces are considered. Moreover, this process may be automated by tools like Keras Tuner or Optuna, which effectively explore the hyperparameter space and recommend the best configuration for the best performance.

Performance Evaluation Metrics

To assess the performance of the CNN algorithm for skin cancer detection, several proven metrics are utilized, namely, accuracy, precision, recall, and F1-Score. Accuracy provides an overall idea of the percent of right predictions-a percept of overall general performance. Nevertheless, since skin cancer image datasets used are highly imbalanced in this context, it would not be sensible or practical to rely only upon these measures. Therefore, alongside accuracy, other important measurements considered are precision, recall, and F1 score. Precision quantifies how many of those positive predictions-e.g., true cases of melanoma-are indeed correct, a critical measure to avoid false positives in medical diagnostics. Recall, on the other hand, is a metric that measures the capability of the model to identify all actual positive instances, such that no malignant cases go unnoticed. The F1-score is the harmonic mean of precision and recall; hence, it provides a balanced measure of the model's ability to correctly identify positives and minimize false negatives. Finally, ROC-AUC, or Receiver Operating Characteristic - Area Under the Curve, will be used to assess the model's discriminatory ability across all thresholds where higher AUC means better overall performance in classification. These are complementary and provide a complete metric to address how well the model did, in particular, to handle what's usually imbalanced classes of very rare presence of melanoma as compared to benign lesions.

RESULTS

Model Performance

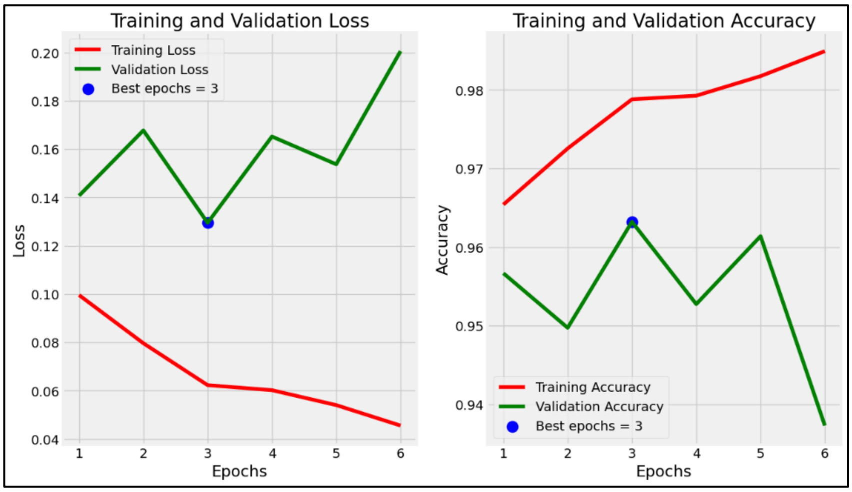


Figure 11: Visualizes Training Validation & Accuracy

These two plots show the training and validation loss and accuracy curves of some machine learning models throughout 6 epochs. The plot on the left shows the training loss trending downward linearly; the validation loss is initially going down, reaches a minimum around epoch 3, and then starts to increase. This means that by epoch 3, this model begins to overfit the training data. The chart on the right shows smoothly increasing training accuracy and validation accuracy which follows suit until a maximum around Epoch 3 before slowly starting its decline. Further evidence of said overfitting behavior concerning loss curves was provided earlier. The best performance is achieved at epoch 3, where the validation loss is minimum and the validation accuracy is maximum. These plots highlight the importance of monitoring both training and validation metrics to prevent overfitting and identify the optimal number of training epochs.

Table 1: Displays CNNs Classification Report

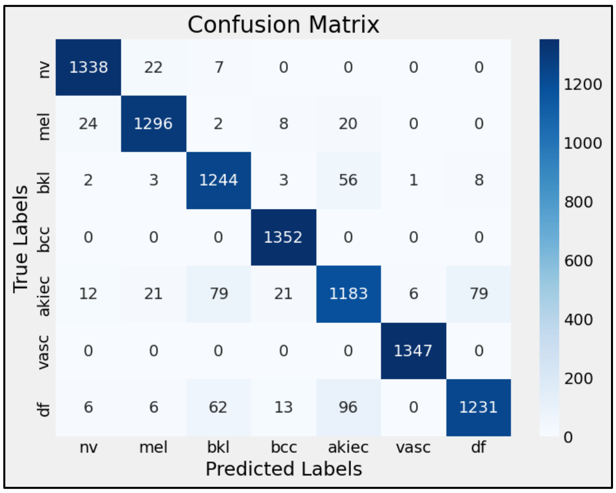
	precision	recall	f1-score	support
nv	0.97	0.98	0.97	1367
mel	0.96	0.96	0.96	1350
bkl	0.89	0.94	0.92	1317
bcc	0.97	1.00	0.98	1352
akiec	0.87	0.84	0.86	1401
vasc	0.99	1.00	1.00	1347
df	0.93	0.87	0.90	1414
accuracy			0.94	9548
macro avg	0.94	0.94	0.94	9548
weighted avg	0.94	0.94	0.94	9548

The table above depicts the classification report of the skin cancer detection model performance on this dataset, containing eight classes of skin lesions, namely, nv (melanocytic nevus), mel (melanoma), bk

THE USA JOURNALS
THE AMERICAN JOURNAL OF MEDICAL SCIENCES AND PHARMACEUTICAL RESEARCH
(ISSN – 2689-1026)
VOLUME 06 ISSUE12

(benign keratosis), bcc (basal cell carcinoma), akiec (actinic keratosis), vasc (vascular lesion), and df (dermatofibroma). It could be realized from this report that the model obtained a very high precision, recall, and F1 score over all classes, with a general accuracy of 94% for this multi-class problem. This model was very good, both in precision since it correctly identifies the actual positive cases and in recall, where it does not have false positives. Also, the macro and weighted average scores are very good in general. However, the slight overperformance of classes such as bcc and vasc concerning others like bkl and akiec could be an area of improvement.

Table 2: Exhibits CNNs Confusion Matrix



Confusion Matrix

True Labels \ Predicted Labels	nv	mel	bkl	bcc	akiec	vasc	df
nv	1338	22	7	0	0	0	0
mel	24	1296	2	8	20	0	0
bkl	2	3	1244	3	56	1	8
bcc	0	0	0	1352	0	0	0
akiec	12	21	79	21	1183	6	79
vasc	0	0	0	0	0	1347	0
df	6	6	62	13	96	0	1231

The confusion matrix offered a detailed breakdown of the classification performance across multiple classes, particularly showing the true labels versus the predicted labels. It managed to classify 1,338 instances of the class " melanocytic nevus " correctly but mislabeled 22 as "melanoma" and 7 as "benign keratosis." For the class "melanaoma", the model had a true positive count of 1,296 but misclassified 8 instances as "melanocytic nevus" and 20 as "benign keratosis." In this class, "df" also showed very good performance with 1,231 correct predictions, though there are still misclassifications, especially in its instances identified as "actinic keratosis)" with 96 and "benign Keratosis " with 13. In summary, most classes have relatively high values on the diagonal, but the misclassifications do raise concerns that the network may need improvement, particularly between similar classes like "bkl" and "mel."

DISCUSSION

Clinical Implications

The developed proposed CNN model for skin cancer detection has great potential to support human clinical decision-making in all dermatology. This developed model automates the various analyses of dermoscopy images, hence acting as just an adjunct tool for active dermatologists, which shall enable fast and accurate skin lesion assay. It is especially valued in those settings where dermatologic overload or access to dermal specialists is poor outside major cities in the USA, The high precision and recall rates of the model can reduce false positives and false negatives, hence the early detection of malignant cases and avoidance of unnecessary biopsies or interventions on benign cases. Moreover, the ability of CNNs to detect subtle patterns that may be imperceptible to the human eye further enhances diagnostic accuracy. This means that clinicians better-informed decisions, at least a reduction in diagnostic uncertainty-something so crucial in diseases like melanoma.

Results have shown that this CNN can easily be integrated into diagnosis workflows in normal dermatological practice to offer a second opinion or even a pre-screening tool for dermatologists. Teledermatology, just now gaining steam in the USA and even more so post-COVID-19 pandemic, may be greatly helped by such technology. This will enable patients to upload photos of skin lesions from the comfort of their homes; the CNN model will scan them for flagging probable malignant lesions for further evaluation. Moreover, the tool assists in triaging cases, as it prioritizes patients with high-risk lesions to immediate care. These applications will not only raise the efficiency of dermatological services but also ensure equal access to quality care, especially for patients in remote locations.

Integration into USA healthcare systems

These highlighted aspects of the CNN model could provide transformative benefits in their incorporation within healthcare systems in the United States. Early diagnosis, mainly of skin cancer, has improved survival rates due to interventions at an early stage. The model can ensure regularity in diagnosis and consistency to a greater degree with diagnostic errors that are well-known in dermatology based on subjective human judgments. Furthermore, this model can be incorporated into EHRs in a way that automatic analysis of dermoscopic images at routine checkups reduces the workload of dermatologists, hence saving time to be spent on consulting patients. This could make the health and medical care provided cheaper since several biopsies and treatments would not be carried out.

Nevertheless, the integration of such AI models as CNNs into clinical workflows is not without its challenges. One major consideration is to ensure interoperability with existing healthcare systems, such as EHR platforms used across hospitals and clinics. The model needs to be validated across diverse populations and different imaging devices to ensure its generalizability and reliability in real-world scenarios. Resistance to adopting AI tools from clinicians, often driven by concerns about trust, job security, or unfamiliarity with the technology, must also be addressed. These barriers can be minimized with extensive training programs and well-articulated guidelines on the effective use of AI tools. Besides, regulatory compliance will have to be maintained with organizations such as the FDA, and ethical standards followed to nurture trust and accountability in deploying AI models.

Limitations and Challenges

While the CNN model has a lot of advantages, some ethical and technical challenges need consideration for successful deployment. Some of the ethical concerns involve using patient data to train AI models, including privacy, consent, and security. Compliance with regulations such as HIPAA is very important in the USA, and anonymization of datasets is required for the protection of patient identity. Moreover, it is important to have no datasets that induce bias, in which changes in minority groups are disproportionately affected, for a dataset to be representative of the broader population. For example, algorithms for skin cancer diagnosis trained only on light-skinned populations have performed very poorly on darker skin, underlining the need to consider inclusive datasets.

Another limitation pertains to the quality and interpretability of the model. CNNs are often described as "black-box" systems, and their decision-making processes are not always transparent. This lack of interpretability may pose challenges in clinical settings where understanding the rationale for a diagnosis is important for patient care. Besides, changes in image quality, lighting, and resolution across datasets could affect model performance and raise questions about its robustness in real-world applications. Generalizability remains another concern since models trained on one dataset may not perform well on completely different datasets, thus requiring extensive validation before clinical deployment.

Future Research Directions

Overcoming pinpointed challenges by future research will be vital to enhance the effectiveness and applicability of the CNN model. The most viable way to do this might be using larger datasets that are more diverse, with images from people of different ethnicities, age brackets, and geographical regions. This would go a long way in enhancing the model's generalization ability across different patient populations and reducing the risk of bias. The model could also apply knowledge, with advances in transfer learning, from pre-trained architectures, thus limiting the need for large datasets and accelerating development. Further, the model may be improved by using complementary data from other modalities, such as genetic information or patient histories, which would enhance not only its diagnostic performance but also yield a broader view of the health status of a patient.

Another exciting area for exploration is real-time skin cancer detection. Further work may, therefore, be directed at an improved CNN model, one that, through portable devices or using a phone application, instantly provides analyses of skin lesions. That will put patients in the position to self-manage their status of skin health and take timely medical advice whenever necessary. Another very promising area is the integration of AI models with personal treatment planning. These CNNs may further develop the characteristics of lesions and relate them to patient outcomes to predict the most effective treatments in any given case, advancing precision medicine in dermatology. Ultimately, ongoing research and collaboration between clinicians, researchers, and technologists will be required to unlock the full potential of CNN models to transform skin cancer detection and treatment.

Impact on the USA Health Care System

Improved Diagnostic Precision and Early Detection: With greatly enhanced accuracy and timeliness, CNNs may bring a paradigm shift in the diagnosis of skin cancer in the USA. The subtle patterns and features that may be missed by the human naked eye are picked up through dermoscopic image analysis by the CNN model. The resultant increased accuracy can lead to the early detection of skin cancer, thereby allowing timely intervention that will improve patient outcomes.

Reduced Diagnostic Errors: One of the significant benefits of AI-driven CNN models is identified in the reduction of diagnostic errors. These models will tend to provide consistent and objective analyses that reduce the human error factor, which might be caused by tiredness, inexperience, or an element of subjective judgment. That would increase reliability for a more perfect diagnosis that can effectively be treated.

Efficiency in Clinical Workflows: The incorporation of CNNs into the clinical workflow may provide efficiency in diagnosis, reducing the burden on dermatologists and other healthcare professionals. With the automation of image analysis, these models give a preliminary diagnosis that may liberate clinicians to attend to more complex cases and interpersonal interactions with patients. The resultant increase in efficiency can then lead to shorter wait times, improved patient satisfaction, and better use of healthcare resources.

Cost-Effectiveness: Early detection and proper diagnosis of skin cancer will lead to long-term cost savings. The reason is that early detection and treatment prevent expensive surgeries and radiation therapies associated with the advanced stages of skin cancer. It could also reduce burdens on healthcare professionals and raise efficiency to lower healthcare costs overall. Personalized Treatment Plans AI-powered CNN models can help in gleaning valuable insights into the patient-specific characteristics that will allow the development of personalized treatment plans. Thus, these models can identify risk factors and predict disease progression by analyzing the patient's medical history, lifestyle factors, and imaging data. The information may further be used to tailor the appropriate treatment strategies by recommending certain medications, therapies, or surgical procedures for optimal outcomes for the patients.

CONCLUSION

The principal aim of this research project is to devise, curate, and propose a deep-learning CNN methodology for skin cancer detection in the USA. The dataset for the current research project was retrieved from the Kaggle website, particularly, The ISIC 2016 Skin Cancer Dataset contained dermoscopic images that were used for skin cancer classification. In this dataset, there were 1271 images of two classes of skin cancer, namely Malignant and Benign. These images were then gathered from the ISIC archive. The dataset was then divided into a training set consisting of 1022 images and a test set consisting of 249 images. The CNN proposed for this work is a deep-learning architecture designed to address skin cancer detection through dermoscopic images. The model follows a sequential architecture with multiple layers dedicated to the extraction of hierarchical features from input images. To assess the performance of the CNN algorithm for skin cancer detection, several proven metrics are utilized, namely, accuracy, precision, recall, and F1-Score. The model obtained a very high precision, recall, and F1-score over all classes, with a general accuracy of 94% for this multi-class problem. This model was very good, both in precision since it correctly identifies the actual positive cases and in the recall, where it does not have false positives. The developed proposed CNN model for skin cancer detection has great potential to support human clinical decision-making in all dermatology. This developed model automates the various analyses of dermoscopy images, hence acting as just an adjunct tool for active dermatologists, which shall enable fast and accurate skin lesion assay. Results have shown that this CNN can easily be integrated into diagnosis workflows in normal dermatological practice to offer a second opinion or even a pre-screening tool for dermatologists.

REFERENCES

1. Alam, S., Hider, M. A., Al Mukaddim, A., Anonna, F. R., Hossain, M. S., khalilor Rahman, M., & Nasiruddin, M. (2024). Machine Learning Models for Predicting Thyroid Cancer Recurrence: A Comparative Analysis.
2. Al Amin, M., Liza, I. A., Hossain, S. F., Hasan, E., Haque, M. M., & Bortty, J. C. (2024). Predicting and Monitoring Anxiety and Depression: Advanced Machine Learning Techniques for Mental Health Analysis. British Journal of Nursing Studies, 4(2), 66-75.
3. Bhowmik, P. K., Miah, M. N. I., Uddin, M. K., Sizan, M. M. H., Pant, L., Islam, M. R., & Gurung, N. (2024). Advancing Heart Disease Prediction through Machine Learning: Techniques and

Insights for Improved Cardiovascular Health. British Journal of Nursing Studies, 4(2), 35-50.

4. Bortty, J. C., Bhowmik, P. K., Reza, S. A., Liza, I. A., Miah, M. N. I., Chowdhury, M. S. R., & Al Amin, M. (2024). Optimizing Lung Cancer Risk Prediction with Advanced Machine Learning Algorithms and Techniques. Journal of Medical and Health Studies, 5(4), 35-48.
5. Dutta, S., Sikder, R., Islam, M. R., Al Mukaddim, A., Hider, M. A., & Nasiruddin, M. (2024). Comparing the Effectiveness of Machine Learning Algorithms in Early Chronic Kidney Disease Detection. Journal of Computer Science and Technology Studies, 6(4), 77-91.
6. Ghosh, H., Rahat, I. S., Mohanty, S. N., Ravindra, J. V. R., & Sobur, A. (2024). A Study on the Application of Machine Learning and Deep Learning Techniques for Skin Cancer Detection. International Journal of Computer and Systems Engineering, 18(1), 51-59.
7. Hasan, E., Haque, M. M., Hossain, S. F., Al Amin, M., Ahmed, S., Islam, M. A., ... & Akter, S. (2024). CANCER DRUG SENSITIVITY THROUGH GENOMIC DATA: INTEGRATING INSIGHTS FOR PERSONALIZED MEDICINE IN THE USA HEALTHCARE SYSTEM. The American Journal of Medical Sciences and Pharmaceutical Research, 6(12), 36-53.
8. Hider, M. A., Nasiruddin, M., & Al Mukaddim, A. (2024). Early Disease Detection through Advanced Machine Learning Techniques: A Comprehensive Analysis and Implementation in Healthcare Systems. Revista de Inteligencia Artificial en Medicina, 15(1), 1010-1042.
9. Hossain, S., Miah, M. N. I., Rana, M. S., Hossain, M. S., Bhowmik, P. K., & Rahman, M. K. (2024). ANALYZING TRENDS AND DETERMINANTS OF LEADING CAUSES OF DEATH IN THE USA: A DATA-DRIVEN APPROACH. The American Journal of Medical Sciences and Pharmaceutical Research, 6(12), 54-71.
10. Hossain, M. S., Rahman, M. K., & Dalim, H. M. (2024). Leveraging AI for Real-Time Monitoring and Prediction of Environmental Health Hazards: Protecting Public Health in the USA. Revista de Inteligencia Artificial en Medicina, 15(1), 1117-1145.
11. Islam, M. Z., Nasiruddin, M., Dutta, S., Sikder, R., Huda, C. B., & Islam, M. R. (2024). A Comparative Assessment of Machine Learning Algorithms for Detecting and Diagnosing Breast Cancer. Journal of Computer Science and Technology Studies, 6(2), 121-135.
12. Jaber, N. J. F., & Akbas, A. (2024). Melanoma skin cancer detection based on deep learning methods and binary Harris Hawk optimization. Multimedia Tools and Applications, 1-14.
13. Lilhore, U. K., Simaiya, S., Sharma, Y. K., Kaswan, K. S., Rao, K. B., Rao, V. M., ... & Alroobaea, R. (2024). A precise model for skin cancer diagnosis using hybrid U-Net and improved MobileNet-V3 with hyperparameters optimization. Scientific Reports, 14.
14. Musthafa, M. M., TR, M., V, V. K., & Guluwadi, S. (2024). Enhanced skin cancer diagnosis using optimized CNN architecture and checkpoints for automated dermatological lesion classification. BMC Medical Imaging, 24(1), 201.
15. Nancy, V. A. O., Prabhavathy, P., Arya, M. S., & Ahamed, B. S. (2023). Comparative study and analysis on skin cancer detection using machine learning and deep learning algorithms. Multimedia Tools and Applications, 82(29), 45913-45957.
16. Nasiruddin, M., Dutta, S., Sikder, R., Islam, M. R., Mukaddim, A. A., & Hider, M. A. (2024). Predicting Heart Failure Survival with Machine

Learning: Assessing My Risk. Journal of Computer Science and Technology Studies, 6(3), 42-55.

17. Obayya, M., Arasi, M. A., Almalki, N. S., Alotaibi, S. S., Al Sadig, M., & Sayed, A. (2023). Internet of things-assisted smart skin cancer detection using metaheuristics with deep learning model. Cancers, 15(20), 5016.

18. Pant, L., Al Mukaddim, A., Rahman, M. K., Sayeed, A. A., Hossain, M. S., Khan, M. T., & Ahmed, A. (2024). Genomic predictors of drug sensitivity in cancer: Integrating genomic data for personalized medicine in the USA. Computer Science & IT Research Journal, 5(12), 2682-2702.

19. Rahman, A., Karmakar, M., & Debnath, P. (2023). Predictive Analytics for Healthcare: Improving Patient Outcomes in the US through Machine Learning. Revista de Inteligencia Artificial en Medicina, 14(1), 595-624.

20. Saleh, N., Hassan, M. A., & Salaheldin, A. M. (2024). Skin cancer classification based on an optimized convolutional neural network and multicriteria decision-making. Scientific Reports, 14(1), 17323.

21. Shah, A., Shah, M., Pandya, A., Sushra, R., Sushra, R., Mehta, M., ... & Patel, K. (2023). A comprehensive study on skin cancer detection using artificial neural network (ANN) and convolutional neural network (CNN). Clinical eHealth.

22. Zareen, S. S., Sun, G., Kundi, M., Qadri, S. F., & Qadri, S. (2024). Enhancing Skin Cancer Diagnosis with Deep Learning: A Hybrid CNN-RNN Approach. Computers, Materials & Continua, 79(1).

23. Zihad, F. (2023, October 17). Skin Cancer Dataset ISIC 2016. Kaggle. https://www.kaggle.com/datasets/mdforiduzzamanzihad/skin-cancer-dataset-isic-2016