Proc. of the 18th IEEE Int. Conf. on Emerging Technologies, FAST-NUCES, Peshawar, Pakistan, 6-7 November 2023

Convolutional Neural Networks for Nail Disease Detection: A Promising Approach in Dermatology

Abdullah Ajmal, Wakeel Ahmad, Syed M Adnan

Department of Computer Science, University of Engineering and Technology, Taxila, Pakistan
{abdullah.ajmal@students.uettaxila.edu.pk, wakeel.ahmad@uettaxila.edu.pk, syed.adnan@uettaxila.edu.pk}

2023 18th International Conference on Emerging Technologies (ICET) | 979-8-3503-2817-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICET59753.2023.10374663

***Abstract*—The aim of this study is to use Deep Learning (DL) techniques to classify and detect human nail diseases. Early and reliable detection is crucial in aiding timely interventions and suitable treatment for nail diseases and their profound impact on a person's well-being. Using a diverse dataset of nail images, a CNN model was developed and trained to achieve the study's objectives. For the model to be able to accurately detect and classify different diseases, the dataset was carefully collected to include various types of nail diseases. In order to improve the model's performance and robustness, the nail images were preprocessed, and data augmentation techniques were applied. The model's assessment encompassed 8 distinct nail diseases, resulting in an impressive accuracy rate of 98.44%. Additional evaluation metrics such as precision, recall, and F1-score were also computed, yielding values of 99.22%, 98.44%, and 99.02% respectively. The achieved outcomes were compared with state-of-the-art (SOTA) techniques, affirming the superiority of the proposed model. This study emphasizes the potential benefits of DL techniques in enhancing healthcare practices, enhancing dermatological diagnostics, and improving the overall well-being of patients suffering from nail diseases.**

***Index Terms*—Nail Disease, Deep Learning, Nail abnormalities, Computer-aided diagnosis, Dermatology, Medical imaging, CNN**

I. Introduction

Toes and fingers are protected from harm by the nail, which is a hardened protein called Keratin. An individual's nail unit can often provide a direct indication of their overall health. There are many diseases that can be associated with discolorations, hyperpigmentation, malignancies, and any anomalies in the nail's texture, including malnutrition psoriatic arthritis, anemia, diabetes, melanoma, fungal infections, and liver diseases. Thus, nail diseases can have a major influence on a person's physical well-being. Nail disorders can include inflammatory, congenital, and hereditary conditions as well as tumors and trauma [1]. Traditional methods of diagnosis, mostly rely on a subjective visual assessment by medical professionals, which is prone to human error, or by available scanning techniques in hospitals are X-Rays, Pathological tests, and Magnetic Resonance Imaging (MRI) [2]–[4] etc. which are lengthy and costly procedures. Identifying nail diseases still remains one of the most difficult and unexplored areas of research. To the year 2003 and prior, no substantial research had been conducted on the detection of diseases in nails using DL techniques [5]. But the advancements in DL techniques in recent years, particularly CNN, have demonstrated promising results in various image classification tasks. Using a CNN-based model, we propose an innovative method for the detection and categorization of nail diseases. The objective of our work was to create an automated system that can assist dermatologists and healthcare professionals in diagnosing nail diseases with greater accuracy and efficiency by using deep learning techniques. To learn discriminative features that differentiate between different nail disease classes, the proposed system utilizes a dataset of nail images and a modern CNN architecture. We affirm the effectiveness and durability of our methodology, highlighting its potential for practical implementation in clinical settings. Following is the organization of the remainder of the paper: Section I presents the Introduction. Section II presents the related work. Proposed methods and materials are presented in Section III. The results of the experiments are discussed in Section IV, and the conclusion is presented in Section V.

II. Literature Review

In [6], the authors employed techniques of image processing on the nails for the development of a system for early-stage disease detection. Specifically, they focused on utilizing nail pigment as a distinguishing feature, achieving an accuracy rate of 65% in their experiments. By exploring unconventional features such as nail pigment in disease detection methodologies, this study offers insights into the growing body of knowledge directed at developing innovative approaches for early diagnosis. In their research, [7] introduced a DL-based approach to identifying nail diseases. Their study presented a unique integrated framework that fused CNNs for feature extraction and classification of nail diseases using image data. The proposed algorithm demonstrated an impressive accuracy of 84.58%. The authors of [8] propose a novel approach. To facilitate early detection of Terry's Nail, the researchers used a model of transfer learning, specifically TensorFlow Inception-V3. An accuracy rate of 95.24% was achieved by the developed system. The study offers promising results for the early detection and intervention of nail diseases using transfer learning methods. [9] describes OnyxRay, a mobile application that uses custom vision machine learning techniques to diagnose a variety of nail diseases. Based on uploaded images, the app classifies nail diseases using image processing methods and a CNN model. An image dataset of 3,000 nail images representing 11 different types of nail diseases was used to train the CNN model. 91.4% of nail diseases can be detected using the app, based on the results of the study. In a similar study [10], DL algorithms were used to develop a system that is automated for the detection of skin diseases. Eczema, psoriasis, tinea corporis, urticaria, and vitiligo were the five common skin disorders they classified. In addition to clinical images, patient information was also incorporated into the proposed system, which brought in impressive results. An accuracy of 97.5% was obtained for this multiclass classification where the precision was 97.7% and sensitivity was 97.7%. These outcomes highlight the exceptional diagnostic performance of the developed system for the five skin diseases. The study underscores the DL algorithm's potential in classifying diverse skin diseases accurately, emphasizing the significance of incorporating clinical images and patient information in the development of such systems. In [11], a Novel DL-based methods are presented for the detection and classification of microscopic brain tumors. This method utilizes a 3D CNN architecture for feature extraction. A selection method based on correlation is used to validate the selected features, which are then classified by sending them to a feed-forward neural network. On three datasets of BraTS from 2015, 2017, and 2018, the proposed method achieves impressive accuracies of 98.32%, 96.97%, and 92.67%. Results Based on the comparative analysis, the proposed method is demonstrated to be as efficient and accurate as existing techniques while offering comparable performance. Through the application of DL methodologies, this research presents a promising advance in the field of microscopic brain tumor detection and classification. In [12], transfer learning is discussed as a method of reducing training data size, minimizing time requirements, and reducing computational costs when building DL architectures. In this study, nine different tomato leaf diseases, including healthy leaves, were classified. In order to extract features, they used five different deep network architectures, namely DenseNet121-Xception, MobileNet, Xception, ShuffleNet, and ResNet50. According to the results, the parameters used in the five CNN models differed significantly in terms of accuracy and parameter usage. Despite having a greater number of parameters, Densenet-Xception shows the highest recognition accuracy of 97.10%. ShuffleNet, on the other hand, achieves an accuracy of 83.68% with fewer parameters. Comparative analyses presented in this paper provide crucial insight into the development of an intelligent system for the diagnosis of tomato diseases using smartphones. In [13], CNNs were used to develop a model for detecting multiple-lesion skin conditions. 16,543 non-standardized images were used for the training of VGG-16, which were divided into training, validation, and test sets. A clinical database was used to obtain the images, which included eczema, rosacea, psoriasis, cutaneous T-cell lymphoma, and acne. A CNN model was found to be highly sensitive and specific for differentiating different skin conditions. The test demonstrated that, for example, rosacea was identified against acne with an accuracy being 85.42% while a specificity of 89.53%, cutaneous T-cell lymphoma from eczema with an accuracy of 74.29% and specificity of 84.09%, and eczema was distinguished against psoriasis with a specificity of 73.57% and sensitivity of 81.79%. This study [14] addresses the critical issue of skin disease diagnosis, a pivotal facet of public health. Timely and precise detection is essential for effective treatment and recovery. Harnessing recent advances in convolutional neural networks (CNNs), the research targets the diagnosis of two prevalent skin conditions - Eczema and Psoriasis. The study deploys five cutting-edge CNN architectures, subjecting them to rigorous 10- fold cross-validation. Notably, the Inception ResNet v2 architecture, in conjunction with the Adam optimizer, attains an impressive validation accuracy of 97.1%. This study [15] addresses the critical issue of Parkinson's disease (PD), characterized by the progressive loss of motor function. The authors propose hybrid deep neural networks combining convolutional neural networks with long short-term memory, creating parallel model. These architectures effectively capture structural features of EEG signals, significantly enhancing the classification accuracy of individuals with PD. The results are promising, with the model achieving an impressive 98.6% accuracy in a challenging 3-class classification scenario. This advancement holds great promise for significantly improving the accuracy of PD diagnosis. There has been significant progress in dermatological image classification, but the classification of nail diseases, especially within the context of a restricted dataset, is an area that warrants further investigation. The purpose of this study is to fill this gap by providing a specialized approach to accurately identify nail diseases, thus contributing to the dermatological diagnostics. A brief summary of these studies is presented in Table I.

TABLE I: Brief Summary of SOTA Techniques

Ref	Year	Model	Classes	Accuracy %
[6]	2016	ESDDS	05	65.00
[7]	2017	Hybrid CNNs	11	84.58
[8]	2019	Inception-V3	02	95.24
[9]	2020	Custom Vision API	12	80.00
[12]	2020	ShuffleNet	09	83.68
[13]	2020	VGG-16	05	88.64
[10]	2021	Mobilenet-V2	05	97.50
[16]	2022	CNN-MDD	05	95.27
[14]	2023	Inception ResNet v2	02	97.10
[15]	2023	Hybrid CNN	03	98.60

III. Materials and Methods

The CNN model used in the proposed system for human nail disease detection comprises of an input layer, a classification layer, and multiple hidden layers. These hidden layers include convolution layers, pooling layers, and fully connected layers, among others. The structure of the deep neural network is illustrated in Figure 1.

A. Dataset

A dataset from Kaggle [17] is used to train and evaluate the proposed model. There are 340 images in this dataset,

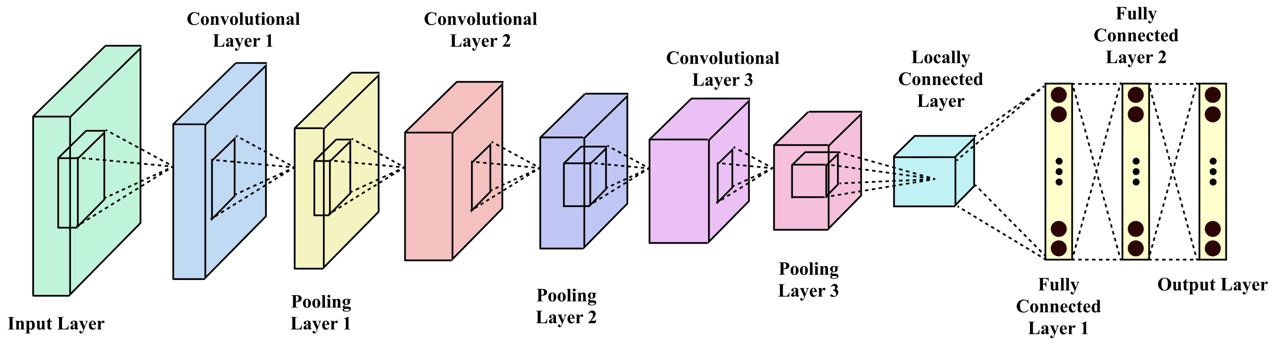

Fig. 1: Convolution Neural Network's Architecture

Fig. 2: Sample images from dataset

TABLE II: Detail of Dataset

Diseases	No. of Images	After Augmentation
Alopecia Areata	47	376
Beau's Lines	42	336
Bluish Nail	50	400
Clubbing	40	320
Darier's Disease	47	376
Eczema	45	360
Koilonychia	38	304
Leukonychia	31	248
Total images	**340**	**2720**

representing eight different classes of nail diseases. Table II summarizes how images are distributed between the classes. Training and testing sets are created to ensure proper evaluation. 80% of the images are from the training set, while 20% are from the testing set. Figure 2. shows sample images from data set.

B. Image Preprocessing

Preprocessing is a vital first step. As part of this model, noise reduction, contrast adjustment, and image normalization were performed. Through these processes, CNN will be able to provide more accurate and reliable predictions for nail diseases by improving nail image quality.

C. Data augmentation

In order to increase a Deep CNN model's performance with limited training data, techniques like image augmentation are commonly used. We have enhanced our dataset firstly by rescaling the images by dividing the pixel values by 255 to create a range of 0 to 1 to ensure that the model could effectively learn from the images. In addition, we employed random vertical and horizontal flips to enhance the model's ability to generalize by introducing variations in the orientation of the nail images. Furthermore, random rotations with a degree of 0.2 were applied. This augmentation simulates different angles from which nails may be captured. These augmentations effectively expanded the dataset from 340 to 2720 images and enhanced its generalizability. As a result of these techniques, robust features are learned through variations in the images. Figure 3. shows sample images through data augmentation.

D. CNN Architecture

There are several layers in the architecture of CNNs, with each layer serving a distinct purpose in the analysis and classification of images. Among these layers, we have segments, features, edge detection, and computation layers that are used to perform tasks such as segmentation and edge detection. A CNN's architecture is composed of layers that enable the network to learn features from the input images, as well as to extract those features from the input images. Figure 3. shows the layout of the CNN architecture employed in this study, which shows the layers of the network and demonstrates how information flows between the layers. A CNN's convolution layer is a fundamental component that generates feature maps by applying convolution filters [18]. In these filters, the kernel pixels are convolved with the corresponding image pixels by computing their dot products. Our model is composed of four convolutional layers, each of which employs a small filter of size 3x3 for the extraction of features. Each convolution layer is activated with the Rectified Linear Unit (ReLU). CNN training is accelerated with this activation function, which introduces non-linearity into the network and aids in capturing complex patterns and representations in data. The ReLU function, defined by Eq. 1.

$$f(x) = \max(0, x) \tag{1}$$

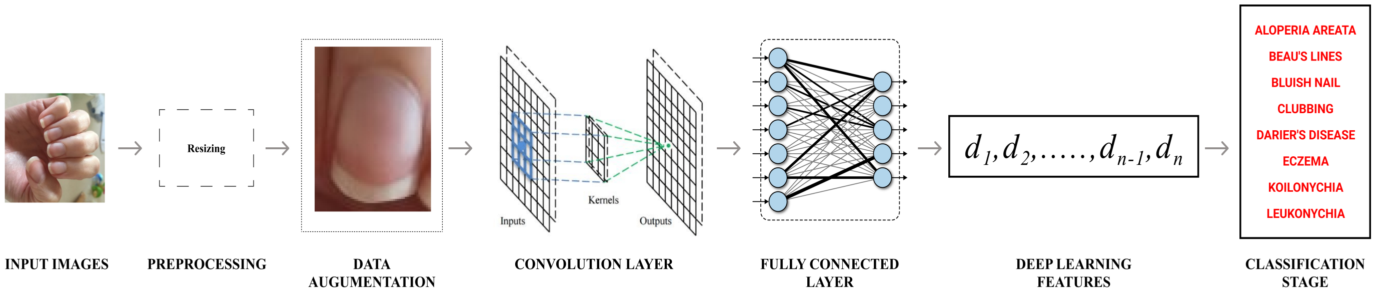

Fig. 3: Proposed Model's Architecture

Here, 'x' is the input to the activation function 'f' on channel 'x'. CNN can learn intricate relationships and enhance its overall learning capacity by leaving positive values the same and changing negative values to zero. Following the Convolution Layer is Pooling layer. Convolution layer outputs are used as inputs in this layer. In essence, the Pooling layer down samples the information by reducing its spatial dimensions.

An example of a pooling method that is commonly used is max-pooling [19], which applies a max filter to non-overlapping sub-regions within an image. The maximum value within each sub-region is selected and retained during max-pooling. The Pooling layer reduces the overall size of the data by retaining the most prominent features. This down sampling process makes the model translationally invariant, which means it can recognize patterns regardless of where they lie in the input image. Our model uses max pooling layers of size 2x2, which allow us to detect patterns in the input image. A CNN's Fully Connected layer is critical in the classification of images. In this layer, visual patterns and semantic information are encoded using feature vectors taken from previous layers. The network learns complex relationships between features and target classes by connecting neurons in the previous layer to neurons in the Fully Connected layer [20]. It produces an output vector representing predicted class probabilities by computing weights and biases during the forward pass. The outputs layer serves as the final layer in the CNN model, assigning labels to the result values obtained from the previous layer. As a result of the information captured in the previous layers, the CNN model accurately detects and classifies images given as input, providing disease-specific outputs such as clubbing, eczema, and leukocytosis. In this study, dataset was divided into two parts. The first was a training set and the second was a test set, after applying resizing and data augmentation techniques to the images. CNN was modeled using the training set, and the evaluation of its performance was determined by the test set. Our data training and testing methods enabled us to evaluate whether the CNN model was accurate and effective in classifying diseases based on its predictions. The approach allowed us to assess the model's capability to generalize well to new images and measure its performance on unseen data.

IV. Results and Discussion

A. Evaluation Metrics

In order to comprehensively measure the performance of our CNN model for nail disease classification, we applied well-known evaluation metrics. These metrics include Accuracy, Precision, Recall, and F1-score. Accuracy calculated in Eq. 2, is one of the fundamental evaluation metrics, that determines whether the model predicts correctly in general. It quantifies the proportion of total samples in the dataset to the correctly classified samples in the dataset. The accuracy score being higher indicates that the model is effectively classifying the diseases, providing a reliable measure of its performance. However, a complete picture of the model's ability to handle imbalanced datasets is not depicted by the accuracy alone. The precision calculated in Eq. 3 indicates how well the model can identify positive samples. Recall, which is sometimes also known as sensitivity or true positive rate, as shown in Eq.4 can be calculated by dividing the number of true positives by the number of false negatives. Based on all positive samples, it represents the proportion of correct positive classifications. The F1 score which is basically putting the precision and recalls into a single measure, allows for a balanced evaluation of the model. It is calculated as the harmonic mean of precision and recall.

$$Accuracy = \frac{TruePositive + TrueNegative}{TotalPositive + TotalNegative} \tag{2}$$

$$PrecisionRate = \frac{TruePositive}{TruePositive + FalsePositive} \tag{3}$$

$$RecallRate = \frac{TruePositive}{TruePositive + FalseNegative} \tag{4}$$

$$F1Score = 2 \times \frac{PrecisionRate \times RecallRate}{PrecisionRate + RecallRate} \tag{5}$$

B. Results

The model we proposed demonstrated excellent performance, achieving an accuracy of 98.44%. This indicates that the model is capable of classifying different nail conditions correctly. Moreover, it demonstrated that the model was capable of accurately identifying positive cases and avoiding

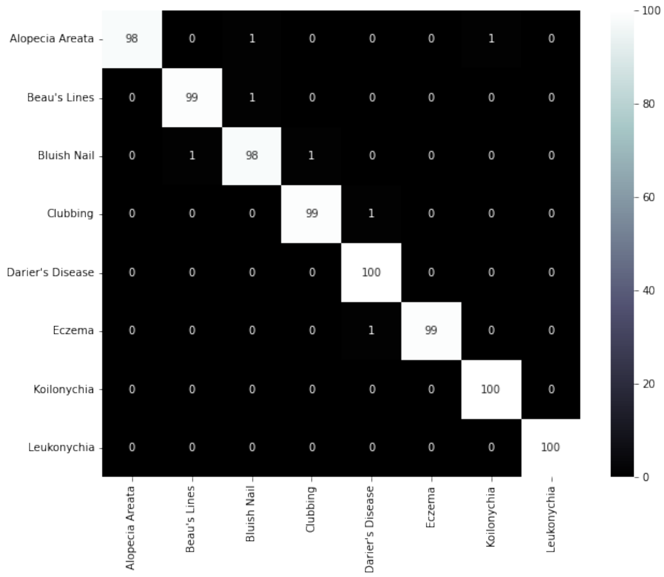

Fig. 4: Confusion Matrix

false negatives with precision and recall values of 99.22% and 98.44%, respectively. Combined precision and recall, the F1 score was calculated to be 99.02%, which indicates a balanced performance. Table III shows the Performance evaluation results.

TABLE III: Performance Evaluation

Parameters	Values Obtained
Accuracy	0.9844
Precision	0.9922
Recall	0.9844
F1-Score	0.9902

A confusion matrix was constructed as shown in Figure 4. to gain further insight into the model's performance. For each class, the confusion matrix offers a detailed breakdown of predicted and true labels. Analyzing the performance of the model is made easier by visualizing the distribution of samples across different classes.

Training of the model was monitored in terms of accuracy and loss. Over the training epochs, Figures 5 and Figure 6 demonstrate the changes in accuracy and loss. With each additional training epoch, the accuracy plot shows an increasing trend in the model's performance. A similar trend is apparent on the loss plot, indicating that the model learned the underlying patterns and minimized its loss. The achieved outcomes were also compared with state-of-the-art (SOTA) techniques, affirming the superiority of the proposed model as shown Figure 7. Our CNN-based approach to classify nail diseases surpasses prior studies with 98.44accuracy across 8 distinct classes, demonstrating its effectiveness. As opposed to [08] and [10], whose accuracies were 95.24% on binary class and 99.5model performs well across a broader range of conditions. As a result of rescaling, strategic flips and rotations, and the fine-tuning of data augmentation techniques,

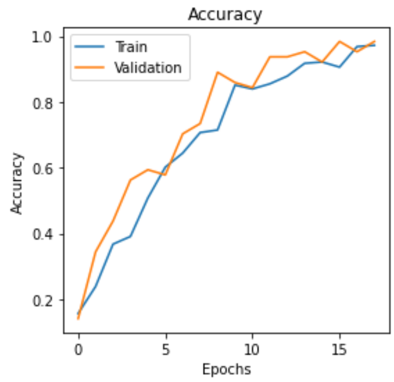

Fig. 5: Accuracy of the Model over Epochs

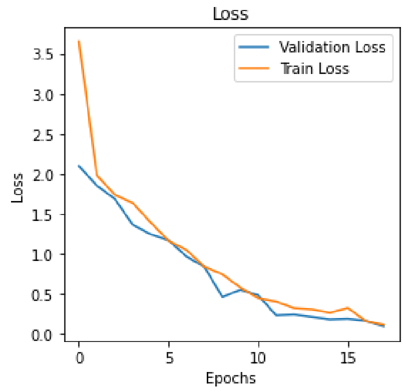

Fig. 6: Loss of the Model over Epochs

this success has been achieved. The streamlined approach ensures high accuracy as well as computational efficiency, demonstrating its practicality for dermatological applications.

V. Conclusion

In this study, a Deep Learning (DL) model was developed and trained to accurately classify nail diseases. In identifying and categorizing various nail diseases, the model achieved an impressive 98.44% accuracy with 99.22% precision, 98.44recall, and 99.02% F1 score, demonstrating its effectiveness. The

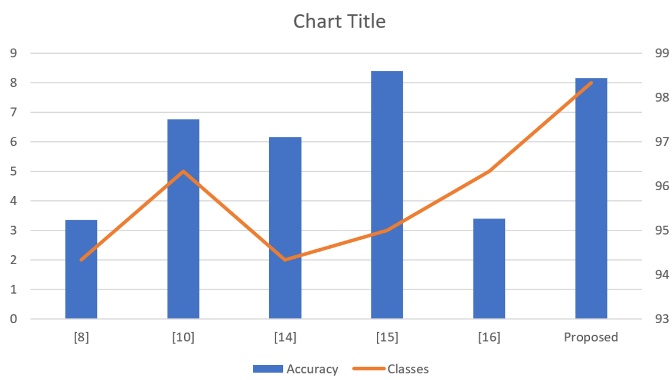

Fig. 7: Comparision with SOTA Techniques

model's performance was further enhanced by incorporating advanced techniques such as data augmentation and image preprocessing. DL provides a reliable tool for diagnosing and treating nail conditions, and thus the results of this study have important implications for healthcare professionals. With high precision, recall, and F1 scores, the model allows healthcare professionals to make accurate classification decisions. Furthermore, the confusion matrix provides valuable insights into the model's performance and application possibilities. Developing and integrating the model into clinical practice can be the focus of future research. A wider adoption of the model in healthcare settings can be achieved through continuous improvement of its performance, expansion of the dataset, and validation across diverse populations. With the power of DL in nail disease classification, healthcare professionals will be able to benefit from a more effective, reliable diagnostic tool that will ultimately improve the quality of patient care and outcome.

REFERENCES

[1] A. J. Wulkan and A. Tosti, "Pediatric nail conditions," *Clinics in dermatology*, vol. 31, no. 5, pp. 564–572, 2013.

[2] J. Velasco, C. Pascion, J. W. Alberio, J. Apuang, J. S. Cruz, M. A. Gomez, B. Molina Jr, L. Tuala, A. Thio-ac, and R. Jorda Jr, "A smartphone-based skin disease classification using mobilenet cnn," *arXiv preprint arXiv:1911.07929*, 2019.

[3] Z. Liu, H. Liu, Y. Xie, Y. Yao, X. Xing, and H. Ma, "Smart image follow-up of black pigmentation on the nail with convolutional neural networks," *Available at SSRN 3834276*.

[4] H. M. Sufian and G. P. Abebe, "Disease identification using finger nail image processing and ensemble nearest neighbor classifiers of color features," Ph.D. dissertation, 2021.

[5] C. Fletcher, R. Hay, and N. Smeeton, "Observer agreement in recording the clinical signs of nail disease and the accuracy of a clinical diagnosis of fungal and non-fungal nail disease," *British Journal of Dermatology*, vol. 148, no. 3, pp. 558–562, 2003.

[6] T. S. Indi and Y. A. Gunge, "Early stage disease diagnosis system using human nail image processing," *IJ Information Technology and Computer Science*, vol. 7, no. 7, pp. 30–35, 2016.

[7] R. Nijhawan, R. Verma, S. Bhushan, R. Dua, A. Mittal *et al.*, "An integrated deep learning framework approach for nail disease identification," in *2017 13th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*. IEEE, 2017, pp. 197–202.

[8] M. Yani, S. M. Budhi Irawan, S, and M. Casi Setiningsih, ST, "Application of transfer learning using convolutional neural network method for early detection of terry's nail," in *Journal of Physics: Conference Series*, vol. 1201, no. 1. IOP Publishing, 2019, p. 012052.

[9] S. L. Pinoliad, D. A. N. Dichoso, A. R. Caballero, and E. M. Albina, "Onyxray: A mobile-based nail diseases detection using custom vision machine learning," in *Proceedings of the 5th International Conference on Information and Education Innovations*, 2020, pp. 126–133.

[10] K. A. Muhaba, K. Dese, T. M. Aga, F. T. Zewdu, and G. L. Simegn, "Automatic skin disease diagnosis using deep learning from clinical image and patient information," *Skin Health and Disease*, vol. 2, no. 1, p. e81, 2022.

[11] A. Rehman, M. A. Khan, T. Saba, Z. Mehmood, U. Tariq, and N. Ayesha, "Microscopic brain tumor detection and classification using 3d cnn and feature selection architecture," *Microscopy Research and Technique*, vol. 84, no. 1, pp. 133–149, 2021.

[12] H. Hong, J. Lin, and F. Huang, "Tomato disease detection and classification by deep learning," in *2020 International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*. IEEE, 2020, pp. 25–29.

[13] K. Thomsen, A. L. Christensen, L. Iversen, H. B. Lomholt, and O. Winther, "Deep learning for diagnostic binary classification of multiple-lesion skin diseases," *Frontiers in medicine*, vol. 7, p. 574329, 2020.

[14] M. Sazzadul Islam Prottasha, S. Mahjabin Farin, M. Bulbul Ahmed, M. Zihadur Rahman, A. Kabir Hossain, and M. Shamim Kaiser, "Deep learning-based skin disease detection using convolutional neural networks (cnn)," in *The Fourth Industrial Revolution and Beyond: Select Proceedings of IC4IR+*. Springer, 2023, pp. 551–564.

[15] K. Li, B. Ao, X. Wu, Q. Wen, E. Ul Haq, and J. Yin, "Parkinson's disease detection and classification using eeg based on deep cnn-lstm model," *Biotechnology and Genetic Engineering Reviews*, pp. 1–20, 2023.

[16] P. Muthukannan *et al.*, "Optimized convolution neural network based multiple eye disease detection," *Computers in Biology and Medicine*, vol. 146, p. 105648, 2022.

[17] "Reuben industriustech n dataset," https://www.kaggle.com/datasets/reubenindustrustech/n, accessed: 26-March-2023.

[18] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights into imaging*, vol. 9, pp. 611–629, 2018.

[19] N. Murray and F. Perronnin, "Generalized max pooling," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 2473–2480.

[20] G. Chartrand, P. M. Cheng, E. Vorontsov, M. Drozdzal, S. Turcotte, C. J. Pal, S. Kadoury, and A. Tang, "Deep learning: a primer for radiologists," *Radiographics*, vol. 37, no. 7, pp. 2113–2131, 2017.