Deep Learning based Classification of Human Nail Diseases using Color Nail Images

Senar Ali Yamaç, Orhun Kuyucuoğlu, Şeyma Begüm Köseoğlu and Sezer Ulukaya
Electrical and Electronics Engineering Department
Trakya University
Edirne, Turkey
{saliyamac,orhunkuyucuoglu,sbegumkoseoglu,sezerulukaya}@trakya.edu.tr - ORCID: 0000-0003-0473-7547

***Abstract*—When the disorders that occur in the fingernails and toenails are not noticed early, they can turn into diseases that affect human life. These diseases in our hands and feet, which are mostly used organs in our daily work, also negatively affect the quality of life. In this study, it is aimed to detect 5 different nail diseases using deep learning architectures. Within the scope of the study, the performance of the 6 most recent deep learning architectures was compared with each other. Although the number of pictures in the open-access database used in the study is low, the obtained results seem to be successful.**

***Keywords*—Convolutional Neural Network (CNN); Data Augmentation; Nail diseases; Telehealth; Transfer learning.**

I. Introduction

In the healthcare field, many diseases can be identified by observing the differences in color and shape of human nails by medical experts. Early detection of diseases plays a significant role in human health. Changes in the color of the nails, in the nail tissue, and deformities are the symptoms of a problem in the body related to some diseases [1]. When classifying nail abnormalities, the color, texture, size, and shape of the nail are taken into account. Problems in the liver, lungs, and heart may manifest as changes in the nails [2]. After examining it with the naked eye, medical doctors evaluate it with clinical examination, dermatoscopy, diagnostic imaging, microbiological tests, and histopathological examination [3].

Today, artificial intelligence and machine learning not only benefit certain technological fields ([4], [5], [6]) but also contribute to dermatology in the medical field [7]. In this way, very small color and shape changes on the nail can be detected automatically using deep learning and image classification algorithms [8]. Early detection of many diseases has been aimed by using image processing techniques and progress has been made in this regard [2]. In this research, Convolutional Neural Network (CNN) models were utilized to classify and detect nail diseases from color nail images.

In this work, 5 types of nail diseases were aimed to be detected. These are called Blue nail, Clubbing, Splinter Hemorrhage, Muehrcke's Lines, and Onycholysis diseases. Figure 1 shows examples of related nail diseases that are also included in the publicly available dataset. Blue nail disease can occur due to a type of circulatory disorder that emerges when the oxygen level in the blood drops below normal. It may be due to disorders in the lungs, heart, blood cells, or blood vessels and should be checked cautiously [9]. In clubbing disease, deformation occurs in the fingers and nails. It can be seen in lung and heart diseases, as well as in cirrhosis and chronic bronchitis. The appearance of the nails is wide and has an oval shape during this disease [10]. Muehrcke's Lines disease is seen as double white horizontal lines on the nails. It appears in people with low albumin levels, kidney disease, malnutrition, and liver disease. It usually occurs on the second, third, and fourth nails of the fingers [10]. Onycholysis is described as the isolation of the nails from the nail ground. There are differences in nail color in the separated parts. It can occur in cases of nail trauma, psoriasis, and thyroid disease. Onycholysis may also be seen in patients with hyperthyroidism [10]. Splinter Hemorrhage is a brown-red, linear hemorrhage. These bleedings are observed under the nail. Usually, splinter bleeding is caused by nail trauma. It is most commonly seen in people with pericardial inflammation, hypertension, lung disease, and diabetes. It is also possible to be seen in people with diseases such as onychomycosis and psoriasis [11].

Recently, many researchers have been able to detect diseases in an early phase using artificial intelligence and image analysis techniques. In the next section, studies in the literature on artificial intelligence-based nail diseases detection are briefly explained.

Fig. 1. From left to right, sample images of Blue nail, Clubbing, Muehrcke's Lines, Onycholysis, and Splinter Hemorrhage diseases.

II. Literature Review

The work of [8] used a few Convolutional Neural Network (CNN) models for a deep learning-based solution. VGG16, ResNet50, VGG19 and DenseNet121 models were tested. The highest performance was reached with the VGG16 method with an accuracy rate of 92 %. Three types of nail disease have been identified. 192 melanoma, 248 yellow nails, and

2022 45th International Conference on Telecommunications and Signal Processing (TSP) | 978-1-6654-6948-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/TSP55681.2022.9851300

600 healthy nail images were used. In this work, they aimed to develop a user-friendly mobile application.

Five different deep learning models such as AlexNet, VGG16, GoogleNet, ResNet50 and DenseNet201 were used to evaluate the performance of the deep CNN approaches in the work of [2]. In the related study, four types of nail diseases were detected. A total of 280 images were used. They achieved the highest accuracy rate of 96.39 % with ResNet50 and DenseNet201 models.

In the work of [12], a transfer learning approach was performed using MobileNetV2. An expanded dataset of 7500 images was created. 4 skin disorders (Sebaceous cysts, Eczema, Guttate Psoriasis, Psoriasis) and 2 nail diseases (Yellow nail syndrome and Paronychia) were identified using deep learning. The employed model reached an accuracy rate of about 92.50 %. They have developed a web application called DermaDoc.

VGG16, which is a deep learning model, was used in the study of [13]. A classification model for 10 dermatological diseases was utilized. Only 2 of these classes belong to the nail disease group. 80 % of the 5500 images obtained from Dermnet datasets were used for training purposes and 20 % for testing purposes. Image processing approaches were also performed on disease images. In this study, an average accuracy of 76.30 % was obtained with fine-tuning setup.

The number of nail diseases is 3, 4, 2, 2 and the number of deep learning methods is 4, 5, 1, 1 in [8], [2], [12] and [13], respectively. This study aims to design a CNN that can detect certain nail diseases from nail images. In this study, a dataset consisting of 187 training and 47 testing images belonging to 5 different classes was used. Compared to related works on nail diseases, our work has a limited dataset, while the number of disease classes is high enough. Even though this limitation, an accuracy of 89 % has been achieved. The original contribution of the current work with respect to the literature is comparing several state-of-the-art deep learning models under the same experimental setting and modifying the best model (ResNet101V2) to further improve the accuracy.

III. Methodology

A. Dataset and Preprocessing Steps

The number of data is so important in deep learning projects. Because more data means a more comprehensive network. Access to publicly available data is just as difficult. We used the nail dataset that is provided on Kaggle in our study [14]. As a preliminary study, we have chosen five classes (Clubbing, Muehrcke's lines, Splinter hemorrhage, Onycholysis, Blue nail) from the provided dataset where the number of classes is comparable with the literature. There are a total of 234 images for five classes in the database. The dimensions of the images are in the range of 35x46 and 526x443. In the preprocessing step, we resize the images as 128x128. Then, we convert the NxM images to 128x128x3 arrays.

B. Data Augmentation and Train/Test Splitting

Data augmentation is described as a group of techniques that construct synthetic images from the available images. The synthetically generated dataset typically comprises minor changes in the original dataset where the estimations of the model are expected to be unvarying [15]. The synthetic dataset may also show combinations between out-of-distribution samples that would be very hard to make inferences. The data augmentation technique is one of the most effective approaches to impact the performance of the training of neural networks by feeding additional images or data [15].

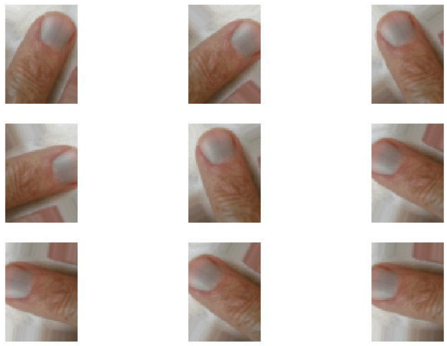

Fig. 2. Sample images after data augmentation step

We also applied data augmentation because the employed dataset was insufficient. In the data augmentation process, flipping, shifting, and zooming methods were utilized. Figure 2 represents examples of nail images after data augmentation. First, we split the dataset into 187 train and 47 test images by considering their weights and belonging only to either train or test set. After separating the dataset as train and test, data augmentation was performed merely on the training set. After the data augmentation process, the number of pictures in the training part has increased 7 times in total including the original images. The overall number of training images is 1309. The grouping of the images as train and test sets is shown below in Table I.

TABLE I. Distribution of Images

	Train	**Test**	**Total**
Muehrcke's Lines	26	7	33
Blue Nail	40	10	50
Clubbing	32	8	40
Onycholysis	40	10	50
Splinter Hemorrhage	49	12	61
Total	187	47	234

C. Model Architecture

Our model uses transfer learning method based on CNNs. We compared the deep learning models such as VGG16, DenseNet169, MobileNet, Xception, InceptionResNetV2, and

ResNet101V2 with each other aiming to get better performance. The reason for choosing the employed CNN architectures is to test the performance of the state-of-the-art deep learning models. To classify five different nail diseases, different parameter combinations were tried and the code was written in Python.

1) CNN: Deep CNNs have reached significant performance in various computer vision duties recently [16].

CNNs are deep learning approaches that get pictures as input and convolve them with filters to reveal features. An MxM picture is convolved with a dxd filter and this results in learning a similar attribute on the complete picture [17]. The mask moves after every step and the attributes are learned by the feature maps. These attribute maps hold the local receptive field of the image and process with shared weights [17], [18]. In our study, we preferred this method as a baseline model.

2) Transfer Learning: Transfer learning is a method of transferring the information of a pretrained system to a new domain (information transfer) instead of creating a new model from scratch. Transfer learning is inspired by the ability of humans for transferring knowledge among tasks. This approach focuses on leveraging the information from the original space to increase the learning capacity or to decrease the quantity of labeled samples needed in a target space [19]. More specifically, transfer learning uses pre-trained weights in all layers of CNN, except for fully connected layers that are added later. This process is provided by the freezing method. The knowledge of low-level features such as edges and corners is transferred to the model. However, the complex features specific to the problem are learned by the model during the training phase of the fully connected layers added later.

There are many deep learning models such as VGG16, VGG19, DenseNet, MobileNet, ResNet etc. Especially, these models are preferred for transfer learning schemes when the dataset is insufficient. These models are often used because deep learning models are trained with millions of data. The initial training of our CNNs mentioned was done with the ImageNet dataset.

3) ResNet101V2: ResNet consists of block structures consisting of convolution and pooling layers. It uses average pooling structures instead of fully connected layers and has 101 layers. In our study, we added 3 fully connected (FC) parts at the end of the ResNet101V2 model. The layers of our ResNet101V2 model are frozen, which prevents the layers from updating their weights. While we are training our model, only the weights of the layers we added later are updated. This approach is also known as the fine-tuning technique. The softmax activation operation is employed in the last output part of the model.

D. Model Training and Experimental Results

The training parameters for the employed method are selected as follows; the number of batch size is 32, the number of epochs is 50, and the learning rate of 0.001 (1e-3). To compile the system, we have utilized the Adam optimization technique and cross-entropy loss. The comparison of the accuracy of transfer learning models is shown below in Table II. As a result of the experiments, the accuracy of our model is 89 % when the ResNet101V2 architecture is employed. The InceptionResNetV2 and VGG16 models show the lowest performances on this classification task, respectively. Accuracy and loss graphics for the best model are presented in Figures 3 and 4, respectively.

TABLE II. ACCURACY OF DIFFERENT DEEP LEARNING ARCHITECTURES

Architecture	**Accuracy**
VGG16	70 %
DENSENET169	85 %
MOBILENET	85 %
XCEPTION	74 %
INCEPTIONRESNETV2	68 %
RESNET101v2	89 %

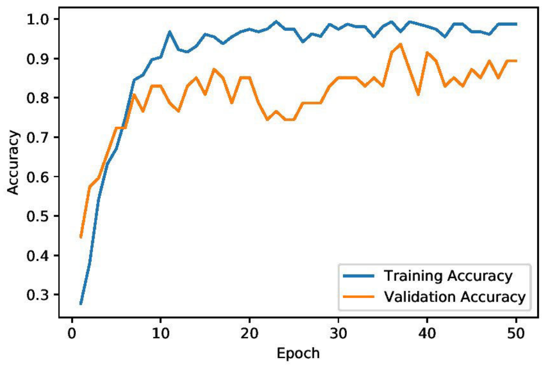

Fig. 3. Accuracy graph for the best model (ResNet101v2)

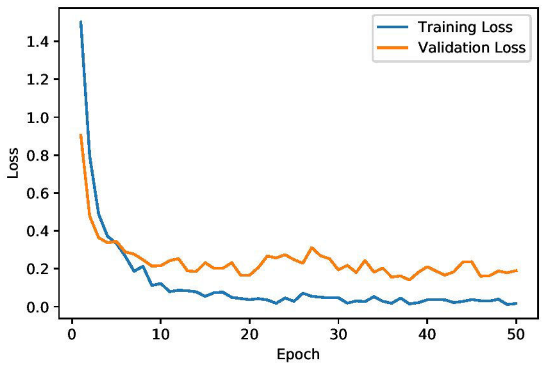

Fig. 4. Loss graph for the best model (ResNet101v2)

A confusion matrix was created to observe in which classes our model was more successful. The confusion matrix of the best model is presented in Figure 5. Rows represent actual diseases and columns represent predicted diseases. If the rows and columns intersect in the same disease for all test samples, this means the related disease was predicted correctly.

The performance metrics were computed after the classification step. These metrics were precision, recall, F1-score, and

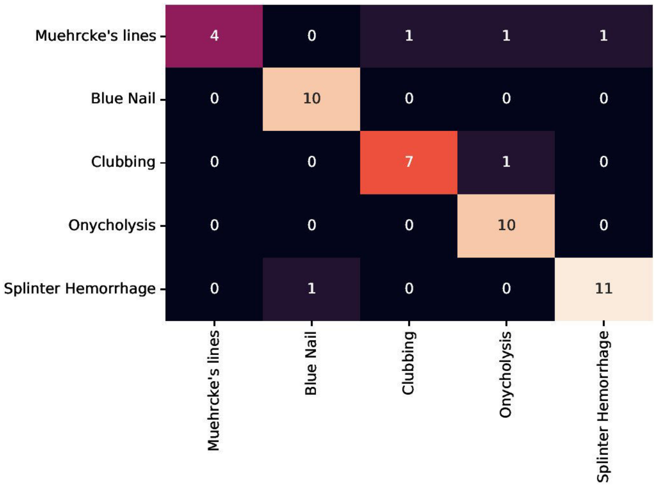

Fig. 5. Confusion matrix for the best model (ResNet101v2).

accuracy. The classification summary of the best-performed model is shown in Table III. According to Table III and Figure 5, Muehrcke's Lines disease is the most confused class and Blue Nail and Onycholysis diseases are the easily classified groups by our model.

TABLE III. DETAILED CLASSIFICATION METRICS

Classes	**Precision**	**Recall**	**F1-score**
Muehrcke's Lines	1.00	0.57	0.73
Blue Nail	0.91	1.00	0.95
Clubbing	0.88	0.88	0.88
Onycholysis	0.83	1.00	0.91
Splinter Hemorrhage	0.92	0.92	0.92
Performance Metrics			
accuracy			0.89
macro average	0.91	0.87	0.88
weighted average	0.90	0.89	0.89

The AUC values of the models for VGG16, Resnet101V2, MobileNet, Xception, InceptionResNetv2 and DenseNet169 are 90.25, 95.49, 93.95, 89.87, 87.60, 94.17, respectively. Moreover, the proposed model is also compared to EfficientNetB5 architecture which is used as baseline CNN and an AUC score of 53.19 is obtained for EfficientNetB5.

IV. Conclusion

In our study, the ResNet101V2 model, which we achieved the best accuracy among different transfer learning models, has been used to classify 5 different nail diseases. Although the number of images is insufficient and some of the samples are of low quality an acceptable accuracy of 89 % has been achieved. Compared to other studies in the literature, the number of classes is more besides contains fewer images. However, considerable accuracy has been achieved.

For future work, we intend to increase the performance of the classification system by combining the predictions of more than one transfer learning architecture in a hybrid fusion technique. Moreover, in future work, we also aim to extend the number of disease classes and pictures and show the effect of using different color spaces. Also, an early detection system for nail diseases may be developed by turning lightweight MobileNet into a mobile application due to its fast operation, which is a noteworthy model when compared to the ResNet101V2 model.

References

[1] 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.

[2] J. Abdulhadi, A. Al-Dujaili, A. J. Humaidi, and M. A.-R. Fadhel, "Human nail diseases classification based on transfer learning," *ICIC Express Letters*, vol. 15, no. 12, pp. 1271–1282, 2021.

[3] U. Wollina, P. Nenoff, G. Haroske, and H. A. Haenssle, "The diagnosis and treatment of nail disorders," *Deutsches Ärzteblatt International*, vol. 113, no. 29-30, p. 509, 2016.

[4] Z. Cömert, A. Şengür, Ü. Budak, and A. F. Kocamaz, "Prediction of intrapartum fetal hypoxia considering feature selection algorithms and machine learning models," *Health Information Science and Systems*, vol. 7, no. 1, pp. 1–9, 2019.

[5] K. M. Sünnetci, M. Ordu, and A. Alkan, "Gait based human identification: a comparative analysis," *Computer Science*, no. Special, pp. 116–125, 2021.

[6] L. Pham, A. Schindler, M. Schutz, J. Lampert, S. Schlarb, and R. King, "Deep learning frameworks applied for audio-visual scene classification," in *Data Science–Analytics and Applications*. Springer, 2022, pp. 39–44.

[7] M. M. Azad, A. Ganapathy, S. Vadlamudi, and H. Paruchuri, "Medical diagnosis using deep learning techniques: A research survey," *Annals of the Romanian Society for Cell Biology*, vol. 25, no. 6, pp. 5591–5600, 2021.

[8] M. Mehra, S. D'Costa, R. D'Mello, J. George, and D. Kalbande, "Leveraging deep learning for nail disease diagnostic," in *2021 4th Biennial International Conference on Nascent Technologies in Engineering (ICNTE)*. IEEE, 2021, pp. 1–5.

[9] H. Pandit and D. Shah, "A system for nail color analysis in healthcare," in *2013 International Conference on Intelligent Systems and Signal Processing (ISSP)*. IEEE, 2013, pp. 221–223.

[10] R. S. Fawcett, S. Linford, and D. L. Stulberg, "Nail abnormalities: clues to systemic disease," *American Family Physician*, vol. 69, no. 6, pp. 1417–1424, 2004.

[11] R. N. Saladi, A. N. Persaud, D. Rudikoff, and S. R. Cohen, "Idiopathic splinter hemorrhages," *Journal of the American Academy of Dermatology*, vol. 50, no. 2, pp. 289–292, 2004.

[12] M. Begum, A. Dhivya, A. J. Krishnan, and S. Keerthana, "Automated detection of skin and nail disorders using convolutional neural networks," in *2021 5th International Conference on Trends in Electronics and Informatics (ICOEI)*. IEEE, 2021, pp. 1309–1316.

[13] A. K. Sah, S. Bhusal, S. Amatya, M. Mainali, and S. Shakya, "Dermatological diseases classification using image processing and deep neural network," in *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*. IEEE, 2019, pp. 381–386.

[14] "Nail dataset," *Kaggle*. Accessed April 21, 2022. [Online]. Available: https://www.kaggle.com/datasets/reubenindustrustech/nail-dataset-new

[15] C. Shorten, T. M. Khoshgoftaar, and B. Furht, "Text data augmentation for deep learning," *Journal of Big Data*, vol. 8, no. 1, pp. 1–34, 2021.

[16] P. Chen, S. Liu, H. Zhao, and J. Jia, "Gridmask data augmentation," *arXiv preprint arXiv:2001.04086*, 2020.

[17] R. Chauhan, K. K. Ghanshala, and R. Joshi, "Convolutional neural network (cnn) for image detection and recognition," in *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)*. IEEE, 2018, pp. 278–282.

[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, no. 4, pp. 611–629, 2018.

[19] W. Ying, Y. Zhang, J. Huang, and Q. Yang, "Transfer learning via learning to transfer," in *International Conference on Machine Learning*. PMLR, 2018, pp. 5085–5094.