

Automatic Detection of Lupus Butterfly Malar Rash Based on Transfer Learning

Jhonatan Souza*, Tiago Mota de Oliveira*, Claudemir Casa*, André Roberto Ortoncelli*†

* Departamento de Informática, Universidade Federal do Paraná, Curitiba, Brazil

† Coordenação de Engenharia de Software, Universidade Tecnológica Federal do Paraná, Dois Vizinhos, Brazil

Abstract—This work presents an approach to the automatic detection of Butterfly Malar Rash (BMR) in images. BMR is a Lupus symptom characterized by a reddish facial rash that appears symmetrically in the cheeks and the back of the nose. The proposed approach is based on Transfer Learning, a popular approach in Deep Learning that consists in the use of pre-trained models as the starting point for computer vision and natural language processing tasks. To perform the experiments, a database was created with images manually collected from the Instagram social network, searching for images with #butterflyrash. We evaluated the proposed approach with eight Convolutional Neural Networks (CNN) architecture. The experimental results are good results, with a precision of up to 0.957.

Index Terms—Lupus diagnosis, Skin lesions, Deep Learning, Computer Vision

I. INTRODUCTION

Lupus Erythematosus (LE) is an inflammatory and autoimmune disease, in which the body develops antibodies against its own cells, which can affect joints, skin, kidneys, blood cells, brain, heart, and lungs. The number of symptoms that can occur makes diagnosis difficult.

There is no cure for Lupus, that more severe cases can cause death, but patients can have a prolonged survival if they receive the proper treatment [1], [2]. In this context, the precise diagnosis of lupus is extremely important.

The diagnosis of Lupus is based on eleven complex criteria, ranging from clinical tests to the individual's report. These criteria were first published in 1971, but have undergone several revisions since then [3].

Based on these criteria, the diagnosis of Lupus can be slow and last for months. Therefore, develop techniques that facilitate and streamline lupus diagnosis is very important.

In this context, this work presents an approach to the automatic detect Butterfly Malar Rash (BMR) - a visual symptom of Lupus, characterized as a reddish facial rash similar to the wings of a butterfly, that appears symmetrically in the cheeks and the back of the nose.

Few studies describe computational methods for automatic detection of facial skin lesions that are symptoms of Lupus. [4], [5]. In [4], unsupervised learning was used to detect BMR in images generated artificially by Generative Adversarial Networks (GANs). In [5], Transfer Learning was explored to detect facial Malar Rash with a model trained to detect Lupus skin rashes, but not specifically BMR.

The face area analyzed is one of the main differences in our approach to the related works. In [5], only the skin lesion

area is used. In [4] the entire face area is used. Our method also analyzes a large part of the face area, but unlike [4], we do not use the regions of the eyes and forehead, because the BMR does not appear in those regions.

The proposed approach to BMR detection uses Transfer Learning to transfer information from a neural network that has already been trained, to solve a different task. We use a model pre-trained with the ImageNet database [6] to create a new model trained in our database for BMR detection.

Our experimental database consists of 905 images of BMR. The database images it was manually collected from the Instagram social network, searching for images with #butterflyrash.

We need to produce our own database, because we do not find in the literature, a database that meets our needs.

To evaluate the proposed approach, experiments were performed with eight pre-trained models with eight CNN architectures: i) Resnet-50 [7]; ii) Interception V3 [8]; iii) Inception-Resnet V2 [9]; iv) Densenet [10]; v) VGG-16 [11]; vi) Xception [12]; vii) Mobilenet [13]; and viii) NasNetLarge [14].

In the experiments, we evaluate the precision of the proposed approach. The experimental results are good. The best results were obtained with Densenet-121 architecture, with a precision of 0.957.

The remaining of this article is organized as follows. Section II has related works and foundations. Details about the database and the proposed approach are in Section III. Results are in Section IV. Finally, the Section V concludes the paper.

II. RELATED WORKS AND FOUNDATIONS

This Section presents important concepts for understanding the proposed approach. Subsection II-A has concepts of CNN and the details of the architectures used in our approach. The Subsection II-B as Transfer Learning concepts. Subsection II-C has a literature review related to lupus automatic detection methods.

A. Convolutional Neural Networks

CNN is a Deep learning algorithm that, through an input image, can assign weights, which means that the network learns in order to differentiate one from the other. This of neural network consists of an input layer that will be an image, several convolutional layers, and an output layer as a classifier.

The first use of CNN was in 1988 by Yann LeCun [15] for document recognition and since then CNN has become more

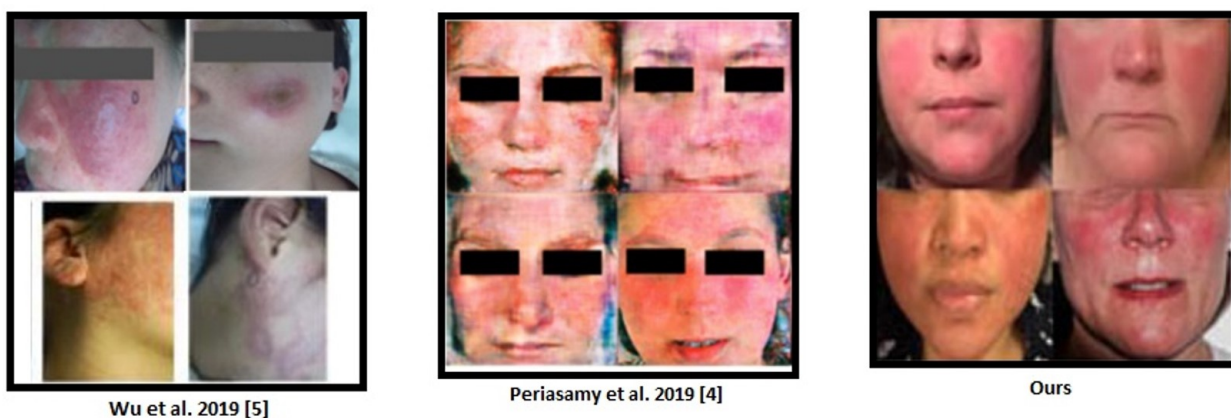


Fig. 1: Examples of the face area used by the lupus skin rash detection methods

complex and has been applied to a wide variety of problems, such as classifying diseases to autonomous cars and bringing great results.

CNN proved to be a great method for working with images. For over the years, CNN's has been bringing excellent results, through its models. Whether for activities such as image classification, detection and recovery [16]–[18].

One of the prominent research fields nowadays is CNN. Several applications of CNN for disease classification can be highlighted: Covid-19 [19], [20], pulmonary nodes [21], skin cancer [22], [23] and Alzheimer's [24].

Our work proposes an approach to BMR detection that uses a CNN model pre-trained to create a new model trained to detect BMR. In our experiments, we explore eight CNN architectures, that was pre-trained with the ImageNet database [6]. Characteristics of the explored CNN architectures as follows:

- 1) *Resnet-50*: Resnet-50 is a minor variation of ResNet 152, and has 48 layers. Being the convolutional layers: 1 max polling and 1 Average Pool layer. The Resnet architecture uses the concept of residual block, which apply shortcuts between the layers and add the values of initial inputs of the layers and the function ReLU of output [7].
- 2) *Inception V3*: The Inception V3 architecture has a differential through architectures called inceptions. Those who are extractors of convolutional characteristics, with the function of learning with few parameters. Inceptions modules can facilitate the mapping process between channels and spatial correlations, by factoring out the series of operations by examining them [8].
- 3) *Inception-Resnet V2*: The Inception-Resnet V2 architecture is a combination of Resnet and Inception, which is capable of using residual connections while maintaining the diversity of scale of the network. In order to improve results [5], [9].
- 4) *Densenet*: Densenet has a structure that aims to work, by adding connections between the layers. And the output results are added from the inputs to the subsequent

layers. And that way the architecture can improve performance with fewer parameters [5], [10].

- 5) *VGG16*: Developed by Simonyan and Zisserman [11], VGG16 is an architecture that has 16 convolutional layers, and 138 million parameters. With six blocks of various layers, the first being made by a combination of layers of convolution. And the last fully connected block. The first convolution block has two layers, with 64 neurons, 3x3 convolutional filters and 2x2 max polling. And subsequent blocks increase the number of filters per block.
- 6) *Xception*: Xception is a combination of the Inception structural idea with the concept of depthwise separable convolutions [12], each filter will convolve with an input channel individually. In the common convolution the filter converts with all input channels, adding the results to improve performance during the convolution process by decreasing network operations during training.
- 7) *Mobilenet*: Proposed by Google in 2017 [13], for mobile applications, it aims to reduce the size and complexity of layers with a focus on efficiency. Like Xception, this CNN structure is based on depthwise separable convolutions to make your architecture lighter.
- 8) *NasNetLarge*: Unlike other architectures, Nasnet was made through recursive stages called blocks. This structure was designed to learn the ideal set of data of interest. As it is a costly approach when the data set is large, a project for a new search space was proposed [14], which allows the transfer of learning from a small dataset to a large dataset.

B. Transfer learning

Transfer Learning is a Deep Learning technique used to transfer information from a neural network that has already been trained, for a given activity, to solve a different task. Through this knowledge transfer procedure, it can help reduce training time, improve network accuracy and work with a smaller database for new training [25].

The Transfer Learning method is the use of pre-trained neural networks to make up for the lack of training data sets. With this, trained networks are used to extract characteristics used in fine tuning, a method for adjusting the network parameters. Incrementally adapting pre-trained resources to new data [26].

In order to solve complex problems with little data, and to reduce training time, the Transfer learning technique is used frequently. This method has several uses, such as to solve problems related to image classification. Characteristic vectors are used, which are generated by a deep neural network, trained to recognize characteristics of an already trained database. After the first training, this network can be used as an entrance to a new neural network using a new database.

With the growth of the internet large databases with images are being created for the most varied applications. But there is a great lack of database for various diseases such as lupus.

Transfer Learning has been showing great results by optimizing training time and generating new models with a small amount of data.

C. Lupus Automatic Detection

Lupus is a chronic autoimmune disease that causes the immune system to attack its own tissues, the symptoms of the disease are varied and similar to other diseases causing difficulties in the diagnosis [4].

In addition to BMR, other symptoms are considered in the Lupus diagnosis, such as, photosensitivity, oral ulcers, arthritis, serositis, renal disorder, neurologic disorder, hematologic disorder, immunologic disorder, antinuclear antibody, and discoid rash skin that can appear on different body parts [27], [28].

There are several studies in the literature that explore machine learning-based approaches to the Lupus diagnosis [29], [30], but few studies are related to Malar Rash detection [4], [5] on images.

The [5]'s method, like our approach is based on transfer learning. In [5], has performed studies with a large clinical image dataset of skin diseases (including Malar Rash) from different body parts. In [4] it is demonstrated the use of artificially generated BMR images generated from Generative Adversarial Networks to train a model that differentiates Lupus from its other counter skin diseases using a Neural Network Classifier.

Despite the small amount of work related to facial Lupus Malar Rash detection on images, it is possible highlight recent works related to skin rash detection [22], [23], [31], [32], which is a concept directly related to our work.

III. PROPOSED APPROACH

We propose an approach to BMR detection based on Transfer Learning. This Section presents the proposed approach. Subsection III-A presents details about the database. Subsection III-B presents the steps of the proposed approach.

A. Database

Since we did not find literature in a public database with the necessary characteristics to perform our experiments, a manual search was necessary to produce our experimental database. The images were manually downloaded from the Instagram social network. We use the #butterflyrash to locate images with BMR.

Our experimental database, it is composed of 905 images of BMR, being 227 images of male and 678 images of female faces. The greater number of female faces is justified by the fact that lupus is more common in this gender [33].

Since our method considers only the face region, we manually segment this area in all images. We do not use the regions of the eyes and forehead, because the BMR does not appear in those regions.

B. Steps of the proposed approach

The proposed approach to detect BMR uses a classifier created with a neural network based on a pre-trained model. The training process consists of five steps. Details of each one of the steps as follow:

- 1) **Data acquisition:** This step represents the process that we execute to produce our database - presented in Subsection: III-A.
- 2) **Pre-processing:** All the images in our database it was pre-processed. The pre-processing is done with the following steps:
 - a) each one image was adjusted to the a standard size: 224x224. We resize the entire image to this size, because this is the pattern of the ImageNet database.
 - b) the face it was manually segmented in the image.
- 3) **Data augmentation:** data augmentation is applied to generate new images and increase our database [34].
- 4) **Pre-training:** In this step a model it was trained with the ImageNet database [6]. We performed eight experiments to validate our approach with each one of the CNN architectures that were presented in the Subsection II-A. The ImageNet is one of the largest image databases, with 1.281.167 images divided into 1.000 classes, and a set of 50.000 images for tests.
- 5) **Final training:** Global Average Pooling (GAP) was used, as pre-trained networks have many parameters in the last layers. For this reason, a reduction in dimensionality was carried out. And the GAP connected to the last layer of the pre-trained model was applied. Two other 1024-d fully connected layers were connected with the ReLU activation function, and then another layer was connected 512-d by the same activation function. And finally, the last layer with 2 neurons with the softmax activation function. In order to return the probability of each of the two classes, positive or negative for the BMR.
- 6) **Malar Rash Detection:** the last trained model it was used to BMR in an image.

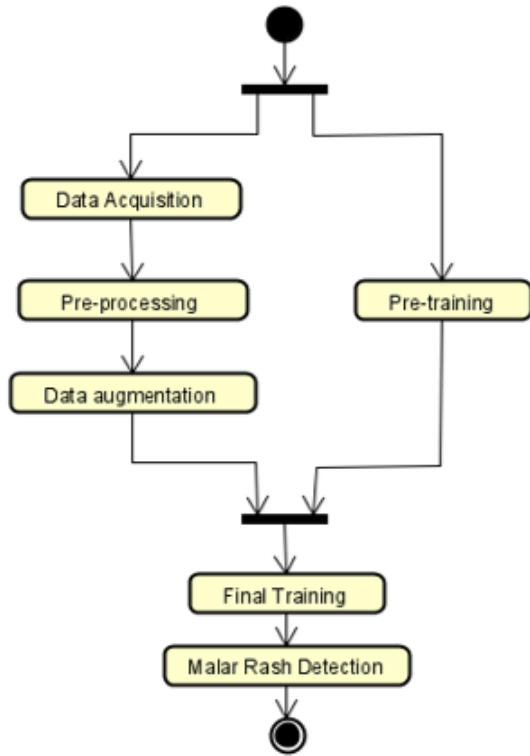


Fig. 2: Activity diagram of the proposed approach

IV. RESULTS AND DISCUSSIONS

This section aims to present the results of the experiments carried out to evaluate the proposed approach. Details of the evaluation metric are in the Subsection IV-A. The experimental results as in the Subsection IV-B. Finally, the Subsection has the analysis of the results.

A. Evaluation Metrics

We use three measures to evaluate experimental results: precision (the fraction of relevant instances among the retrieved instances), recall (the fraction of the total relevant instances that were actually retrieved) and F1-Score (or f-measure, is a harmonic mean between precision and recall). Four parameters it was used to compute these measures:

- Number of True Positives (NTP): numbers of images with BMR in which BMR was detected;
- Number of False Positives (NFP): number of images with BMR in which BMR was not detected;
- Number of False Negatives (NFN): numbers of images without BMR in which BMR was detected;
- Number of True Negative (NTN): numbers of images without BMR in which BMR was not detected;

The precision, recall and F1-Score are computed with the Equations 1, 2, 3.

$$precision = \frac{NTP}{NTP + NFP} \quad (1)$$

$$recall = \frac{NTP}{NTP + NFN} \quad (2)$$

$$F1 - Score = 2 * \frac{precision * recall}{precision + recall} \quad (3)$$

B. Experimental Results

We execute eight experiments instances to evaluate the proposed approach. In each instance, it was used a different CNN architecture to train the pre-trained model. Details of each of the CNN architectures used were in Subsection II-A.

Table I as the recall, precision and F1-score of all experimental instances and also the average result.

Net	Recall	Precision	F1-Score
Resnet50	0.912	0.909	0.91
Inception-Resnet V2	0.826	0.868	0.84
InceptionV3	0.839	0.863	0.85
Densenet-121	0.941	0.957	0.94
Xception	0.826	0.876	0.85
NASNetLarge	0.651	0.717	0.68
VGG16	0.871	0.891	0.88
Mobilenet	0.898	0.893	0.89
Average Result	0.845	0.871	0.85

TABLE I: Experimental results

C. Analysis of results

In the mean, our experiments showed recall, precision and F1-score greater than 0.84. Densenet-121 presented results greater than 0.94 - the best experimental results.

Our results are similar to those of [5] - a recent method for cutaneous lesion detection. On average, our results were higher, which indicates that our approach got good results.

V. CONCLUSION

This work presents an approach to automatically detect BMR. The approach is relevant because BMR is one of the symptoms of Lupus, which is a disease that is difficult to diagnose because lupus is a disease that is difficult to diagnose because its diagnosis is based on several criteria.

Our approach combines a strategy to obtain a set of images and a method for BMR detection based on Transfer Learning that is pre-trained with ImageNet database. The pre-trained model is used to create the first layer of the final model, that is trained with our database.

Experiments were carried out with eight models pre-trained with eight CNN architectures. The experimental results it is good, reaching an accuracy of 0.957 with the Densenet-121.

The main difficulty we find to develop this work is the lack of a large database with BMR images. For this reason, we created our database with images from the social network Instagram.

We pretend, in future works, apply our approach to develop a mobile app to streamline the process of diagnosing lupus. The results with the Mobilenet architecture, suggest that the development of this application is viable.

REFERENCES

- [1] V. P. Bernardes, L. D. B. Oliveira, and C. Marcon, "Lupus eritematoso sistêmico juvenil: Diagn [ostico de doença crônica e dinâmica familiar," *Barbarói*, pp. 75–90, 2011.
- [2] D. J. Wallace, *The lupus book: A guide for patients and their families*. Oxford University Press, 2019.
- [3] C. Yu, M. E. Gershwin, and C. Chang, "Diagnostic criteria for systemic lupus erythematosus: a critical review," *Journal of Autoimmunity*, vol. 48, pp. 10–13, 2014.
- [4] P. Periasamy and V. L. Byrd, "Generative adversarial networks for lupus diagnostics," in *Practice and Experience in Advanced Research Computing on Rise of the Machines (learning)*, 2019, pp. 1–8.
- [5] Z. Wu, S. Zhao, Y. Peng, X. He, X. Zhao, K. Huang, X. Wu, W. Fan, F. Li, M. Chen *et al.*, "Studies on different cnn algorithms for face skin disease classification based on clinical images," *IEEE Access*, vol. 7, pp. 66 505–66 511, 2019.
- [6] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *IEEE Conference on Computer Vision and Pattern Recognition*. Ieee, 2009, pp. 248–255.
- [7] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *International Conference on Machine Learning*, 2010.
- [8] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818–2826.
- [9] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *AAAI Conference on Artificial Intelligence*, 2017.
- [10] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *IEEE Conference on Computer Vision and Pattern recognition*, 2017, pp. 4700–4708.
- [11] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *International Conference on Learning Representations*, 2015.
- [12] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 1251–1258.
- [13] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. A. Mobilenets, "Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.
- [14] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning transferable architectures for scalable image recognition," in *IEEE Conference on Computer Vision and Pattern recognition*, 2018, pp. 8697–8710.
- [15] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [16] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-scale video classification with convolutional neural networks," in *IEEE conference on Computer Vision and Pattern Recognition*, 2014, pp. 1725–1732.
- [17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, 2012, pp. 1097–1105.
- [18] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1915–1929, 2012.
- [19] X. Wang, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, W. Liu, and C. Zheng, "A weakly-supervised framework for covid-19 classification and lesion localization from chest ct," *IEEE Transactions on Medical Imaging*, 2020.
- [20] S. Hu, Y. Gao, Z. Niu, Y. Jiang, L. Li, X. Xiao, M. Wang, E. F. Fang, W. Menpes-Smith, J. Xia *et al.*, "Weakly supervised deep learning for covid-19 infection detection and classification from ct images," *IEEE Access*, 2020.
- [21] A. E. Jatobá, L. L. Lima, and M. C. Oliveira, "Pulmonary nodule classification with 3d convolutional neural networks," in *Anais do XV Workshop de Visão Computacional*. SBC, 2019, pp. 67–72.
- [22] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.
- [23] kG Utsch, C. dos Santos, and J. Samatelo, "Convolutional neural network for skin lesion classification," in *Workshop de Visão Computacional*. SBC, 2018, pp. 105–110.
- [24] S. Basaia, F. Agosta, L. Wagner, E. Canu, G. Magnani, R. Santangelo, M. Filippi, A. D. N. Initiative *et al.*, "Automated classification of alzheimer's disease and mild cognitive impairment using a single mri and deep neural networks," *NeuroImage: Clinical*, vol. 21, p. 101645, 2019.
- [25] R. Ribani and M. Marengoni, "A survey of transfer learning for convolutional neural networks," in *Conference on Graphics, Patterns and Images - Tutorials (SIBGRAP-T)*. IEEE, 2019, pp. 47–57.
- [26] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
- [27] E. M. Tan, A. S. Cohen, J. F. Fries, A. T. Masi, D. J. Mcshane, N. F. Rothfield, J. G. Schaller, N. Talal, and R. J. Winchester, "The 1982 revised criteria for the classification of systemic lupus erythematosus," *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, vol. 25, no. 11, pp. 1271–1277, 1982.
- [28] M. C. Hochberg, "Updating the american college of rheumatology revised criteria for the classification of systemic lupus erythematosus," *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, vol. 40, no. 9, pp. 1725–1725, 1997.
- [29] S. Gomathi and V. Narayani, "A proposed framework using cac algorithm to predict systemic lupus erythematosus (sle)," in *World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave)*, 2016, pp. 1–6.
- [30] S. Balderas-Díaz, K. Benghazi, G. Prados, and E. Miró, "Designing configurable and adaptive systems in ehealth," in *Workshop on ICTs for Improving Patients Rehabilitation Research Techniques*. New York, NY, USA: Association for Computing Machinery, 2015, p. 118–121.
- [31] T. A. Rimi, N. Sultana, and M. F. A. Foyzal, "Derm-nn: Skin diseases detection using convolutional neural network," in *International Conference on Intelligent Computing and Control Systems*. IEEE, 2020, pp. 1205–1209.
- [32] 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.
- [33] S. Z. Y. Wasef, "Gender differences in systemic lupus erythematosus," *Gender Medicine*, vol. 1, no. 1, pp. 12–17, 2004.
- [34] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, p. 60, 2019.