

Using Deep Learning to Detect Melanoma in Dermoscopy Images

Julie Ann A. Salido and Conrado Ruiz Jr.

Abstract—Melanoma is a common type of cancer that affects a significant number of people. Recently, deep learning techniques have been shown to be highly accurate in classifying images in various fields. This study uses deep learning to automatically detect melanomas in dermoscopy images. First, we preprocess the images to remove unwanted artifacts, such as hair, and then automatically segment the skin lesion. We then classify the images using a convolutional neural network. To evaluate its effectiveness, we test this classifier using both preprocessed and unprocessed images from the PH² dataset. The results show an outstanding performance in terms of sensitivity, specificity, and accuracy. In particular, our approach was 93% accurate in identifying the presence or absence of melanoma, with sensitivities and specificities in the 86%–94% range.

Index Terms—Deep learning, dermoscopy image, image processing, melanoma detection

I. INTRODUCTION

Skin cancer is one of the most common malignancy types. In the US alone, over 5 million cases have been diagnosed every year [1]. Melanoma is one of the most common and fatal types of skin cancer and involves the unrestrained growth of pigment-producing cells. In the US, it is responsible for 4% of all cancer deaths and 6 out of every 7 skin cancer-related deaths [2]. It is estimated that 9,730 people will die from melanoma in 2017 [3].

Melanoma can affect anyone although Caucasians are more likely to suffer from it than other races. However, it is less obvious in people with dark skin, so it is often diagnosed at a later stage when the disease is more advanced. This is because many people have the mistaken impression that people of color cannot get skin cancer [4]. It caused the most cancer-related deaths globally in 2016, with an estimate of 76,380 people having invasive melanomas, of which approximately 46,870 were male and 29,510 were female [5].

Melanoma is highly treatable if it is detected early, but advanced melanoma can spread to the lymph nodes and other organs, which can be fatal. Medical experts and professional equipment are crucial for early and accurate melanoma detection. In contrast, more limited access to such expert opinions makes it a challenge to provide adequate levels of care to the populations that are at the risk of this disease.

Usually, patients initially see a skin abnormality. Then medical experts use dermoscopy for diagnosis. This is a

high-resolution skin-imaging process that reduces skin surface reflections, allowing doctors to examine the deeper underlying structures. They are used to non-invasively evaluate [6] *in vivo* the colors and microstructures of the epidermis, dermoepidermal junction, and papillary dermis. This has opened up a new avenue for examining pigmented skin lesions and especially identifying the early stages of melanoma [7], [8]. Using this approach, specially trained medical experts have demonstrated diagnostic accuracies as high as 75%–84% [9]. However, the diagnostic performance drops significantly if the doctors have not been adequately trained [10], [11].

To address the issues caused by limited access to specialists, especially in developing countries, there has been considerable research focusing on developing automated image analysis systems that can detect skin diseases based on dermoscopy images. There have been several recent publications reviewing the different techniques used [9], [12] as well as dermoscopy papers developing diagnostic criteria for early melanoma detection [13]–[16]. However, these criteria still involve dermoscopy image characteristics that can only be assessed by dermatologists or medical specialists.

In this paper, we proposed a method of detecting and removing hair from dermoscopy images and present a way of classifying skin lesions using deep learning. We then evaluate both the hair detection algorithm and the classifier using images from the PH² dataset [1].

The rest of the paper is organized as follows. In Section II, we discuss some related literature. In Section III, we present our proposed methods before implementing and evaluating them in Section IV and analyzing the results numerically. Finally, in Section V, we present our conclusions and plans for future work.

II. RELATED WORK

Deep learning techniques attempt to enable computers to learn from a large number of examples. Deep learning models automatically categorize input datasets, such as images, audio, or documents, directly. They can yield excellent and up-to-date classifications that can sometimes beat human assessment.

Deep learning uses neural network architectures with several layers that are trained with large datasets, with the most popular type [17] being convolutional neural networks (CNNs). Fig. 1 shows a fully connected neural network, a characteristic of the CNN. In this study, we use preprocessed dermoscopy images as input and obtain the classification result (i.e., skin disease type) as output.

CNNs have been shown to be very effective for a number

of complex image recognition tasks; they have also been used to extract features. For example, AlexNet uses a pretrained CNN to extract features for training an image classifier [18]. In addition, several pretrained networks are currently gaining popularity, most of which have been trained on the ImageNet [19] dataset, which contains 1.2 million training images that are taken from the Internet and divided into ~1000 object categories. AlexNet uses ImageNet as an image source since it is reputable and widely used, with freely available training datasets and benchmarks. Fig. 2 shows the ImageNet CNN architecture used by AlexNet for image classification.

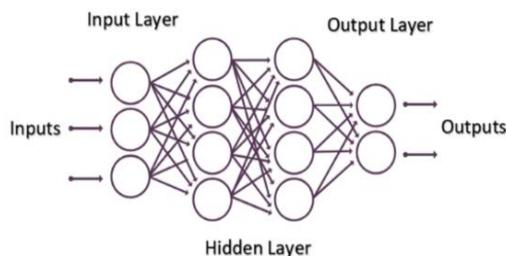


Fig. 1. Neural network with fully connected layers of nodes.

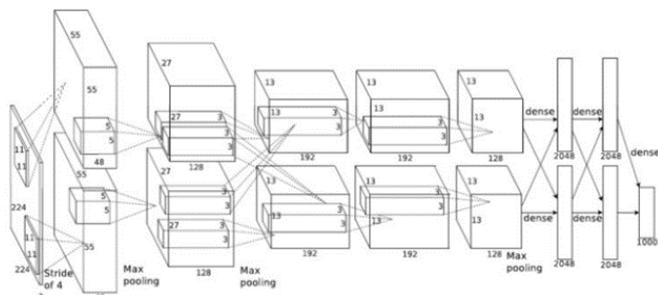


Fig. 2. ImageNet CNN architecture used by AlexNet [19].

As noted above, AlexNet is a pretrained ImageNet-based deep CNN for image classification [20]. It has been able to learn a promising set of abstract features for a comprehensive range of images, as shown by the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [21]. It has 25 layers. The first layer (Image Input), receives $227 \times 227 \times 3$ images with “zero-center” normalization, while the last layer (Classification Output) calculates the network’s performance based on the given targets and outputs as well as optional performance weights and other parameters. It has 5 convolution layers (Layers 2, 6, 10, 12, and 14) and 3 4096×4096

fully connected layers. It can also be used for transfer learning wherein a pretrained model is utilized as a base to gain knowledge about other tasks and perform feature extraction. It is often quicker and easier to refine a network using transfer learning than to train a new network from scratch.

III. SKIN LESION CLASSIFICATION VIA DEEP LEARNING

This section describes our approach to classifying skin lesions in dermoscopy images using deep learning. This begins by preprocessing the images (using hair removal and inpainting) and then classifying them using deep learning. Fig. 3 shows a flowchart of the proposed approach.

Here, we outline the steps involved in applying our methods to the PH² dataset, created by the Automatic computer-based Diagnostic system for Dermoscopy Images (ADDI) project [1].

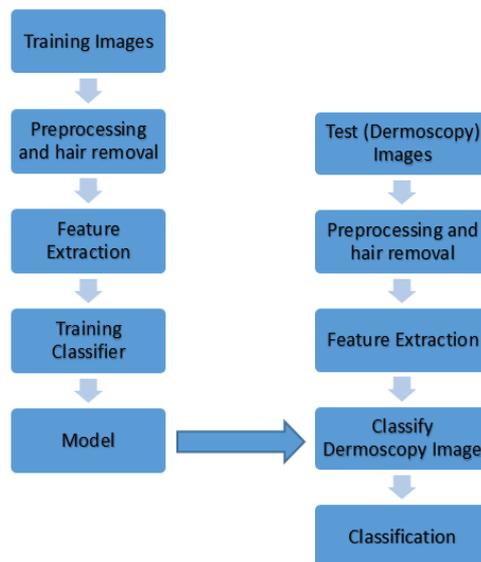


Fig. 3. Flowchart showing the steps involved in our approach.

A. Image Preprocessing

First, we preprocessed the dermoscopy images. Fig. 4 shows some example images from the dataset. We converted the images into 24-bit RGB bitmap files and rescaled them to a size of 227×227 to make them compatible with AlexNet.

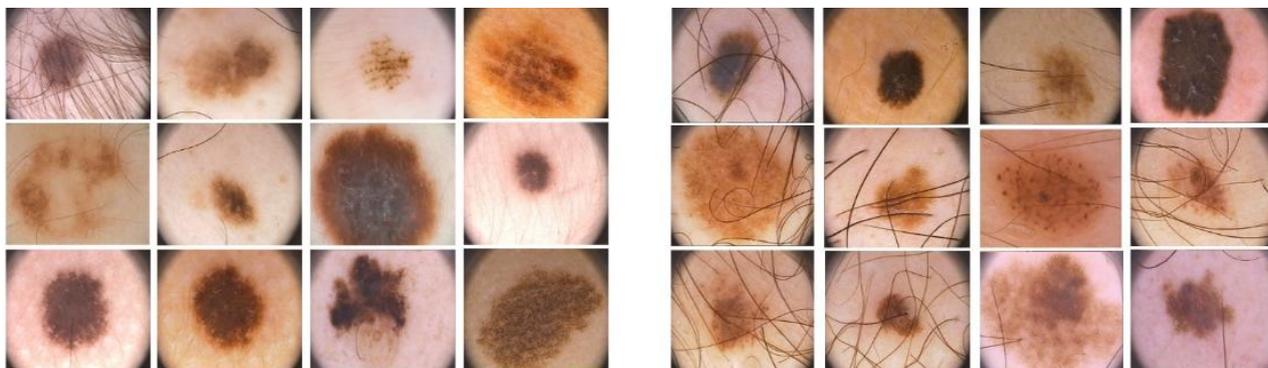


Fig. 4. Sample dermoscopy images from the PH² dataset [1].

We then performed hair removal on all images in the dataset, as described in our previous work [22]. First, we

individually applied a median filter to each RGB channel. Then, a morphological bottom-hat filter is applied to each

color space of the filtered image using a disk-shaped structuring element with a size of 5. We then performed binary image conversion on each channel with a threshold of 5%. Next, we multiplied the binarized hair arrays for each RGB channel together element-wise before again performing binary image conversion with a threshold of 20%. Subsequently, we performed morphological erosion and dilation on the binary images using square structuring

elements with sizes of 1 and 3, respectively. We then identified small objects in the binary images to obtain hair masks. We used these to remove less than 300 pixels from the original images and filled in the removed hair pixels based on adjacent non-hair pixels using a harmonic inpainting technique from [23], as discussed in [24]. Fig. 5 shows some examples of the resulting preprocessed dermoscopy images.



Fig. 5. Sample preprocessed dermoscopy images from the PH² dataset.

B. Transfer Learning and Classification

Transfer learning involves taking a pretrained model and using it as a starting point for a new task. In this case, we used AlexNet as the pretrained network and refined it to create a network specifically for dermoscopy images.

This fine tuning involved not only retraining the AlexNet-based classifier to deal with new dermoscopy images but also adjusting the pretrained network’s weights via backpropagation. AlexNet can be fine-tuned by keeping some of the earlier layers fixed and only adjusting certain higher level parts of the network (in this case, the last fully-connected layer, i.e., Layer 23). This approach was inspired by the idea that the initial features extracted by AlexNet are fairly generic and will still be useful for edge and object detection, while later layers become progressively more specific to the current classification task and must therefore be retrained to classify the dermoscopy image dataset.

Fig. 6 shows a visualization of the first convolutional layer’s weights [18]. This is an $11 \times 11 \times 3$ layer comprising 96 kernels, of which the first 48 are color-agnostic and the other 48 are color-specific. This image was produced by obtaining the second convolutional layer’s weights. Then, the image was scaled and resized.

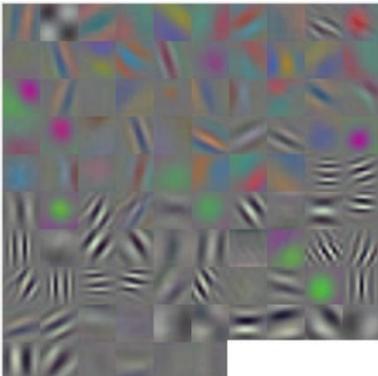


Fig. 6. The 96 kernels of the first convolutional layer’s weights [18].

We retrained the last fully connected layer to extract features relevant to the required number of dermoscopy image categories. We also retrained the last layer (Layer 25) to classify the images using these categories.

C. Network Training

To train the network, we used a stochastic gradient descent (SGD) with momentum (SGDM) solver. Here, the momentum is the gradient step taken during the previous training iteration. The standard SGD algorithm updates the parameter vector $\theta_{\ell+1}$, gradually reducing the error by way of the loss function gradient $\nabla E(\theta)$ [25] as follows:

$$\theta_{\ell+1} = \theta_{\ell} - \alpha \nabla E(\theta_{\ell}), \quad (1)$$

where ℓ is the number of iterations, $\alpha > 0$ is the learning rate, and $E(\theta)$ is the loss function. The SGD algorithm calculates the gradient and thus updates the parameters for each iteration using a mini batch size of 64 in order to reduce the loss function after every iteration.

One issue with the standard SGD algorithm is that it may oscillate across the sharp sides of a ravine on either side of the optimal path. To prevent this type of oscillation, we added a momentum term to the parameter update equation [26] as follows:

$$\theta_{\ell+1} = \theta_{\ell} - \alpha \nabla E(\theta_{\ell}) + \gamma (\theta_{\ell} - \theta_{\ell-1}), \quad (2)$$

where γ is the contribution of the previous gradient to the current iteration and can range between 0 (no contribution) and 1 (full contribution).

D. Classification Performance Evaluation

The classification performance was evaluated in terms of sensitivity, specificity, and accuracy [27]. These are defined in terms of the numbers of true positives (TPs), true negatives (TNs), false negatives (FNs), and false positives (FPs). A TP is a case wherein the diagnostic test indicates the presence of

a disease in a patient known to have a specific disease. Likewise, a TN is a case wherein a test indicates that the disease-free patient has a disease. Meanwhile, an FP is a case wherein the test falsely indicates that a disease-free patient has a disease, and an FN is a case wherein the test falsely indicates that the disease is absent.

Sensitivity shows how good the test is at detecting the disease. It is the proportion of patients with the disease that are correctly identified by the test as follows:

$$\text{Sensitivity} = \frac{TP}{TP+FN}. \quad (3)$$

Likewise, specificity shows how good the test is at identifying normal (disease-free) patients. It is the proportion of such patents that are correctly identified by the test as follows:

$$\text{Specificity} = \frac{TN}{TN+FP}. \quad (4)$$

Finally, accuracy is the proportion of test results that are correct (either positive or negative) as follows:

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)}. \quad (5)$$

E. Experimental Setup

First, the full PH² image set (containing a total of 200 images) was divided into the training, validation, and testing sets, with 70% being used for training, 20% for testing, and 10% for validation.

We then applied the SGDM algorithm to the training dataset using a mini-batch size of 64, a maximum of 1000 epochs, and a γ parameter of In addition, we set the initial learning rate to 0.001 as the only consequence of the learning rate being too small is that training may take an unnecessarily long time, but if it is too high, then training may yield a suboptimal outcome.

We repeated this process twice, once with 2 classes (melanoma and non-melanoma) and then with 3 classes (atypical nevus, common nevus, and melanoma). We also exposed the models to both the original and preprocessed (i.e., with hair removed) images. After training, we retained both *DermoscopyNet* models for testing and validation.

All elements of our method were implemented using MATLAB 9.2 on a computer equipped with 2.8 GHz Intel Core i7-7700HQ CPU with 20 GB RAM and an NVIDIA GeForce GTX 1050 GPU. The computer was installed with the Windows 10 Home (64-bit) operating system.

IV. RESULTS AND EVALUATION

The proposed hair removal method took an average of 53.59 ms per image, including harmonic inpainting. Training the model using 160 images took an average of 1723.185 and 1588.05 s for 2 and 3 categories, respectively, while training with 320 images took 3937.15 s for 3 categories.

A. Classification as Melanoma or Non-Melanoma

The first model used 2 categories: melanoma and

non-melanoma. Tables I and II summarize the classification performance for this model with hair removal for the testing and validation datasets, respectively, indicating that it was 93% accurate for both categories during testing and 95% accurate during validation, giving overall accuracies of 92.5% and 95% during testing and validation, respectively. Tables III and IV present the corresponding confusion matrices.

TABLE I: CLASSIFICATION RESULTS FOR MELANOMA AND NON-MELANOMA WITH HAIR REMOVAL (TESTING)

Skin Disease	Sensitivity	Specificity	Accuracy	Overall Accuracy
Melanoma	86%	94%	93%	92.5%
Non-melanoma	94%	86%	93%	

TABLE II: CLASSIFICATION RESULTS FOR MELANOMA AND NON-MELANOMA WITH HAIR REMOVAL (VALIDATION)

Skin Disease	Sensitivity	Specificity	Accuracy	Overall Accuracy
Melanoma	80%	100%	95%	95%
Non-melanoma	100%	80%	95%	

TABLE III: CONFUSION MATRIX FOR MELANOMA (M) AND NON-MELANOMA (NM) WITH HAIR REMOVAL (TESTING)

		Actual Class	
		M	NM
Predicted Class	M	6	2
	NM	1	31

TABLE IV: CONFUSION MATRIX FOR MELANOMA (M) AND NON-MELANOMA (NM) WITH HAIR REMOVAL (VALIDATION)

		Actual Class	
		M	NM
Predicted Class	M	4	0
	NM	1	15

TABLE V: CLASSIFICATION RESULTS FOR MELANOMA AND NON-MELANOMA WITHOUT HAIR REMOVAL (TESTING)

Skin Disease	Sensitivity	Specificity	Accuracy	Overall Accuracy
Melanoma	55%	93%	83%	82.5%
Non-melanoma	93%	55%	83%	

TABLE VI: CLASSIFICATION RESULTS FOR MELANOMA AND NON-MELANOMA WITHOUT HAIR REMOVAL (VALIDATION)

Skin Disease	Sensitivity	Specificity	Accuracy	Overall Accuracy
Melanoma	100%	89%	90%	90%
Non-melanoma	89%	100%	90%	

Likewise, Tables V and VI summarize the model's classification performance without hair removal during testing and validation, respectively, indicating that it was 90% accurate for both categories during validation but only 82.5%

accurate during testing. Tables VII and VIII present the corresponding confusion matrices.

TABLE VII: CONFUSION MATRIX FOR MELANOMA (M) AND NON-MELANOMA (NM) WITHOUT HAIR REMOVAL (TESTING)

		Actual Class	
		M	NM
Predicted Class	M	6	2
	NM	5	27

TABLE VIII: CONFUSION MATRIX FOR MELANOMA (M) AND NON-MELANOMA (NM) WITHOUT HAIR REMOVAL (VALIDATION)

		Actual Class	
		M	NM
Predicted Class	M	2	2
	NM	0	16

TABLE IX: CLASSIFICATION RESULTS FOR MELANOMA AND NON-MELANOMA WITH AND WITHOUT HAIR REMOVAL (TESTING)

Skin Disease	Sensitivity	Specificity	Accuracy	Overall Accuracy
Melanoma	100%	96%	96%	96.25%
Non-melanoma	96%	100%	96%	

TABLE X: CLASSIFICATION RESULTS FOR MELANOMA AND NON-MELANOMA WITH AND WITHOUT HAIR REMOVAL (VALIDATION)

Skin Disease	Sensitivity	Specificity	Accuracy	Overall Accuracy
Melanoma	100%	91%	93%	92.5%
Non-melanoma	91%	100%	93%	

TABLE XI: CONFUSION MATRIX FOR MELANOMA (M) AND NON-MELANOMA (NM) WITH AND WITHOUT HAIR REMOVAL (TESTING)

		Actual Class	
		M	NM
Predicted Class	M	13	3
	NM	0	64

TABLE XII: CONFUSION MATRIX FOR MELANOMA (M) AND NON-MELANOMA (NM) WITH AND WITHOUT HAIR REMOVAL (VALIDATION)

		Actual Class	
		M	NM
Predicted Class	M	5	3
	NM	0	32

Tables IX and X summarize the combined classification performance of the model both with and without hair removal during testing and validation, respectively, indicating that it

was 96% accurate for both categories during testing and 93% accurate during validation. The accuracy is higher here due to the increased number of dermoscopy images used for training. Tables XI and XII present the corresponding confusion matrices.

B. Classification into 3 Categories

Next, we tested the 3-category model, which considered atypical nevus, common nevus, and melanoma. Tables XIII and XIV summarize the classification performance without hair removal during testing and validation, respectively, indicating that it was able to diagnose melanoma with over 93% accuracy; however, it performed poorly for the other 2 skin disease categories, leading to an overall accuracy of 70% for both datasets. Tables XV and XVI present the corresponding confusion matrices.

TABLE XIII: CLASSIFICATION RESULTS FOR ATYPICAL NEVUS, COMMON NEVUS, AND MELANOMA WITHOUT HAIR REMOVAL (TESTING)

Skin Disease	Sensitivity	Specificity	Accuracy	Overall Accuracy
Atypical nevus	65%	78%	73%	70%
Common nevus	75%	75%	75%	
Melanoma	73%	100%	93%	

TABLE XIV: CLASSIFICATION RESULTS FOR ATYPICAL NEVUS, COMMON NEVUS, AND MELANOMA WITHOUT HAIR REMOVAL (VALIDATION)

Skin Disease	Sensitivity	Specificity	Accuracy	Overall Accuracy
Atypical nevus	63%	75%	70%	70%
Common nevus	67%	82%	75%	
Melanoma	100%	94%	95%	

TABLE XV: CONFUSION MATRIX FOR ATYPICAL NEVUS (A), COMMON NEVUS (C), AND MELANOMA (M) WITHOUT HAIR REMOVAL (TESTING)

		Actual Class		
		A	C	M
Predicted Class	A	10	5	1
	C	4	11	1
	M	2	0	6

TABLE XVI: CONFUSION MATRIX FOR ATYPICAL NEVUS (A), COMMON NEVUS (C), AND MELANOMA (M) WITHOUT HAIR REMOVAL (VALIDATION)

		Actual Class		
		A	C	M
Predicted Class	A	5	2	1
	C	3	5	0
	M	0	0	4

Tables XVII and XVIII summarize the model's classification performance with hair removal during testing and validation, respectively, indicating that it was able to diagnose melanoma with 95% accuracy during validation,

and 90% during testing. However, again, the low overall accuracies (70% and 67.5% during validation and testing, respectively) mean that the model was ineffective for predicting the 3 categories using only 200 dermoscopy images with hair removal. Tables XIX and XX present the corresponding confusion matrices.

TABLE XVII: CLASSIFICATION RESULTS FOR ATYPICAL NEVUS, COMMON NEVUS, AND MELANOMA WITH HAIR REMOVAL (TESTING)

Skin Disease	Sensitivity	Specificity	Accuracy	Overall Accuracy
Atypical nevus	63%	75%	70%	
Common nevus	69%	79%	75%	67.5%
Melanoma	75%	94%	90%	

TABLE XVIII: CLASSIFICATION RESULTS FOR ATYPICAL NEVUS, COMMON NEVUS, AND MELANOMA WITH HAIR REMOVAL (VALIDATION)

Skin Disease	Sensitivity	Specificity	Accuracy	Overall Accuracy
Atypical nevus	63%	75%	70%	
Common nevus	71%	77%	75%	70%
Melanoma	80%	100%	95%	

TABLE XIX: CONFUSION MATRIX OF ATYPICAL NEVUS (A), COMMON NEVUS (C), AND MELANOMA (M) WITH HAIR REMOVAL (TESTING)

		Actual Class		
		A	C	M
Predicted Class	A	10	5	1
	C	4	11	1
	M	2	0	6

Tables XXI and XXII, show the combined classification performance both with and without hair removal during testing and validation, respectively, indicating that the proposed system was 96% accurate for melanoma, 86% accurate for common nevus, and 88% accurate for atypical nevus during testing. In addition, it was 95% accurate for melanoma, 90% accurate for common nevus, and 90% accurate for atypical nevus during validation. The overall accuracy was excellent for all 3 categories, with average accuracies of 90% and 92% during testing and validation, respectively. Tables XXIII and XXIV present the corresponding confusion matrices.

Figs. 7 and 8 show some examples of images that were incorrectly and correctly classified, respectively, by the 2-category model.

TABLE XX: CONFUSION MATRIX FOR ATYPICAL NEVUS (A), COMMON NEVUS (C), AND MELANOMA (M) WITH HAIR REMOVAL (VALIDATION)

		Actual Class		
		A	C	M
Predicted Class	A	5	2	1
	C	3	5	0
	M	0	0	4

TABLE XXI: CLASSIFICATION RESULTS FOR ATYPICAL NEVUS, COMMON NEVUS, AND MELANOMA WITH AND WITHOUT HAIR REMOVAL (TESTING)

Skin Disease	Sensitivity	Specificity	Accuracy	Overall Accuracy
Atypical nevus	78%	93%	86%	
Common nevus	87%	88%	88%	85%
Melanoma	100%	96%	96%	

TABLE XXII: CLASSIFICATION RESULTS FOR ATYPICAL NEVUS, COMMON NEVUS, AND MELANOMA WITH AND WITHOUT HAIR REMOVAL (VALIDATION)

Skin Disease	Sensitivity	Specificity	Accuracy	Overall Accuracy
Atypical nevus	93%	88%	90%	
Common nevus	83%	95%	90%	87.5%
Melanoma	88%	97%	95%	

TABLE XXIII: CONFUSION MATRIX FOR ATYPICAL NEVUS (A), COMMON NEVUS (C), AND MELANOMA (M) WITH AND WITHOUT HAIR REMOVAL (TESTING)

		Actual Class		
		A	C	M
Predicted Class	A	10	5	1
	C	4	11	1
	M	2	0	6

TABLE XXIV: CONFUSION MATRIX FOR ATYPICAL NEVUS (A), COMMON NEVUS (C), AND MELANOMA (M) WITH AND WITHOUT HAIR REMOVAL (VALIDATION)

		Actual Class		
		A	C	M
Predicted Class	A	5	2	1
	C	3	5	0
	M	0	0	4



Fig. 7. Example of incorrectly classified dermoscopy images.



Fig. 8. Example of correctly classified dermoscopy images.

V. CONCLUSIONS

In this paper, we discussed a simple algorithm for hair removal and a deep learning-based approach for skin lesion classification using dermoscopy images.

The hair removal process used morphological operators and inpainting, and this simple approach was shown to be as effective as other methods at detecting, removing, and correcting hairs in dermoscopy images [22]. This is an important preprocessing step in the identification and classification of melanomas because hair adds extraneous features to the images. Herein, we showed that such preprocessing increases the classification accuracy of our model and thus helps in detecting melanomas.

In future work, we plan to explore different types of skin lesion images to better assess our lesion classification model. This will be done by considering other datasets or using images from the Internet. It would also be interesting to investigate other training algorithms for classification, and it may be useful to perform skin detection, especially when handling the varying skin colors of people of different ethnicities.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous referees for their valuable comments and helpful suggestions. Ms. Salido acknowledges the Commission on Higher Education (CHED) in collaboration with the De La Salle University (DLSU) and Aklan State University for funding support through the CHED K-12 Transition (CHED K-12) Program. Dr. Ruiz received funding from the University Research Coordination Office (URCO) of DLSU.

REFERENCES

- [1] T. Mendonça, P. M. Ferreira, J. S. Marques, A. R. Marcal, and J. Rozeira, "PH 2-A dermoscopic image database for research and benchmarking," in *Proc. 35th Annual International Conference of the IEEE on Engineering in Medicine and Biology Society (EMBC)*, July 2013, pp. 5437-5440.
- [2] R. P. E. Losina, A. Walensky, F. C. Geller, L. L. Beddingfield, B. Wolf, A. Gilchrist, and K. A. Freedberg, "Visual screening for malignant melanoma: A cost-effectiveness analysis," *Archives of Dermatology*, vol. 143, no. 1, pp. 21-28, 2007.
- [3] American Cancer Society. Cancer Facts & Figures 2017, Atlanta: American Cancer Society, 2017.
- [4] N. Howlader, A. M. Noone, M. Krapcho *et al.*, "SEER cancer statistics review," National Cancer Institute, April 2015.
- [5] Skin Cancer Foundation. (June 2017). [Online]. Available: <http://www.skincancer.org/skin-cancer-information/melanoma>.
- [6] R. H. Johr, "Dermoscopy: Alternative melanocytic algorithms—the ABCD rule of dermatoscopy, menzies scoring method, and 7-point checklist," *Clinics in Dermatology*, vol. 20, no. 3, pp. 240-247, 2002.
- [7] G. Argenziano, I. Zalaudek, R. Corona, F. Sera, L. Cicale, G. Petrillo, and H. P. Soyer, "Vascular structures in skin tumors: A dermoscopy study," *Archives of Dermatology*, vol. 140, no. 12, pp. 1485-1489, 2004.
- [8] H. Pehamberger, A. Steiner, and K. Wolff, "In vivo epiluminescence microscopy of pigmented skin lesions. I. Pattern analysis of pigmented skin lesions," *Journal of the American Academy of Dermatology*, vol. 17, no. 4, pp. 571-583, 1987.
- [9] A. A. Abder-Rahman and T. M. Deserno, "A systematic review of automated melanoma detection in dermoscopic images and its ground truth data," in *Proc. SPIE, Medical Imaging 2012: Image Perception, Observer Performance, and Technology Assessment*, 2012, vol. 8318.
- [10] H. Kittler, H. Pehamberger, K. Wolff, and M. Binder, "Diagnostic Accuracy of dermoscopy," *The Lancet Oncology*, vol. 3, no. 3, pp. 159-165, 2002.
- [11] M. Binder, M. Schwarz, A. Winkler, A. Steiner, A. Kaider, K. Wolff, and H. Pehamberger, "Epiluminescence microscopy, a useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists," *Archives of Dermatology*, vol. 131, no. 3, pp. 286-291, 1995.
- [12] N. K. Mishra and M. E. Celebi. "An overview of melanoma detection in dermoscopy images using image processing and machine learning," eprint arXiv: 1601.07843 2016.
- [13] F. Nachbar, W. Stolz, T. Merkle, A. B. Cognetta, T. Vogt, M. Landthaler, and G. Plewig, "The ABCD rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions," *Journal of the American Academy of Dermatology*, vol. 30, no. 4, pp. 551-559, 1994.
- [14] N. R. Abbasi, H. M. Shaw, D. S. Rigel, R. J. Friedman, W. H. McCarthy, I. Osman, and D. Polsky, "Early diagnosis of cutaneous melanoma: revisiting the ABCD criteria," *Jama*, vol. 292, no. 22, pp. 2771-2776, 2004.
- [15] A. M. Menzies, L. E. Haydu, L. Visintin, M. S. Carlino, J. R. Howle, J. F. Thompson, and G. V. Long, "Distinguishing clinicopathologic features of patients with V600E and V600K BRAF-mutant metastatic melanoma," *Clinical Cancer Research*, vol. 18, no. 12, pp. 3242-3249, 2012.
- [16] G. D. Leo, G. Fabbrocini, C. Liguori, A. Pietrosanto, and M. Sclavenzi, "ELM image processing for melanocytic skin lesion diagnosis based on 7-point checklist: A preliminary discussion," in *Proc. Int. Symp. Meas. for Res. and Ind. Appl.*, Budapest, Hungary: IMEKO, 2004, pp. 474-479.
- [17] Deep Learning. (June 25, 2017). [Online]. Available: <https://www.mathworks.com/discovery/deep-learning.html>.
- [18] Inc. MathWorks, *MATLAB: The Language of Technical Computing, Desktop Tools and Development Environment*, version 7, vol. 9, MathWorks, 2005.
- [19] J. Deng, W. Dong, R. Socher, L.-J. Li *et al.*, "ImageNet: A large-scale hierarchical image database," *IEEE Computer Vision and Pattern Recognition (CVPR)*, 2009.
- [20] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929-1958, 2014.
- [21] O. Russakovsky, J. Deng, H. Su *et al.*, "ImageNet large scale visual recognition challenge," *International Journal of Computer Vision (IJCV)*, vol. 115, no. 3, pp. 211-252, 2015.
- [22] J. A. A. Salido, and C. Ruiz Jr., "Using morphological operators and inpainting for hair removal in dermoscopic images," in *Proc. The Computer Graphics International Conference*, ACM, June 2017, p. 2.
- [23] T. F. Chan and J. Shen, "Nontexture inpainting by curvature-driven diffusions," *Journal of Visual Communication and Image Representation*, vol. 12, no. 4, pp. 436-449, 2001.
- [24] S. Parisotto and C. Schönlieb, "MATLAB Codes for the Image Inpainting Problem, GitHub repository," *MATLAB Central File Exchange*, September 2016.
- [25] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, NY, 2006.
- [26] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*, The MIT Press, Cambridge, Massachusetts, 2012.
- [27] W. Zhu, N. Zeng, and N. Wang, "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations," *NESUG Proceedings: Health Care and Life Sciences*, Baltimore, Maryland, 2002.



Julie Ann A. Salido has been a PhD student in Computer Science at De La Salle University (DLSU) since 2016. She obtained her MSc in computer science from the Department of Computer Studies, Algorithm and Complexity Laboratory at the University of the Philippines, Diliman in 2015 focusing on bioinformatics, information technology, and applied computer science.

She is currently a recipient of a CHED K-12 Transition Graduate School Scholarship for PhD in computer science. She was also a recipient of the Science and Engineering Government Scholarship Program of Commission and Higher Education (June 2010–May 2012) at the University of the Philippines, Diliman. She has been an assistant professor at Aklan State University since 2016. She has also served as chair of Monitoring and Evaluation, ICT Coordinator, and Instructor at Aklan State University. She has conducted research on computer graphics, life sciences, and bioinformatics, which has been published.



Conrado R. Ruiz Jr. obtained his bachelor's degree in computer science with a specialization in software technology (cum laude) from DLSU in 1999. He received his master's and doctorate degrees from the School of Computing of the National University of Singapore in 2001 and 2015, respectively. His several works in the areas of computer graphics, computational art, and multimedia information retrieval have been published.

His articles have also been published in IEEE Transactions on

Visualization and Computer Graphics, Visual Computer and Multimedia Tools, and Applications Journal. He has also presented at various conferences, such as Computer Graphics International, ACM SIGGRAPH Asia, EuroGraphics, and ACM I3D. He was a BPI Science Awardee in 1999 and received a special citation from the National Academy of Science and Technology during the Talent Search for Young Scientists in 2007. He is currently an associate professor at the College of Computer Studies of DLSU.