Melanoma Detecting using Convolutional Neural Network and Transfer Learning

Nhu Quynh Duong
Department of Computer Science
Golisano College of Computing and Information Sciences
Rochester Institute of Technology
Rochester, NY 14586
qnd5128@rit.edu

***Abstract*—TODO – The abstract goes here. It should be about 250 words, not much more. The purpose of the Abstract is to get people interested in reading your paper. Do not reveal the final conclusions, but hint that you did good things here. Explain why someone would want to read you paper.**

The purpose of an abstract is to get someone TO WANT TO READ YOUR PAPER. It is Marketing. It lets the user decide if they care to read any more. It hints at the conclusion, but does not give away the punchline.

***Index Terms*—Keyword1; Keyword2; Keyword3**

Abstract

This is a different way to do an abstract.

I. Introduction

Melanoma is one of the deadliest forms of skin cancer. If detected early, it can have a survival rate to 90% [1]. However, detecting this skin cancer is challenging due to its high visual similarity to other benign skin lesions [2]. Although melanoma is not as common as other skin cancer types, it accounts for the majority of skin cancer deaths [3]. The most challenging part about diagnosing skin cancer is differentiating between melanoma and nonmelanoma like common and atypical nevi [1], [3]. This is due to minimal distinctions between malignant and benign lesions, leading to inaccurate diagnostic in visual inspections by dermatologists [1]. Due to the high importance of early detection, the development of reliable automated systems for melanoma detection has been a critical area of research.

In recent years, Convolutional Neural Networks (CNNs) have evolved as a powerful tool for visual pattern recognition and classification tasks, especially medical image analysis [4], [5]. CNNs are effective for skin lesion classification because they can automatically learn hierarchical features from raw image data and have been shown to outperform traditional feature engineering approaches [6]. These networks have been successfully applied to melanoma detection, where they extract important features like color, texture, and lesion shape which are crucial for accurate diagnosis [5]. However, training CNNs from scratch is often impractical for medical imaging tasks due to limited availability of labeled datasets [4].

Transfer Learning is a technique where pre-trained models are fine-tuned for a new task [7]. Therefore, it addresses the challenge of having to train a CNN model completely from scratch which is resource intensive. This method is particularly useful for skin lesion classification, where the annotated datasets are limited [2]. Pre-trained models like AlexNet and VGG16, which have been successfully used for image classification tasks, can be fine-tuned for specific tasks such as melanoma detection [8]. Studies have shown that fine-tuned pre-trained models achieve high accuracy even on relatively small medical datasets [2], [5], [9].

The International Skin Imaging Collaboration (ISIC) dataset has been extensively used, developed, and heavily researched for melanoma detection. It contains over 400,000 public dermoscopic images, making it one of the largest publicly available datasets for skin lesion analysis [10]. This dataset is larger than other publicly available datasets like Ph2 [11] and Dermofit Image Library. Therefore, this paper will use the data from the ISIC 2019 challenge which includes datasets from previous years (2017 and 2018) [12]–[14].

In this paper, I will apply AlexNet and VGG16 for skin lesion classification using transfer learning. The methodology will involve pre-processing steps such as removing duplicates, cropping and normalization, followed by data augmentation techniques like flipping, resizing and rotating to diversify and increase the size of the training data. Then, I will train the data using both AlexNet and VGG16 to compare performance and accuracy between these models. Finally, I will evaluate the performance of these models using metrics like accuracy, sensitivity, specificity, F1-score, AUC-ROC, and confusion matrix.

The rest of this paper is as follows: Section 2 will provide an overview of the related work in melanoma detection using different deep learning techniques. Section 3 presents the methodology for this project including details about the dataset, pre-processing and data augmentation techniques, and model parameters and architectures. Section 4 will go over the evaluation metrics measured, and Section 5 discusses the performance results and compares them with existing methods from previous work. Section 6 will talk about my thoughts throughout this project and discuss potentials for future work. Finally, Section 7 provides a summary of the paper with insights into the applicability of CNN-based systems for melanoma detection and recommendations for future work.

II. Background

Deep learning, particularly Convolutional Neural Networks (CNNs), has greatly advanced melanoma detection by improving skin lesion classification. Transfer learning further enhances this by utilizing pre-trained models, making it possible to achieve high accuracy even with limited medical data. This section explores key models like AlexNet, ResNet, EfficientNet, and VGGNet, and discusses hybrid techniques that combine multiple methods for better performing melanoma classification.

A. Melanoma Detection using Deep Learning

Deep learning models, particularly Convolutional Neural Networks (CNNs), have been widely used for skin lesion classification due to their accuracy in image recognition. They can detect features and patterns from large amounts of images to outperform other neural networks in image classification. They can extract features from images without the need of immense effort on manual feature engineering [15]. Deep learning is much more efficient in data processing due to its nature of passing images through layers of the network and each layer processes important parts of the image to eventually identify the desired output [15]. There are various pre-trained neural network models that work well in image recognition tasks such as ResNet, AlexNet, EfficentNet [16], ImageNet, and VGGNet [8].

B. Transfer Learning and Pre-trained Models

Pre-trained models have become more prevalent in the medical imaging field. Pre-trained network trains on extremely large data and process large amounts of information [8]. This is also called transfer learning where models are pre-trained for a specific task based on selected features [7]. Transfer learning is useful for skin lesion classification because of its advantage on medical images. Most of these datasets are limited and briefly annotated, thus making it very difficult to train new machine learning models. Training ML models takes time, resources, and high expertise which would be quite expensive. With pre-trained models, they already have the features selected, tasks, functions and weights so fine-tuning it to new tasks would be more efficient [7]. Pre-trained CNNs, initially trained on large datasets for image classification like ImageNet, have been fine-tuned for medical tasks such as skin lesion classification [2].

1) AlexNet: One of the most widely used pre-trained models is AlexNet, introduced by Krizhevsky et al. [17], which has been applied in various domains including skin lesion classification. For example, Hosny et al. [2] utilized AlexNet for classifying different skin cancer like melanoma, common nevus, and atypical nevus. The model was fine-tuned by replacing the last layer with a softmax classifier. Through transfer learning, the model achieved an accuracy of 98.61% on the ph2 dataset [2]. The ph2 dataset is a small set of 200 dermoscopic images that is well annotated [11]. This demonstrates the model's effectiveness of pre-trained networks on small medical datasets.

2) ResNet: ResNet and its variants have also been widely used in skin lesion classification. The depth of the network and the residual connections allow it to overcome the vanishing gradient problem of image classification, which makes it suitable for complex medical imaging tasks [18]. In the study by Budhiman et al. [9], multiple variants of ResNet was used to compare and classify melanoma based on the ISIC 2018 dataset. For skin cancer classification, deep networks like ResNet can learn and abstract features from dermoscopic images that are not easily recognizable by human vision [18]. The deeper layers of ResNet are capable of extracting high detailed features such as texture, color, and patterns in skin lesions, which are crucial for accurate diagnosis [9]. In the study, Budhiman et al. [9] applied ResNet architectures with varying depths (ResNet 50, 40, 25, 10, and 7) and found that it's most accurate using ResNet-50. It was able to achieve high classification accuracy of 82.9% [9].

3) EfficientNet: Another significant model in transfer learning is EfficientNet, which uses a balanced scaling method that optimizes network depth, width, and resolution. The model address the computational limitation of deep learning models, particularly in tasks like image recognition [16]. In the study by Venugopal et al. [19], EfficientNet were used on the ISIC 2019, ISIC 2020, and HAM10000 datasets to achieve accuracy of 99.23% on the ISIC 2020 dataset. The key to its success and high accuracy is due to its scaling method to ensure the model can learn more detailed and complex patterns in dermoscopic images without being computationally expensive or lead to over-fitting [19].

4) VGGNet:

C. Hybrid Techniques

While some papers focus on one particular machine learning technique to perform Melanoma classification, other papers combine various techniques to build a more robust classification model. These methods leveraged the strengths of each approach to improve the robustness, accuracy, and generalization of classification by addressing limitations of each model. By fusing techniques such as CNNs, machine learning algorithms, and ensemble methods, hybrid techniques have shown significant improvements in skin lesion classification and segmentation tasks [20].

In Listing 1 a HelloWorld program is shown.

Listing 1. writer.c.

```
int main() {
  char * msg = "Hello World\n";
  write(STDIN_FILENO, msg, strlen(msg));
  exit(0);
}
```

III. Methodology

This section will go over the steps taken to prepare the data, choose the hyperparameters and model architectures, and the training process. I will have system diagrams of the 2 models I will be using AlexNet and VGG16. I will also insert

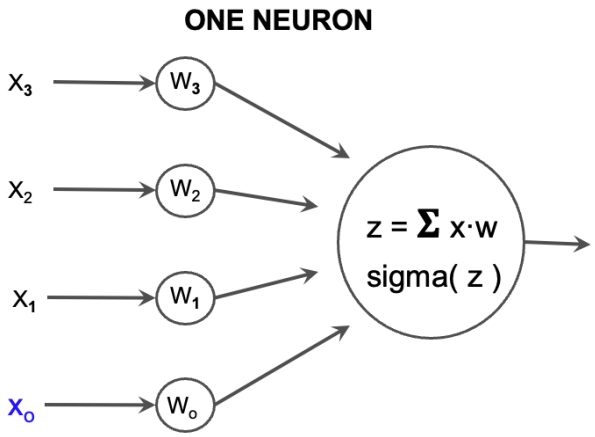

Fig. 1. Captions should say what we want the user to see. Try to get a figure in the upper right hand side of the first page. That way the reader can see results of your work. LABEL the caption to refer to the figure later on.

dermoscopy image examples from the dataset. I plan to do a methodology diagram consisting of all the steps that I took to train the models.

A. Dataset

- **Description of the ISIC 2019 Dataset**: I will give a brief introduction of the dataset and why I chose this one compared to other publicly available datasets.
- **Data attributes**: I will describe the structure of the data (number of images, types of skin lesions, spread of data, annotations, and data source)

B. Data Handling

- **EDA**: I will describe some of the data discovery that I found. Maybe difficulties working with data, data cleaning work needed, duplicates, etc.
- **Data prep-rocessing techniques**: cropping, normalization, and removal of duplicate diagnosis. I will also describe how I resized the dataset to fit with the selected models' dimension requirements.
- **Data Augmentation techniques**: I will go over techniques like flipping, horizontal and vertical flipping to increase size of dataset.

C. Model Selection

- **Transfer Learning and CNN**: Talk about the use of pre-trained models and fine-tuning process specific to melanoma classification (modifying the final softmax layer, adjusting learning rates, etc.)
- **AlexNet**: Discuss choosing AlexNet and its architecture (insert diagram for AlexNet). Talk about the modification to AlexNet.
- **VGG16**: Discuss choosing VGG16 and its architecture (insert diagram for VGG16). Talk about the modification to VGG16. Discuss the differences between the 2 and hypothesize which one might perform better and why.

TABLE I
AN EXAMPLE OF A TABLE. WITH THE CAPTION ABOVE THE TABLE.

Heading One	Heading Two
Data One	Data Two
Three	Four

D. Training Process

- **Hyperparameter**: I will discuss the hyperparameter choices like batch size, learning rate, and epoch size.
- **Mini-batch Gradient Descent**: I will talk about optimization techniques to improve model performance.
- **Loss Function**: I will talk about the loss function.
- **Regularization**: If I choose to use regularization techniques then I will mention it here.

Diagrams of experiment. System architecture of each selected models and an overall training process diagrams. *PLEASE NOTE* All figures should have captions under them. AND the captions should tell the reader what to see in the figure.

Tables. AND, tables have a caption above them, not below them.

Using the package for urls, you can include URLs like this... https://cs.rit.edu or whatever.

IV. Evaluation

A. Metrics

I will have a table of accuracy, sensitivity, specificity, and F1-score for both models and compare side by side. I may use some performance metrics from other studies that use ML techniques and compare it with CNN models.

- Accuracy: I'll talk about the accuracy output of both models.
- Sensitivity: talk about model's ability to identify melanoma
- Specificity: talk about the model's ability to identify non-melanoma
- F1 Score: Measure the precision and recall balance of both models
- AUC-ROC: Show the graph for this metric

B. Model Performance

In this section I will be discussing both the AlexNet and VGG16 performance through various metrics. If one significantly performs better than the other, then I will discuss the reasoning based on their architectures.

V. Discussion

In this section, I will talk about my challenges throughout the process. Any difficulties or unique findings I may have. I will go through interesting findings from my research. I want to talk about the choice of datasets and models based on all the background research I did.

I will also mention the potential use of this in medical settings. I want to emphasize how this may benefit dermatologists and reduce the diagnostic errors.

VI. Future Work

A. Model Improvements

In this section I will be discussing other CNN models that I did not used because they are much bigger and more complex to train (EfficientNet, ResNet, MobileNet). I will talk about the potentials of other CNNs and how they could further improve the performance and accuracy. I will go over the potential of ensemble learning by combining multiple models to improve the classification and reduce false positives or negatives.

B. Dataset Alternatives

For this section, I will discuss the other datasets available such as PH2 and Dermofit Library. Also, I can talk about the ISIC 2024 Challenge dataset which has 400,000 images and would be great dataset to train large models. The annotations for these images are much more detailed compared to the 2019 one but I did not choose it due to its large size and resource limitation.

VII. Conclusion

This section is where I will be summarizing the key findings of my paper such as performance of the proposed models and comparison with other existing techniques. I will also include the application of this research in the medical field and how these findings assist dermatologists in early melanoma detection. I will give some dataset and model limitations which could affect the accuracy or performance results. Lastly, I will review the future work for further improvements such as using other models or combining techniques.

Acknowledgment

I wish to really thank my capstone advisor for all his time and help....

Thank HP or Minseok.

Thank your course advisor.

Other professors.

Thank your parents.

References

[1] N. Codella, J. Cai, M. Abedini, R. Garnavi, A. Halpern, and J. R. Smith, "Deep learning, sparse coding, and svm for melanoma recognition in dermoscopy images," in *Machine Learning in Medical Imaging. MLMI 2015*, ser. Lecture Notes in Computer Science, L. Zhou, L. Wang, Q. Wang, and Y. Shi, Eds., vol. 9352. Springer, Cham, 2015, pp. 118–126. [Online]. Available: https://doi.org/10.1007/978-3-319-24888-2_15

[2] K. M. Hosny, M. A. Kassem, and M. M. Foaud, "Skin cancer classification using deep learning and transfer learning," in *2018 9th Cairo International Biomedical Engineering Conference (CIBEC)*, 2018, pp. 90–93.

[3] A. C. Society, "Cancer facts and figures 2018," https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2018/cancer-facts-and-figures-2018.pdf, 2018, accessed: 10 Oct 2024.

[4] L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. Heng, "Automated melanoma recognition in dermoscopy images via very deep residual networks," *IEEE Transactions on Medical Imaging*, vol. 36, no. 4, pp. 994–1004, 2017.

[5] J. Kawahara, A. BenTaieb, and G. Hamarneh, "Deep features to classify skin lesions," in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, 2016, pp. 1397–1400.

[6] B. Harangi, "Skin lesion classification with ensembles of deep convolutional neural networks," *Journal of Biomedical Informatics*, vol. 86, pp. 25–32, 2018.

[7] J. Murel and E. Kavlakoglu, "What is transfer learning?" https://www.ibm.com/topics/transfer-learning, February 12 2024.

[8] A. Kalaivani and S. Karpagavalli, "Deep neural network optimization for skin disease classification forecast analysis," in *2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS)*, vol. 1, 2022, pp. 708–713.

[9] A. Budhiman, S. Suyanto, and A. Arifianto, "Melanoma cancer classification using resnet with data augmentation," in *2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, 2019, pp. 17–20.

[10] N. C. F. Codella, D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti, S. W. Dusza, A. Kalloo, K. Liopyris, N. Mishra, H. Kittler, and A. Halpern, "Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic)," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018, pp. 168–172.

[11] T. Mendonça, P. M. Ferreira, J. S. Marques, A. R. S. Marcal, and J. Rozeira, "Ph2 - a dermoscopic image database for research and benchmarking," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013, pp. 5437–5440.

[12] P. Tschandl, C. Rosendahl, and H. Kittler, "The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," p. 180161, 2018.

[13] N. C. F. Codella, D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti, S. W. Dusza, A. Kalloo, K. Liopyris, N. Mishra, H. Kittler, and A. Halpern, "Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic)," 2017.

[14] C. Hernández-Pérez, M. Combalia, S. Podlipnik, N. C. F. Codella, V. Rotemberg, A. C. Halpern, O. Reiter, C. Carrera, A. Barreiro, B. Helba, S. Puig, V. Vilaplana, and J. Malvehy, "Bcn20000: Dermoscopic lesions in the wild," *Scientific Data*, vol. 11, p. 641, 2024.

[15] J. Holdsworth and M. Scapicchio, "What is deep learning?" https://www.ibm.com/topics/deep-learning, June 17 2024.

[16] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," *CoRR*, vol. abs/1905.11946, 2019. [Online]. Available: http://arxiv.org/abs/1905.11946

[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, p. 84–90, may 2017. [Online]. Available: https://doi.org/10.1145/3065386

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015. [Online]. Available: http://arxiv.org/abs/1512.03385

[19] V. Venugopal, N. Raj, M. Nath, and N. Stephen, "A deep neural network using modified efficientnet for skin cancer detection in dermoscopic images," *Decision Analytics*, vol. 8, p. 100278, 06 2023.

[20] B. Shetty, R. Fernandes, A. Rodrigues, R. Chengoden, S. Bhattacharya, and K. Lakshman, "Skin lesion classification of dermoscopic images using machine learning and convolutional neural network," *Scientific Reports*, vol. 12, 10 2022.