Melanoma Classification Using Deep Learning

Ramiz Akhtar
University of Virginia
409 13th St. NW Apt. 3D
Charlottesville, VA 22903
rsa5wj@virginia.edu

Rayaan Faruqi
University of Virginia
1815 Jefferson Park Avenue Apt. 10
Charlottesville, VA 22903
raf9dz@virginia.edu

Kunaal Sarnaik
University of Virginia
43288 John Danforth Court
Ashburn, VA 20147
kss7yy@virginia.edu

Motivation

According to the Center for Disease Control (CDC), melanoma is the most serious and deadly type of skin cancer.[1] The incidence and mortality of invasive melanoma in the United States has risen steadily, with the lifetime risk of developing the disease skyrocketing from 0.03% in 1930 to 1.82% in 2017.[2] Furthermore, from 2012 to 2016, approximately 77,698 new cases of melanoma occurred in the United States each year; previous studies have also found that melanoma is more likely to metastasize relative to other types of skin cancers.[3] Given that the global incidence and mortality of melanoma are expected to continue rising at an alarming rate, diagnosing and treating the disease early and efficiently is an urgent priority.[4] However, as can be observed in **Figure 1** below, distinguishing melanoma not only from other types of skin diseases, but also from benign skin moles (i.e. nevi), at the dermatologist-level is an arduous process that must account for various factors.

Figure 1: Depicted are three skin lesions, the similarities of which make it difficult to diagnose and treat melanoma early in its progression. **a.)** Skin lesion positive for melanoma. **b.)** Benign nevus negative for melanoma. **c.)** Skin lesion not positive for melanoma, but positive for seborrheic keratosis.

Compounded by the clinical complexities involved when determining the specific stage of a tumor, this difficulty in diagnosing melanoma results in many patients not noticing a potentially fatal tumor appearing on their skin.[5] Moreover, the current process for diagnosis, which involves presenting to a series of doctors, can not only be expensive but also intimidating for patients. Considering the social, economic, and political determinants of healthcare that additionally belabor the process, the patient may become discouraged and delay diagnosis until a later date. Worse yet, they may delay seeking diagnosis until experiencing an adverse outcome as a result of their endorsed malignancy. Consequently, patients who discover they have invasive melanoma at a later stage often experience worse clinical and financial outcomes, including metastasis, poor prognosis, depression, and higher insurance costs.[6]

Thus, providing diagnostic information for a potentially fatal melanoma tumor via a smartphone application through utilization of machine learning principles may make early detection more feasible,

accessible, and economical by way of reduced treatment costs. Furthermore, such an application may also lead to improved clinical outcomes when it comes to the metastasis and prognosis of an invasive melanoma tumor. Ultimately, by making the first step for diagnosis more convenient, effective, equitable, and accessible, clinical outcomes for patients may improve substantially.

Background

Melanoma arises from genetic mutations in melanocytes, which are the upper layer cells that give skin its pigmentation.[7,8] Exposure to ultraviolet (UV) radiation from the sun causes greater production of melanin by the melanocytes, effectively darkening the skin to protect it from the UV radiation.[8] In excess, however, UV radiation triggers mutations in the melanocytes, thereby causing uncontrollable cell growth.[8] Melanomas may also be hidden and develop in areas that are not exposed to UV light such as between fingers and on the scalp. Therefore, other unknown factors beyond UV light may also contribute to an individual's risk of developing melanoma.[1]

Considering that melanoma is the most deadly type of skin cancer, early diagnosis becomes paramount to a positive clinical outcome. However, melanoma is difficult to diagnose because it can develop anywhere on the human body.[1] Melanoma is often found in the form of a mole but can also occur in normal-looking skin, emphasizing that it can progress undetected. Intuitively, melanoma in moles is comparatively easier to detect than in normal-looking skin.[1] Initial symptoms of melanoma in moles include asymmetric shape, irregular border, uneven distribution of color, and time-based evolution of any of the aforementioned characteristics.[1]

Finally, there are several clinical, financial, and psychosocial implications of an untimely melanoma diagnosis that can result in detrimental consequences for the patient. For instance, similar to any other type of cancer, the progression of invasive melanoma is much more difficult to manage clinically in latter stages as the likelihood of metastasis and tumor recurrence increases substantially.[9] Furthermore, a late diagnosis can result in severe financial distress for the patient due to the sheer urgency, unpredictability, and expense associated with emergent treatment procedures.[10] Additionally, there are severe mental and psychiatric factors that toll patients while undergoing palliative care treatment associated with melanoma; these include substance abuse, depression, and anxiety.[11]

Related Work

The utilization of deep learning in health care, with applications including imaging diagnosis, digital pathology, prediction of hospital admission, drug design, and others, has been an emerging area of study.[12] With recent leaps in computational power and rapid advancements in artificial intelligence, many investigators have leveraged deep learning techniques to help mitigate one of the most prevalent diseases in the word: cancer.[13] Specifically, a hot topic has been classifying images of tumors using machine learning as a method of delivering efficient, timely, and accurate diagnoses that can decrease the number of adverse outcomes, treatments costs, and downstream complications commonly associated with cancer progression.[14]

The most substantial amount of work within this area in recent years has been breast cancer. In June, 2020, Sichuan and Yibin University researchers Lui et al. implemented a novel classification method to classify pathological images of breast cancer through utilization of deep convolutional neural networks (CNNs), DeepBC integrated Inception, ResNet, AlexNet, and feature extraction.[15] They achieved 92% and 96.43% accuracy in classifying patients and images, respectively, for both benign and malignant tumors across several types commonly associated with breast cancer, such as ductal and lobular

carcinomas. Their findings suggest robustness and generalizability of their method compared to the state-of-the-art, oncologist-level classification of breast cancer tumors, further reporting positive downstream outcomes which include increased diagnosis efficiency, reduction in pathologist workload, and greater aversion to the possibility of misdiagnosis.

Methods of classifying skin cancer tumors through utilization of deep learning have also been implemented; however, developments in this area are much more primitive. Esteva et al. at Stanford University demonstrated a dermatologist-level classification of skin cancer using deep CNNs trained on a dataset of 129,450 clinical images consisting of 2,032 different diseases.[16] However, the Stanford team reported several shortcomings in their findings, primarily noting a lack of analysis regarding the sensitivity and specificity of their design which can result in downstream socioeconomic complications such as mass hysteria from a high false-positive rate.

The work discussed in this proposal intends to build off of previous skin disease assessment models such as that of Esteva et al., yet focus the effort more towards optimizing sensitivity, specificity, and generalizability of the resulting deep CNN model.[17–19]

Claim / Target Task

The overarching goal of this project is to utilize a supervised multi-class classification algorithm to assess the properties of an image containing indications to either 1) melanoma, 2) seborrheic keratosis, or 3) non-melanoma skin (i.e. nevi). An 11-gigabyte kaggle dataset provided by user Pablo Lopez Santori that consists of labeled 2,750 images across these three classifications will be used for training, validation, and testing.[20]

The primary focus of this classification algorithm will be to minimize the false-negative rate as false negatives can result in severe consequences for a patient including treatment costs, adverse outcomes, and exacerbations of psychosocial symptoms. Secondarily, there will be an attempt to minimize the false-positive rate to decrease potential complications including mass hysteria and individual stress. If successful, a novel CNN will be created to provide a computational-based diagnosis that will be less costly relative to traditional diagnoses through a healthcare provider with regards to both time and money (**Figure 2**).

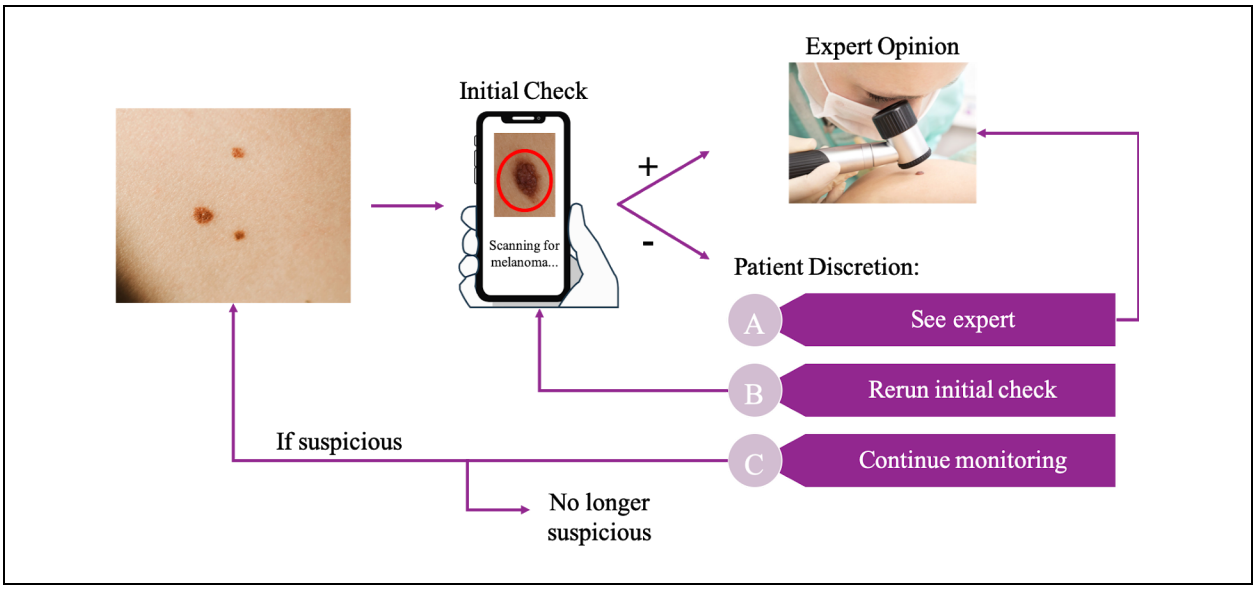

Figure 2: Depicted is an intuitive schematic showcasing a potential future timeline of diagnosis for melanoma and other skin conditions using models such as the one proposed. Patients will have the ability to scan their skin lesion of interest with the mobile application through utilization of the deep convolutional neural network and then choose to act on the pre-diagnosis based on the possible options outlined above. Patients may choose inexpensive, in-home checks utilizing image classification models such as the one proposed. Depending on the results, patients may choose to seek an expert's opinion or make a clinical decision at their own discretion. Furthermore, as precision medicine procedures develop, future medical diagnoses may involve machine learning algorithms acting as cursory opinions and pre-diagnoses. Both cases necessitate that such algorithms and classification models are made as error-free as possible; an important indicator of an error-free model includes optimizing the model's specificity. In other words, robustness and usability of the model in real-world scenarios heavily depends on minimizing the number of false negatives. The threshold in the proposed model will lean edge cases towards a positive diagnosis for either melanoma or seborrheic keratosis, with the justification being that false positives are typically less harmful than false negatives in regard to long-term health outcomes.

Proposed Solution

We propose using TensorFlow with the Keras API in Python to create a deep CNN with a k-nearest neighbor (kNN) algorithm appended at the end instead of a typical final dense layer. The layers in the deep CNN will consist of a combination of Conv2D, Activation, MaxPooling2D, Flatten, Dense, and Dropout layers. We will iteratively determine the best combination of these layers through model testing and validation in addition to a careful sensitivity analysis as described below. A Conv2D layer is essentially a spatial convolution over images that creates a convolution kernel. The kernel is convolved with the input layer, resulting in a tensor of outputs. An activation function will ensure that outputs are in an appropriate input form for the next layer of the model through normalization, determining which neurons in any given layer are activated. MaxPooling2D layers perform downsampling of the input representation based on the *pool_size* parameter it receives. Flatten layers put the model into a 1D array. Finally, dense layers and dropout layers are used to ensure that all neurons are fully connected and that the model does not overfit through regularization, respectively. The kNN algorithm will be implemented in place of the typical final dense layer in an effort to increase algorithmic diversity and take advantage of the prior work done by the CNN in identifying features and reducing image classification complexity.

In order to further minimize the possibility of overfitting in our model, we will use data augmentation to artificially boost diversity of the image dataset and "warp" the image dimensions. We believe this will be well suited for skin-disease classification, as melanomas and other skin disorders may come in varying shapes and sizes with especially aberrant borders, complex pathoanatomy, and variant shades of color. Effectively, this will allow the model to reduce dependence on size of the skin lesion in any given image while still accounting for aberrant border patterns, thus helping classify images that originate from a variety of sources and allowing for increased flexibility in image capture distance from the subject.

In order to maximize the specificity of our model, we will iteratively alter the layers of the deep convolutional neural network outlined previously and investigate which model will minimize the false-negative rate. This, in addition to achieving an overall accuracy of greater than 90%, will be the primary endpoint that we intend to achieve in our proposed solution. Moreover, we will also attempt to iteratively alter the layers so that the sensitivity is maximized, until doing so will achieve a suboptimal model specificity. Both of these approaches will be carried out through alteration of the model's threshold in terms of what characterizes a positive or negative diseased skin lesion such that edge cases will tend towards a positive melanoma or seborrheic keratosis classification.

The proposed solution will hopefully be a highly robust and generalizable deep CNN that can eventually be implemented in a mobile operating system, such as iOS or Android, to classify skin lesion images as melanoma positive skin, seborrheic keratosis positive skin, or non-melanoma, non-seborrheic keratosis nevi. Another endpoint that we hope to achieve with this proposed solution is transferability to other skin diseases for image classification, such as those of eczema, dermatitis, and psoriasis, with the intended goal of opening future avenues of investigation and research related to efficient, accurate, and low-cost dermatologist-level skin disease classification through utilization of deep CNNs.

References

1. Melanoma - Symptoms and causes. Mayo Clinic. Accessed October 11, 2020. https://www.mayoclinic.org/diseases-conditions/melanoma/symptoms-causes/syc-20374884
2. Glazer AM, Winkelmann RR, Farberg AS, Rigel DS. Analysis of Trends in US Melanoma Incidence and Mortality. *JAMA Dermatol*. 2017;153(2):225. doi:10.1001/jamadermatol.2016.4512
3. Melanoma Incidence and Mortality, United States–2012–2016 | CDC. Published September 16, 2020. Accessed October 31, 2020. https://www.cdc.gov/cancer/uscs/about/data-briefs/no9-melanoma-incidence-mortality-UnitedStates-2012-2016.htm
4. American Cancer Society. What Is Melanoma Skin Cancer? | What Is Melanoma? Published August 14, 2019. Accessed October 11, 2020. https://www.cancer.org/cancer/melanoma-skin-cancer/about/what-is-melanoma.html
5. Marghoob AA, Scope A. The complexity of diagnosing melanoma. *J Invest Dermatol*. 2009;129(1):11-13. doi:10.1038/jid.2008.388
6. Melanoma - Diagnosis. Cancer.Net. Published June 25, 2012. Accessed October 31, 2020. https://www.cancer.net/cancer-types/melanoma/diagnosis
7. Domingues B, Lopes JM, Soares P, Pópulo H. Melanoma treatment in review. *ImmunoTargets Ther*. 2018;7:35-49. doi:10.2147/ITT.S134842
8. Halpern A, Marghoob A, Reiter O. Melanoma. The Skin Cancer Foundation. Accessed October 31, 2020. https://www.skincancer.org/skin-cancer-information/melanoma/
9. Voss RK, Woods TN, Cromwell KD, Nelson KC, Cormier JN. Improving outcomes in patients with melanoma: strategies to ensure an early diagnosis. *Patient Relat Outcome Meas*. 2015;6:229-242. doi:10.2147/PROM.S69351
10. Financial help for people who have skin cancer. Accessed October 31, 2020. https://www.aad.org/public/diseases/skin-cancer/types/common/melanoma/financial-help
11. Fawzy FI, Fawzy NW, Hyun CS, et al. Malignant melanoma. Effects of an early structured psychiatric intervention, coping, and affective state on recurrence and survival 6 years later. *Arch Gen Psychiatry*. 1993;50(9):681-689. doi:10.1001/archpsyc.1993.01820210015002
12. Zhu W, Xie L, Han J, Guo X. The Application of Deep Learning in Cancer Prognosis Prediction. *Cancers*. 2020;12(3). doi:10.3390/cancers12030603
13. Landhuis E. Deep learning takes on tumours. *Nature*. 2020;580(7804):551-553. doi:10.1038/d41586-020-01128-8
14. Nadeem MW, Ghamdi MAA, Hussain M, et al. Brain Tumor Analysis Empowered with Deep Learning: A Review, Taxonomy, and Future Challenges. *Brain Sci*. 2020;10(2). doi:10.3390/brainsci10020118
15. Wenzhong L, Huanlan L, Caijian H, Liangjun Z. *Classifications of Breast Cancer Images by Deep Learning*. Radiology and Imaging; 2020. doi:10.1101/2020.06.13.20130633
16. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115-118. doi:10.1038/nature21056
17. Chan S, Reddy V, Myers B, Thibodeaux Q, Brownstone N, Liao W. Machine Learning in Dermatology: Current Applications, Opportunities, and Limitations. *Dermatol Ther*. 2020;10(3):365-386. doi:10.1007/s13555-020-00372-0
18. Ray R, Abdullah AA, Mallick DK, Ranjan Dash S. Classification of Benign and Malignant Breast Cancer using Supervised Machine Learning Algorithms Based on Image and Numeric Datasets. *J Phys Conf Ser*. 2019;1372:012062. doi:10.1088/1742-6596/1372/1/012062
19. Dhahri H, Al Maghayreh E, Mahmood A, Elkilani W, Faisal Nagi M. Automated Breast Cancer Diagnosis Based on Machine Learning Algorithms. Journal of Healthcare Engineering. doi:https://doi.org/10.1155/2019/4253641
20. Melanoma Detection Dataset. Accessed October 31, 2020. https://kaggle.com/wanderdust/skin-lesion-analysis-toward-melanoma-detection