Using CNNs and Machine Learning to Diagnose Skin Cancer Moles

Anika Sivasankar

Massachusetts Academy of Math and Science at Worcester Polytechnic Institute

Worcester, MA

Executive Summary

A skin cancer diagnosis often involves a very time-consuming process that requires extracting mole samples, testing in a lab, and weeks of waiting for results. This delay can affect any medical treatment the patient might have to undergo. Machine Learning models offer a more promising alternative to this process by providing accurate diagnoses quickly. Individuals could submit an image of their mole to the model and receive a diagnosis almost instantly on whether the mole is benign or malignant.

A specific type of Machine Learning model must be used: a Convolutional Neural Network (CNN). Convolutional Neural Networks are a deep learning approach used specifically for image recognition tasks. CNNs are chosen because they have high accuracy with recognizing features within medican images, a characteristic essential for distinguishing between benign and malignant conditions (Li et al., 2023).

Some issues, such as image quality and lack of diverse skin tone representation within the images, are flaws that exist in some pre-existing Machine Learning models that impact wide use in the medical community (Daneshjou et al., 2022). Addressing these flaws is essential to creating a model that is accurate and practical for real-world application by dermatologists and other healthcare workers looking for a model to help make their jobs easier and more efficient.

Therefore, this proposed project aims to develop a Convolutional Neural Network-based Machine Learning model to classify skin cancer moles as benign or malignant based off of an image given to the model by the user.

Using CNNs and Machine Learning to Diagnose Skin Cancer Moles

When a questionable skin cancer mole is developed, the typical process would involve going to a healthcare provider, undergoing a biopsy to remove the mole from the skin, and waiting for the sample to be analyzed in a lab. This more traditional method can be quite a lengthy process—taking weeks just to receive results. For those patients who are waiting for answers and the results, this waiting period can be extremely stressful and could even delay any kinds of treatment if the mole is found to be malignant. Time is an especially valuable element in certain cases of skin cancer, where early detection is often the key to an effective treatment plan for the patient. Now, imagine the level of convenience that an online, image-based tool could give to the public, instantly analyzing and assessing an uploaded image of a mole, determining if it's benign or malignant in just a matter of seconds. A tool like this could create faster decision-making, leading to earlier treatment for the patient.

Overall, skin cancer can be one of the deadliest cancers once it metastasizes. Early detection of melanoma is extremely crucial. When caught before it reaches the lymph nodes, about 99% of patients survive for at least five years (Cleveland Clinic). Once the cancer starts to metastasize, the 99% five-year survival rate can fall as low as 27%, which is why it is important to diagnose skin cancer before it gets too out of hand and becomes nearly incurable.

In recent years, Machine Learning models have been slowly incorporated into medical practice and healthcare, including in cancer diagnosis (Cruz & Wishart, 2007). Machine Learning technology and Artificial Intelligence have shown evidence to significantly reduce man-made errors that doctors might typically make (Choudhury & Asan, 2020). No matter how much training doctors receive before treating patients, they are still human; therefore, errors can still occur. Machine Learning models can be a great solution to problems physicians encounter often in their careers. With ML models being able to know patient data and history of the patient it is aiming to treat, the model could provide the best and most accurate personalized treatment plans for the patient, something that even doctors could have trouble doing. And overall, incorporating ML and AI into medican diagnisus can increase the efficiency and accuracy of the diagnosis (Khalifa & Albadawy, 2024).

There have been many studies that used Machine Learning technology to efficiently diagnose skin cancer and differentiate the cancer as benign or malignant, using Convolutional Neural Networks, a type of ML model used for image recognition and analyzing images (Yamashita & Nishio, 2018). The model created during this study followed a similar layout to the one seen in Figure 1. The dataset images were fed into the model, went through image preprocessing methods, were put in a pre-trained model, and were classified from there, using Grad-CAM visualization techniques. One specific study from 2023 addressed challenged in diagnosing skin cancer using both Machine Learning and Deep Learning using CNNs Melarkode et al., 2023). However, one point that this paper wanted to incorporate into future methods was NGS, to increase the level of efficiency in data output the Machine Learning model ends up giving to the user (Melarkode et al., 2023). However, there are some flaws in those already existing models, such as not giving the best and most accurate diagnoses when it comes to different levels of quality in the images, or even in images with skin-color diversity.

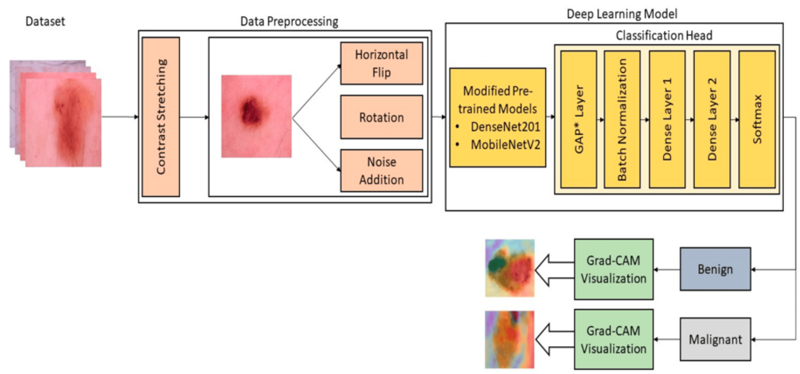


Figure 1: Example of how a ML model to Classify Skin Cancer Lesions works. (Rehman et al., 2022)

The overall results of this model will ultimately be to develop a more reliable tool for users to use to diagnose skin cancer moles as benign or malignant, regardless of the quality of the image that the user decides to submit.

Section II: Specific Aims

The goal of this proposal is to develop a tool to help in the diagnosis of skin cancer or questionable moles. These technologies can make diagnoses more accurate and efficient, helping doctors find and treat problems more quickly and improving the overall care patients receive. The overall engineering goal of this project is to create a Machine Learning Model using CNNs to accurately and efficiently diagnose whether a mole represents benign or malignant skin cancer while addressing any flaws in past ML models that aim to perform the same goal. The rationale is that when a skin cancer mole is developed, you would have to go to a healthcare provider and have a biopsy done to remove the mole and have an analysis done, which can give a diagnosis. But this process can take an extremely long amount of time, and for patients who might have benign moles and are waiting for their results that say so, that delay can push back any kinds of treatment steps. The work this project proposes will allow more efficient diagnosis while still having the same, if not a better level of accuracy.

Specific Aim 1: Use Python to create a base-level Convolutional Neural Network Machine Learning model, as well as find a good quality and well-designed publicly accessible dataset that can be used in the model training process.

Specific Aim 2: Train the model using 70% of the dataset found, making sure it covers all the aspects and factors needed to determine if a skin cancer mole is benign or malignant.

Specific Aim 3: With the remaining 30% of the dataset, test out the model to see if it can accurately diagnose the picture of the skin cancer mole fed to the model as benign or malignant.

People who use this ML model are expected to get an accurate diagnosis in a very short period, whether the image of the skin cancer mole that they inputted into the model is benign or malignant, as well as potentially an explanation of why that diagnosis was made.

Section III: Project Goals and Methodology

Relevance/Significance

Skin cancer is one of the most common and potentially deadly forms of cancer, especially if not detected early. Traditional diagnostic methods, such as biopsies, can take weeks to provide results, causing delays in treatment and increased stress for patients (American Cancer Society, n.d.).

A Machine Learning model that can quickly and accurately diagnose skin cancer from an image could help address this challenge. By reducing the wait time for a diagnosis and providing a tool accessible to more people, this technology could play a significant role in early detection. Early diagnosis is critical, as the survival rate for melanoma drops significantly once it spreads (Cleveland Clinic). Developing a model that is inclusive and accurate-can make a meaningful difference in improving patient outcomes and saving lives.

Innovation

The innovation in this project is solving problems that current Machine Learning models for skin cancer diagnosis haven't fully fixed. While there are ML models that can identify if a mole is benign or malignant, many struggle with things like not addressing images of people with different skin tones, using high-quality data, or explaining how they made their decisions (Patel et al., 2023). This project aims to improve on these by building a model that uses a more diverse data set addresses all skin tones, and potentially includes explanations for its diagnoses so doctors and patients can understand the results. This combination of improvements isn't that common in pre-existing models, which makes this approach very unique and useful.

Methodology

The planned Machine Learning model for diagnosing whether questionable skin cancer moles are benign or malignant will work by analyzing an uploaded image of a mole and giving a diagnosis almost instantly. The programming language that this model will be coded in is Python. Python was chosen for the coding language because it has a high level of flexibility and a lot of support for machine learning libraries. The first phase of this project involves building the model and preparing it to ultimately be trained with a publicly accessible dataset from Kaggle that includes various images of skin cancer moles (Tschandl, Rosendahl, & Kittler, 2018). These images will be of skin cancer moles that are both benign (non-cancerous) and malignant (cancerous). The representation of both benign and malignant skin cancer mole images in the dataset allows for the model to learn how to accurately distinguish between the two types of cases of skin cancer moles.

After the initial training, the model will undergo a testing phase, where it will be trained with a wide range of additional skin cancer mole images. These test images will be chosen in a way that will guarantee that there is variety in quality, ranging from some more high-resolution, clear images to some lower-quality, grainier ones. Variation in image quality is necessary because, in the real world, users may input images of skin moles that vary widely in clarity (Alshahrani et al., 2024). Some of those images may be very sharp and detailed, whereas other images might be more blurry or pixelated. Since the model has no control over the quality of images it receives, this testing with different quality levels of images will prepare the model to handle any input that the user might give it-and still deliver an accurate diagnosis efficiently.

Justification and Feasibility

The methods used in this project, especially using a Convolutional Neural Network (CNN) to diagnose skin cancer, are critical to achieving the goal of creating a fast, accurate, and reliable tool to diagnose. CNNs are highly effective for image analysis tasks, and studies (Li et al. 2023) have shown they can accurately differentiate between benign and malignant skin moles, making them a strong choice for this kind of project.

Summary of Preliminary Data.

The first model that was worked on was designed to classify skin lesions into different categories, like benign and malignant types of skin cancer. This model is important because it shows that by distinguishing between lesion types, there is hope that doctors can prioritize care for more high-risk cases while monitoring less serious conditions, potentially saving lives and resources (Daneshjou et al., 2023). Using a Convolutional Neural Network (CNN), this model was able to group images from the dataset into specific types of skin cancer lesions (like "mel," "nv," and "bkl") based on their features (such as color, shape, and border of the mole), making it easier to understand the variety of skin lesions. The data collected from this “preliminary” model serves as the preliminary data of this project.

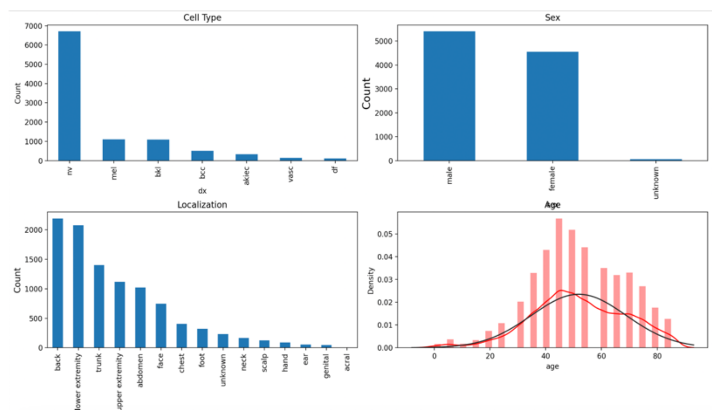


Figure 2: Dataset Characteristics. The top-left graph shows the distribution of cell types, the top-right graph displays the distribution by sex, the bottom-left graph illustrates the distribution by localization, and the bottom-right graph depicts the age distribution (Kaggle 2019).

Expected Outcomes

Once approximately 70% of the dataset is used to train the Machine Learning model to differentiate benign and malignant skin cancer moles, the remaining 30% of the dataset will be used to test the model to make sure it can accurately and efficiently determine if the given image of the skin cancer mole represents benign cancer or malignant cancer. The images of skin cancer moles that will be given to the Machine Learning model during the testing phase will be similar to the images seen in the figure to the right. The preliminary data of this project will be the diagnoses that the model gives once it is fed the testing data. This knowledge will be used to see if the model can provide an accurate diagnoses in an efficient amount of time. If it does, the model will be given additional images to test and see if it can still give accurate diagnoses. This training/testing pattern will continue until the model is able to take any kind of image of a skin cancer mole and correctly determine if it is benign or malignant.

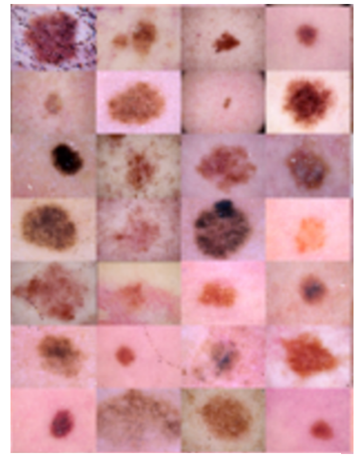


Figure 3: Examples of images of skin cancer mole that will be given to ML model. Scott Mader, Kaggle - (Tschandl, Rosendahl, & Kittler, 2018).

Potential Pitfalls and Alternative Strategies.

Some pitfalls that might be encountered during the testing process are running into situations where the model incorrectly diagnoses the skin cancer mole. If this situation occurs during the testing process, an alternative strategy of using new training data will be used. The reason for this is the images used to train the model might have been of better or worse quality than the image that was used to test the model. Therefore, incorporating varying levels of image quality into the training data is essential when we run into this issue.

Section III: Equipment

The equipment that will be used for this specific project is a MacBook Pro, using different web-based applications, as well as Visual Studio Code to create the Machine Learning model using Python and a dataset of images of skin cancer moles from Kaggle.

Section V: Ethical Considerations

An important thing to consider when creating this model is making sure the patient's data, specifically the image, that they're imputing into the model is not going into the wrong hands. Therefore, an audit and monitoring log can be implemented to analyze the data access events effectively, seeing who is accessing the data at what point (Tremblay, 2023).

Section VI: Timeline

Publicly Accessible Dataset to train the model will be selected - by November 1

Solidify program being used to create model - by November 15

Complete emails to mentors - by November 25

Set up calls with professors and mentors - by November 28

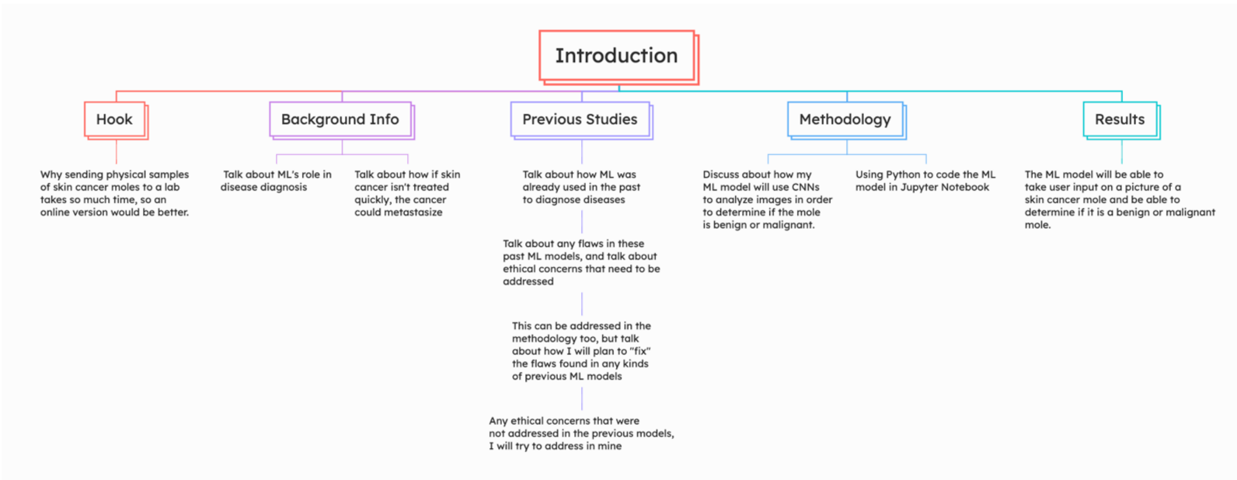
Complete base level coding of model - by December 3

Contact Microsoft and Mass General for more data to access - by December 20

Start using those datasets to continue training the model - by (TBD, depending on when Microsoft and Mass General respond)

Section VII: Appendix

This chart shows what key points will be addressed in the introduction.



Section VIII: References

Alshahrani, M., Al-Jabbar, M., Ebrahim Mohammed Senan, Ibrahim Abdulrab Ahmed, & Saif, M. (2024). Analysis of dermoscopy images of multi-class for early detection of skin lesions by hybrid systems based on integrating features of CNN models. *PloS One*, *19*(3), e0298305–e0298305. https://doi.org/10.1371/journal.pone.0298305

American Cancer Society. (n.d.). *Waiting for your biopsy or cytology test results*. American Cancer Society. https://www.cancer.org/cancer/diagnosis-staging/tests/biopsy-and-cytology-tests/waiting-for-your-biopsy-or-cytology-test-results.html

Cancer Research UK. (n.d.). *Tests for skin cancer*. Cancer Research UK. https://www.cancerresearchuk.org/about-cancer/skin-cancer/getting-diagnosed/tests

Choudhury, A., & Asan, O. (2020). Role of Artificial Intelligence in Patient Safety Outcomes: Systematic Literature Review. *JMIR Medical Informatics*, *8*(7). https://doi.org/10.2196/18599

Cruz, J. A., & Wishart, D. S. (2007). Applications of Machine Learning in Cancer Prediction and Prognosis. *Cancer Informatics*, *2*, 59. https://pmc.ncbi.nlm.nih.gov/articles/PMC2675494/

Daneshjou, R., Vodrahalli, K., Novoa, R. A., Jenkins, M., Liang, W., Rotemberg, V., Ko, J., Swetter, S. M., Bailey, E. E., Gevaert, O., Mukherjee, P., Phung, M., Yekrang, K., Fong, B., Sahasrabudhe, R., Allerup, J. A. C., Okata-Karigane, U., Zou, J., & Chiou, A. S. (2022). Disparities in dermatology AI performance on a diverse, curated clinical image set. *Science Advances*, *8*(32). https://doi.org/10.1126/sciadv.abq6147

Khalifa, M., & Albadawy, M. (2024, March 5). *AI in Diagnostic Imaging: Revolutionising accuracy and efficiency*. Computer Methods and Programs in Biomedicine Update. https://www.sciencedirect.com/science/article/pii/S2666990024000132

Li, M., Jiang, Y., Zhang, Y., & Zhu, H. (2023). *Medical image analysis using deep learning algorithms*. *Frontiers in Public Health*, 11, Article 1273253. https://doi.org/10.3389/fpubh.2023.1273253

Melarkode, N., Srinivasan, K., Qaisar, S. M., & Plawiak, P. (2023). AI-powered diagnosis of skin cancer: A contemporary review, open challenges, and future research directions. *Cancers, 15*(4), 1139. https://doi.org/10.3390/cancers15041139

Patel, R., Foltz, E. A., Witkowski, A., & Łudzik, J. (2023). Analysis of Artificial Intelligence-Based Approaches Applied to Non-Invasive Imaging for Early Detection of Melanoma: A Systematic Review. *Cancers*, *15*(19), 4694–4694. https://doi.org/10.3390/cancers15194694

Tremblay, T. (2023, November 20). *Data Access Audit: How to Create Data Access Logs for Auditing*. Kohezion. https://www.kohezion.com/blog/data-access-audit

Tschandl, P., Rosendahl, C., & Kittler, H. (2018). *The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions* [Data set]. Harvard Dataverse. https://doi.org/10.7910/DVN/DBW86T

Waiting for Your Biopsy or Cytology Test Results. (n.d.). https://www.cancer.org/cancer/diagnosis-staging/tests/biopsy-and-cytology-tests/waiting-for-your-biopsy-or-cytology-test-results.html

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: An overview and application in radiology. *Insights into Imaging, 9*(4), 611–629. https://doi.org/10.1007/s13244-018-0639-9