



# Discussion Post

**In this discussion thread, you'll post your capstone ideas and obtain feedback from your peers. Given the large class size, I'll grade at the group level, but you are encouraged to post ideas while forming groups. Each group is required to give a brief description of a project idea or two, along with an expected outcome, and also respond to the post of two peer groups.**

**Our discussion post (aim for 3 paragraphs)**

**State idea (our group, faculty mentors, and our idea)**

**Background of our problem**

**What is currently going on,**

**And what we aim to achieve!**

Our group (Reagan Esteves, Jamie Hartman, Charles Tyndal, and Chenyu Zhao) is working with Dr. Ding and Dr. Hamm on Skin of Color Disease Diagnosis. This project highlights a current issue where there is a lack of data and research on skin disease for people of color. Currently, there are no significant AI-recognizable images for people with darker skin. This racial disparity in images leads AI to misdiagnose or even ignore skin diseases in those with darker skin colors.

In sub-Saharan Africa, skin diseases are extremely high, especially in children. When these diseases are left untreated, they can lead to devastating chronic disabilities and impairments. These Neglected Tropical Diseases (NTDs) are co-endemic throughout West Africa. However, early detection and medical care can prevent the further spreading of these NTDs.

Our proposal is to teach AI to recognize trademark signs of skin disease present in darker skin colors. The images of skin will be coming from sub-Saharan Africa where we will use knowledge from computer vision to first enhance pictures of skin conditions in order to have a proper diagnosis. We will also crop images to the most optimal size for skin disease identification. The area with skin problems will have a boundary box for AI to have an easier way to identify the skin disease location. If needed, the images will be categorized based on quality to determine which ones will need to be further enhanced or cropped. Enhancing the picture may remove interference with the identification of skin diseases.

One of the challenges will be to detect multiple diseases in a single image and effectively use AI to differentiate the two. Another challenge is cropping and highlighting the boundaries of the skin disease area.

# 9/7 Zoom Questions & Notes

**Questions for Ding and Hamm**

1. **Cross-comparison (different races)**
2. **How to divide work for this project**
3. **Look into the current data set (don’t know quality or quantity)**
4. **Is the data like already identified the skin disease**
5. **Doing multiple identifications?**
6. **Unbalanced data, accuracy issues depending on race**
7. **Alg to use, direct, need a new model?**
8. **Challenge in data or machinery model**
9. **Feasibility - is feasible or not?**

# **Milestone 1**

Due Sunday, Sep 11, 2022

This milestone requires you to write a brief description of your project, goals, and how you will accomplish them.

The report should be well organized, clearly written, and easy to follow, with relevant and legitimate information presented. It should have good mechanics (grammar, punctuation, spelling, etc.); 11-12 point font, with standard margins and 1-1.5 spacing; filler whitespace and text are avoided. This document should be 2-3 pages of text not including references. As a general guideline for this milestone and other reports, it is better to turn in a great concise document that is a few lines short of the requirement that a long watered-down document; this short(er) document should still satisfy all other quality criteria to receive full credit.

**Components of your writeup**

1. A clear and concise description of the problem you’re solving (this part will later become the abstract of your final report).

* It should be clear from the description what your final product is planned to be (a program implementing a computational model, a paper describing a new algorithm, a paper describing the study of the problem with your algorithm, etc.)
* The description should be accessible for a computer science reader unfamiliar with your interdisciplinary area.

b) Motivate your problem with context and examples. At minimum:

* The context, a sample situation, or a class of situation where the problem occurs.
* The audience is the potential user of your software and the reader of your paper.
* Need, why doing the work you do is important? What would be the positive effect of having accomplished what you plan for yourself?
* Analogous situations, if applicable: does this problem or similar problem appear in other areas of life?
* Novelty/uniqueness of your problem. Are you analyzing new data with the existing computational method? Are you developing a new computational method to look at the existing data? There should be a novel component in your capstone project, describe it.

c) Describe previous work of others that’s relevant to your project. You should:

* Describe the existing work that has already been done to accomplish your or a similar task. If applicable, provide comparisons of your approach with those solutions.
* Start collecting references and cite all sources that you have consulted. These references will form the foundation for later stages of writing up your capstone work. Published sources should contain standard bibliographical info (at minimum authors, title, venue/book/journal, year, pages). Online sources should contain working links. For theoretical results, only the articles published in peer-reviewed conferences and journal papers, and books, are allowed sources. For practical projects, other public sources (such as Wikipedia, StackOverflow, Reddit, TechCrunch, Wired, and the like) are allowed as long as you have critically assessed them and reasonably believe them to be true.

Jamie has worked with some computer vision in the past related to analyzing an image and figuring out the defects, labeling them, and categorizing them. The company that he works for has not launched its product yet so he cannot provide any more specific examples. Charles, Chenyu, and Reagan are currently taking classes that relate to the topic at hand with the faculty mentors.

d) Describe the challenges of the problem and your approach to them, and why it might work. You should:

* Describe the envisioned challenges for solving your problem.
* Technologies/tools/approaches. Although this is an early phase and the technicalities may appear later or change, at this point you should have an overall idea of the techniques and tools (software packages, mathematical methods) necessary for your work. Describe them.
* Why your particular approach is desirable or unique. Relate your approach to the descriptions of previous work outlined above.
* If your project relies on analyzing data, describe how you will obtain it and the nature of the data source (your advisor, another Tulane affiliate, data will be handed to you, you will get open dataset online, you will generate synthetic data that mimics accurate data, and so on).

e) Description of feasibility as discussed with the faculty mentor who will oversee your project, along with a meeting plan for the semester. It is best to go over your Milestone 1 document with your mentor several times to ensure that you are all on the same page. Remember that they will grade your milestones together with me and that your meeting plan will comprise 20% of your grade.

Our faculty mentors, Dr. Ding and Dr. Hamm believe this project can be furthered in steps with the aid of multiple students working on separate steps. Afterward, we come together to unify the work done for these steps to create a program that will efficiently and impartially diagnose skin diseases in darker skin tones. Our meeting plan with Dr. Ding and Dr. Hamm will consist of meeting every two weeks. We will consult with our faculty advisors about progress, how it aligns with the capstone project, and questions or if we need help.

f) Roles of each member in your group and an accounting of how much overall work each member is taking on.

We are working together yayyyyyy!!!!!

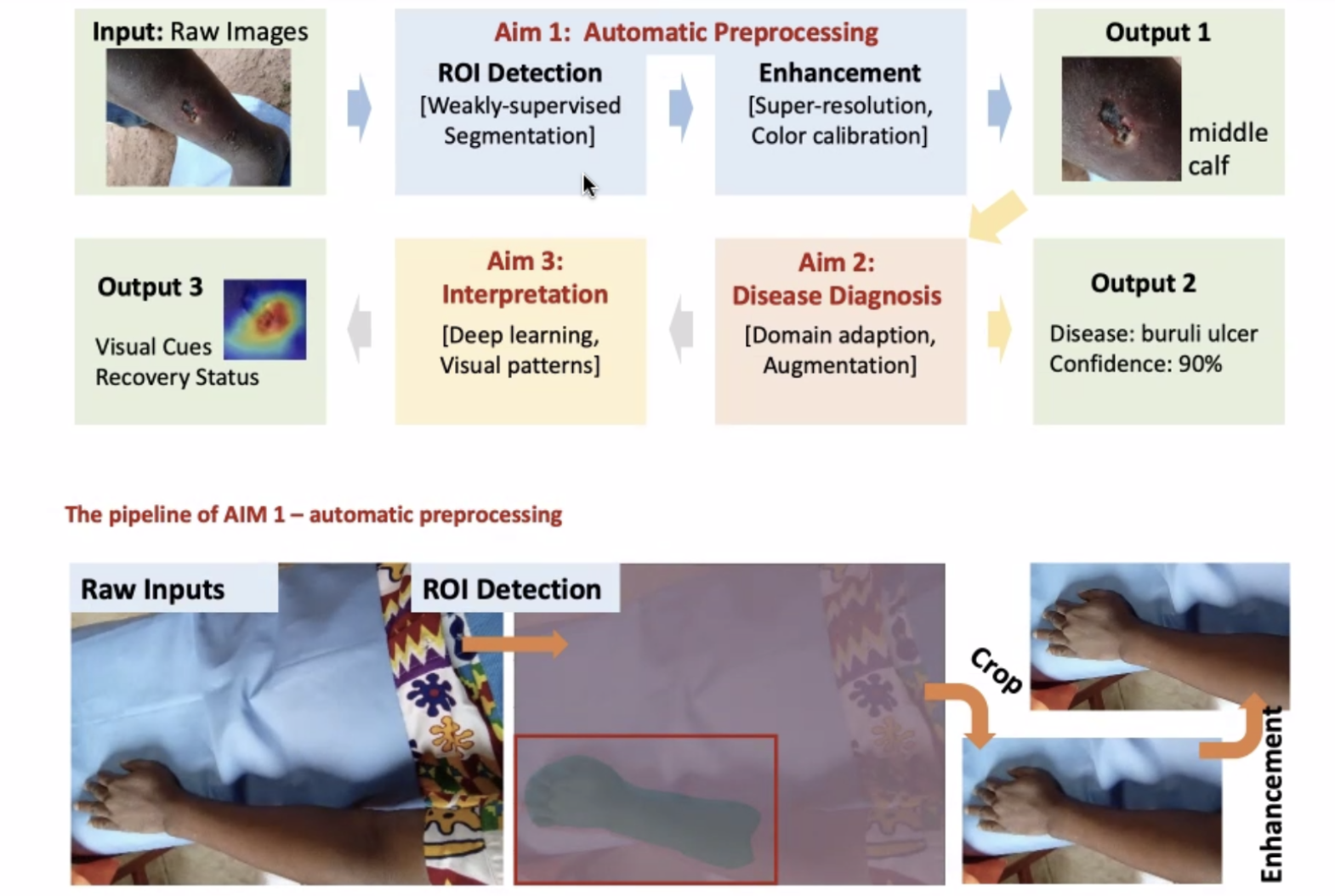
Jamie: Project Manager and CV Enhancement lead

Charles: Taking a lead on teaching machine learning and properly feeding data-sets

Chenyu: Social media

Reagan: observing

G) Bibliography



Reagan Esteves, Jamie Hartman, Charles Tyndal, and Chenyu Zhao

CMPS-4010-01

Dr. Zheng

September 21, 2022

**Milestone 1**

1. **Description of Skin Disease Diagnosis for Darker Skin Tones and Past Experiences**

In current medical studies, there is a large data disparity between light and dark skin tone patients. As a result, many darker skin tone patients have a higher chance of getting misdiagnosed or having their symptoms ignored due to the lack of research and data [6][8]. Our proposal is to teach AI to recognize trademark signs of skin disease present in darker skin colors. The images of skin will be coming from sub-Saharan Africa where we will use knowledge from computer vision to first enhance pictures of skin conditions in order to have a proper diagnosis. We will also crop images to the most optimal size for skin disease identification. The area with skin problems will have a boundary box for AI to have an easier way to identify the skin disease location. If needed, the images will be categorized based on quality to determine which ones will need to be further enhanced or cropped. Enhancing the picture may remove interference with the identification of skin diseases. We will achieve this with computer vision and machine learning algorithms implemented with Python.

Currently, this problem occurs worldwide; however, it is most prevalent in Sub-Saharan Africa. Our program will be catered to healthcare professionals helping them do their job more efficiently and accurately. This work is necessary because of the racial disparity that is evident in AI recognition of skin diseases [1][2]. With this project, we will be able to help AI recognize skin conditions in people of color, and, in turn, be able to revert or prevent damage done to those affected [2]. As previously mentioned, this problem occurs worldwide, and this program can be helpful to healthcare providers across the globe [4]. The project may aid in reducing misdiagnosis, reducing the wait time for a diagnosis, and getting treatment quicker for patients [7]. We are utilizing new data with some limited computational methods. We are aiming to expand upon the current progress so that the recognition of skin disease through AI learning is more equitable and accurate [2][5].

Jamie has worked with some computer vision in the past related to analyzing an image and figuring out the defects, labeling them, and categorizing them. The company he works for has not launched its product yet so he cannot provide any more specific examples. Charles, Chenyu, and Reagan are currently taking classes that relate to the topic at hand with the faculty mentors.

1. **Similar Research and Challenges**

In the research “Skin Disease Recognition Method Based on Image Color and Texture Features” conducted by Li-sheng Wei, Quan Gan, and Tao Ji, they observed skin problems by using image color techniques to reduce noise in the images before using the GLCM tool to analyze the skin texture. Their research targeted four specific skin diseases: herpes, paederus dermatitis, and psoriasis on light skin tones. Although our project uses similar strategies such as image filtering, our filtering is focused on sharpening low-quality pictures and contrast. The sharpening and enhancement of the photos allow a higher accuracy in determining the skin problem with AI. Moreover, our project uses AI’s ability for skin disease identification based on a given data set of skin disease images on darker skin tones [3].

With this research comes open-source data found on Github. Lydia Zoghbi and Anirudh Topiwala uploaded code that uses segmentation on skin diseases or trauma. This code can detect diseases and trauma in dark and light skin tones; however, it only focuses on abdominal skin rather than any part of the skin. This limited data set will need to be expanded upon so that segmentation can be used on any portion of the skin. [12]

One of the major challenges is enhancing the skin images sent from African hospitals to ensure more accurate identification of the skin disease. Because many African hospitals do not have up-to-date technology and resources, many of the pictures of the skin are not clear which will make skin disease identification harder. Many of the images will be taken from phones with the affected area not always perfectly in frame or focus. Moreover, another issue to consider is how the lighting and angle of the picture taken will disrupt the AI’s ability to identify where the patient’s skin is located and the disease. To resolve the low-resolution pictures, our first aim is to use image filtering and crop the image to the point of interest and remove the background of the image to ensure that it does not disrupt the AI identification[9][10]. Unlike previous research, this project is focusing on image filtering darker skin tones rather than lighter skin tones. Instead of using skin texture to detect skin problems such as in the research “Skin Disease Recognition Method Based on Image Color and Texture Features”, this project uses a dataset of already diagnosed skin disease images. The already diagnosed skin disease images are then used by the AI algorithm to determine skin disease based on how similar the images of the skin problem are [3].

Another major challenge is ensuring the image data are equally spread among the skin disease identification needed [4]. Currently, the data set of images given are from multiple hospitals in Africa and it is not known how evenly spread the images are for detecting skin problems. The possible lack of data for certain skin diseases that need identification is a major challenge in terms of ensuring the accuracy of the AI’s ability to determine the correct diagnosis. A possible solution to combat the AI’s accuracy is to gather more data to balance the number of images needed per skin disease identification. Another possible solution is to have the AI algorithm present an accuracy percentage and a similarity percentage to illustrate how confident the AI is in determining what skin disease it is and the possible similarity the image has to other skin diseases.

1. **Feasibility and Meeting Plan**

Our faculty mentors, Dr. Ding and Dr. Hamm believe this project can be broken into three major aims: automatic preprocessing, disease diagnosis, and interpretation. The project is further broken down into smaller steps where each group member will take up a step. Afterward, the steps are unified to create a program that will efficiently and impartially diagnose skin diseases in darker skin tones. Our mentors have advised us to primarily focus on the first aim of the overall project, that being taking in images and making them clear for the AI to be able to attempt to read. Our meeting plan with Dr. Ding and Dr. Hamm will consist of meeting about every two weeks. Because the data contains patient information, we will need to complete training sent from the hospitals before we are allowed to access the data. We will also complete a similar training from Tulane. We will consult with our faculty advisors about progress, how it aligns with the capstone project, and questions or if we need help.

1. **Roles of Each Member, Personal Challenges, and Solutions**

Charles, Reagan, Chenyu, and Jamie will split the work for this project into two separate groups: one for segmentation and the other for ROI detection. The project will be in Python programming language. Some tools we will use are Google Colab and local servers to run larger-scale inputs. Chenyu and Reagan are going to focus on image enhancement. They will utilize the information we learn in Dr. Ding’s Computer Vision class to greatly enhance the image quality of cropped images. The enhancement will be conducted through a series of image filtering to give contrast between the person’s body part and the background. The image filtering will allow an easier ROI detection which Jamie and Charles will focus mainly. Jamie and Charles will also assist in cropping the images to display only the affected skin area. Charles is going to focus on what he has learned in his Machine learning class and Jamie is going to apply the Computer Vision information she has learned in Dr. Ding’s class.

As mentioned, the main challenge we may face in our project is ensuring the images would be enhanced to the required standard as many of the images are taken with poor-quality cameras or cell phone images. Moreover, because the images are of darker skin tones, it is harder to determine the area of the skin disease without enhancing the images. There is also an angle and lighting issue in the images taken. To combat the image clarity issue, we plan on applying multiple filters to increase the contrast and sharpen the image. For instance, we may first apply a Gaussian filter to first smooth the image and then apply the sharpen filter to further increase the clarity of the image. Another way to ensure the image’s clarity is by applying Fourier Transform to reduce the noise in the image. The multiple filters and Fourier Transform will help in creating more clarity of the area of the skin disease against darker skin tone. Similarly, another issue would be accidentally oversharpening the images which may affect what is detected. A possible solution would be to limit the filtering to a couple of filters and run tests after each additional filter we apply to ensure the images are not oversharpened.

After we render better quality of these images with the use of filters, the secondary challenge will be determining the area of the skin with the skin disease because the data set we are given is just images of the whole limb with a simple disease label. One possible way to resolve the challenge is using edge detection to take the background out and just have the limb. Color calibration may also be used for a better view of the background vs the body limb and to find the specific area the skin disease is located at. Superpixels may be used to group areas of skin disease. If colors do not match with what is expected of the normal skin tone, then the grouping of the superpixels will be able to indicate that and we can find the diseased spot.

**Bibliography**

[1] R. Fathy and J. B. Lipoff. Lack of skin color in Google image searches may reflect under-representation in all educational resources. Journal of the American Academy of Dermatology, 86(3):e113–e114, 2022.

[2] J. A. Diao and A. S. Adamson. Representation and misdiagnosis of dark skin in a large-scale visual diagnostic challenge. Journal of the American Academy of Dermatology, 86(4):950–951, 2022.

[3] V. Anand, S. Gupta, and D. Koundal. Skin disease diagnosis: challenges and opportunities. In Proceedings of Second Doctoral Symposium on Computational Intelligence, pages 449–459. Springer, 2022.

[4] J. Lester, J. Jia, L. Zhang, G. Okoye, and E. Linos. Absence of images of skin of color in publications of COVID-19 skin manifestations. The British journal of dermatology, 183(3):593–595, 2020.

[5] E. Akuffo-Addo, M. N. Nicholas, and M. Joseph. COVID-19 Skin Manifestations in Skin of Colour. Journal of Cutaneous Medicine and Surgery, 26(2):189–197, 2022.

[6] C. Cassius, L. Frumholtz, A. de Masson, O. Dadzie, A. Petit, et al. Under-representation of people of African ancestry in publications on the cutaneous manifestations of COVID-19: coincidence or physiology? Journal of the European Academy of Dermatology and Venereology, 2021.

[7] R. Pangti, S. Gupta, N. Nischal, and A. Trikha. Recognizable vascular skin manifestations of SARS CoV-2 infection are uncommon in patients with skin-of-color. Clinical and Experimental Dermatology, 2020.

[8] N. Kluger and M. Samimi. Is there an under-representation of skin of color images during the COVID-19 outbreak? Medical hypotheses, 144:110270, 2020.

[9] A. Bhardwaj and P. P. Rege. Skin lesion classification using deep learning. In Advances in Signal and Data Processing, pages 575–589. Springer, 2021.

[10] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun. Dermatologist-level classification of skin cancer with deep neural networks. nature, 542(7639):115–118, 2017.

[11]Li-sheng Wei, Quan Gan, Tao Ji, "Skin Disease Recognition Method Based on Image Color and Texture Features", *Computational and Mathematical Methods in Medicine*, vol. 2018, Article ID 8145713, 10 pages, 2018. <https://doi.org/10.1155/2018/8145713>

[12]Topiwala Anirudh, Al-Zogbi Lidia, and Fleiter Thorsten, Krieger Axel, “Adaptation and Evaluation of Deep Leaning Techniques for Skin Segmentation on Novel Abdominal Dataset”, *2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE)*},p. 752--759, IEEE,2019.