Project Design Phase-I
Proposed Solution

Date	19 September 2022
Team ID	PNT2022TMID13957
Project Name	AI-based localization and classification of skin disease with erythema
Maximum Marks	2 Marks

Proposed Solution

S. No	Parameter	Description
1.	Problem Statement (Problem to be solved)	Although computer-aided diagnosis (CAD) is used to improve the quality of diagnosis in various medical fields such as mammography and colonography, it is not used in dermatology, where noninvasive screening tests are performed only with the naked eye, and avoidable inaccuracies may exist. our study shows that CAD may also be a viable option in dermatology by presenting a novel method to sequentially combine accurate segmentation and classification models. Given an image of the skin, we decompose the image to normalize and extract high-level features. Using a neural network-based segmentation model to create a segmented map of the image, we then cluster sections of abnormal skin and pass this information to a classification model. We classify each cluster into different common skin diseases using another neural network model. Our segmentation model achieves better performance compared to previous studies, and achieves a near-perfect sensitivity score in unfavorable conditions. Our classification model is more accurate than a baseline model trained without segmentation, while also being able to classify multiple diseases within a single image. This improved performance may be sufficient to use CAD in the field of dermatology
2.	Idea / Solution description	The segmentation and classification of skin diseases has been gaining attention in the field of artificial intelligence because of its promising results. Two of the more prominent approaches for skin disease

		segmentation and classification are clustering algorithms and support vector machines (SVMs). Clustering algorithms generally have the advantage of being flexible, easy to implement, with the ability to generalize features that have a similar statistical variance. On experimenting with various clustering algorithms, such as fuzzy c-means, improved fuzzy c-means, and K-means, achieving approximately 83% true positive rates in segmenting a skin disease. implemented an ISODATA clustering algorithm to find the optimal threshold for the segmentation of skin lesions. An inherent disadvantage of clustering a skin disease is its lack of robustness against noise. Clustering algorithms rely on the identification of a centroid that can generalize a cluster of data. Noisy data, or the presence of outliers, can significantly degrade the performance of these algorithms. Therefore, with noisy datasets, caused by images with different types of lighting, non-clustering algorithms may be preferred; however, implemented an improved version of the fuzzy clustering algorithm using the RGB, HSV, and LAB colour spaces to create a model that is more robust to noisy data. SVMs have gained attention for them effectiveness in high-dimensional data and their capability to decipher "…subtle patterns in noisy and complex
3.	Novelty / Uniqueness	we present a method to sequentially combine two separate models to solve a larger problem. In the past, skin disease models have been applied to either segmentation or classification. In this study, we sequentially combine both models by using the output of a segmentation model as input to a classification model. In addition, although past studies of non-CNN segmentation models used innovative preprocessing methods, recent CNN developments have focused more on the architecture of the model than on the preprocessing of data. As such, we apply an innovative preprocessing method to the data of our CNN segmentation model. The methods described above lack the ability to localize and classify multiple diseases within one

		image; however, we have developed a method to address this problem. Our objective is two-fold. First, we show that CAD can be used in the field of dermatology. Second, we show that state-of-the-art models can be used with current computing power to solve a wider range of complex problems than previously imagined. We begin by explaining the results of our experimentation, followed by a discussion of our findings, a more detailed description of our methodology, and finally, the conclusions that can be drawn from our study
4.	Social Impact / Customer Satisfaction	People may be affected with erythema and still will not know the seriousness they will not know that you must see a doctor for consulting and treat it at the start. our project will use Computer-aided diagnosis (CAD) is a computer-based system that is used in the medical imaging field to aid healthcare workers in their diagnoses. CAD has become a mainstream tool in several medical fields such as mammography and colonography. However, in dermatology, although skin disease is a common disease, one in which early detection and classification is crucial for the successful treatment and recovery of patients, dermatologists perform most noninvasive screening tests only with the naked eye. So, this may be a great idea that will help our society
5.	Business Model (Revenue Model)	As our main objective was to demonstrate the viability of CAD, the performance was mostly determined using pixel-level sensitivity rather than the Intersection over Union or the Dice coefficient metrics that are often used to measure segmentation performance. Moreover, we mainly focused on the true positive rates of segmentation, represented by the sensitivity metric. This is because our aim was to create a screening test method to help healthcare workers make a more accurate diagnosis by preventing abnormal skin from being overlooked. Nevertheless, we also measured the performance of our model using the specificity, Dice coefficient, and Harsdorf distance to provide a more complete performance comparison. We measured

		these metrics by comparing the output from our U-Net model to an image that was masked by professional dermatologists. Going through each pixel, if a pixel of the U-Net output was black and the pixel of the dermatologist-masked image at the same location was black, this is seen as a true negative. If both were white, this was seen as a true positive. If the U-Net output was black but the dermatologist mask was white, this was seen as a false negative, and the converse was a false positive
6.	Scalability of the Solution	We have shown that even without a large dataset and high-quality images, it is possible to achieve sufficient accuracy rates. In addition, we have shown that current state-of-the-art CNN models can outperform models created by previous research, through proper data preprocessing, self-supervised learning, transfer learning, and special CNN architecture techniques. Furthermore, with accurate segmentation, we gain knowledge of the location of the disease, which is useful in the preprocessing of data used in classification, as it allows the CNN model to focus on the area of interest. Lastly, unlike previous studies, our method provides a solution to classify multiple diseases within a single image. With higher quality and a larger quantity of data, it will be viable to use state-of-the-art models to enable the use of CAD in the field of dermatology