CPS 4801 Final Project Proposal

Amir Zaman
Dept. of Computer Science
Kean University
Union, United States
zamanam@kean.edu

Ayan Desai
Dept. of Computer Science
Kean University
Union, United States
desaiay@kean.edu

Chigozie Ofodike
Dept. of Computer Science
Kean University
Union, United States
ofodikech@kean.edu

I. Introduction

Skin cancer remains a significant health concern globally, with early detection playing a pivotal role in successful treatment outcomes. In this study, we focus on leveraging machine learning techniques to aid in the classification of skin lesions as malignant or benign. To facilitate our research, we have chosen the Skin Cancer Malignant vs Benign dataset available on Kaggle. The selection of this dataset was guided by its straightforward nature and ease of comprehension, making it conducive for educational purposes within our infrastructure. Moreover, its modest size enables efficient processing and experimentation. By working with this dataset, we aim to gain practical experience in applying learned concepts from our coursework to real-world data, thereby bridging the gap between theoretical knowledge and practical implementation. Our ultimate objectives include mastering the application of various artificial intelligence methods and independently exploring novel techniques pertinent to skin cancer detection.

II. Dataset Introduction

The dataset comprises images depicting skin moles, meticulously categorized into two classes: malignant, indicative of skin cancer, and benign, representing non-cancerous lesions. Each image undergoes preprocessing, primarily involving meticulous labeling to facilitate model training. These annotated images serve as the foundation for our classification endeavors, enabling the development and evaluation of machine learning algorithms for skin lesion diagnosis.

III. Literature Review

Recent advancements in the field of dermatology have witnessed the integration of artificial intelligence (AI) techniques to enhance skin cancer detection accuracy. Notably, Barata et al. (2023) employed a reinforcement learning model to augment the performance of existing supervised learning algorithms in skin lesion classification. Their approach yielded promising results, mitigating overconfidence while maintaining high marker sensitivities. This study underscores the potential of AI-driven decision support systems in improving diagnostic accuracy and patient outcomes in dermatological practice [1].

IV. Existing Methods To Be Used

Our methodology encompasses the utilization of established machine learning models and evaluation metrics tailored to skin lesion classification. Specifically, we plan to employ decision tree and ResNet-50 pre-trained classification models. Evaluation will be conducted using metrics such as Intersection over Union (IoU), precision, and recall to assess model performance comprehensively. Furthermore, we will incorporate uncertainty quantification techniques, including Entropy methods for model assessment to enhance the robustness and reliability of our findings.

V. Experiment/Simulation Plans

Our experimental workflow entails the classification of skin lesion images, followed by partitioning into distinct training and testing cohorts. The RGB pixel intensities within the images will serve as feature attributes for model training. Subsequently, both decision tree and ResNet-50 models will undergo training using the labeled dataset. Performance evaluation will be conducted on the test set, utilizing IoU, precision, and recall calculations to gauge classification accuracy and efficacy. Through meticulous experimentation and analysis, we aim to delineate the comparative strengths and limitations of the employed models in skin lesion classification tasks.

References

[1] Barata, C., Rotemberg, V., Codella, N.C.F. et al. A reinforcement learning model for AI-based decision support in skin cancer. Nat Med 29, 1941–1946 (2023). https://doi.org/10.1038/s41591-023-02475-5https://doi.org/10.1038/s41591-023-02475-5