Advanced Programming

Project Proposal

Skin Cancer Recognition

Teacher: Yeleu Sultanmurat

Group: IT-2209

Group Members: Aiteni Tolegen, Ilham Assimov, Erniyaz Ashuov

January 23, 2024

Skin Cancer Recognition with Machine Learning

Problem Description:

The goal of this project is to develop a machine learning model capable of recognizing skin cancer, with a focus on addressing its impact on the population of Kazakhstan. Skin cancer is a growing concern globally, and early detection is crucial for effective treatment. This model aims to provide a reliable and accessible tool for preliminary skin cancer screening.

Background Information:

Cancer is a significant health concern in Kazakhstan, being the third leading cause of premature death with over thirty thousand people diagnosed annually. Among the most common cancers are lung, skin, breast, and stomach. A study published in the "Asian Pacific Journal of Cancer Prevention" emphasizes the impact of breast cancer in Kazakhstan, highlighting it as one of the leading causes of mortality. Additionally, the World Health Organization (WHO) provides comprehensive data on cancer in Kazakhstan, including statistics and information on various types of cancer, which helps understand the broader context of cancer incidence and mortality rates in the country.

The high incidence of skin cancer specifically underscores the urgent need for improved diagnostic tools. Modern technologies like machine learning offer promising solutions for early and more efficient detection, which is crucial for effective treatment and management of skin cancer. This background illustrates the importance of developing innovative approaches to cancer detection, particularly for skin cancer, in Kazakhstan.

Link to "Asian Pacific Journal of Cancer Prevention": https://pubmed.ncbi.nlm.nih.gov/37898839/

Link to WHO: https://www.who.int/publications/m/item/cancer-kaz-2020

Available Solutions

Advancements in technology have led to the development of innovative solutions for skin cancer detection, leveraging artificial intelligence to enhance accuracy and accessibility. Notable examples include:

1. DermaSensor: A handheld device approved by the U.S. Food and Drug Administration (FDA), DermaSensor uses AI-powered technology to detect skin cancer. It operates by emitting light onto suspicious moles or lesions and analyzing the reflected light with an AI algorithm to identify the presence of cancer.
   Link: https://www.reuters.com/business/healthcare-pharmaceuticals/us-fda-clears-dermasensors-ai-powered-skin-cancer-detecting-device-2024-01-17/

2. MIT's AI Tool for Melanoma Detection: Developed by researchers at the Massachusetts Institute of Technology (MIT), this system utilizes deep convolutional neural networks (DCNNs) to

identify and screen for early-stage melanoma. The AI model is trained on wide-field images of patients' skin, effectively distinguishing suspicious pigmented lesions with high sensitivity. Link: https://news.mit.edu/2021/artificial-intelligence-tool-can-help-detect-melanoma-0402

3. FDA-Cleared Non-Invasive Skin Cancer Detector: This AI-powered device is specifically designed for primary care use, enabling non-invasive detection of skin cancer. It assesses the cellular and subcellular features of skin lesions, identifying common skin cancers such as melanoma, basal cell carcinoma, and squamous cell carcinoma, based on an algorithm cleared by the FDA. Link: https://www.medpagetoday.com/dermatology/skincancer/108301

Data Acquisition:

The data for this project will be sourced from the International Skin Imaging Collaboration (ISIC) dataset, accessible through the ISIC Challenge 2020 and the ISIC Archive. This extensive dataset offers a comprehensive collection of over 20,000 to 30,000 skin lesion images, which are invaluable for training and validating our machine learning model. These images encompass a wide range of skin types, conditions, and lesions, providing a diverse and robust dataset for developing an accurate and reliable skin cancer detection model. The availability of such a large number of images is crucial for enhancing the performance of the model and ensuring its effectiveness in identifying various forms of skin cancer.

Link: https://challenge2020.isic-archive.com/

Link: https://challenge.isic-archive.com/data/#2019

Solution Description:

We're developing a machine learning model using TensorFlow in Google Colab, programmed with Python, to detect skin cancer from images. This model will be part of a web application where users can upload skin photos for instant cancer analysis. The app could potentially diagnose other skin diseases, making skin cancer screening more accessible and efficient.

Tech Stack:

Python: For data processing and model development.

Libraries: Pandas, Numpy, and Matplotlib for data manipulation and visualization.

TensorFlow: For building and training the machine learning model.

JavaScript/Flask: To develop the web application frontend and possibly to fetch API data.

OpenCV: Potentially used for image processing and analysis.

Additional Information:

Our project will use photographic data with key details like age, diagnosis, image type, and more. The dataset includes age approximation, benign/malignant status, lesion ID, and sex, crucial for training our

machine learning model. This detailed approach is designed to improve the accuracy and reliability of detecting skin cancer.