Finalized Title:

AI-Driven Image Segmentation for Accurate Cancer Cell Identification and Classification

Problem Statement:

Accurate diagnosis of cancer significantly relies on the precise identification and classification of cancer cells within medical images such as histopathological slides. Manual examination of these images by pathologists is not only time-consuming but also subject to human error and variability. With the increasing incidence of cancer worldwide, there is an urgent need for automated, reliable, and efficient diagnostic tools to assist medical professionals in analyzing medical images.

The objective of this project is to develop an AI-driven image segmentation algorithm that can accurately identify and classify different types of cancer cells in medical images. By leveraging advanced deep learning techniques, this tool aims to enhance the accuracy of cancer diagnoses, reduce the workload on pathologists, and ultimately contribute to better patient outcomes.

Algorithm of the Project:

1. **Data Acquisition and Preprocessing:**
    - **Dataset Collection:**
        - Obtain a comprehensive dataset of medical images containing various types of cancer cells from reputable sources like The Cancer Imaging Archive (TCIA) or the NCI's Genomic Data Commons.
    - **Data Preprocessing:**
        - Resize images to a consistent resolution suitable for the model.
        - Normalize pixel values to standardize the input data.
        - Apply data augmentation techniques (rotation, flipping, scaling) to increase dataset diversity and prevent overfitting.

Dataset Authentication:

- **Source Verification:**

    - Utilize publicly available and peer-reviewed medical imaging datasets to ensure data reliability and credibility.

- **Ethical Compliance:**
    - Adhere to all ethical guidelines, including obtaining necessary permissions for data usage.
    - Ensure all patient data is anonymized to protect privacy, complying with HIPAA regulations or equivalent.
- **Data Integrity:**
    - Confirm that all images are annotated and validated by certified medical professionals or pathologists.
    - Maintain documentation of data provenance and any preprocessing steps applied.

Expected Output

Functional Segmentation Algorithm

- **Trained Model**: A robust deep learning model capable of accurately segmenting cancer cells in medical images.
- **Segmentation Accuracy**: High precision in identifying and delineating cancer cells from surrounding tissues.

Classification Results

- **Cancer Type Classification**: Accurate classification of segmented cells into various cancer types.
- **Performance Metrics**: High accuracy, precision, recall, and F1-score in classification tasks.

Performance Report

- **Evaluation Metrics**: Comprehensive metrics including accuracy, precision, recall, F1-score, and confusion matrix.
- **Benchmark Comparison**: Comparative analysis with existing methodologies or benchmarks to highlight improvements and innovations.

Documentation and Deployment

- **Project Report**: Detailed documentation outlining the methodology, experiments, results, and conclusions.
- **User Interface**: A deployable application or web interface for end-users to interact with the model and visualize results.

- **Codebase**: Well-documented codebase for reproducibility and further development.
- **Extended Capabilities**: Ideas for expanding the model to handle more cancer types or integrate with other diagnostic tools.
- **Clinical Impact**: Discussion on the potential impact on clinical practices and patient care, including benefits and limitations.

Team Members: **Team Lead**

S Karthikeya-2320030097 D Mahith-2320030165

M Sathvik Reddy-2320030162

D Mahith-2320030165