

Data Collection and Preprocessing Phase

Date	15 October 2024
Team ID	LTVIP2024TMID24772
Project Title	Implementation of Deep Learning Techniques to Detect Malaria
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Template

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

Data Collection Plan Template

Section	Description
Project Overview	This project aims to implement deep learning techniques to detect malaria from blood smear images. The primary objective is to develop a CNN-based model that can accurately classify images as either parasitized (infected) or uninfected using a publicly available dataset. The project will involve image preprocessing, feature extraction, and model training to achieve high accuracy in identifying malaria-infected cells.
Data Collection Plan	The data for this project will be collected from publicly available datasets. Specifically, the Kaggle dataset "Cell Images for Detecting Malaria" will be used as the primary source of blood smear images. The dataset contains images of parasitized and uninfected cells that are labeled for supervised learning.

Raw Data Sources Identified	The raw data is obtained from the Kaggle dataset: Cell Images for Detecting Malaria . This dataset contains around 27,000 images (13,799 parasitized and 13,812 uninfected). The images are labeled, providing a binary classification task. Each image is in PNG format, and the dataset is well-suited for training convolutional neural networks (CNNs) to detect malaria-infected cells.
-----------------------------	---

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Dataset 1	The dataset contains blood smear images used for detecting malaria. It includes both parasitized and uninfected cell images, totaling around 27,000 labeled images.	https://www.kaggle.com/iarunava/cell-images-for-detecting-malaria	Image	708 MB	Public