

Data Collection and Preprocessing Phase

Date	9 July 2024
Team ID	SWTID1720043892
Project Title	WCE Curated Colon Disease Using Deep Learning
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	The project aims to develop a deep learning model for curated colon disease classification from medical imaging data to assist in early detection, treatment planning, and improving patient outcomes.
Data Collection Plan	<ol style="list-style-type: none"> 1. Obtain access to WCE images or videos from popular open-source sites such as Kaggle, GitHub, and UCI repositories. 2. Look for datasets specifically related to colon diseases detected through WCE. 3. Prioritize datasets with diverse patient demographics (age, ethnicity, etc.).

Raw Data Sources Identified	<p>1.WCE Curated Colon Disease Dataset on Kaggle:</p> <p>2.Multiclass Classification of Gastrointestinal Colorectal Cancer:</p> <p>3.GitHub Repository for WCE Curated Colon Disease Classification</p>
--------------------------------	--

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
WCE Curated Colon Disease Dataset on Kaggle	A curated dataset specifically for colon disease classification using WCE images.	Link of Dataset 1	jpg files	2 GB	Public