

Data Collection and Preprocessing Phase

Date	15 June 2025
Team ID	SWTID1749820017
Project Title	Dog Breed Identification using Transfer Learning
Maximum Marks	2 Marks

Data Collection Plan

Section	Description
Project Overview	The Dog Breed Classifier project aims to develop a deep learning-based system that can identify the breed of a dog from an image. The system will assist pet owners, vets, and animal shelters in quickly and accurately determining a dog's breed to enable better care and classification.
Data Collection Plan	Data will be sourced from publicly available and verified datasets that contain labeled images of various dog breeds. The selected sources include benchmark datasets used in academic and professional machine learning tasks.
Raw Data Sources Identified	Kaggle Dog Breed Identification Dataset: A high-resolution image dataset with labeled breed classes, useful for training and testing deep learning models.

Raw Data Sources

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dog Breed Identification/Stanford dog dataset	Contains 20,000+ labeled dog breed images used in Kaggle competitions.	https://www.kaggle.com/competitions/dog-breed-identification/dataset?select=train or http://vision.stanford.edu/aditya86/ImageNetDogs/	Image	750.43 MB	Public (Kaggle account)