

## Data Collection and Preprocessing Phase

<b>Data Source</b>	<b>Data Quality Issue</b>	<b>Severity</b>	<b>Resolution Plan</b>
Dataset	Image Variation	High	<p>Collect images from diverse sources (different cameras, lighting conditions, angles).</p> <p>Implement data augmentation techniques (rotation, scaling, cropping) during preprocessing.</p>
Dataset	Occlusion	Moderate	<p>Include images with partial occlusion, and/or train the model to be robust to it.</p>
Dataset	Insufficient Resolution	Moderate	<p>Establish a minimum resolution threshold for images. Use superresolution techniques, if feasible, to enhance the resolution of some images.</p>

Date	28 January 2026
Student Name	Mubina Iqbal Mulla
Project Title	Uncovering The Hidden Treasures Of The Mushroom Kingdom: A Classification Analysis
Maximum Marks	2 Marks

**Data Quality Report Template**

Dataset	Unbalanced Classes	High	Employ stratified sampling to ensure proportional representation of each mushroom species. Use data augmentation for minority classes. Explore the use of weight loss functions during training.
---------	--------------------	------	--