

Configuration Manual

MSc Research Project
MSc in Artificial Intelligence

Bintong Chen
Student ID:23135808

School of Computing
National College of Ireland

Supervisor: Abdul Shahid

National College of Ireland
Project Submission Sheet
School of Computing



Student Name:	Bintong Chen
Student ID:	23135808
Programme:	MSc in Artificial Intelligence
Year:	2025
Module:	MSc Research Project
Supervisor:	Abdul Shahid
Submission Due Date:	11/08/2025
Project Title:	Beyond Text: How Images Compensate for Missing Quantitative Data in Multimodal Nutritional Estimation
Word Count:	602
Page Count:	3

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Bintong Chen
Date:	12th September 2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	<input type="checkbox"/>
Attach a Moodle submission receipt of the online project submission , to each project (including multiple copies).	<input type="checkbox"/>
You must ensure that you retain a HARD COPY of the project , both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	<input type="checkbox"/>

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Bintong Chen
23135808

1 Introduction

This document outlines the configuration details and step-by-step instructions for replicating the implementation of the project “Beyond Text: How Images Compensate for Missing Quantitative Data in Multimodal Nutritional Estimation”.

The workflow includes:

- (1) Data preprocessing
- (2) Model training/fine-tuning and evaluation using a text-based pipeline
- (3) Model training/fine-tuning and evaluation using an image-based pipeline
- (4) Model training/fine-tuning and evaluation using a multimodal pipeline

To run the project, the minimum hardware requirement is a computer with Python, Jupyter Notebook, and pip (Python package installer) installed.

2 Project Structure

The **nutrition_estimator_project** folder contains two subfolders and four Jupyter Notebook files. The **raw_data** folder includes raw data from the Nutrition5K project, such as dish metadata (CSV files) and food images (PNG files). The **processed_data** folder contains the cleaned and transformed data generated during preprocessing.

There are four main Jupyter Notebooks:

- (1) **data_preprocessing.ipynb**: for preparing the dataset
- (2) **estimation_with_text_input.ipynb**: for running the text-based pipeline
- (3) **estimation_with_images.ipynb**: for running the image-based pipeline
- (4) **modular_multimodal.ipynb**: for running the multimodal pipeline

To replicate the experiment, run the notebooks in the order listed above.

The project also includes a **requirements.txt** file, which lists all necessary Python libraries and dependencies.

Figure 1 illustrates the research architecture, showing detailed processing stages.

3 Instructions to Reproduce the Results

3.1 Download and Replace Raw Data

Before proceeding, you must download the raw dataset required for preprocessing and model training.

- (1) Download the dataset from the following link:

https://drive.google.com/file/d/1zHtD3L7wq-oViWDrso1sV-XxZumewCED/view?usp=drive_link

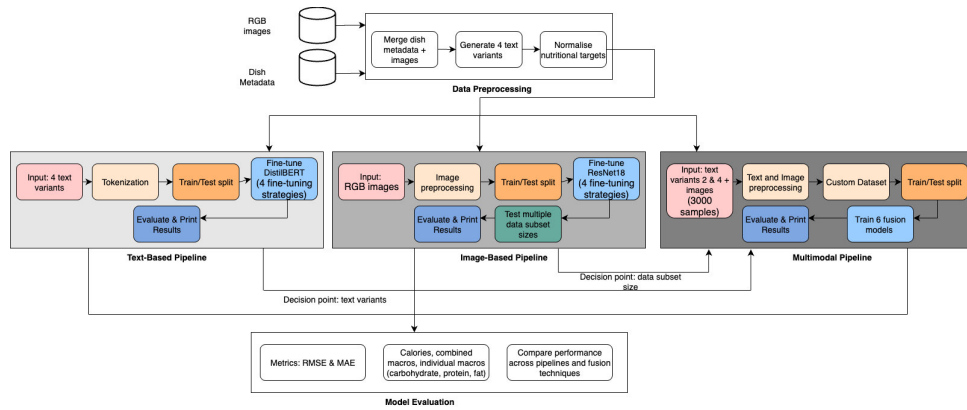


Figure 1: Research architecture showing the three parallel pipelines (text-based, image-based, and multimodal) with their respective processing stages.

(2) Replace the existing (empty) `raw_data` folder in the project directory with the downloaded one.

This folder contains essential image and metadata files from the Nutrition5K dataset. Skipping this step will cause file not found errors in all notebooks.

3.2 Environment Setup

To install the required dependencies, run the following command in the terminal to install all dependencies.

```
pip install -r requirements.txt
```

3.3 Data Preprocessing

Run the `data_preprocessing.ipynb` notebook to prepare the dataset. This notebook performs the following tasks:

- (1) Loads raw metadata and image paths from the `raw_data/` folder
 - (2) Generates four text variants for each dish (with/without mass data, and different ingredient orders)
 - (3) Scales and normalizes the target nutrition values (calories, fat, carbs, protein)
 - (4) Saves the processed data and scaler files in the `processed_data/` folder
- This step must be completed before running any of the model training notebooks.

3.4 Text-Based Pipeline

To reproduce the results of the text-based pipeline, run the `estimation_with_text_input.ipynb` notebook.

This notebook performs the following:

- (1) Loads the processed dataset with selected text variant
- (2) Fine-tunes a DistilBERT model to predict nutrition values
- (3) Evaluates performance across different text formats (with/without mass data)
- (4) Tests four fine-tuning strategies to analyze their effect on model performance
- (5) Reports RMSE and MAE for calories and macronutrients

3.5 Image-Based Pipeline

To reproduce the results of the image-based pipeline, run the `estimation_with_images.ipynb` notebook.

This notebook performs the following:

- (1) Loads RGB dish images and their corresponding nutrition labels
- (2) Preprocesses images (resizing, normalization)
- (3) Fine-tunes a ResNet18 model for regression
- (4) Tests different training subset sizes to explore the effect of data volume
- (5) Evaluates model performance with various fine-tuning configurations
- (6) Reports MAE/RMSE for each nutritional target

3.6 Multimodal Pipeline

To reproduce the results of the image-based pipeline, run the `modular_multimodal.ipynb` notebook.

This notebook performs the following:

- (1) Loads both image and text features from the prepared dataset
- (2) Implements multiple fusion techniques, including: Early fusion, Late fusion, Attention-based fusion (global weights and cross-modal), and Gated fusion (scalar and per-feature weights)
- (3) Evaluates multimodal model performance
- (4) Reports MAE/RMSE for each nutritional target