SCHOOL OF COMPUTING AND INFORMATION TECHNOLOGY

UNIVERSITY OF WOLLONGONG AUSTRALIA

INDIVIDUAL Assignment Coversheet

*This form is to be completed by students submitting **online copies** of essays or assignments for a Faculty of the Arts, Social Sciences and Humanities subject for the School of Geography and Sustainable Communities, School of Education, and School of Health and Society.*

PLAGIARISM

Deliberate plagiarism may lead to failure in the subject. Plagiarism is cheating by using the written ideas or submitted work of someone else. The University of Wollongong has a strong policy against plagiarism. See Acknowledgement Practice/Plagiarism Prevention Policy at **http://www.uow.edu.au/about/policy/UOW058648.html**

Student Name: Jeslyn Ho Ka Yan________________________ **7-digit UOW ID:** 8535383______________

Subject Code & Name: CSCI218 & Foundation of Artficial Intelligence____________________________

Assignment Title: INDIVIDUAL ASSIGNMENT____

Tutorial Group: T02____

(T02, T03, T04, T05)

Tutor's Name: Cher Lim____

Assignment Due Date: 9TH FEB 2025__________________________

DECLARATION

I certify that this is entirely my own work, except where we have given fully documented references to the work of others, and that the material contained in this assignment has not previously been submitted for assessment in any formal course of study. I understand the definition and consequences of plagiarism.

ACKNOWLEDGEMENT

The marker of this assignment may, for the purpose of assessing this assignment, reproduce this assignment and provide a copy to another member of academic staff. If required to do so, we will provide an electronic copy of this assignment to the marker and acknowledge that the assessor of this assignment may, for the purpose of assessing this assignment:
a) Reproduce this assignment and provide a copy to another member of academic staff; and/or
b) Communicate a copy of this assignment to a plagiarism checking service such as Turnitin (which may then retain a copy of this assignment on its database for the purpose of future plagiarism checking).

Student Signature: ____ **Date:** 3 Feb 2025_______

[Insert e-signature or type name]

Introduction

(**A short introduction of the assignment and the dataset**)

Automated flower classification is widely used in multiple fields, such as floriculture, botany, and online commerce. This assignment explores various machine learning techniques to categorize images of flowers into five distinct types: daisy, tulip, rose, sunflower, and dandelion. The dataset utilized in this study is sourced from **Kaggle**, containing over 4,000 images with different resolutions.

This report focuses on three primary classification approaches:

1. **k-Nearest Neighbors (k-NN):** A simple classification algorithm that determines the class of an image based on the labels of its nearest neighbors.

2. **Multi-Layer Perceptron (MLP):** A neural network model with multiple layers, which captures complex patterns in the dataset.

3. **Convolutional Neural Network (CNN):** A deep learning-based image classifier designed for pattern recognition in images.

Each model undergoes training, validation, and testing, with the dataset split into **60% training, 20% validation, and 20% testing**. Various evaluation metrics, such as accuracy, precision, recall, and F1-score, are analyzed alongside confusion matrices to identify challenging classifications. Additionally, this study explores how feature extraction techniques, model architectures, and hyperparameter tuning impact classification performance. Lastly, training time and inference speed are considered to evaluate the computational efficiency of each approach.

Task 1: k-Nearest-Neighbour Classifier

(**Label CLEARLY your answer to each question**. Any unlabeled answers will **NOT** be graded)

Answers:

1.1 Color Histogram Sizes

To enhance feature extraction, color histograms were computed for the images using different bin sizes:

- **4 bins ([4,4,4]):** Coarse-level features with reduced detail.
- **6 bins ([6,6,6]):** The default configuration with a balanced feature representation.
- **8 bins ([8,8,8]):** A finer-level representation capturing more details.

Each image was resized to **150×150 pixels**, and the histograms were normalized.

Results:

Histogram Size	Validation Accuracy (Best k)	Precision	Recall	F1 Score
4 bins	`k=9`, **0.4844**	0.4781	0.4676	0.4637
6 bins (default)	`k=11`, **0.4762**	0.5008	0.4815	0.4815
8 bins	`k=15`, **0.4403**	0.4549	0.4725	0.4549

Observations:

- Smaller histograms (4 bins) **simplify computations** but slightly reduce accuracy due to feature loss.
- The **default size (6 bins)** strikes a balance between accuracy and computational efficiency.
- Larger histograms (8 bins) capture more **intricate details** but **increase computational cost**.

1.2 Finding the Optimal K

To determine the **optimal value of k**, multiple odd values from 1 to 15 were tested using the validation set.

Best K Values:

K	Validation Accuracy (4 bins)	Validation Accuracy (6 bins)	Validation Accuracy (8 bins)
1	0.4241	0.3975	0.4021
3	0.4461	0.4311	0.4090
5	0.4762	0.4426	0.4276
7	0.4751	0.4380	0.4322
9	0.4844	0.4496	0.4287
11	0.4774	0.4762	0.4311
13	0.4716	0.4519	0.4334
15	0.4809	0.4739	0.4403

Key Insights:

- **k = 11** yielded the best validation accuracy (**0.4762**) for histogram size **6**.
- For larger histogram sizes (8 bins), a higher **k (15)** provided marginal improvement.

1.3 Classification Metrics and Confusion

Task: Report classification metrics and the most confusing classes**.**

Results on Test Set Performance (K=optimal)

Histogram Size	Test Accuracy	Precision	Recall	F1 Score
4 bins	0.4676	0.4781	0.4676	0.4637
6 bins	0.4815	0.5008	0.4815	0.4815
8 bins	0.4549	0.4725	0.4549	0.4549

Observation:

- The **confusion matrix** showed that **dandelion and daisy** were the most confused classes.
- Possible reasons:
  - **Similar colour features** (yellow petals).
  - **Background clutter** in dataset images.

1.4 Average Inference Time

We measured the time taken to classify a single sample across 10 runs.

Results:

Histogram Size	Average Inference Time (seconds)
4 bins	0.0018
6 bins	0.0032
8 bins	0.0059

- Smaller histograms (4 bins) had faster inference due to reduced feature dimensionality.
- Larger histograms (8 bins) were slower due to increased computational complexity.

1.5 Correctly and Incorrectly Classified Images

Visualization:

Correctly Classified Images:

Incorrectly Classified Images:

Observation:

- Correct classifications were achieved for flowers with distinct color patterns (e.g., sunflowers).
- Misclassifications occurred for flowers with overlapping features (e.g., daisies and tulips).

Task 2: Multi-layer Perceptrons

(**Label CLEARLY your answer to each question**. Any unlabeled answers will **NOT** be graded)

Answers:

2.1 Design of MLP Structures

Nine MLP structures were designed based on the rules provided:

MLP Structure	Number of Hidden Layers	Number of Neurons in Each Hidden Layer
1	1	[149]
2	1	[108]
3	1	[216]
4	2	[149, 108]
5	2	[216, 5]
6	2	[108, 216]
7	3	[149, 108, 5]
8	3	[216, 108, 5]
9	3	[216, 216, 5]

The design ensures a balance between model complexity and the dataset's size by using empirical rules for determining the number of neurons in each layer.

2.2 Selection of Optimal Network Architecture

Process

1. Each structure was trained for **10 epochs** on the training set.
2. The validation accuracy was used to identify the best-performing architecture.
3. Structures with higher validation accuracy were considered optimal.

Quantitative Results

MLP Structure	Validation Accuracy
1	58.98%
2	57.59%
3	59.79%
4	61.30%
5	57.01%
6	61.41%
7	60.02%
8	61.88% (Best)
9	55.04%

Best Network Architecture:

The best validation accuracy (**61.88%**) was achieved by **MLP Structure 8**, which has three hidden layers [216, 108, 5].

2.3 Evaluation of the MLP Classifier

Test Set Performance:

The selected structure ([216, 108, 5]) was evaluated on the test set:

- **Accuracy:** 56.94%
- **Precision:** 58.43%
- **Recall:** 56.94%
- **F1 Score:** 56.07%

Confusion Matrix

The confusion matrix highlights the performance of the classifier for each class:

Most Confused Classes: The MLP classifier most frequently confused **dandelions** with **daisies**, indicating a challenge in distinguishing these classes.

2.4 Training and Inference Time

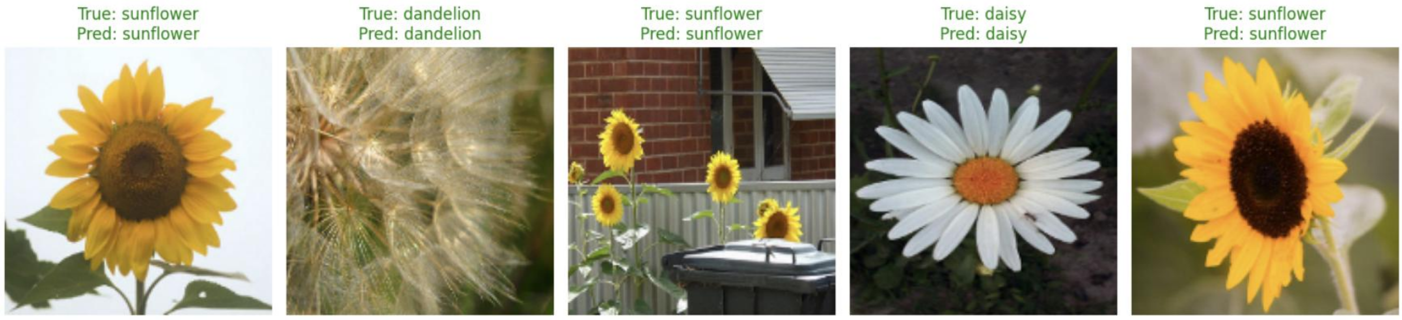
- **Training Time**: Approximately 6.2 seconds (10 epochs for the best model).
- **Average Inference Time**: 0.0002 seconds per image.

The model is efficient for both training and prediction, suitable for moderate-scale datasets.

2.5 Visualization of Predictions

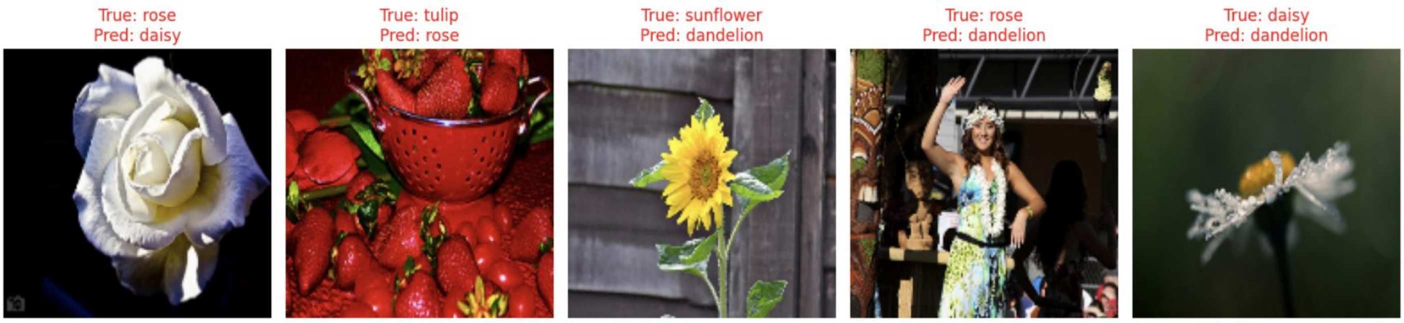
Correctly Classified Images

Below are examples of correctly classified images:



Incorrectly Classified Images

Below are examples of incorrectly classified images:



Task 3: Convolutional Neural Network

(**Label CLEARLY your answer to each question**. Any unlabeled answers will **NOT** be graded)

Answers:

3.1 Preprocessing and Building CNN Model

The **CNN model** implemented in this study consists of several layers to efficiently process image data:

- **Convolutional Layers:**
  - The first **Conv2D** layer applies **32 filters** with a kernel size of **3×3**.
  - The second **Conv2D** layer applies **64 filters** with a kernel size of **3×3**.
- **Pooling Layers:**
  - **MaxPooling2D layers** follow each convolutional layer to reduce spatial dimensions.
- **Flatten Layer:**
  - Converts extracted feature maps into a single-dimensional vector.
- **Fully Connected Layers:**
  - A dense layer with **128 neurons** and **ReLU activation**.
  - A **Dropout layer (0.5 dropout rate)** to prevent overfitting.
  - The final output layer consists of **5 neurons** (corresponding to the five flower categories) with a **softmax activation function**.

Dataset Processing:

- **Image Preprocessing:**
  - Each image was resized to **150×150 pixels**.
  - Pixel values were **normalized** between 0 and 1.
- **Data Splitting:**
  - **Training Set:** 60%
  - **Validation Set:** 20%
  - **Test Set:** 20%

3.2 Training CNN Model

The CNN model was trained using 10 epochs with the Adam optimizer and a batch size of 32.

Training Results

Epoch	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	32.43%	1.6115	42.76%	1.2853
5	69.61%	0.7752	64.08%	0.9681
10	92.90%	0.2111	65.59%	1.3742

Classification Metrics on Test Set

The CNN classifier was evaluated on the test set, and the following metrics were obtained:

- **Accuracy:** 60.53%
- **Precision:** 61.41%
- **Recall:** 60.53%
- **F1 Score:** 60.64%

Common Classification Errors:

- The confusion matrix analysis revealed that the model frequently misclassified tulips as roses due to their similar petal shapes and colors.

3.3Training and Validation Metrics

Visualization

1. **Confusion Matrix**
    - The confusion matrix highlights the performance of the model across the five flower classes.
    - **Significant Misclassifications:** Tulip and Rose are the most confused classes.
2. **Training and Validation Metrics**
    - **Loss vs. Epochs:**
        - Training loss steadily decreased across epochs.
        - Validation loss increased after Epoch 5, indicating overfitting.
    - **Accuracy vs. Epochs:**
        - Training accuracy improved significantly over epochs.
        - Validation accuracy plateaued after Epoch 5.

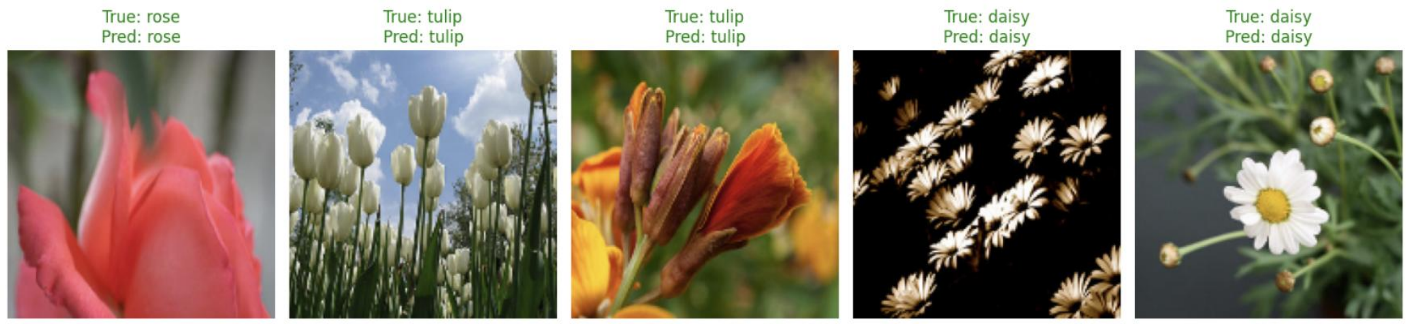
3.4 Training and Inference Time

- **Training Time:** 1,370.69 seconds
- **Average Inference Time per Sample:** 0.0131 seconds

3.5 Visualization of Predictions

Correctly Classified Images

- Five correctly classified images were visualized, showcasing accurate predictions across various classes.
- **Observation:** Correct classifications corresponded to flowers with distinct features (e.g., Sunflower and Dandelion).



Incorrectly Classified Images

- Five incorrectly classified images were visualized, highlighting the model's challenges with visually similar classes (e.g., Tulip vs. Rose).
- **Observation:** Misclassifications often involved subtle differences in petal shapes and colors.