Report

- Submit a brief report (2-4 pages) summarizing:
    - Evaluation Reports (F1-score, Accuracy, Loss, etc.)
    - Challenges faced.
    - Why is a CNN better at classifying images than a standard neural network?

This report evaluates the performance of a basic dense neural network versus a convolutional neural network (CNN) for classifying images of flowers. The dataset includes 4,000 labeled images divided equally across five flower types. The dense model served as a baseline, while the CNN, designed with hierarchical convolutional layers and enhanced using data augmentation, significantly improved classification accuracy.

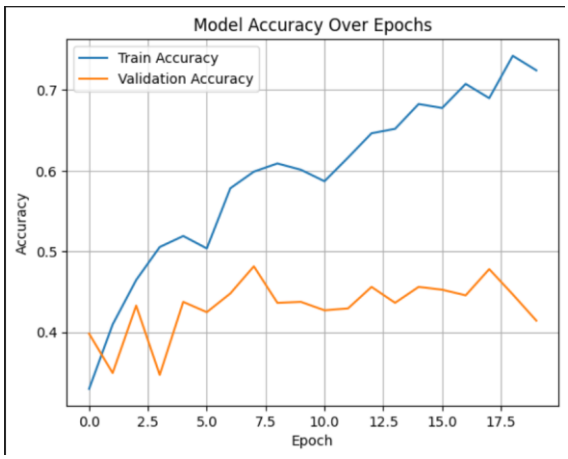
PERFORMANCE EVALUATION

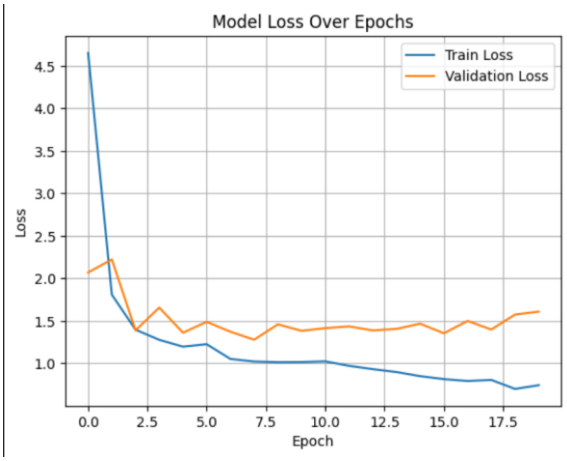
The convolutional neural network performs better on all standards as compared to the foundational regression model which was run for 20 epoches and gave the following results:

class	precision	recall	F1-score
Daisy	0.354	0.490	0.411
Dandelion	0.469	0.181	0.260
Rose	0.336	0.662	0.446
Sunflower	0.664	0.524	0.586
Tulip	0.435	0.325	0.372
Macro Avg	0.452	0.436	0.415

The macro-average F1-score was **0.415**, suggesting moderate overall classification effectiveness across the five flower categories. Some

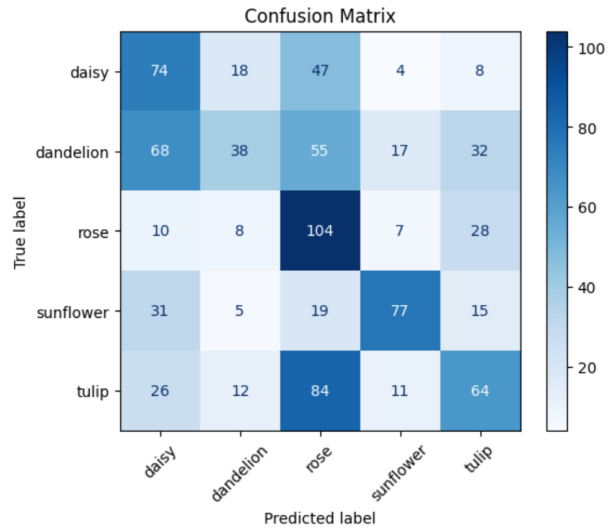
classes showed however a weaker performance showing possibilities of class confusion





As assumed the accuracy curve follows an upward learning graph while the loss decreases significantly, this model gave unexpectedly good results with the training accuracy and validation accuracy varying quite close to 0.74-0.75 and 0.42-0.44 at the end of the complete training. Usually baseline models for object classifications give a reasonably low accuracy on test/validation data about 0.25. The loss on the training data and validation data was quite high about 0.7 and 1.60.

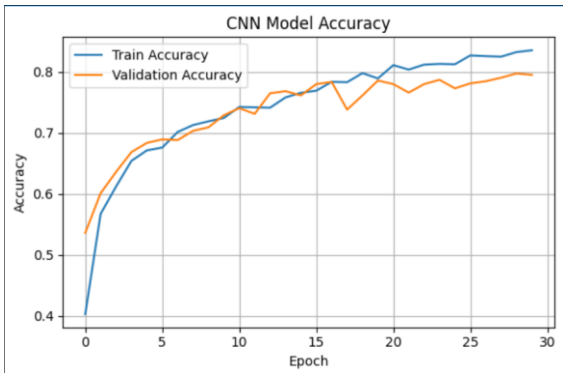
Based on the following confusion matrix:

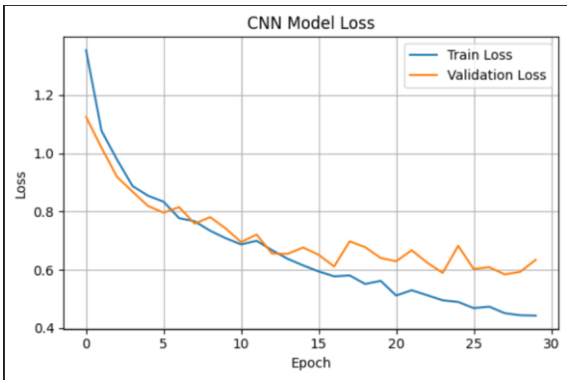


The true labels as supplied by the dataset are along the y axis and the predictions made by the model along the x axis the higher values along the main diagonal shows the number of correct predictions.

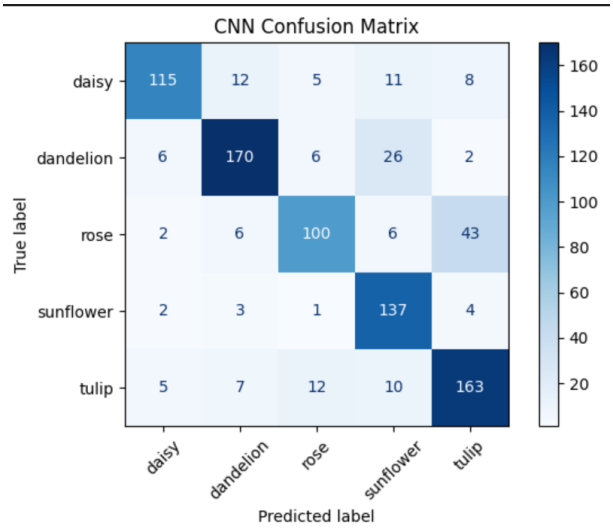
class	precision	recall	F1-score
Daisy	0.884	0.733	0.802
Dandelion	0.850	0.836	0.843
Rose	0.738	0.575	0.646
Sunflower	0.656	0.922	0.766
Tulip	0.705	0.798	0.749
Macro Avg	0.767	0.773	0.761

The **macro-averaged F1-score of 0.761** indicates a well-balanced model across all classes. Compared to the earlier dense neural network, the CNN architecture shows significant improvement in class-wise consistency





this model gave unexpectedly good results with the training accuracy and validation accuracy varying quite close to 0.82-0.83 and 0.79 at the end of the complete training of 30 epoches. Usually convolutional neural network for object classifications give a higher than expected accuracy on test/validation data which would be at about 0.65-0.70.



Challenges Faced:

- Setting up the dataset structure and linking it properly in Colab
- Managing memory and resizing images during preprocessing
- Poor normalization initially caused the CNN to perform worse than the dense model (accuracy: 0.224)
- Misclassification due to overlapping visual features — in one case, all predictions were mapped to “dandelion”
- Slower training rates for CNNs, especially on CPU
- Saving/loading models (.h5) for later evaluation
- Interpreting accuracy/loss plots and identifying signs of overfitting

WHY CNNS?

Convolutional neural network’s identify structures in an image by identifying individual features, convolution itself helps derive/convert available information to a more useful state convolving the image pixels with a kernel helps in edge detection may it be horizontal vertical or circular edges these done independently are framed together helping to identify shape, thus helps a convolutional neural network see an object as a cluster of features. A dense model or a

baseline neural network works on idea of regression it uses weights and parameters trying to readjust the weights so as to match the accompanying label with the dataset, it tries to guess what might be happening with the image itself, so a baseline model thinks like this flower has a bunch of white pixels then it must be say daisy while it might be a white rose on the other hand the convolutional network recognises the shape of the petals of the flower based on the input dataset if there's examples of a white rose it would prefer that over a daisy considering not only the colour but also the structure/features.