DEEP LEARNING FOR IMAGE PROCESSING: CLASSIFYING FLOWERS

OBJECTIVE

- To classify 104 species of flowers from a Kaggle data set with over 7K flower images that are split up on test, train, and validation sets using deep learning techniques.
- The objective is to reach a F1 score as close to 1 as possible (in the range of 0.5-1)

RESULTS

Model 2 using inception reached an F1 score of 0.92

And had an accuracy of 91%

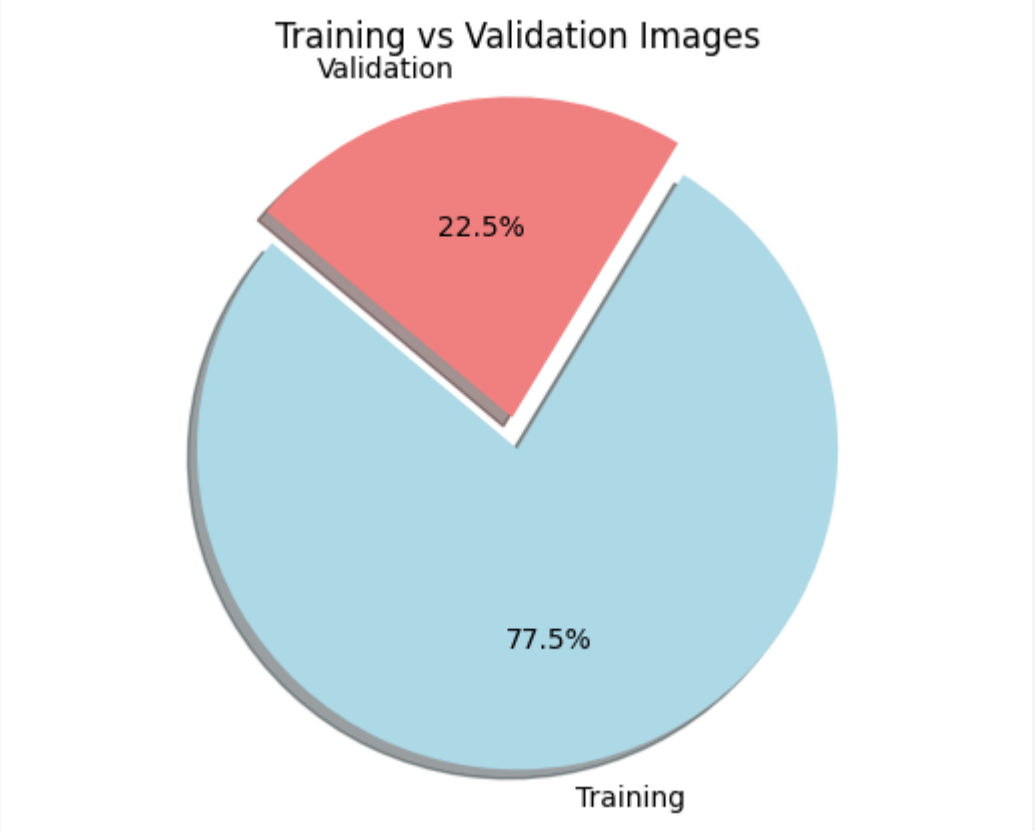
Model 2 was the best Model

EDA

TRAINING VS. VALIDATION IMAGES

- Training data :77.5%
- Validation data: 22.5%

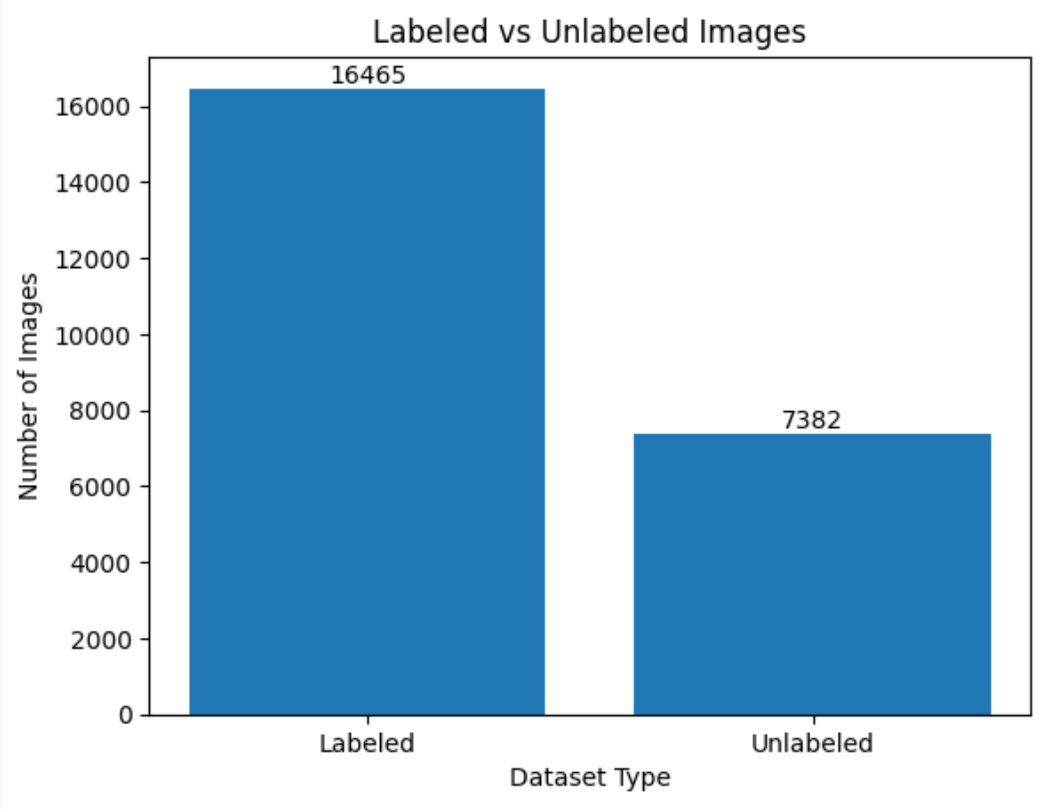
Most of the data was training data.



LABELED VS. UNLABLED IMAGES

- Labeled images: 16,465
- Unlabeled images: 7,382

The majority of the data was labeled.



GRAY SCALE COLOR DISTRIBUTION HISTOGRAMS :4 RANDOM IMAGES

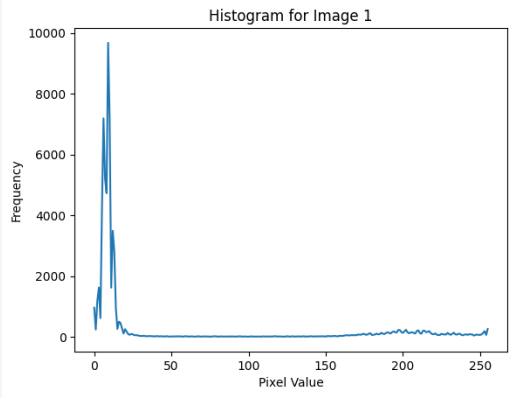
1. Convert images to gray scale

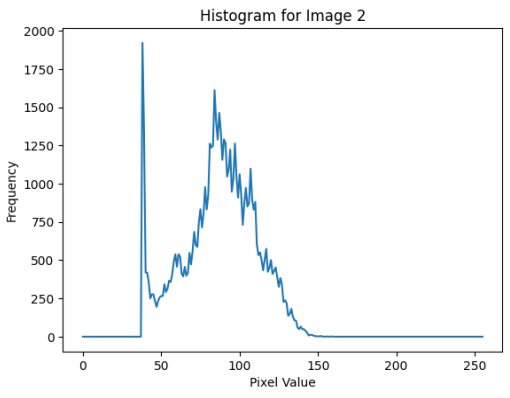
2. Get histogram of color distribution

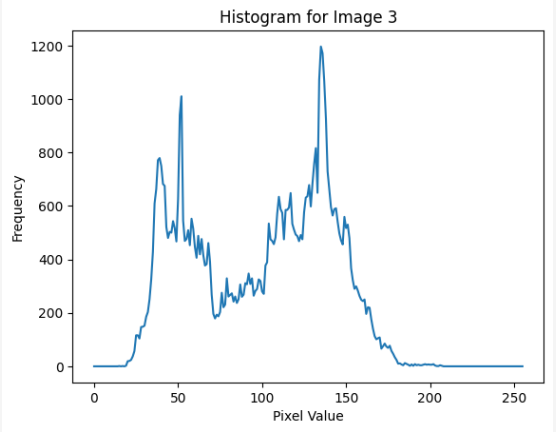
   0: Black
   255: White

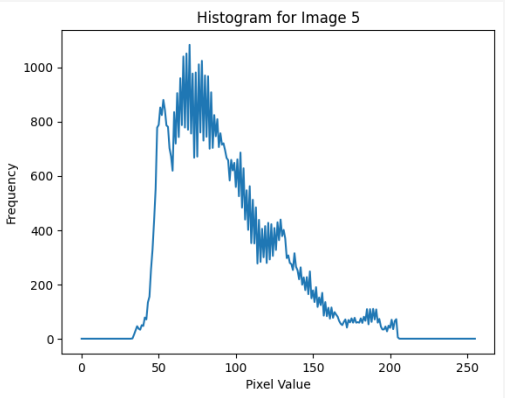
3. View if images had a similar color distribution or if random

Color Distribution was Random



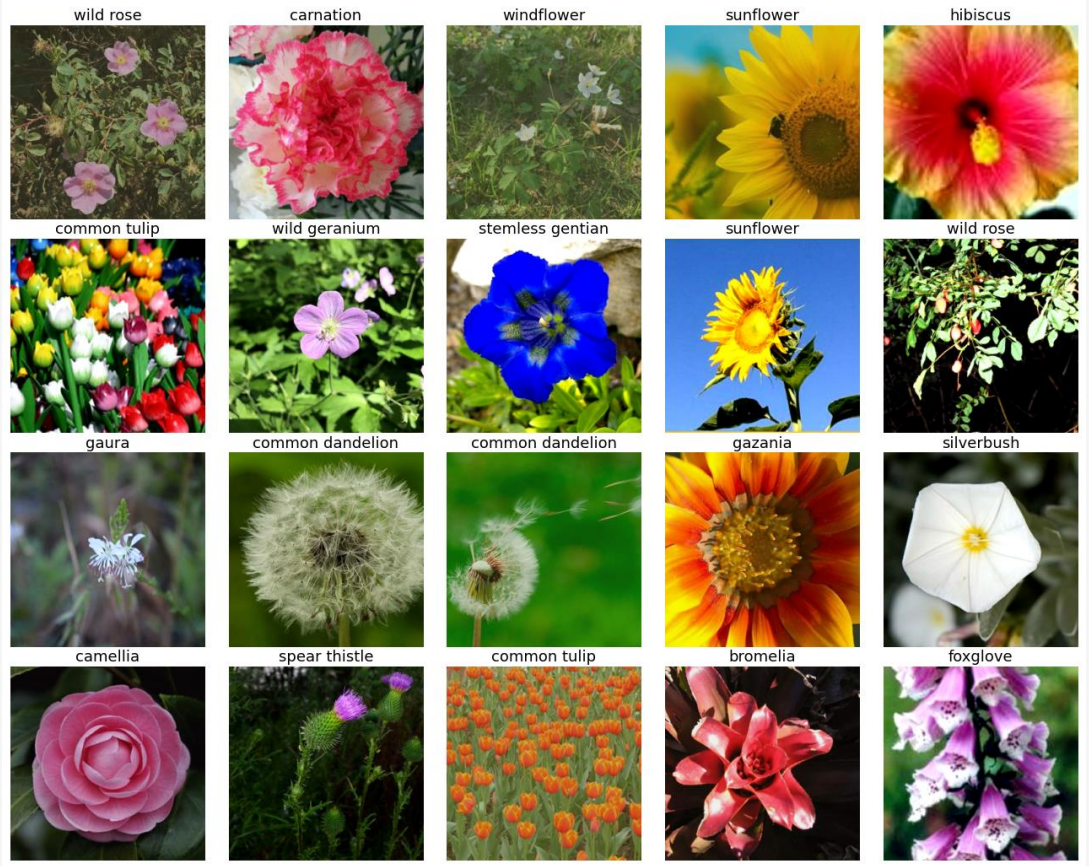






VIEWING DATA

Images are displayed as a grid with the appropriate label on top



AUGMENTING DATA

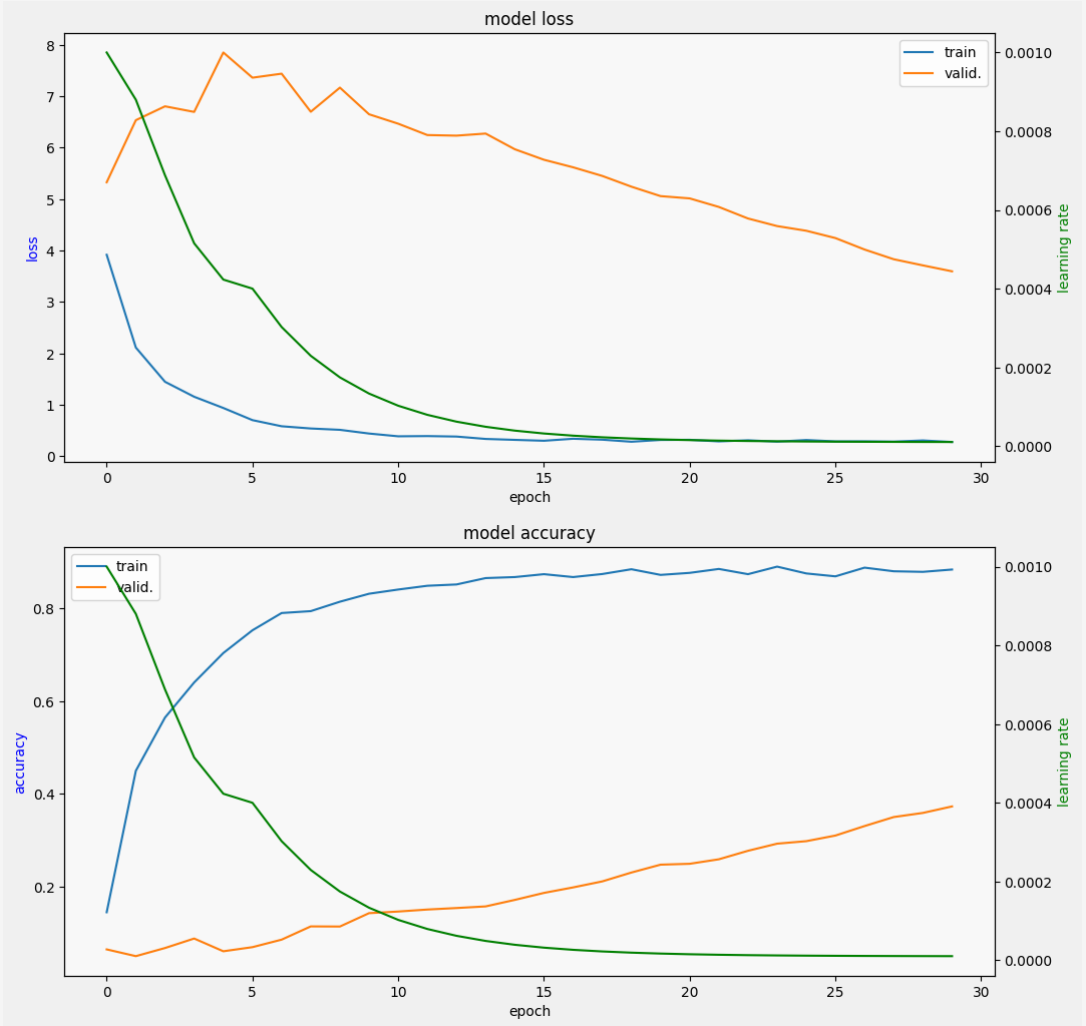
- To increase variety in data i.e. Artificially Increasing data
- Random Left/Right flipping
- Random contrast

MODEL 1

- Sequential ( )
- Mobilenetv2_1.00_224
- Global average pooling
- Yes, Drop out

Model 1 Accuracy

- Starts with LOW accuracy, increases exponentially until it becomes steady.
- Minimum model loss



Results for Model 1

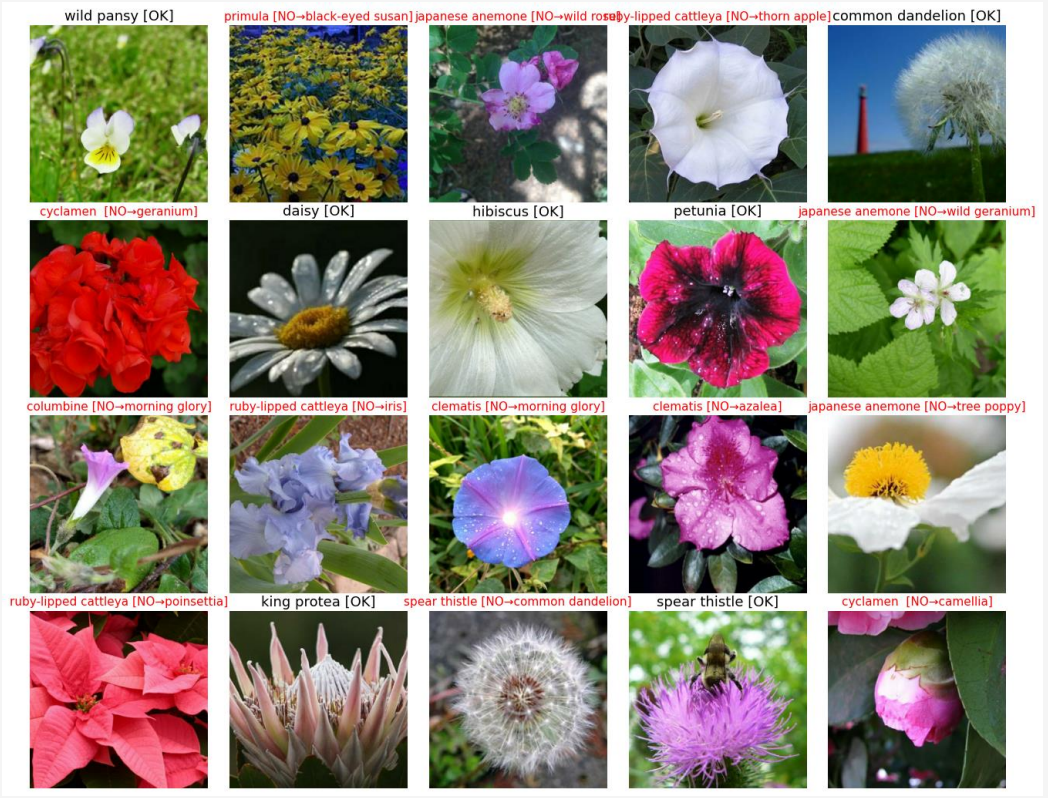
```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 mobilenetv2_1.00_224 (Funct  (None, 16, 16, 1280)     2257984
 ional)

 global_average_pooling2d (G  (None, 1280)             0
 lobalAveragePooling2D)

 dropout (Dropout)           (None, 1280)              0

 dense (Dense)               (None, 104)               133224

=================================================================
Total params: 2,391,208
Trainable params: 2,357,096
Non-trainable params: 34,112
_________________________________________________________________
```



Many obvious errors:
13/25 Images Incorrect

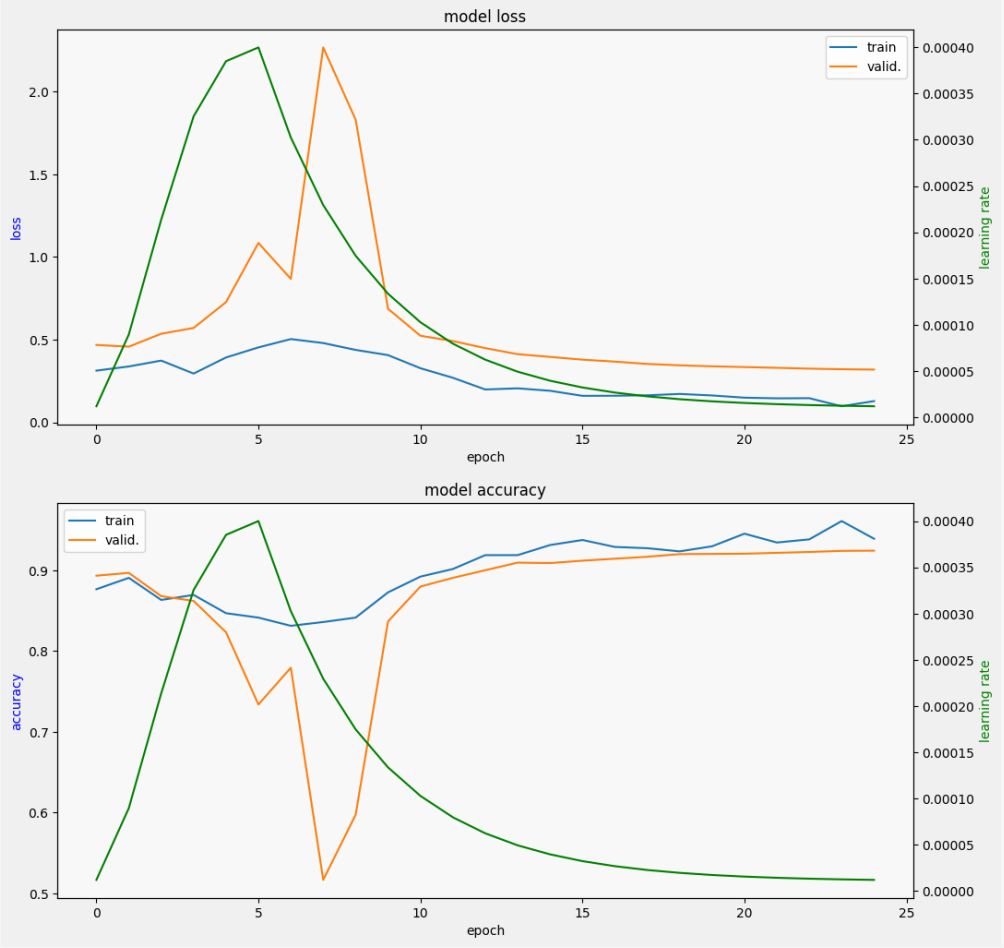
F1:0.40
Prescision:63.8%
Recall: 0.44

MODEL 2

- Sequential ( )
- Inception_resnet_v2
- Global average pooling
- Yes, Drop out

Model 2 Accuracy

- Starts with high accuracy, dips, and then slowly increases until it becomes steady
- High model loss in the beginning



Results for Model 2

Layer (type)	Output Shape	Param #
inception_resnet_v2 (Functional)	(None, 14, 14, 1536)	54336736
global_average_pooling2d_4 (GlobalAveragePooling2D)	(None, 1536)	0
dropout_4 (Dropout)	(None, 1536)	0
dense_4 (Dense)	(None, 104)	159848

Total params: 54,496,584
Trainable params: 54,436,040
Non-trainable params: 60,544

NO obvious errors: 0/25 Images Incorrect

F1:0.92
Prescision:91.4%
Recall: 0.93

CONCLUSION & FUTURE APPLICATIONS

- Best Model: Model 2 (91% accuracy.)
- Field: Flower identification is useful in the field of ecology and agriculture.
- Application: Determining native & invasive species.
- Future improvement:
  - Obtaining Higher accuracy (Try maxpooling instead of Avg)
  - Seed recognition