Weekly Report 2

Comparative Analysis of Image Classification Models for Efficient and Accurate Classification across Diverse Image Types

Student: Yusif Mukhtarov

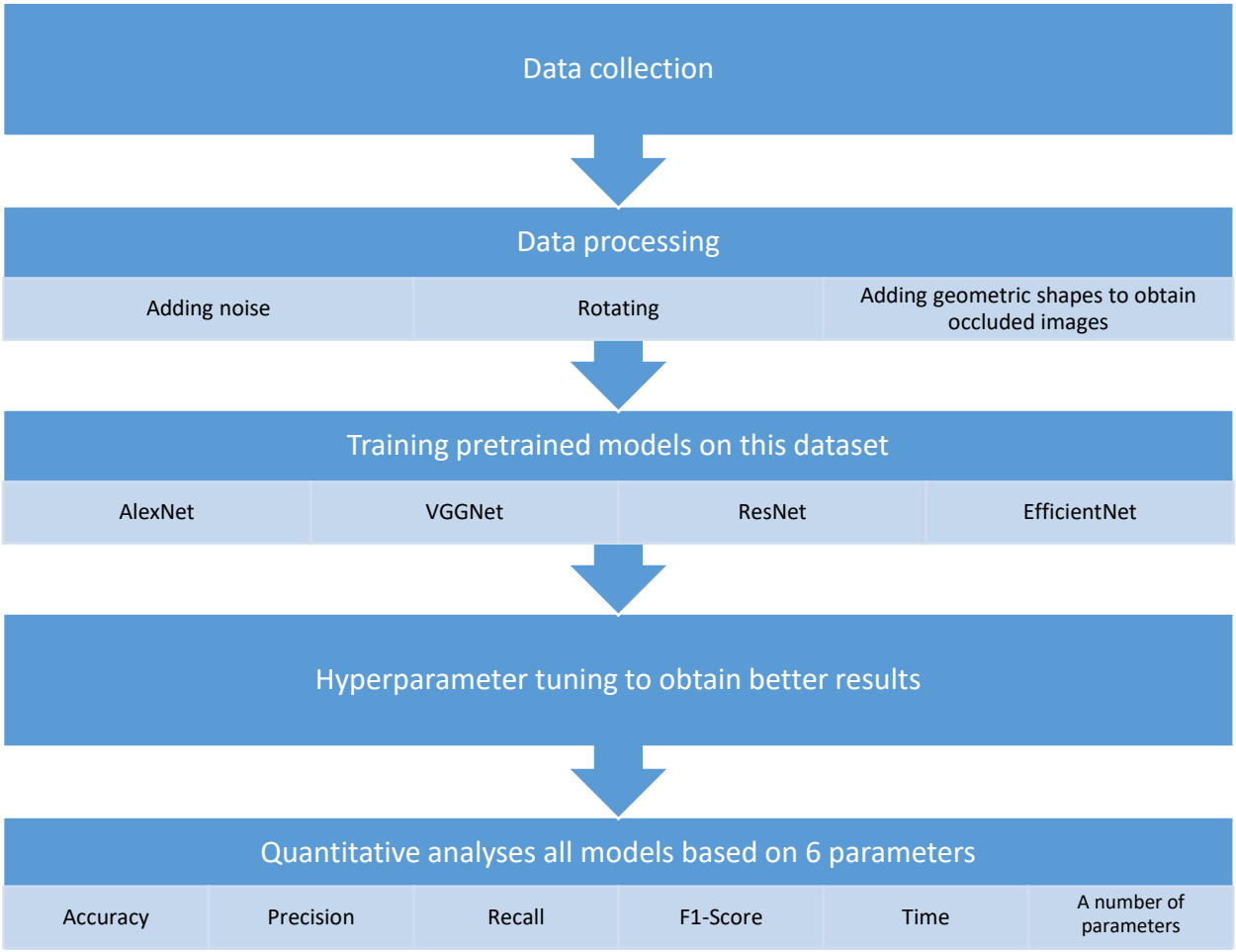
Date: 25/06/2023

Research objective

The fundamental objective of this **quantitative** research is to identify the best image classification models in terms of both time efficiency and accuracy, under different image scenarios such as rotated images, occluded images and noisy images. Several widely recognized models, including AlexNet, VGGNet, ResNet, EfficientNet will be considered for this analysis. **Qualitative** analysis has been made to pick those models since they are considered by experts to be the best.

The dataset chosen for this experiment will be vegetable custom datasets that incorporate the required image conditions. Data preprocessing will be carried out to facilitate the suitable conditions for each model, involving rotating, adding noise to images, and adding geometric figures with different shapes to images to form occluded images.

To achieve an in-depth and unbiased analysis, the study will take into account several performance metrics. Classification accuracy will primarily measure the model's performance. Simultaneously, precision, recall, and F1-Score will also be considered to provide a good understanding of the model's performance. Moreover, the time taken for training and prediction will be recorded to measure the model's time efficiency. Additionally, the model's size in terms of parameters will also be observed. So, our final table with results will include Accuracy, Precision, Recall, F1-Score, Time and number of parameters of each classification models. The plan for this specific research looks like:



Data preprocessing

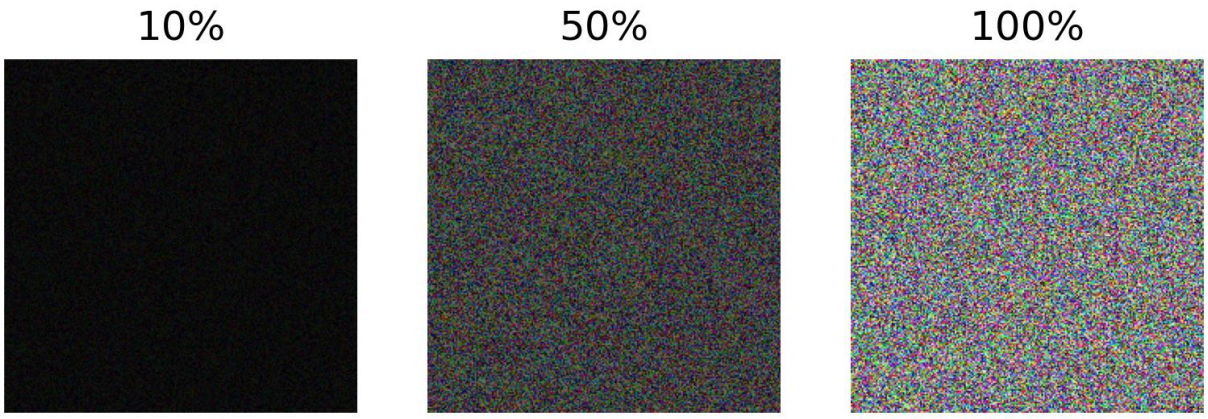
In choosing dataset special attention had to be paid to not use datasets that were used for pretrained models. After careful analyses of several datasets Vegetable Dataset (https://www.kaggle.com/datasets/misrakahmed/vegetable-image-dataset) has been chosen. In this dataset there are 21000 images from 15 classes, where each class contains a total of 1400 images. This dataset is split into three parts, where 70% for training and 15% for testing, and the rest 15% for validation. Although in Kaggle it was said that each class has an equal proportion and image resolution is 224×224 and in *.jpg format, it was detected that there are 9 images with different shapes:

```
 from train dataset:
1 0741.jpgwith shape(210, 224, 3)
2 0176.jpgwith shape(198, 224, 3)
3 0126.jpgwith shape(211, 224, 3)
4 0609.jpgwith shape(200, 224, 3)
5 0430.jpgwith shape(193, 224, 3)
6 0526.jpgwith shape(205, 224, 3)

 from validation dataset:
7 1138.jpgwith shape(187, 224, 3)
8 1150.jpgwith shape(223, 224, 3)

 from test dataset:
9 1246.jpgwith shape(207, 224, 3)
```

Considering that all images will be resized before training and testing it was decided that this anomaly is insignificant. Furthermore, images should be mixed with noise. For that Gaussian noise was opted to be used in 10, 30, 50, 70 and 100% percentages. The picture below shows the difference between different percentages of noise:

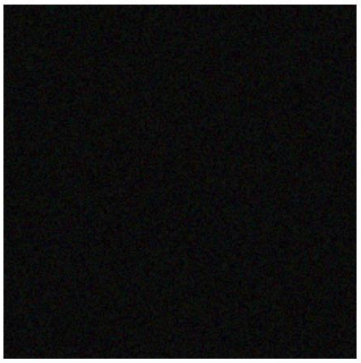


Next pictures will show how the image is corrupted after applying different noise levels:

10%

Original

Gaussian Noise

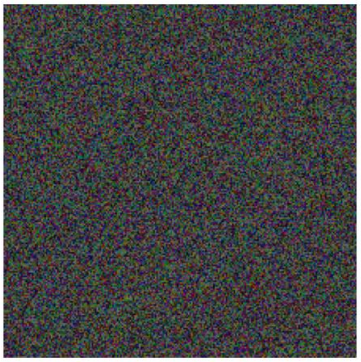


Combined

50%

Original

Gaussian Noise

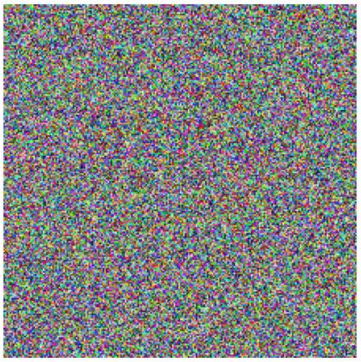


Combined

100%

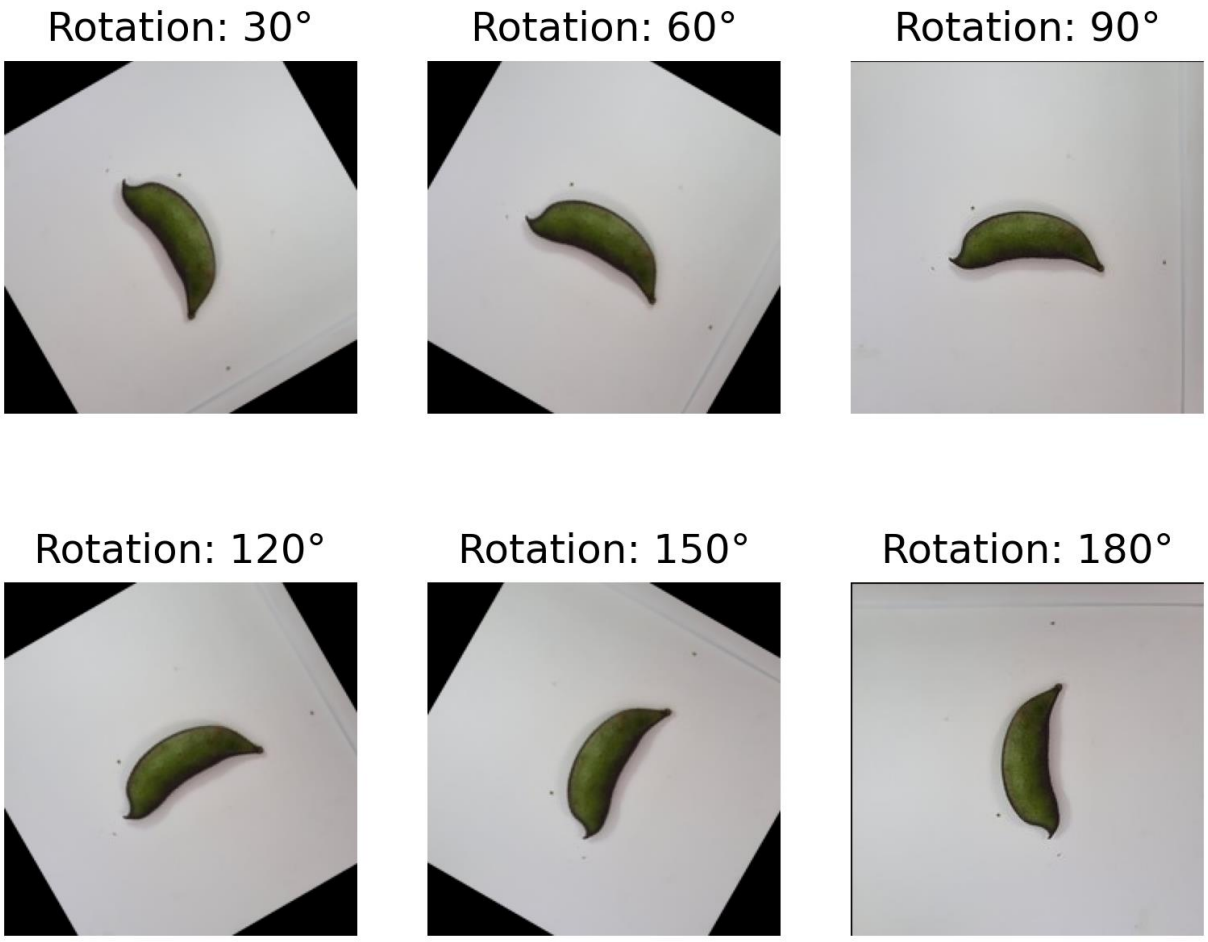
Original

Gaussian Noise



Combined

Moreover, the images were additionally rotated at angles of 30, 60, 90, 120, 150, and 180 degrees without changing its shapes:



Finally, occluded images were gotten by manually adding rectangle to images: