Individual Report

Jinbo Li
Team 04
Professor: Amir Jafari

Abstract

Convolutional Neural Network is one of the common ways applied to analyze visual imagery. The team was planning to build a CNN using TensorFlow, which is suitable for animal classification by finding the best model with the goal of the highest accuracy in testing.

Introduction

Image recognition is a popular technology, which can accurately identify the visual content in the image, and provide a variety of objects, scenes, and concept labels, with the ability of target detection and attribute recognition, to help clients accurately identify and understand the image. The technique has been applied to multiple fields such as text recognition, license plate recognition, and face recognition. The team would explore the topic of computer vision primarily, applying the image classification technique to animal detection. The animal classification could be used in animal management and tag the different animal species, for further application, it could be developed and suitable for animal photo recognition, early childhood education science, and image content analysis, to help people get a better idea of animals and the diversity of nature. The team's objectives are to design a CNN model with 90% accuracy, which will be accomplished by exploring the architecture of the CNN model and comparing the state-of-the-art pre-trained models and the model with the customized architecture. The team is using TensorFlow to implement the network, for the reason that TensorFlow is a powerful and mature deep learning library with strong visualization capabilities, and there are multiple options for advanced model development. With the data augmentation of the training, the data will expand up the training dataset, therefore, the model can be fitted better. The hyperparameters that learning rate, number of epochs, batch size also are determined experimentally to achieve the highest accuracy.

Individual Work

This project is a cooperative work. And I was involved in almost all parts of the collaboration. Like some part of the coding, training model, writing group proposal, writing group report and doing the slides.

Firstly, I was in charge of the understanding of the dataset. I split the original dataset into training dataset, validation dataset, test dataset and did exploratory data analysis of every dataset to check if the dataset is balanced or not. Then I also did a part of data augmentation, like rotating the image.

Then I also did the works of training model. In our group, we divided up that each person did some training models and then compared whose model had the higher scores. Therefore, I customized a model and tried some pretrained models, like the VGG16, EfficientNetB5 and so on to check their performance. And I found my customized model didn't have a very good performance, worse than the pretrained models. And all the pretrained models I tried, I found the VGG16 had a best performance in our dataset.

Architecture of VGG16

There are 13 convolutional layers, 3 fully connected layers, 5 max pooling layers and 1 softmax layer in VGG16. Although summing up to 21 layers in VGG16, it has only 16 weight layers (learnable parameters layers). The input size of VGG16 is 224 by 224 with 3 RGB channels. VGG16 used kernels of 3 * 3 size with a stride size of 1 pixel and max pooling was performed over 2 * 2 pixel windows with stride 2. And it implemented three fully connected layers from which first two were of size 4096 and after that a layer with 1000 channels for 1000-way ILSVRC classification and the final layer is a softmax function.

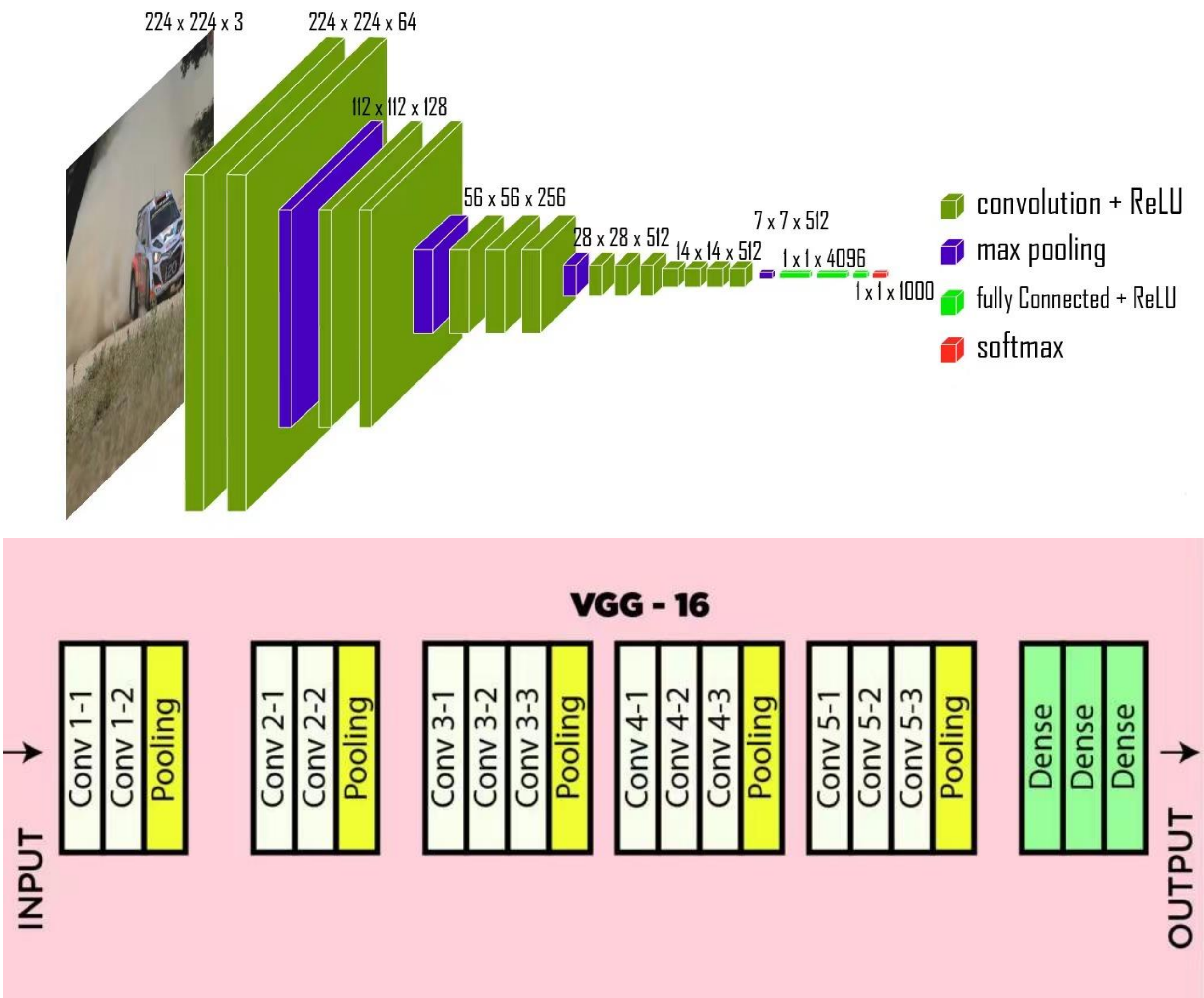

Figure 1. VGG16 model architecture and network

Results

Firstly, I split the dataset into training dataset, validation dataset and testing dataset. Training dataset and validation dataset. Therefore, we were having 10848 of image files in training, 4649 of image files in validation and 1731 of image files in testing.

```
                                                    id  target
0                                cats/cats_00337.jpg    cats
1          spider/OIP-7030ghChVLkW0tCHWK1j-gHaE7.jpeg  spider
2             hen/OIP-d5AYSW_M0rY_3Gw8MzoP9QHaEO.jpeg     hen
3           sheep/OIP-fRihzYpMaAt4htY3lPCezQHaFj.jpeg   sheep
4                              panda/panda_00812.jpg   panda
...                                                ...     ...
10843  spider/ef32b70c20e90021d85a5854ee454296eb70e3c...  spider
10844        spider/OIP-BCyJXTYuCAKPX1M3GQHZZAHaHa.jpeg  spider
10845  sheep/eb3cb00d2cf7033ed1584d05fb1d4e9fe777ead2...   sheep
10846         horse/OIP-dNxYCBzHlD_5vtCGESEEyAHaFj.jpeg   horse
10847           cow/OIP-nGtTEtZy3ZEAguoCKZSvWgHaFO.jpeg     cow

[10848 rows x 2 columns]
                                                id     target
0        spider/OIP-9p5nlorKdSoXY9OBwaih4gHaF6.jpeg     spider
1                               dogs/dogs_00451.jpg       dogs
2         horse/OIP-2eWvMHtLEvEycesigyT1SQHaFc.jpeg      horse
3          dogs/OIP-5c8U_LpbyHZ8cy0CmzM9xwHaKV.jpeg       dogs
4      squirrel/OIP-eJR5RZOJvwgWEGba8QdJ3gHaGB.jpeg   squirrel
...                                             ...        ...
4644                            dogs/dogs_00910.jpg       dogs
4645     spider/OIP-8QzpzRweC4ut5BHr7UL-1AHaFC.jpeg     spider
4646   elephant/OIP-7ZsL7LHkQIXBzT0hLE3jNAHaEK.jpeg   elephant
4647  butterfly/OIP-l_sJfw8ek_UMzcLmOF_fCQHaFE.jpeg  butterfly
4648          hen/OIP-4w-ODk3_r1OkLZV3diLnXQHaJ4.jpeg        hen

[4649 rows x 2 columns]
                                               id    target
0     elephant/OIP-7cwlVygDaV7WRs4iml_ExAHaE8.jpeg  elephant
1                          panda/panda_00127.jpg     panda
2     elephant/OIP-NxWsyjj31g3LQtDrBV3ZoQHaE8.jpeg  elephant
3     squirrel/OIP-Iso-nFbLHfJEtvVQ9izECQAAAA.jpeg  squirrel
4                          panda/panda_00763.jpg     panda
...                                            ...       ...
1726                           monkey/n7149.jpg    monkey
1727     sheep/OIP-0HdY2Xdt_cs602BufH68ZwHaEK.jpeg     sheep
1728       hen/OIP-c_HWl1sSrGFA9blZ5k4n3gHaJc.jpeg       hen
1729    spider/OIP-2Wtd0x41IwBfnEYNsq1WdQHaF8.jpeg    spider
1730                           monkey/n7118.jpg    monkey

[1731 rows x 2 columns]
```
Figure 2. Training, validation and testing dataset

Then I did the exploratory data analysis of each dataset.

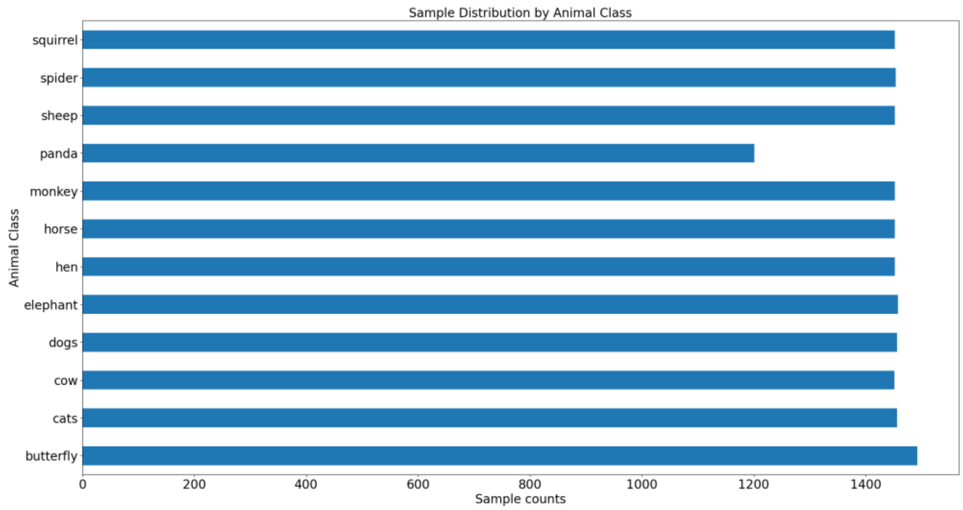

Figure 3. Sample Distribution
This is the overall sample distribution. There are almost 1400 image files for each class.

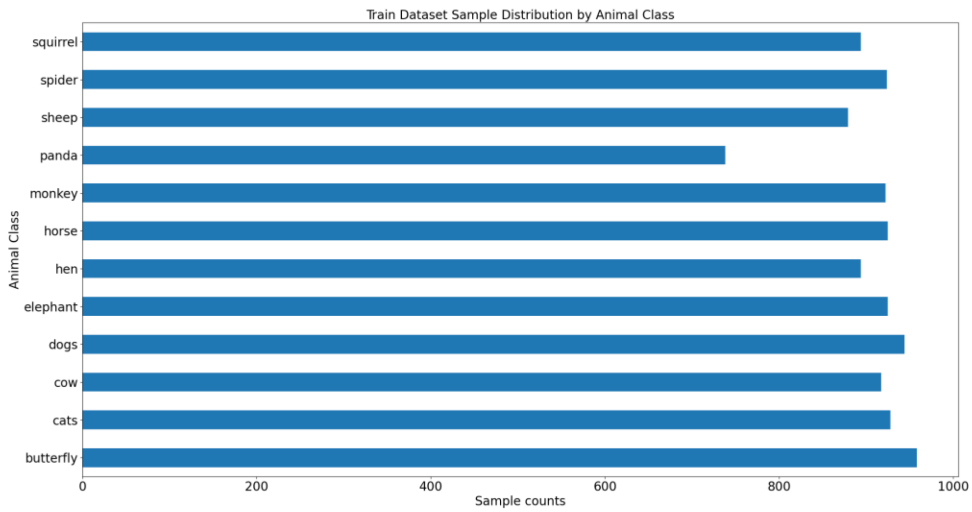

Figure 4. Train dataset Distribution
This is the train dataset sample distribution. There are almost 900 image files for each class.

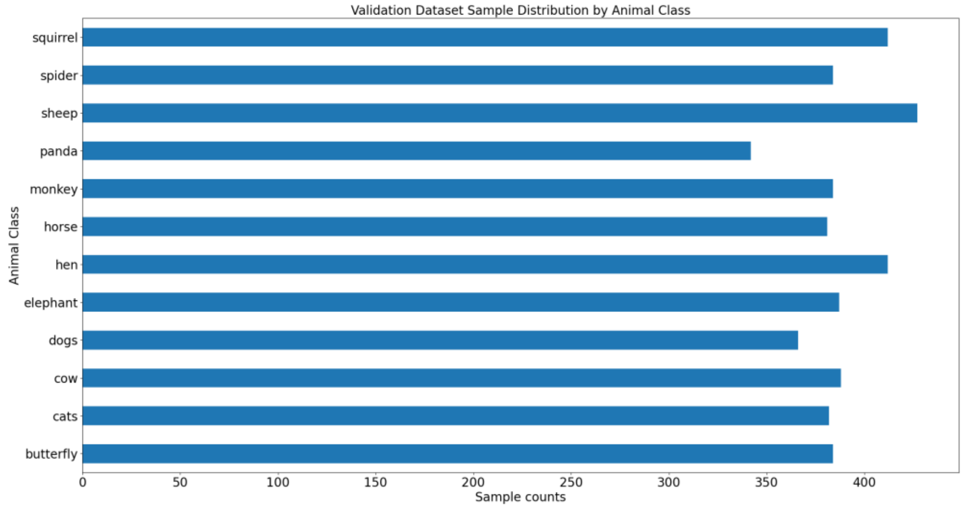

Figure 5. Validation dataset Distribution
This is the validation dataset sample distribution. There are almost 350 image files for each class.

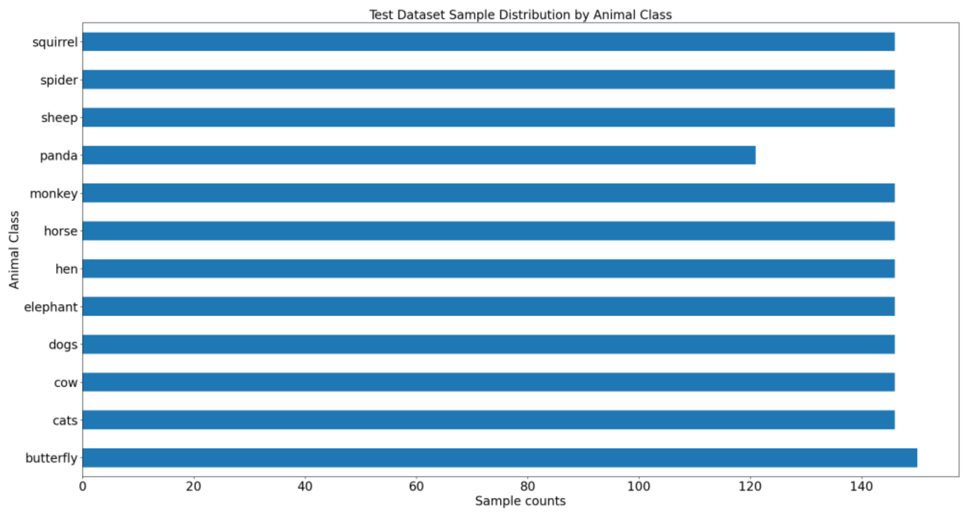

Figure 6. Test dataset Distribution
This is the test dataset sample distribution. There are almost 140 image files for each class. As we can see from these pictures, the datasets are almost balanced.
Then I used pretrained model VGG16 and EfficientNetB5 to check their performance.
And here are the plots.

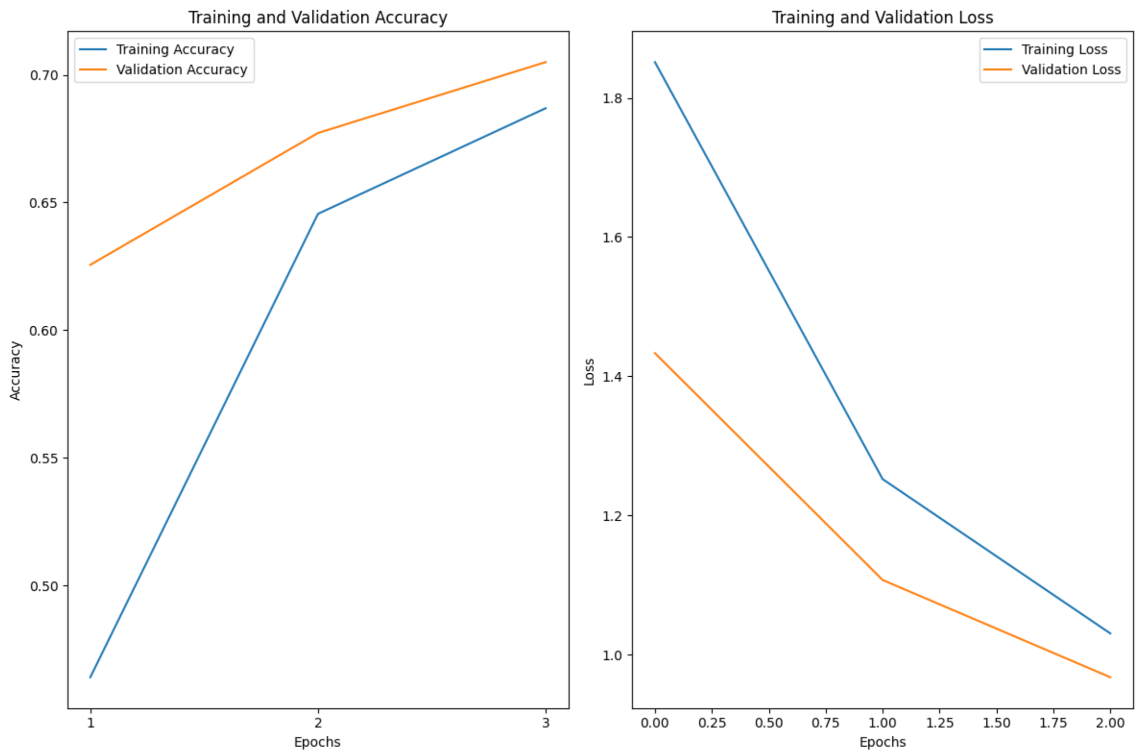

Figure 7. EfficientNetB5

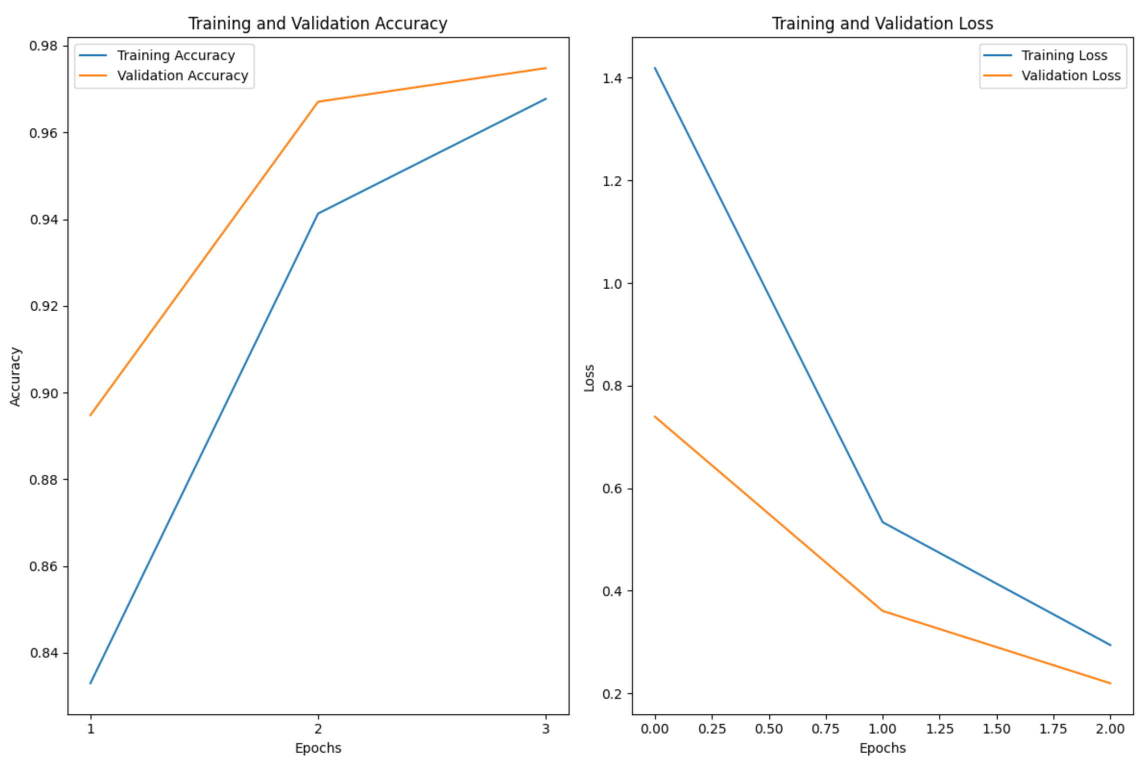

Figure 8. VGG16

Summary

In this project, I learned how to do the data augmentation and how to use CNN architecture to solve the classification problems. Apart from these, I knew lots of pretrained models of CNN through this project. I understood their architectures and how was their performance. Most importantly, I have the experience that how to work with group. And for the future, I think we can transfer our models to other datasets to check the performance. And we can also do more hyperparameters to improve our models or try some other architectures, like RNN, DNN.

Percentage of code that found or copied from the internet
18%

References

SHASHANK GUPTASHASHANK GUPTA 3, pythiestpythiest 4, Rahul Reddy VemireddyRahul Reddy Vemireddy 74544 silver badges66 bronze badges, David MakovozDavid Makovoz 42044 silver badges99 bronze badges, Saghan MudbhariSaghan Mudbhari 5111 silver badge11 bronze badge, Sujit JenaSujit Jena 3122 bronze badges, Lerner ZhangLerner Zhang 34611 silver badge88 bronze badges, goyuiitvgoyuiitv 1133 bronze badges, & Olivier D'AnconaOlivier D'Ancona 10122 bronze badges. (1964, October 1). *Micro average vs Macro average performance in a multiclass classification setting*. Data Science Stack Exchange. Retrieved April 22, 2022, from https://datascience.stackexchange.com/questions/15989/micro-average-vs-macro-average-performance-in-a-multiclass-classification-settin

Nepal, P. (2020, December 8). *VGGNet architecture explained*. Medium. Retrieved April 22, 2022, from https://medium.com/analytics-vidhya/vggnet-architecture-explained-e5c7318aa5b6#:~:text=VGGNet%20is%20a%20Convolutional%20Neural%20Network%20architecture%20proposed%20by%20Karen,network%20depth%20on%20its%20accuracy.

Overfit and underfit: Tensorflow Core. TensorFlow. (n.d.). Retrieved April 22, 2022, from https://www.tensorflow.org/tutorials/keras/overfit_and_underfit

Xu, Y., Noy, A., Lin, M., Qian, Q., Li, H., & Jin, R. (2020, October 3). *WeMix: How to better utilize data augmentation*. arXiv.org. Retrieved April 25, 2022, from https://arxiv.org/abs/2010.01267#:~:text=The%20main%20limitation%20of%20data,of%20existing%20data%20augmentation%20methods.