Classification and visualization on damaged Houses with different neural networks

ST 807 G6 Haoze Wu, Liuding He and Shuyi Pi

Motivation and About the Project

In this work, we need to build a model to find out if a building has been damaged by a hurricane. To measure the damage is labor-intensive and time-consuming and not efficient at all. With a computer model, people can quickly measure the damage by the satellite image of a place.
In this project, we think we not only need to select a model that can classify the image, but we also need to find out which part of the image is the most important part for our model to make the prediction.

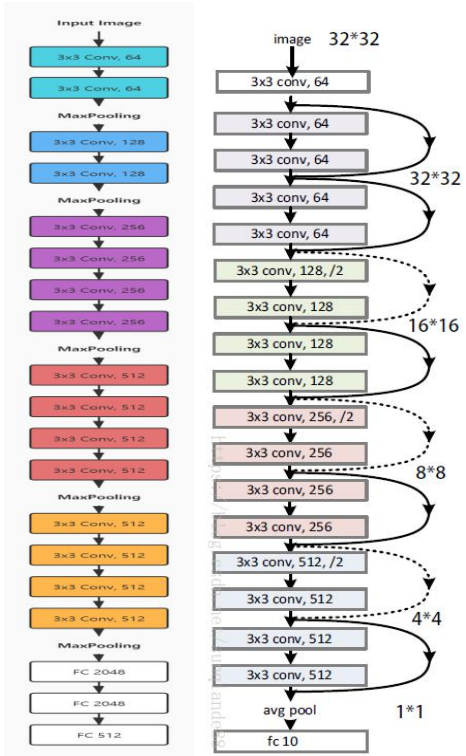
Data and Labels

The data are satellite images from Texas after Hurricane Harvey divided into two groups (damage and no_damage). The goal is to make a model which can automatically identify if a given region is likely to contain flooding damage.
We preprocess the data with keras.preprocessing.image.ImageDataGenerator. This function can help us augment the image data. In our work, we preprocess the image with shifting its width and height, and zoom the image. We also add Gaussian filtering to the image so the image can be blurred, we think it may increase the robust of our model.

References

Cao, Quoc Dung, and Youngjun Choe. "Building damage annotation on post-hurricane satellite imagery based on convolutional neural networks." *Natural Hazards* 103.3 (2020): 3357-3376.
He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).
Zeiler, M.D. & Fergus, R.. (2013). Visualizing and understanding convolutional networks. European Conference on Computer Vision(ECCV). 8689. 818-833.

Model

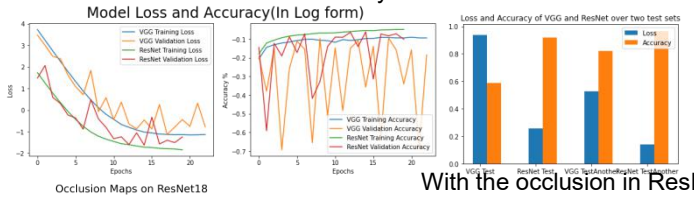


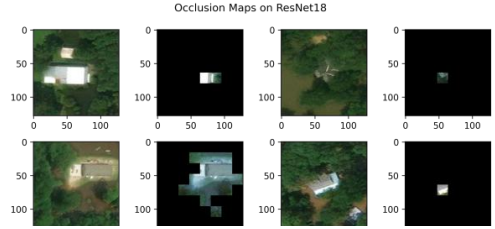
VGG-19 ResNet-18

In our work, we have tried two different models, which is VGG-19 and ResNet-18. The scale of the two model are similar, and they are not so large so most of our computer can easily train them.
The VGG-19 net is a simple convolution neural network with a greater number of conv layer than some simple net.
The ResNet net use a method called "shortcut connection" and it use a global average pooling layer to take place of the max pooling layer, so there is a smaller number of parameters in this type of network.
After training the network, we also use a method of occlusion map to recognize which part of the image can help the network to classify the image. The detail of this method is blur some part of the image, use the model to predict the image again and find out if the result have changed greatly.

Results

We notice that during the training process, the VGG-19 net has a fluctuation on its validation accuracy. And the ResNet-18 has a great validation accuracy and test accuracy.





With the occlusion in ResNet-18, we found out that when we blur different part of the image, only with some major part of the building being occluded, the prediction will show a great difference with that with the full image. But the VGG-19 model can't focus on the building.

Conclusion and Future Work

In this project, we use two model to predict the satellite images from Texas after Hurricane Harvey. We compare the effectiveness between VGG-19 and ResNet-18. With the Residual block, the neural network can easily converge, and we can get a better performance in both of the testing set. When the traditional full connect layer is replaced by some convolution layer and average pooling, there is a smaller number of parameters than using VGG-19 net. In this dataset, totally we have find out that using a neural network with Residual block and average pooling have a better performance.
In the future, we think we can use the occlusion map result and feature map result to improve the structure of our network, so it may be more robust. The structure of the neural network can also be improved. Some other network structure like DenseNet may also perform well in this dataset. What's more, we can also do some research on the different method of data augmentation so that our model can be more robust.
The occlusion map in our VGG-19 net doesn't work well in the occlusion map, because it can't show the important part of the image. When the structure of the network is changed, the occlusion map will show the significant part of the image more clearly but not like our result.
We also believe that in the future, some different method can be used to not only recognize the image with the full building in the image, but also recognize the image with occlusion. When the building is partly displayed or totally hide, there may be a model that can finish the task by the surroundings of the building.