# Scene Classification and GradCam Visualization

Shiyue Wang, Yuqi Wang, Olivia Qu, Wenlu Dong

1. Motivation

Nowadays, Microsoft AI For Earth, has created the most detailed United states forest map using satellite imagery and AI, which would essentially be a game changer in reducing deforestation, pests and wildfires.

Based on this, We wish to use a deep learning model wich will be train based on Convolutional Neural Networks (CNNs) and Residual Blocks to detect the type of scenery in images.

This project could be practically used for detecting the type of scenery from the satellite images. In addition, this project will cover the use of a technique known as Grad-Cam to observe and explain how AI models think. In detail, How can we understand what our convolutional neural network (black box) sees and understands when making a decision?
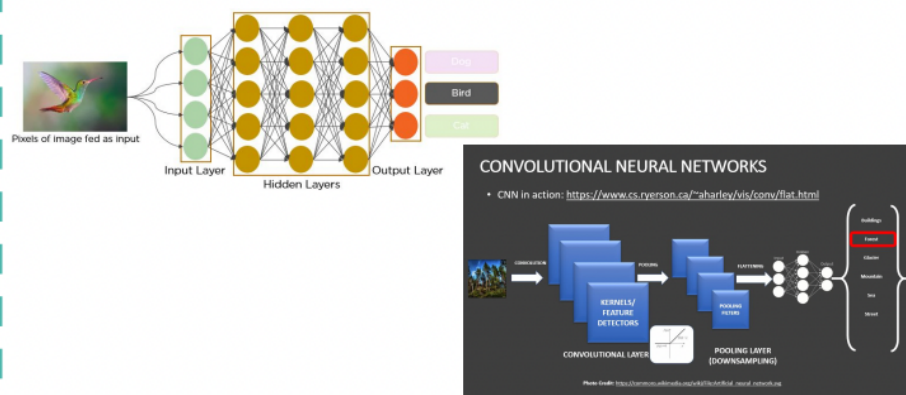
2. Dataset Overview

- 14034 training and 3000 test images with highly accurate, detailed and consistent annotations from a wide range of natural scenes from all around world.
- The dataset have 6 categories: street, sea, buildings, forest, mountain, glacier.
- Preprocessed to fit 96*96 dimension.
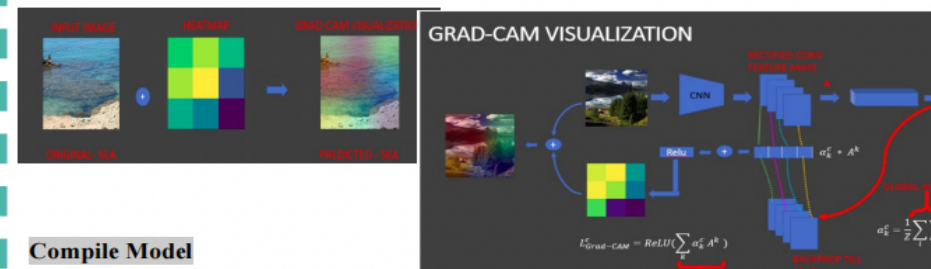
Classified Sample Images
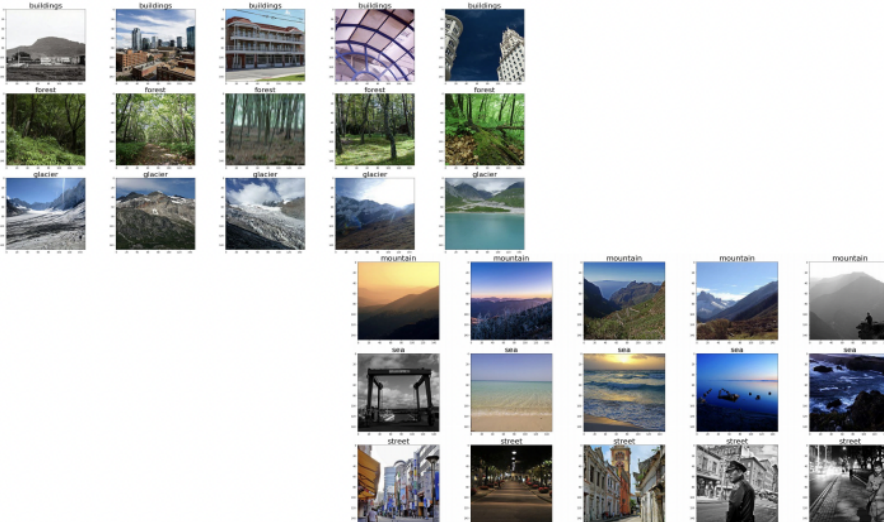
3. Deep Learning

Residual Network

- In Residual Networks we use a technique called skip conection to solve the vanishing gradient problem.
- ResNet help in tracking the vanishing gradient problem using identity mapping.
- ResNet has an incredible performance in the networks with large number of layers.
- Each Res-Block consist of convolution block followed by two identity block
- Each convolutional layer is followed by a batch normalization layer and a ReLU activation function.
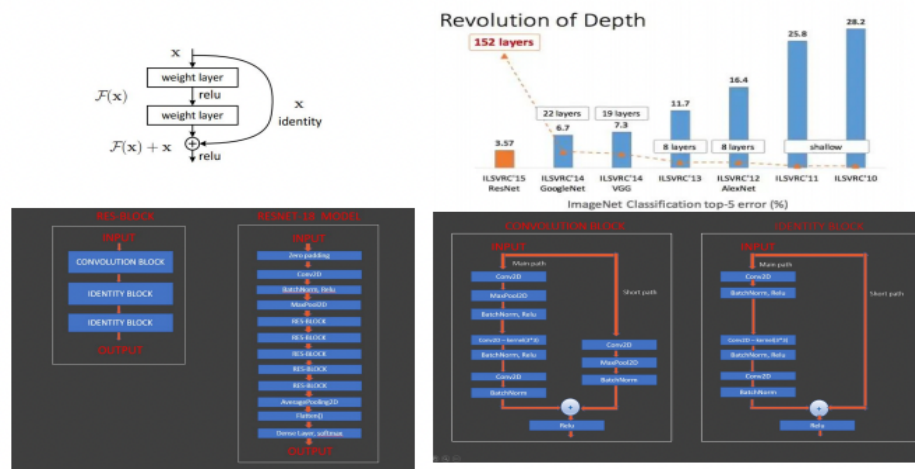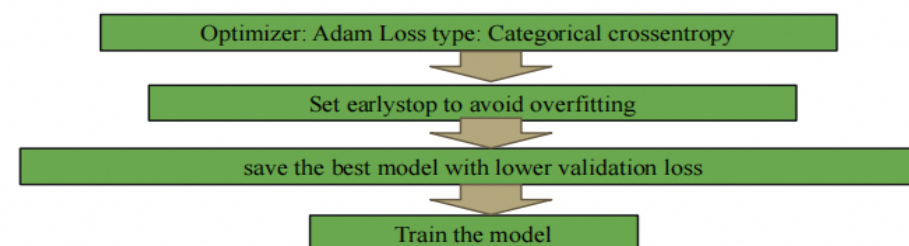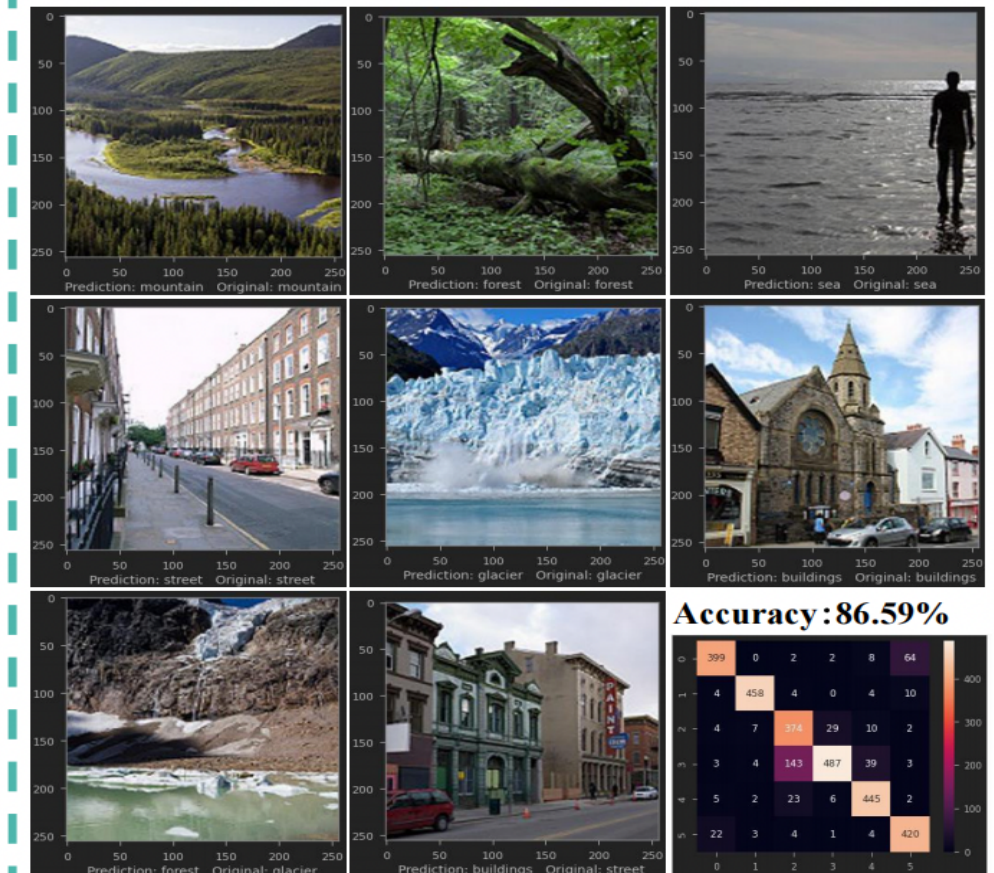

GRAD-CAM VISUALIZATION

- Gradient-Weighted Class Activation Mapping (Grad-CAM) helps visualize the regions of the input that contributed towards making prediction by the model.
- It does so by using the class-specific gradient information flowing into the final convolutional layers of CNN to localize the important regions in the image that resulted in predicting that particular class.


Compile Model

- Optimizer: Adam Loss type: Categorical crossentropy
- Set earlystop to avoid overfitting
- save the best model with lower validation loss
- Train the model

4. Assess the Model

Accuracy: 86.59%

5. Result of Grad-Cam Visulization

- Heat map scaled on the top of the original images.
- The difference change in color shows the regions which model focus on to make the prediction.
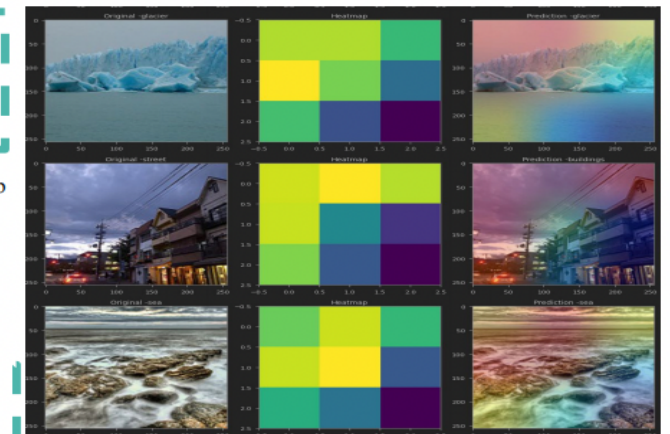
6. Conclusion

- The model performs pretty good with 86.59% accuracy. The confusion matrix also shows this result. Among 6 categories, "forest" got the highest precision (97%) and glacier got the lowest precision (68%). The glacier category could be further improve by tuning the model or more experimentation in image augmentation.
- Grad-Cam can really show how the machine think by showing where portion in the photos its focusing on.
- **Q:** How can we make the Grad-CAM heat-map more precise?