Remote Sensing Image Classification

Atakan Serbes
21200694

Dataset

- NWPU-RESISC45 dataset [Cheng, 2017]
  - Remote Sensing (One of the most comprehensive)
  - 256 x 256
  - 700 img per class
- Object classes are selected, scene classes are discarded (forest, beach, etc.)

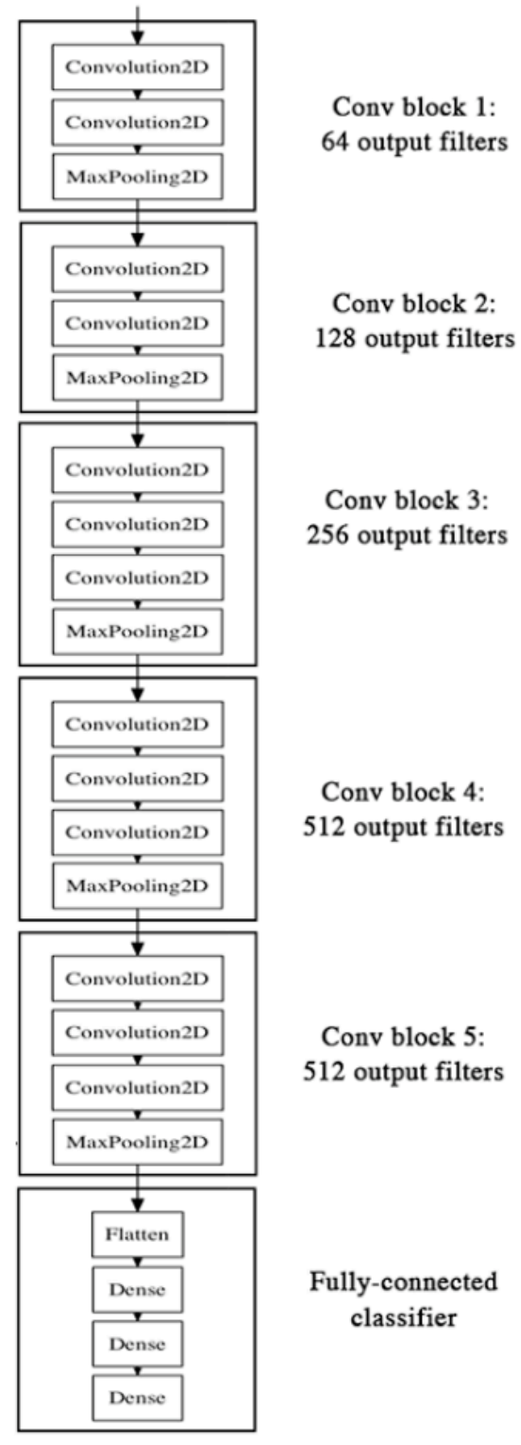
Methodology

- Goal is to learn an image classification model over a set of training images
- 700 image divided as 560:70:70 (train:val:test), (low number of instances)
- Transfer learning using a VGG-16 network trained on ImageNet, as VGG-16 performs better than AlexNet and GoogLeNet in survey results. [Cheng, 2017]

Methodology

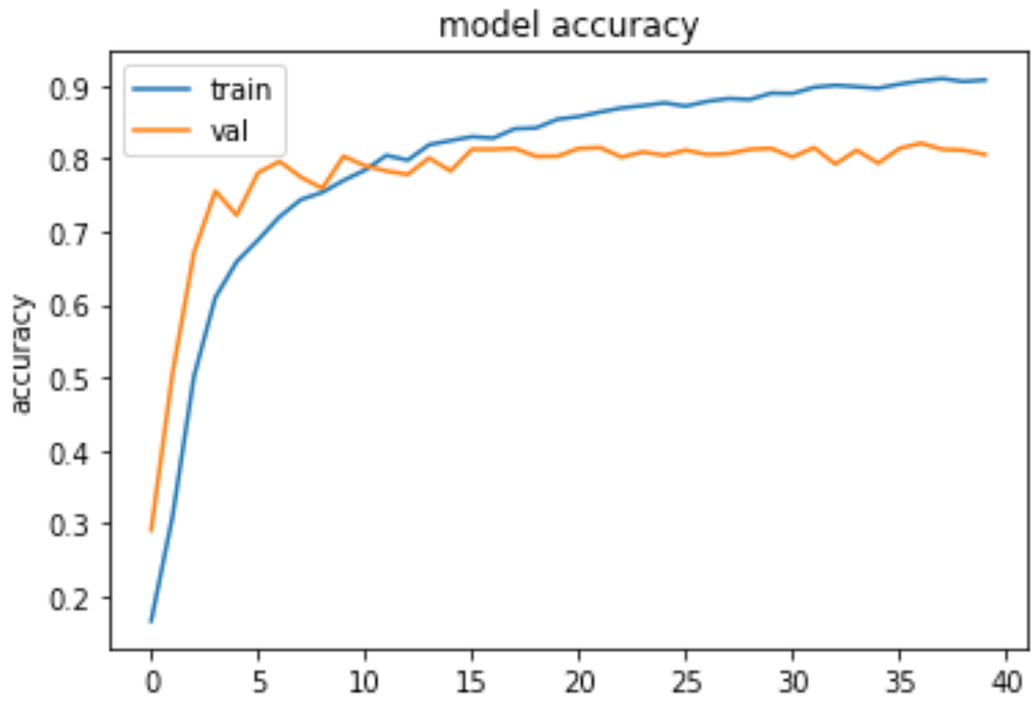
- After feature extraction with VGG-16's convolutional layers, FC layers are added (1 or 2) followed by a softmax of 12 nodes for 12 classes.

- Categorical cross-entropy is used for the loss function, (penalizes predictions that are farther away from the correct prediction.)



Experiments

- VGG16 based model achieves up to 0.81 validation accuracy (SotA in [Cheng, 2017] is 0.90)
- This is achieved without any augmentation

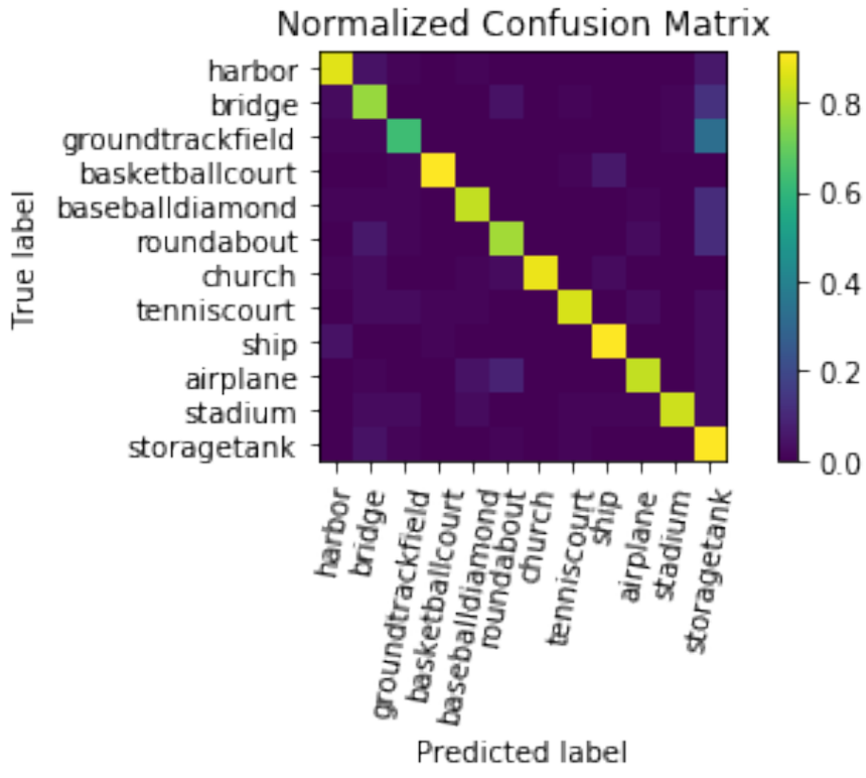


Experiments

- Augmented dataset with because of the flexibility of remote sensing images (from above)
  - Width,height shifts
  - Rotation, flips
- Does not improve results. (0.69 validation accuracy, 0.68 train accuracy)

Experiments

- Results show that it predicts some classes better than others (especially circular shaped objects)
- These features might be relevant to pretrained features



Conclusion

- Developed a classification model based on VGG-16 architecture using very little data

- Models are not very robust, may be due to pretained features and dataset feature differences

- Aerial images are different with different types of edges & low-level features that transfer learning might not accommodate (most images in ImageNet are in up-right position)

- For the future work, I would investigate further architectures and maybe detection