University of Science and Technology of Hanoi

Introduction to Deep learning

Topic: Study the MobileNet for remote sensing image classification

Nguyễn Nữ Lan Anh [BA12-016]

Phùng Nam Khánh [BA12-092]

Lại Phú Huy [BA12-084]

Nguyễn Quốc Dương [BI12-123]

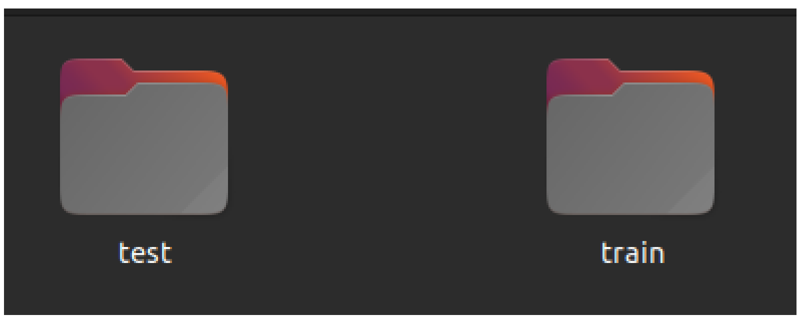
Trần Ngọc Huy [BA12-090]

Hanoi, September 2024

1. Problem :

    Remote sensing image classification is a crucial task in earth observation, environmental monitoring, and urban planning. As the volume and resolution of satellite imagery continue to increase, there is a growing need for efficient and accurate classification methods. This study explores the application of MobileNet, a lightweight convolutional neural network, for remote sensing image classification tasks.

2. Dataset:



    - Dataset contains 2 parts: test and train (each part contains about 300 images)

3. MobileNet model

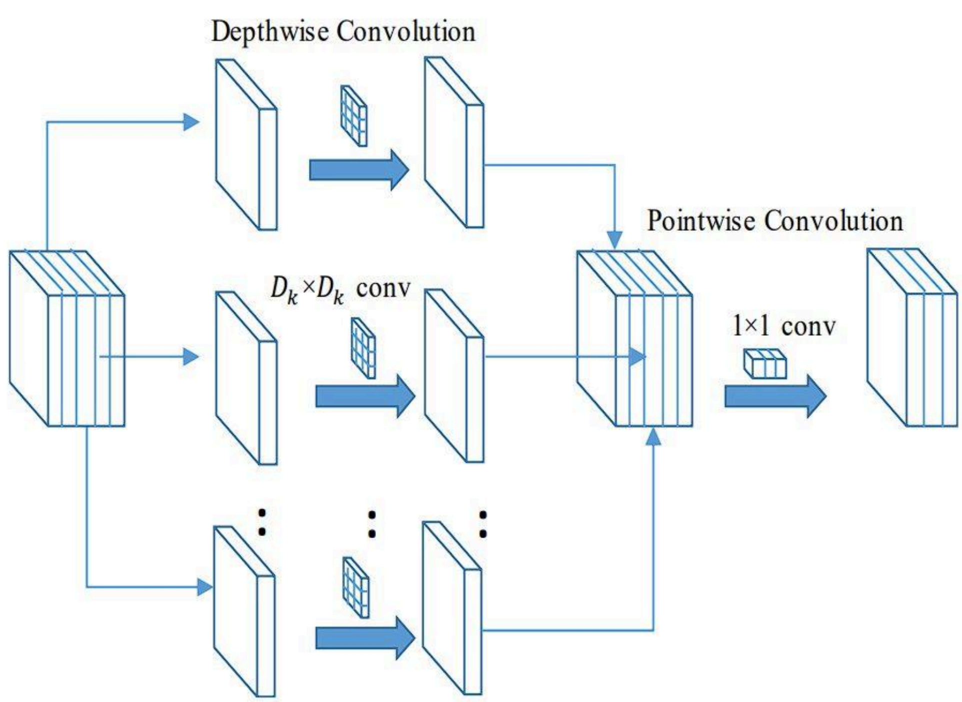
    - MobileNet, developed by Google, is a lightweight convolutional neural network designed for mobile and embedded vision applications. Key features of MobileNet include:

    · Depthwise separable convolutions for reduced computational cost

Steps in Depthwise Convolution

- Separation of Channels: Standard convolution applies filters to each channel of the input image. Whereas depthwise convolution only applies a single filter per input channel.

- Filter Application: As each filter is applied independently, the output is the result of convolving (multiplication and summation) a single input channel with a dedicated filter.
- Output Channels: The output of the depthwise convolution has the same number of channels as the input.
- Reduced Complexity: Compared to standard convolution, the total number of multiplicative operations is reduced.
- For standard convolution the total number of multiplicative operations = $K \times K \times C \times D \times \text{height} \times \text{width}$
- For depthwise convolution, the number of operations = $K \times K \times C \times \text{height} \times \text{width}$



· Width multiplier to adjust the number of channels in each layer

· Resolution multiplier to adjust the input resolution

- For our remote sensing application, we modify the MobileNet-V2 architecture:

· Input size: 256x256x4 (to accommodate the NIR band)

· Output: 6 classes

· Final layer: Global Average Pooling followed by a Dense layer with softmax activation

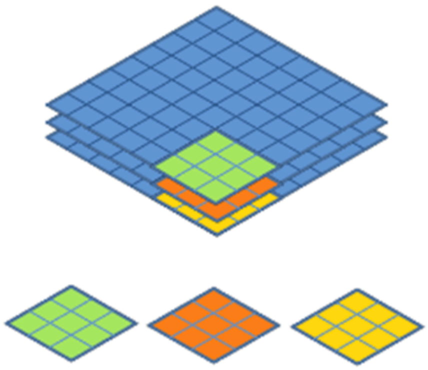
- The implementation uses TensorFlow and the Keras API. Key modifications to the provided code include:

· Adapting the input pipeline in dataset_factory.py for the RemoteSenseNet dataset

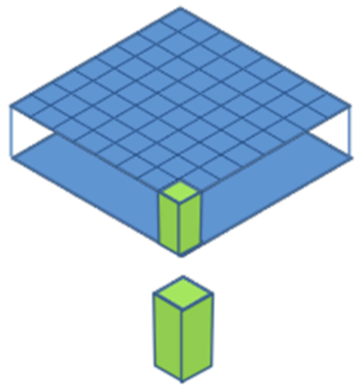
· Modifying preprocessing_factory.py to handle 4-band satellite imagery

· Adjusting train_image_classifier.py for MobileNet and our specific task

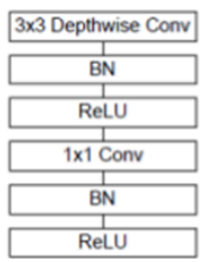
4. Figures



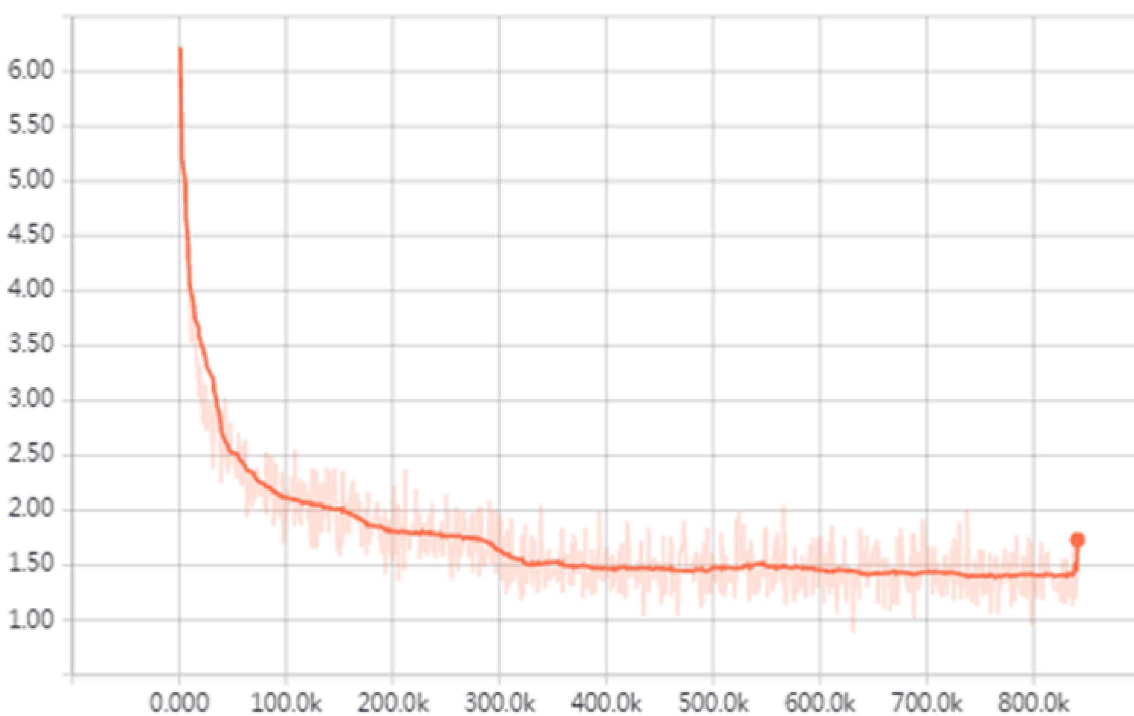
Depthwise Convolutional Filters



Pointwise Convolutional Filters



Depthwise Separable Convolution



4. Potential Modifications for Remote Sensing

- To better adapt MobileNet for remote sensing tasks, consider the following modifications:
- Adjust the input size to match the resolution of the satellite/aerial imagery.
- Modify the final classification layer to match the number of land use/land cover classes in the dataset.
- Experiment with transfer learning by initializing the model with weights pre-trained on a large remote sensing dataset.

5. Conclusion

- MobileNet offers a promising architecture for remote sensing image classification due to its efficiency and small model size. By adapting the provided implementation to work with remote sensing datasets and potentially making task-specific modifications, it can be effectively used for classifying satellite and aerial imagery. Further experimentation with different width multipliers and resolution settings can help find the optimal balance between accuracy and computational efficiency for specific remote sensing applications.
- This report provides an overview of using MobileNet for remote sensing image classification based on the given code and information. For a more comprehensive study, actual implementation, training, and evaluation on a specific remote sensing dataset would be necessary.