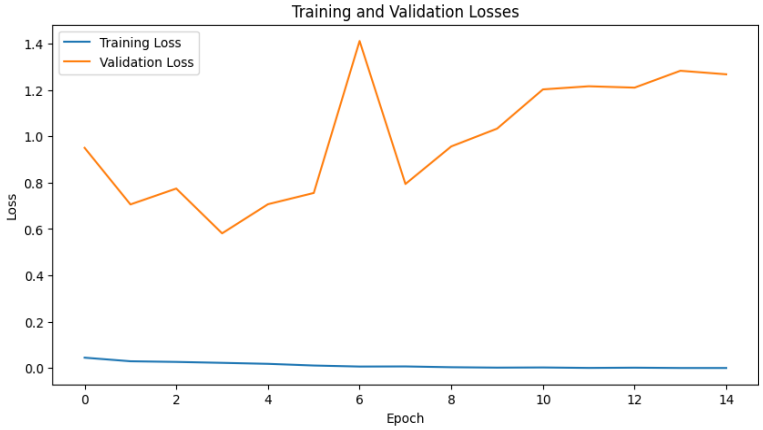
About Dataset

The Intel Image Classification dataset, sourced from Kaggle, comprises approximately 14,000 training images, 3,000 test images, and 7,000 prediction images. It's a multiclass image scene classification dataset featuring six categories: buildings, forest, glacier, mountain, sea, and street. These images represent natural scenes from diverse locations worldwide.

CNN Architecture

Architecture & Specifications: The CNN architecture consists of three convolutional layers, max-pooling layers for feature extraction, and two fully connected layers for classification. It is trained on a dataset with 6 classes: buildings, forest, glacier, mountain, sea, and street. The training setup includes a batch size of 32, 15 epochs, and a learning rate of 0.05. During training, it achieves high accuracy on the training set but slightly lower accuracy on the validation and test sets.



Set	Loss	Accuracy
Training	0.006	99.9%
Validating	1.283	78.3%
Testing	1.211	80.3%

Strengths:
The following are the strengths of this architecture as compared to other models.

1. High Training Accuracy: The model achieves a training accuracy of nearly 99.9%, indicating that it effectively learns to classify the training data.

2. Simple Architecture: The IntelNet architecture is relatively simple, making it easy to understand and implement. It consists of three convolutional layers, which can capture essential features from images.

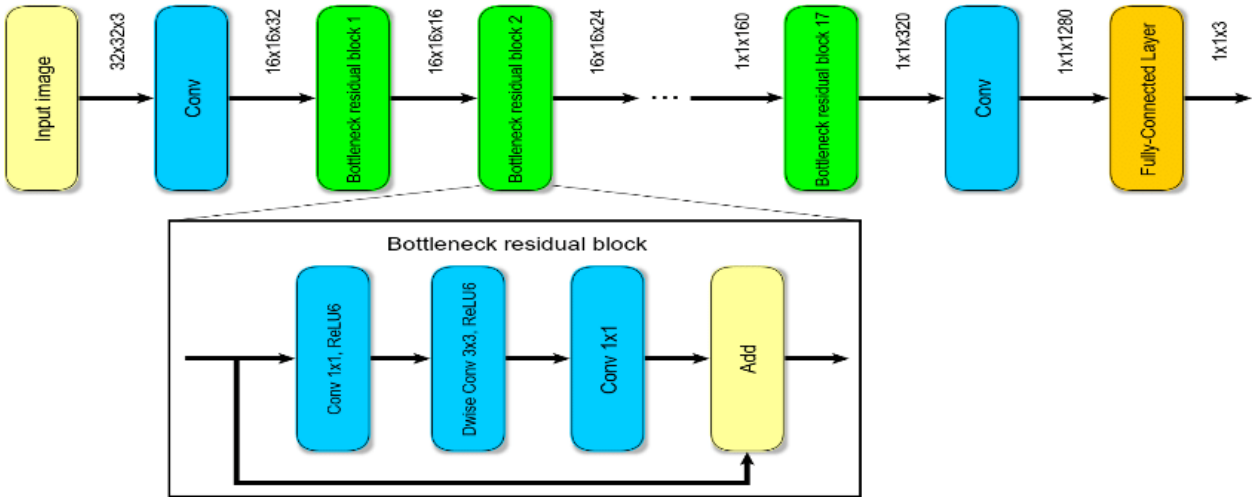
Limitations:
The following are the weaknesses or limitations of this architecture as compared to other models.

1. Overfitting: The model exhibits a significant gap between training and validation/test accuracy, suggesting potential overfitting. It may have learned to memorize the training data rather than generalize to unseen data.

2. Learning Rate: The choice of a high learning rate (0.05) might contribute to training instability and slow convergence. Fine-tuning the learning rate and incorporating learning rate scheduling could improve performance.

3. Model Complexity: The model's simplicity can be a limitation when dealing with complex image data. Adding more convolutional layers or using a pre-trained model (e.g., ResNet or VGG) might lead to better feature extraction.

Comparative Analysis-I: Finetuning MobileNet-V2



Architecture & Specifications: MobileNet-V2 is a lightweight convolutional neural network architecture known for its efficiency. It employs depthwise separable convolutions and linear bottlenecks to reduce model size while maintaining high performance. The fine-tuned MobileNet-V2 model was trained with specific specifications. The training utilized a batch size of 128, ran for 5 epochs, and employed an Adam optimizer with a learning rate of 0.01.

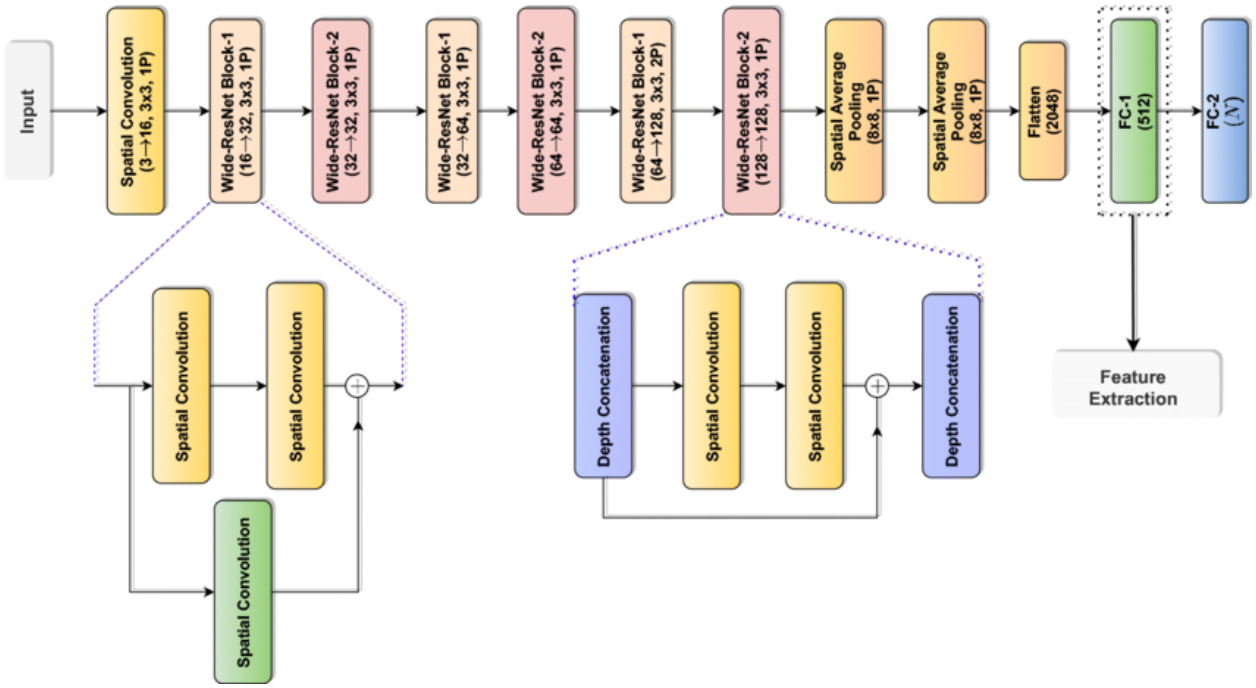
Results:

Set	Loss	Accuracy
Training	0.457	83.8%
Validating	0.485	76.7%
Testing	0.474	81.5%

Strength: MobileNet-V2's efficient architecture led to faster training and superior results in fewer epochs, optimizing time and resources.

Limitation: MobileNet-V2 may plateau at a certain accuracy level and struggle to further improve, especially on challenging datasets.

Comparative Analysis-II: Finetuning Wide ResNet-50



Architecture & Specifications: Wide ResNet-50 is a deep convolutional neural network known for its wide layers, emphasizing feature reuse. It excels in capturing intricate image patterns, making it suitable for complex tasks. The training utilized a batch size of 128, ran for 5 epochs, and employed an Adam optimizer with a learning rate of 0.01.

Results:

Set	Loss	Accuracy
Training	1.582	38%
Validating	0.909	59.8%
Testing	0.914	62.4%

Strength: Wide ResNet-50 excels in image feature extraction and complex tasks, offering robust performance in computer vision applications.

Limitation: Due to its wide architecture, Wide ResNet-50 may not be as well-suited for the Intel Image Classification dataset, which is less complex than tasks that require its extensive feature extraction capabilities.