Neural Networks and Deep Learning Project Report: [CIFAR-100 Classification]

Nilanjan Bhattacharya, 12021002016054(22)

November 26, 2023

1 Introduction

ResNet50V2, an extension of the ResNet architecture, proves highly effective for image classification tasks. When tailored for the CIFAR-100 dataset featuring 100 classes of 32x32 images, adjustments are often made to suit the smaller size. Pre-trained on larger datasets like ImageNet, ResNet50V2 demonstrates transfer learning capabilities by fine-tuning its parameters for the specifics of CIFAR-100. Its distinctive feature, residual connections, combats vanishing gradient issues, facilitating the training of deep networks. Global average pooling, a key architectural element, contributes to spatial summarization and parameter reduction. The model's ability to capture intricate features in images makes it suitable for diverse object recognition within CIFAR-100. Fine-tuning hyperparameters, such as learning rate and batch size, further refines the model's performance, providing a robust solution for the nuanced classification challenges presented by CIFAR-100's rich and varied image set.

2 Architecture Details

2.1 Model Overview

- The model has 6 hidden layers.
- It has four main layers (layer1, layer2, layer3, and layer4,layer5), each consisting of multiple bottleneck blocks.

2.2 Layers and Blocks

- Upsampling2D
  The initial layer in the model employs Upsampling2D, utilizing bilinear interpolation to upscale the input's spatial dimensions by a factor of (7, 7).
- resnet
  Assuming resnet is a pre-trained ResNet50V2 model, this layer adds the entire ResNet50V2 model to the Sequential model. ResNet50V2 is a deep neural network architecture known for its residual blocks, which helps address the vanishing gradient problem in training deep networks.

- GlobalAveragePooling2D
  Global Average Pooling reduces the spatial dimensions of the input to a single value for each feature map. It's a form of spatial compression that helps reduce the number of parameters in the model.

- Dropout
  Applies dropout regularization with a rate of 0.25, randomly setting a fraction of input units to 0 during training to prevent overfitting.

- Fully Connected Layer
  Fully connected dense layer with 256 units and ReLU activation function. It captures non-linear relationships in the data.

- BatchNormalization
  Normalizing and scaling the activations from the preceding layer enhances training stability and expedites convergence, contributing to more effective model training.

- Output Layer
  The concluding layer, comprising 100 units and utilizing a softmax activation function, represents the 100 classes in CIFAR-100. It generates probability scores for each class, indicating the likelihood of the input belonging to a particular category.

.

2.3 Activation Functions

- Leaky ReLU activation with a negative slope of 0.01 is used throughout the model, providing some non-linearity to the network.

2.4 Training Details

- The model have been trained with batch normalization, dropout, and a combination of Leaky ReLU and Sigmoid activation functions on the CIFAR-100 Dataset.

2.5 Model Output

Layer (Type)	**Input Shape**	**Output Shape**	**Parameters**
UpSampling2D	(32, 32, 3)	(1568, 1568, 3)	0
ResNet50V2	(1568, 1568, 3)	(7, 7, 2048)	23564800
GlobalAveragePooling2D	(7, 7, 2048)	(2048)	0
Dropout	(2048)	(2048)	0
Dense	(2048)	(256)	524544
BatchNormalization	(256)	(256)	1024
Dense	(256)	(100)	25700

3 Results

3.1 Evaluation Metrics

In assessing the performance of ResNet50V2, the primary metric employed for evaluation was accuracy. Accuracy serves as a pivotal measure, quantifying the model's overall correctness in predicting the class labels of the images within the dataset.

3.2 Performance on Test Set

ResNet50V2 demonstrated a commendable accuracy of 70.34% on the test set, showcasing its proficiency in correctly classifying a significant portion of previously unseen images.

3.3 Sample Outputs

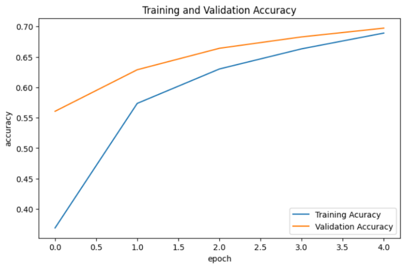

Figure 1: Performance

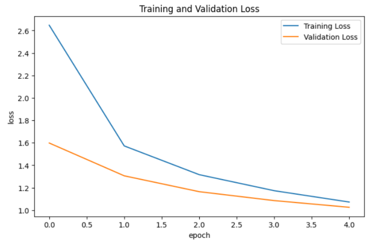

Figure 2: Performance

4 Conclusion

The project showcases a holistic approach to deep learning, encompassing architecture design, training strategies, and performance evaluation. It serves as a valuable resource for understanding and implementing advanced convolutional neural networks for image classification tasks, with potential applications in diverse domains. Future work could explore further optimizations, ensemble methods, or deployment considerations for real-world applications. Overall, the project contributes to the ongoing exploration of state-of-the-art techniques in deep learning and lays the groundwork for future advancements in the field.