Introduction

The purpose of this study was to evaluate the performance of two variations of the VGG-16 model using the CIFAR-100 dataset for image classification tasks. With the dataset consisting of 50,000 training images and 10,000 testing images across 100 classes, our aim was to assess how different modifications to the VGG-16 architecture impacted model performance. By comparing the standard VGG-16 model with a modified version incorporating additional techniques such as batch normalization, dropout layers, and kernel regularization, we sought to identify which approach yielded superior results in terms of accuracy and generalization.

Method

In this study, the CIFAR-100 dataset served as the basis for evaluating the performance of two variations of the VGG-16 model in image classification tasks. The dataset comprises 50,000 training images and 10,000 testing images, distributed across 100 classes. Initial preprocessing involved loading the dataset into respective variables for training and testing images (x_train and x_test) and their corresponding labels (y_train and y_test).

The first model architecture closely adhered to the traditional VGG-16 structure, consisting of convolutional and pooling layers followed by fully connected layers. Each convolutional block contained two convolutional layers with Rectified Linear Unit (ReLU) activation functions, alongside a max-pooling layer. The fully connected layers consisted of two hidden layers with ReLU activation, culminating in an output layer employing softmax activation. In contrast, the second model, a modified version of VGG-16, incorporated additional techniques to enhance performance. This included the integration of batch normalization layers after each convolutional layer, along with dropout layers progressively increased in magnitude to accommodate the increasing complexity of the layers. Furthermore, a kernel regularizer employing L1 and L2 regularization with coefficients of 0.01 was applied to the final convolutional layer to further mitigate overfitting.

Both models were trained using the Adam optimizer and categorical cross-entropy loss function. Training was conducted for a maximum of 20 epochs and 100 epochs, with a batch size of 128 samples. To prevent overfitting, early stopping was implemented. The models' performance metrics, including loss and accuracy, were recorded for both the training and validation sets throughout the training process.

Experiment

In this study, we investigated the performance of the VGG-16 model applied to the CIFAR-100 dataset. The experiment was structured into two main parts to comprehensively assess the model's behavior.

Part 1: Implementation of the VGG-16 Model

Initially, we began by preparing the CIFAR-100 dataset, loading it into variables x_train, y_train, x_test, and y_test, and conducting necessary preprocessing steps to ensure compatibility with the model architecture. Subsequently, we designed the model architecture to closely mirror that of the VGG-16 model. Once the model architecture was defined, it was compiled using the appropriate configurations and trained using the fit method, allowing it to learn patterns and features from the training data.

Part 2: Development of a Modified VGG-16 Model

Following the evaluation of the standard VGG-16 model, we proceeded to create a modified version with slight variations. This modified model retained the basic structure of VGG-16 but incorporated additional techniques such as the inclusion of BatchNormalization() after each convolutional layer and the application of regularization methods to mitigate overfitting. The remainder of the experimental setup remained consistent with the first part, ensuring a direct comparison between the two models.

Results

The outcomes of the experiment provided insights into the performance of both the standard VGG-16 model and the modified version.

In the case of the standard VGG-16 model, the maximum accuracy achieved was 0.8882, with a corresponding minimum loss of 0.3555. However, the percent difference between the training and validation accuracies (81.6%) suggests a notable disparity between the training accuracy and validation accuracy (Fig 1). A graphical representation of the validation loss against the number of epochs revealed a parabolic trend, reaching its minimum at epoch 7.5 before sharply increasing. Similarly, the graph depicting validation accuracy exhibited a steep linear rise until epoch 7.5, after which it plateaued (Fig 2)

On the other hand, the modified VGG-16 model demonstrated improvements in certain aspects. While the maximum accuracy attained was slightly lower at 0.6350, the percent difference between training and validation accuracies (6.6%) was significantly reduced (Fig 3). The validation loss graph displayed a continual exponential decrease over the course of the experiment (Fig 4). The validation accuracy remained suboptimal, reaching only 3/5 of the desired value after 100 epochs.

Analysis

The analysis of the results revealed important insights into the performance and characteristics of the two models. In the case of the standard VGG-16 model, the large difference between training and validation accuracies suggested a focus on details that did not generalize well to unseen data, indicating overfitting. Additionally, the rapid increase in both validation accuracy and loss during certain intervals further underscored the model's failure to recognize crucial patterns.

Conversely, the modified VGG-16 model exhibited more balanced performance, with reduced overfitting and a smoother validation loss curve. However, the inability to achieve desired accuracy levels, even after applying regularization strategies, suggested potential underfitting, indicating a lack of model complexity to capture all underlying patterns effectively.

In conclusion, the study highlighted the importance of balancing model complexity and regularization techniques in achieving optimal performance in image recognition tasks. While overfitting models are unreliable due to their failure to generalize well to unseen data, excessive use of regularization methods can lead to underfitting, resulting in suboptimal performance. Therefore, finding the right balance between complexity and regularization is essential for developing reliable and effective models for image recognition.

Appendix

Fig 1.

```
Epoch 79/100
391/391 [==============================] - 11s 28ms/step - loss: 3.4248 - accuracy: 0.6006 - val_loss: 3.6523 - val_accuracy: 0.5680
Epoch 80/100
391/391 [==============================] - 11s 29ms/step - loss: 3.4148 - accuracy: 0.6019 - val_loss: 3.6475 - val_accuracy: 0.5731
Epoch 81/100
391/391 [==============================] - 11s 28ms/step - loss: 3.4168 - accuracy: 0.6087 - val_loss: 3.6521 - val_accuracy: 0.5677
Epoch 82/100
391/391 [==============================] - 11s 28ms/step - loss: 3.4049 - accuracy: 0.6051 - val_loss: 3.6748 - val_accuracy: 0.5683
Epoch 83/100
391/391 [==============================] - 11s 28ms/step - loss: 3.4306 - accuracy: 0.6063 - val_loss: 3.5997 - val_accuracy: 0.5783
Epoch 84/100
391/391 [==============================] - 11s 28ms/step - loss: 3.3983 - accuracy: 0.6065 - val_loss: 3.7097 - val_accuracy: 0.5702
Epoch 85/100
391/391 [==============================] - 11s 28ms/step - loss: 3.3905 - accuracy: 0.6101 - val_loss: 3.6900 - val_accuracy: 0.5727
Epoch 86/100
391/391 [==============================] - 11s 28ms/step - loss: 3.4077 - accuracy: 0.6114 - val_loss: 3.7387 - val_accuracy: 0.5736
Epoch 87/100
391/391 [==============================] - 11s 28ms/step - loss: 3.3883 - accuracy: 0.6121 - val_loss: 3.7158 - val_accuracy: 0.5711
Epoch 88/100
391/391 [==============================] - 11s 28ms/step - loss: 3.3907 - accuracy: 0.6139 - val_loss: 3.5848 - val_accuracy: 0.5788
Epoch 89/100
391/391 [==============================] - 11s 28ms/step - loss: 3.3782 - accuracy: 0.6117 - val_loss: 3.6777 - val_accuracy: 0.5807
Epoch 90/100
391/391 [==============================] - 11s 28ms/step - loss: 3.3528 - accuracy: 0.6170 - val_loss: 3.6441 - val_accuracy: 0.5838
Epoch 91/100
391/391 [==============================] - 11s 27ms/step - loss: 3.3916 - accuracy: 0.6169 - val_loss: 3.6308 - val_accuracy: 0.5837
Epoch 92/100
391/391 [==============================] - 11s 27ms/step - loss: 3.3840 - accuracy: 0.6180 - val_loss: 3.5307 - val_accuracy: 0.5962
Epoch 93/100
391/391 [==============================] - 11s 28ms/step - loss: 3.3511 - accuracy: 0.6211 - val_loss: 3.6919 - val_accuracy: 0.5808
Epoch 94/100
391/391 [==============================] - 11s 28ms/step - loss: 3.3912 - accuracy: 0.6191 - val_loss: 3.6185 - val_accuracy: 0.5872
Epoch 95/100
391/391 [==============================] - 11s 28ms/step - loss: 3.3664 - accuracy: 0.6211 - val_loss: 3.6344 - val_accuracy: 0.5847
Epoch 96/100
391/391 [==============================] - 11s 28ms/step - loss: 3.3643 - accuracy: 0.6261 - val_loss: 3.6912 - val_accuracy: 0.5771
Epoch 97/100
391/391 [==============================] - 11s 28ms/step - loss: 3.3517 - accuracy: 0.6245 - val_loss: 3.6396 - val_accuracy: 0.5838
Epoch 98/100
391/391 [==============================] - 11s 28ms/step - loss: 3.3575 - accuracy: 0.6257 - val_loss: 3.6648 - val_accuracy: 0.5844
Epoch 99/100
391/391 [==============================] - 11s 28ms/step - loss: 3.3249 - accuracy: 0.6276 - val_loss: 3.6359 - val_accuracy: 0.5853
Epoch 100/100
391/391 [==============================] - 11s 28ms/step - loss: 3.3552 - accuracy: 0.6283 - val_loss: 3.7160 - val_accuracy: 0.5647
Test loss: 3.72
Test accuracy: 0.56
```

Fig2.

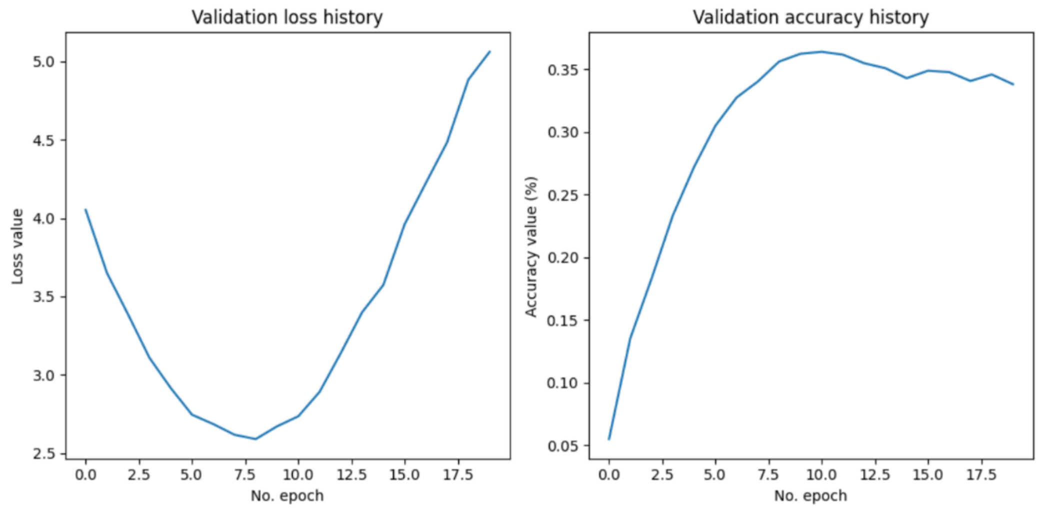
Validation loss history
Loss value
No. epoch
Validation accuracy history
Accuracy value (%)
No. epoch

Fig 3.

```
=================================================================
Total params: 4371844 (16.68 MB)
Trainable params: 4371844 (16.68 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
Epoch 1/20
391/391 [==============================] - 20s 22ms/step - loss: 4.3175 - accuracy: 0.0284 - val_loss: 4.0514 - val_accuracy: 0.0548
Epoch 2/20
391/391 [==============================] - 7s 18ms/step - loss: 3.8592 - accuracy: 0.0919 - val_loss: 3.6493 - val_accuracy: 0.1352
Epoch 3/20
391/391 [==============================] - 8s 21ms/step - loss: 3.4854 - accuracy: 0.1587 - val_loss: 3.3828 - val_accuracy: 0.1828
Epoch 4/20
391/391 [==============================] - 7s 17ms/step - loss: 3.1700 - accuracy: 0.2143 - val_loss: 3.1075 - val_accuracy: 0.2332
Epoch 5/20
391/391 [==============================] - 9s 23ms/step - loss: 2.8965 - accuracy: 0.2669 - val_loss: 2.9149 - val_accuracy: 0.2719
Epoch 6/20
391/391 [==============================] - 8s 20ms/step - loss: 2.6525 - accuracy: 0.3180 - val_loss: 2.7446 - val_accuracy: 0.3047
Epoch 7/20
391/391 [==============================] - 7s 19ms/step - loss: 2.4450 - accuracy: 0.3606 - val_loss: 2.6849 - val_accuracy: 0.3273
Epoch 8/20
391/391 [==============================] - 8s 20ms/step - loss: 2.2328 - accuracy: 0.4046 - val_loss: 2.6165 - val_accuracy: 0.3401
Epoch 9/20
391/391 [==============================] - 7s 17ms/step - loss: 2.0198 - accuracy: 0.4500 - val_loss: 2.5891 - val_accuracy: 0.3560
Epoch 10/20
391/391 [==============================] - 8s 19ms/step - loss: 1.8127 - accuracy: 0.4981 - val_loss: 2.6704 - val_accuracy: 0.3622
Epoch 11/20
391/391 [==============================] - 7s 17ms/step - loss: 1.6015 - accuracy: 0.5498 - val_loss: 2.7346 - val_accuracy: 0.3638
Epoch 12/20
391/391 [==============================] - 7s 19ms/step - loss: 1.3891 - accuracy: 0.5986 - val_loss: 2.8901 - val_accuracy: 0.3615
Epoch 13/20
391/391 [==============================] - 7s 17ms/step - loss: 1.1883 - accuracy: 0.6501 - val_loss: 3.1382 - val_accuracy: 0.3547
Epoch 14/20
391/391 [==============================] - 7s 19ms/step - loss: 1.0016 - accuracy: 0.6995 - val_loss: 3.3982 - val_accuracy: 0.3507
Epoch 15/20
391/391 [==============================] - 7s 18ms/step - loss: 0.8460 - accuracy: 0.7423 - val_loss: 3.5727 - val_accuracy: 0.3427
Epoch 16/20
391/391 [==============================] - 7s 18ms/step - loss: 0.7145 - accuracy: 0.7806 - val_loss: 3.9597 - val_accuracy: 0.3487
Epoch 17/20
391/391 [==============================] - 7s 17ms/step - loss: 0.6167 - accuracy: 0.8088 - val_loss: 4.2238 - val_accuracy: 0.3476
Epoch 18/20
391/391 [==============================] - 7s 18ms/step - loss: 0.5175 - accuracy: 0.8357 - val_loss: 4.4833 - val_accuracy: 0.3405
Epoch 19/20
391/391 [==============================] - 7s 17ms/step - loss: 0.4817 - accuracy: 0.8502 - val_loss: 4.8834 - val_accuracy: 0.3457
Epoch 20/20
391/391 [==============================] - 7s 17ms/step - loss: 0.4259 - accuracy: 0.8655 - val_loss: 5.0612 - val_accuracy: 0.3379
```

Fig 4.