Video Transcript

Title: A Practical Tutorial on Convolutional Neural Networks:
Subtitle: How Network Depth Affects Image Classification on CIFAR-10

Hello everyone.
My name is Sanaan Ashfaq, and this video is an additional explanation provided to give a better practical understanding of my Convolutional Neural Network tutorial. The complete and detailed written explanation is available in the full PDF report, and this video is meant to visually support that report.

In this tutorial, I explain what Convolutional Neural Networks are and how the depth of a CNN affects image classification performance. I demonstrate this using two different CNN models trained on the CIFAR-10 dataset using TensorFlow and Keras.

What is a Convolutional Neural Network

A Convolutional Neural Network, or CNN, is a type of deep learning model that is specially designed for image data. CNNs automatically learn important visual features such as edges, textures, shapes, and object parts directly from raw images. A CNN typically consists of convolutional layers, pooling layers, and fully connected dense layers.

Convolutional layers extract features using filters. Pooling layers reduce the spatial size of the features and help prevent overfitting. Dense layers perform the final classification using the features learned by earlier layers.

Deeper CNNs can learn more complex and hierarchical features than shallow networks, which often leads to better performance.

CIFAR-10 Dataset

For this project, I used the CIFAR-10 dataset. It is a widely used benchmark dataset for image classification. The dataset contains sixty thousand colour images of size thirty-two by thirty-two pixels. These images are divided into ten classes: airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck.

There are fifty thousand training images and ten thousand test images. The dataset was automatically loaded using the built-in Keras dataset loader.

Here, you can see examples of sample images from the training dataset.

Shallow CNN Architecture

The first model is a shallow Convolutional Neural Network. This model consists of a single convolutional layer with thirty-two filters, followed by a max-pooling layer. The output is then flattened and passed to a dense fully connected layer with sixty-four neurons. The final layer is a softmax output layer with ten neurons, one for each class.

This model contains approximately five hundred and twenty-six thousand trainable parameters. It was trained for ten epochs using the Adam optimiser and sparse categorical cross-entropy loss.

Deep CNN Architecture

The second model is a deep Convolutional Neural Network. This model consists of multiple convolutional blocks with increasing numbers of filters. The network includes two convolutional layers with thirty-two filters, followed by pooling, then two convolutional layers with sixty-four filters, followed by pooling, and finally one convolutional layer with one hundred and twenty-eight filters.

After feature extraction, the output is flattened and passed through a dense layer with one hundred and twenty-eight neurons. A dropout layer is also used to reduce overfitting. The final layer is a softmax output layer with ten neurons.

This deep CNN was also trained for ten epochs using the same optimiser and loss function as the shallow model, ensuring a fair comparison.

Training Results of the Shallow CNN

These figures show the training and validation accuracy and loss of the shallow CNN over ten epochs. The training accuracy improves initially but quickly reaches a plateau. The validation accuracy stabilises at around sixty-three percent. The loss curve also shows limited improvement after the early epochs.

Training Results of the Deep CNN

These figures show the training and validation accuracy and loss of the deep CNN over ten epochs. The deep CNN continues learning for more epochs and achieves a significantly higher validation accuracy of approximately seventy-five percent. The loss curve decreases more smoothly and reaches a lower final value, indicating better generalisation.

Model Comparison

This comparison clearly shows that the deep CNN outperforms the shallow CNN in both accuracy and loss. Increasing network depth allows the model to learn richer and more discriminative features from the images.

Confusion Matrix

This figure shows the confusion matrix of the deep CNN on the CIFAR-10 test set. Most of the predictions lie along the main diagonal, indicating that the model correctly classifies most images. Some confusion is observed between visually similar classes such as cats and dogs.

Example Predictions

This figure shows several example predictions produced by the deep CNN. Green labels represent correct predictions, while red labels represent incorrect predictions. The incorrect predictions mainly occur when the images are visually ambiguous or low in resolution.

Final Conclusion

In conclusion, this project demonstrates that deeper Convolutional Neural Networks are more effective at image classification than shallow networks. When both models were trained for the same ten epochs, the deep CNN achieved significantly higher accuracy on the CIFAR-10 dataset.

The full implementation, complete tutorial, saved models, and all results are available in the GitHub repository. This video was provided as an additional explanation to support the written report.

Thank you very much. Sanaan Ashfaq Sign out

GitHub Repository

GitHub Repository Link:
[https://github.com/SanaanAshfaq/cnn_depth_filter_cifar10]

Video

Video Link:
[https://drive.google.com/file/d/1f_-RBEhX6NhgMS_UGIibgss9b9fbdi-G/view?usp=sharing]