Jesus Ocampo

Professor: Patricia McManus

ITAI 1378 Comp Vision

10/10/2024

Reflective Journal

CNN Architecture

In this lab, I used a CNN with multiple convolutional layers, followed by pooling and fully connected layers. CNNs work well for images because they can automatically extract features like edges and textures. The model's convolutional layers helped identify these patterns, which is why CNNs perform better than regular neural networks for this task. The pooling layers reduced the size of the image, making the computations faster. Unlike regular neural networks that flatten the input, CNNs keep the spatial structure intact, so they're better for images.

Model Performance

The CNN achieved a validation accuracy of 92%. It performed well overall but had trouble with images where chihuahuas looked similar to muffins. Some misclassifications happened, especially with tricky angles or colors that were close to each other. Adding data augmentations like random rotation and flips improved it performance. This helped the model learn more and reduced overfitting. The accuracy from data augmentation was noticeable.

Comparison

Compared to the traditional neural network used before, This CNN did much better. The regular neural network couldn't handle image data as well because it flattens everything, losing spatial information. With CNNs, the layers can detect patterns in the image, making the model more accurate. The CNN took longer to train, but the improvement in accuracy was worth it. The

Jesus Ocampo

Professor: Patricia McManus

ITAI 1378 Comp Vision

10/10/2024

previous model had an accuracy of 75%, while the CNN hit 92%. Training time was longer, but the results were far better.

Challenges and Solutions

One issue I faced was the learning rate. Initially, the model didn't converge well. Lowering the learning rate helped it perform better. Overfitting was another problem. I fixed that by adding dropout layers and using more data augmentation. This balanced the model and improved validation accuracy. Organizing the data in the right format for PyTorch was also tricky at first, but once the paths were correct, the rest went smoothly.

Real-World Applications

CNNs like this are used for a lot of real-world tasks. For example, they're used in medical imaging to detect diseases and in self-driving cars for object detection. These models can analyze images faster and more accurately than humans. They're also used in retail, where customers can search for products using images instead of keywords. CNNs are powerful tools in fields like healthcare, transportation, and e-commerce.

Ethical Considerations

There are ethical concerns with image classification models. For example, if the training data isn't diverse, the model could be biased. This can lead to issues in applications like facial recognition, where the model may not work as well for certain demographics. Privacy is another

Professor: Patricia McManus

ITAI 1378 Comp Vision

10/10/2024

concern, especially in surveillance. It’s important to make sure these models are used responsibly and that they’re trained on fair and diverse datasets.

References:

- GeeksforGeeks. (n.d.). Convolutional Neural Networks (CNN) - Introduction to Deep Learning. https://www.geeksforgeeks.org/introduction-to-convolutions-using-python/.
- Microsoft Azure. (n.d.). Deep Learning for Image Classification. Retrieved from https://docs.microsoft.com/en-us/azure/machine-learning/tutorial-deep-learning.