Zaid Alayedy
ITAI 1378
Patricia Mcmanus
20 october/2024

L07 Chihuahua or Muffin with CNN

CNN Architecture:

1 - Briefly describe the CNN architecture and how it differs from the traditional neural network used in the previous workshop.

Convolutional Neural Networks (CNNs) are very useful for things such as image classification and quite different from other neural networks discussed at previous workshops. CNN architecture consists of fundamental components such as convolutional layer, pooling layers, and fully connected layers. Convolutional filters rotate over the input image to pick up local characteristics like edges and pattern which is spatially independent of the shape of the data. When we pool layers, this reduces the dimension of the data, which makes the model more scalable without overfitting. Instead, old school neural networks repackage the input as a 1D vector without retaining important spatial information, and depend entirely on layers that are completely connected (with every neuron wired to all neurons of the previous layer). CNNs are also more parameter efficient as they share weights in convolutional layers while traditional networks have to use separate weights for each connection. With this pattern, CNNs are able to see top-down hierarchical patterns in images, which is why CNNs are much better suited for image analysis than feedforward neural networks.

2- Model Performance:

- Observe and report on the model's performance, including accuracy and any interesting patterns in the misclassifications.

This performance was measured in terms of correctness and false categorisation. Upon training, it produced very good accuracy on the training and validation data. Convolutional layers were used to provide spatial features in the model, and classifications were erroneous when the classes had visually similar features, such as Chihuahuas being mislabeled as muffins. These mistakes were also more common in low-resolution or fuzzy pictures where essential details are unclear. There was also some overfitting, as the model outperformed the validation set only slightly less than the training set, possibly learning specific examples instead of generalizing well.

3- Comparison:

- Compare the CNN with the traditional neural network model in terms of performance and training time.

For the image classification, CNNs are usually much more efficient than the classic neural networks. CNNs specifically learn how to compute grids, like images by storing spatial associations via convolutional layers. So they can get better at picking up features such as edges and textures that will be useful in image recognition than traditional full connected networks, which lose spatial details by flattening the input data.

As for training time, CNNs can be slow to train because the convolutional and pooling layers are more complicated. But they are more parameter efficient since they share weights across the image. Neural networks rely on more parameters in conventional neural networks, for example, where each neuron is linked to each neuron of the adjacent layer. This renders fully coupled networks more inefficient with big high dimensional objects such as images, and can cause performance issues despite potentially shorter training times. To recap, CNNs require more training, but by storing spatial patterns they do a better job at image recognition.

4- Challenges and Solutions:
- Reflect on the challenges you faced during the lab and how you overcame them.

During the lab, I faced a few challenges that required problem-solving. One issue was code formatting, where indentation errors caused problems with commands like !git clone. I fixed this by carefully checking and correcting the indentation. Structuring and training the CNN was another challenge, as it was new to me. I struggled with selecting hyperparameters like learning rate and batch size but improved through trial and error, balancing underfitting and overfitting. Lastly, training the CNN was slow due to memory constraints, so I reduced the batch size and simplified the model to improve efficiency without losing too much accuracy.

5- Real-World Applications:
- Consider potential real-world applications of this type of image classification model.

CNN-based image classification models have various applications. In healthcare settings, they are also used to diagnose diseases on images such as X-rays or MRIs. CNNs also assist object recognition and navigating in autonomous vehicles by distinguishing road signs, pedestrians and hazards. They're also used in security systems to identify faces and track people, in stores to search products and use the visual search capabilities, and in agriculture to detect pests and track plant health. These models help make processes automated and accurate in various sectors.

6- Ethical Considerations:
- Discuss any ethical considerations regarding the development and deployment of such models.

There are ethical questions when it comes to the creation and implementation of CNN-based image classification models. A big one would be biased training data, which can produce unfair or biased results, for example in facial recognition or medicine. Also, privacy, because the models may violate privacy during surveillance or diagnosis. Also, there is the risk of data breaches, especially if it concerns sensitive medical data. And last but not least accountability, transparency of decision making and corrections to blunders or prejudices to be just and to avoid harm.

References:

Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition.