L07 Chihuahua or Muffin with CNN

Martin Demel

Department of Science, Technology, Engineering & Math, Houston Community College

ITAI-1378: Computer Vision Artificial Intelligence

Patricia McManus

October 10th, 2024

**Introduction / Summary of the Laboratory Session**

In this laboratory session, I will walk you through a session on how Convolutional Neural Networks differentiate between images of chihuahuas and muffins. Furthermore, I will compare it to the traditional Neural network used in the previous workshop and explain the benefits of Convolutional Neural Networks. The model in this laboratory work achieved an astonishing 96.88% after 10 epochs, significantly outperforming the traditional neural network used in a previous workshop. Overall, I very much enjoyed the objectives of this laboratory work, and I have learned key concepts of Convolutional Neural Networks. There are still some parts I must learn more about and get a deeper understanding but after completing the assignment I feel more confident than before this laboratory work.

**Understanding the CNN architecture**

With a Convolutional Neural Network, the architecture consisted of three layers, each followed by the ReLU activation function and MaxPooling. I was slightly unsure of the MaxPooling layers and after the research, I understood that MaxPooling reduces the spatial dimensions while preserving important features, thus reducing the computational cost.

Furthermore, the convolutional layers used filters at different scales ranging from 32 to 128 filters respectively to capture specific patterns in the images provided such as textures or possible edges in the images. The output was then flattened and passed through the two connected layers for classification purposes.

The key takeaway for me is that CNN differs from the standard neural network that was used in the previous workshop by the way the image was first flattened into a single vector before being let into the network. On the other hand, the CNN preserves the special relationship in the image through the convolutional layers which makes it ultimately more suitable for image classification.

*This brings me to the point that CNN is more suitable for image data, perhaps it was designed specifically for this purpose as to what I know so far.*

**Model Performance and Comparison**

During the training in the Convolutional Neural Network, I have learned that while both models improved over time, by performing the epochs, the Convolutional Neural Network leveraged its advantage of the convolutional layers to learn special features, which is not possible by the standard Neural Network.

**A collage of pictures of dogs and muffins

Description automatically generated**With 10 epochs, the Convolutional Neural Network achieved a validation accuracy of 96.88%. This tells us important information that the model generalized well to the data that has not previously been seen by the model. As you can see in Figure 1, the model performed exceptionally.

It is important to mention that the Convolutional Neural Network took longer per epoch. That is caused by the additional complexation of the convolutional layers. But still, it was faster in terms of epochs with better accuracy.

Figure 1 – Model Performance

**Challenges and solutions**

During my laboratory work, I have faced several challenges. The main challenge was to correctly determine the input size of the image for the Convolutional Neural Network. To overcome the challenge, I remembered that it is important to review the dataset of files provided for this laboratory work. After examining the images, I have further reviewed the Python code provided for the model architecture. I have also performed several tests with different dimensions like 64x64 or 224x224 but ultimately, I identified that images with dimensions of 128x128 pixels were the perfect and optimal selection for this laboratory work. By doing the experiments I have learned that the dimension of 128x128 pixels was a balance between maintaining image detail and the model could process the data effectively.

*This reinforced for me the importance of understanding the input data and how it can directly impact the model’s performance.*

Furthermore, as you will see in the html file attached to this laboratory work, I have faced a challenge with Index Error. After some review I understood that to fix that, I had to change the code by adding a condition to break the loop once the number of images exceeds the available grid slot.

**Real-world applications**

After seeing how Convolutional Neural Networks can outperform classical Neural Networks, I have to say that I can now see better potential in the Healthcare Industry. This is a key aspect if we think about how important is to have the most accurate identification. As the medical image analysis progresses over time, we will see exponential improvements in accuracy, time to deliver, and other key aspects.

Furthermore, as the computational power improves over time, we will also see improved object detection in autonomous vehicles. CNN can help in object detection and distinguish between pedestrians, cars, obstacles, or simple road work.

**Ethical considerations**

The biggest ethical consideration is data privacy. Proper consent and data anonymization must take place before the model uses the data to learn. Just recently, Meta, formerly Facebook, was accused of using Instagram pictures without prior consent from users.[1]

Furthermore, as I have already touched on in real-world applications, misclassification risk is another consideration that we must not overlook. A simple misclassification can lead to serious consequences. Just a simple wrong image classification can lead to a wrong diagnosis and ultimately a wrong treatment. This can be reassured by thorough validation and testing before deploying into production.

**Conclusion / Personal reflection**

This laboratory work was very exciting for me. There were times when I was confused and unsure how to exactly understand the processing of Convolutional Neural Networks but ultimately, I have found my way by reading some of the articles about Convolutional Neural Networks. [2,3]

This laboratory work further deepened my passion to learn more about the Convolutional Neural Networks and Healthcare applications. It has also impacted me on the professional side since data processing is part of my daily job.

**Resources:**

[1] Growcoot, M. (2024, May 14). *Meta is Using Your Instagram Photos to Train its ‘Amazing’ AI Image Generator*. PetaPixel. <https://petapixel.com/2024/05/14/meta-is-using-your-instagram-photos-to-train-its-amazing-ai-image-generator/>

[2] GeeksforGeeks. (2024, September 9). *Introduction to Convolution Neural Network*. GeeksforGeeks. <https://www.geeksforgeeks.org/introduction-convolution-neural-network/>

[3] Saha, S. (2023, April 8). A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way. *Medium*. <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>