Reflective Journal- Lab 07: Chihuahua or Muffin with CNN

1. **Introduction**:

* **Overview**: Lab 07 continued our exploration of the Chihuahua vs. Muffin workshop, but this time deploying a Convolutional Neural Network (CNN) instead of the traditional Neural Network from Lab 06. According to the notebook, CNN models are “particularly well-suited for image classification tasks because it can learn spatial hierarchies of features directly from the image data.”
* **Purpose**: This reflective journal will be covering building the CNN, loading data, training the model on the data, visualizing the results and then experimenting with changing parameters including modifying the learning rate, adjusting the number of epochs, and adding/ removing layers.

1. **CNN Architecture**:

* The architecture of the model consists of 3 initial, 2-D convolutional layers, comprised of a ReLU or Rectified Linear Unit function (which introduces non-linearity), and a Max Pooling 2D layer (reduces the spatial dimensions of the feature maps). After the high-level features are extracted by the convolutional layers, the fully connected layers work to flatten and then make the final classification decision, the class with the highest score being the model’s prediction.
* **Difference from NN**: Unlike CNNs, which have been specifically designed for image data and utilize convolutional layers to automatically learn relevant features, traditional NNs require manual feature engineering, which can be time consuming and less accurate. While CNNs exploit the spatial relationships between pixels in an image, NNs treat the input as a flattened vector, ignoring the spatial structure of the image, which can limit the model’s ability to learn complex patterns.

1. **Model Performance**: I ran the notebook “as is” initially and the model performed with a perfect validation accuracy of 1.0000, though after some consulting with Google Gemini now know this could be misleading due to overfitting or a small validation set. To have a clearer picture of our model’s accuracy, we could increase the size of the training data set or, according to viso.ai, use data augmentation which “involves artificially expanding the training dataset using various transformations like rotation, scaling, and flipping. This not only diversifies the training data but also helps the model generalize better to new data.”[1]
2. **Comparison**: Unlike the past lab’s NN performance, which depending on the parameters used, performed its best at a 95% accuracy, the CNN performed perfectly with 100% accuracy. The training time however did take much longer for the CNN, though we used more epochs initially than with the NN.
3. **Challenges and Solutions**: Personally, and admittedly unique to my current situation, having enough computational power to train the model. I have been away from Wi-Fi and have relied on my phone’s hotspot to complete the task but managed it without much additional problem.

* **Experimenting With Parameters**: The model performed perfectly with the default settings, so I knew adjusting them would most likely result in poorer performance. I changed the epochs from 10 to 3 and the validation accuracy did lower from 100% to 96.67%, as expected. Next, I modified the learning rate from 0.001 to 0.01, which seems to have led to faster initial progress, but I would like to experiment more with increasing the rate until it shows signs of instability. Lastly, I attempted to remove the third layer of the model’s convolutional layers but could not resolve a runtime error I was getting after repeated attempts. In the future, I would try to instead add a layer and see if this improved the model’s performance.

1. **Real-World Applications**: Today CNNs are used in a vast variety of fields including in text processing (sentiment analysis, topic categorization, language translation), medical imaging analysis (protein structure prediction, genetic data analysis, mutation identification, integration with reinforcement learning, combination with generative models), image classification, object recognition from videos, recommender systems, financial time series analysis, natural language processing(audio processing, speech recognition, sound classification, music composition), and human-computer interfaces.[1]
2. **Ethical Considerations**: With CNNs in particular and their constant evolution and increasingly high accuracy, always being conscious of inherited bias in image classification models which can lead to an unfair and disproportionate representation of certain groups over others has never been more important. Other ethical concerns include privacy and consent, often in the context of video surveillance whether by the government or private parties. Finally, the need for regulation and governance, which many organizations and businesses across the world have implemented by adopting AI frameworks in order to do so.
3. **References**:
4. Klingler, Nico. 2024 January, 2. *Convolutional Neural Networks (CNNs): A 2024 Deep Dive*. viso.ai: Deep Learning.

https://viso.ai/deep-learning/convolutional-neural-networks/