Jimmy Harvin & Charles Wallis
CS4375 - Professor Karen Mazidi
Dec 4, 2022

Image Classification

Sequential Models

Sequential model uses sequential data, where objects in a data set have a certain order. Examples could be DNA sequences, temperature sequence, or audio over time data. The sequential model that is provided by keras adds layers sequentially. These models will take in sequences of data, then output sequences of data. For dog breed identification this may not seem like a good idea intuitively, but since computers can read images as a sequence of numbers by pixels, this can be applied to classifying dog breeds.

RNN

RNN is a type of architecture that deals with sequences of data. CNN signals can be one-dimensional, two-dimensional, and three-dimensional depending on the domain where domain is defined by 'where it is mapped from' and 'where it is mapped to' and since a domain is just a temporary input to X, dealing with sequence data is basically the same as dealing with one-dimensional data. RNN is useful when we are trying to predict what may come after, such as voice or text recognition. LSTM is a type of RNN model that can provide a solution to the vanishing gradient problem, which is a problem that occurs during the process of updating the weight of networks, where the gradient continues to vanish over time and the state of past not having any effect on the current training.

CNN

In CNN, feature maps are extracted from a number of convolutional layers and dimensions are reduced through sampling to only gather important information about the feature map. This is an essential technology used in most computer vision fields, such as image classification. It requires a low computation amount and shows great performance compared to existing mlps. ResNet is often used, because it performs very well when transfer learning is applied to pre-trained weights. In our notebook, ResNet50 is used. It adds residual value for each block bounded by Convolution -> Batch Normalization -> ReLU. Training goes through hundreds of convolution layers yet it performs well because of the residuals. One problem that CNN models have, is that it needs big data input to train. That is why we chose to use ResNet, a pre-trained model, to perform transfer learning and use it to identify dogs.

Transfer Learning

Transfer learning is a technique for reusing a model that has already learned a particular task, to perform another one. These models tend to train faster than building a model from scratch, and often perform better too. This is possible because the network has been trained with a lot of features from many images, and the deeper the network becomes, it will learn more types of features. By utilizing transfer learning, we can bring ResNet directly and enter our data to create a classifier model. We can reset the top layer as we want, and it only trains using the data entered after it has been fully connected. After that, we can fine tune it to increase accuracy.

Performance

The performances of the algorithm itself when we ran the code seemed not quite right, they were very far from what we expected. It definitely is possible that there has been a mistake in how we entered our data to the model. The CNN ResNet50 Transfer learning model performed greatly, but it required a lot of computation amount and time. 5 epochs of it took 3 hours to run (could have been a google colab limitation), and we were not able to allocate enough time to do another run with more epochs for a better result. We definitely saw improvements for different epoch counts though, and it is certain that this model can perform very well after several training epochs.