Identifying Smartphone Models from Images Using Neural Networks.

By Vivian Carlson, Tyler Jedziniak, Aubhro Sengupta, Lukas Simchuk, and Kaitlyn Werho

Introduction

Smartphones have become an important tool used for completing everyday computing tasks. As the years go by, it has been more common to resell used smartphones on the secondary market. The ability to search for smartphone models based on images of that model can be useful for consumers who are looking to purchase used or new smartphones on the secondary market.

The purpose of our project is to identify a smartphone model based on an image of that smartphone model. We have decided to see if this task is possible by first using a simple neural network which will identify the likelihood of an image containing an iPhone model. We are using an image dataset from Kaggle in order to train and test this model.[4] After building this model, we have decided to test a more advanced Image recognition model known as a Residual Neural Network. We then compared the effectiveness of both models.

There has been research in the past related to smartphone models. Akua G. Biney and Harin Sellahewa from the University of Buckingham[1] have made an SVM classifier to identify smartphone models based on the images taken by the smartphone. Additionally, Lanh Tran Van, Sabu Emmanuel, and Mohan S Kankanhalli have created an SVM model for identifying smartphones based on Chromatic Aberration. [5] However instead of an SVM model, we hope to achieve higher recognition accuracies with a convolutional neural network model.

The dataset includes 10200 jpg images of smartphones labeled either an apple or “other” phone. The data is pulled from a prelabeled kaggle dataset and organized into a train and test set. The labels were pulled from a csv file that shows the model of each phone. The images are cropped to a uniform size and applied with a grayscale filter. Each image is composed into a tensor that will be passed into the convolutional network models.

Background

The baseline simple neural network model implemented a design by Tianmei Guo. [6] The architecture of the model features three convolutions and two pooling layers as shown in the figure below. The convolution is important as it filters the input image using a kernel. The output is fed into a ReLu layer to serve as the activation function. The activations are fed through a pooling layer which summarizes the feature map to a lower dimension. This convolution, ReLu, and pooling layer sequence is repeated two times, with a final convolution and ReLu layer. The final output feature map is linearly transformed to a 1x2 dimension prediction vector which represents the probability of the input image being either an Apple or “other” phone. To train the convolutional neural network, the model computes the cross entropy loss of the forward propagation. The gradients are then recomputed using an ADAM optimizer, through back propagation.

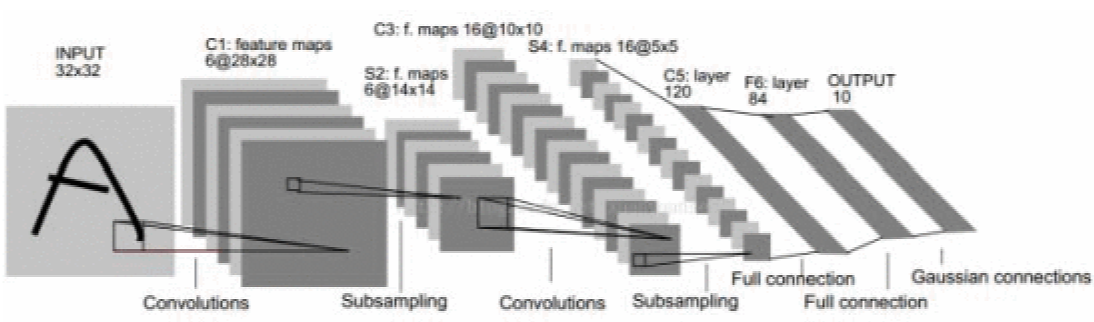

Figure:1 Simple Convolutional Neural Network [6]

A pre-trained residual neural network can be used to achieve even higher accuracy. ResNets mimic pyramid cells in the cerebral cortex by utilizing shortcuts in the neural network which allow features to skip layers. These skip weights are trained along with the weights in the convolutional layers. Skipping layers allows for less layers to be used and speeds training time. The pretrained ResNet is trained over ImageNet, a large database of 14 million hand annotated images. Since not all gradients will be used, on the mobile phone data, most of the layer can be frozen, where the gradients of the layer are restricted to change.

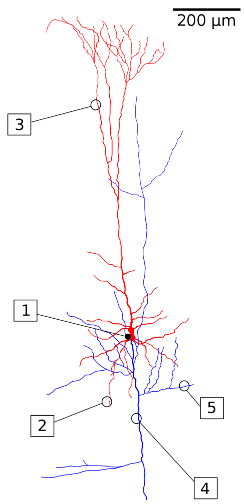

Figure 2: Pyramid Cell
Biological Analogue of ResNet[2]

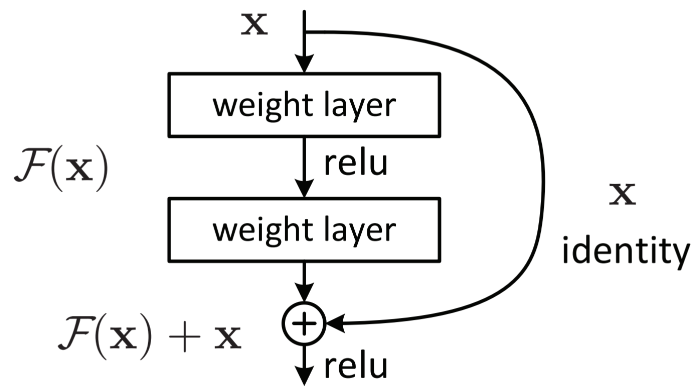

Figure 3: Simple ResNet Building Block Diagram[3]

Discussion and Results

Our simple neural network model functioned reasonably given its simplicity, with a final accuracy on our test set of 69.9%. This value was found to be around the maximum we could achieve by manipulating the learning rate and number of training epochs, and is likely very near the ceiling of the accuracy we could achieve with this network. Accuracy on the dev set varied slightly more between each training of the model, typically ending at around 68-70%, and loss consistently decreased at each epoch, implying that our learning rate was well tuned. This network could be improved by employing more advanced machine learning designs, but as we built this network to provide a simple baseline for comparison, we chose to avoid manipulating anything beyond epochs and learning rate.

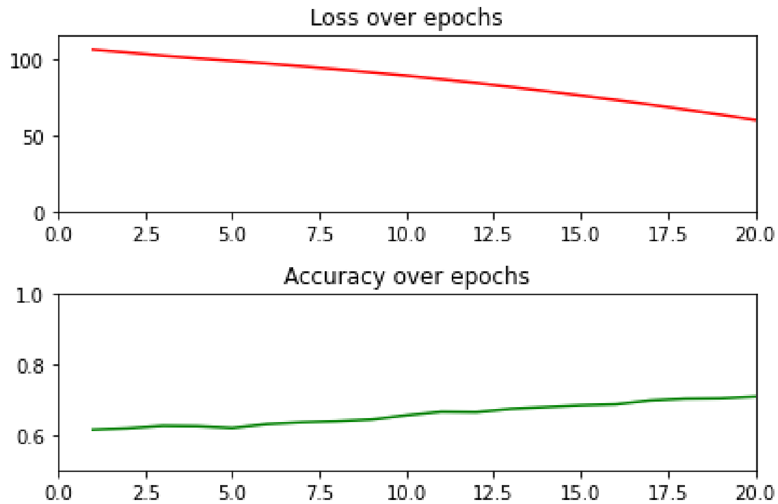

Figure 4: Loss and Dev accuracy for simpleNet

The other approach we took is using the Resnet model bundled in pytorch. We used the pretrained model and trained it on the dataset for 20 epochs. Since the net is too large to run gradient descent for all weights we froze all the layers except the last layer. We used cross entropy loss and an adam optimizer to tune the weights and biases of the last couple layers and got an 84% accuracy.

Conclusion

For our project, we analyzed a dataset of smartphone images to determine if we could make an effective neural network to detect a particular brand of smartphone. We decided to investigate this question by using neural networks to identify iPhones in a smartphone dataset. We decided to observe the effectiveness of a simple neural network and a residual neural network. The results of the simple neural network showed a roughly 65-70% accuracy. The results of the residual neural network showed 84% accuracy. There are many applications for detecting the type of mobile phone in an image. As previously noted, a website that resells used phones could verify the model of phone shown in the image. Another possible application could be identifying phones in criminal investigation footage.

Bibliography

1. A. G. Biney and H. Sellahewa, "Analysis of smartphone model identification using digital images," *2013 IEEE International Conference on Image Processing*, Sep. 2013.
2. Fabuio, Piramidal cell. Wikimedia, 2017.
3. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jul. 2016.
4. L. Jayawardena, "Mobile / Smartphone Images Dataset." Version 1, Nov-2021. <https://www.kaggle.com/lasaljaywardena/mobile-smartphone-images-dataset/version/1>
5. L. T. Van, S. Emmanuel, and M. S. Kankanhalli, "Identifying source cell phone using chromatic aberration," *Multimedia and Expo, 2007 IEEE International Conference on*, Jul. 2007.
6. T. Guo, J. Dong, H. Li, and Y. Gao, "Simple convolutional neural network on Image Classification," *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)(*, Mar. 2017.