Face Mask Recognition

Milestone 1 Deliverable

Team: Emma Ascolese, Stephanie Keene, Brennen Hogan, Ted Donegan

Milestone Accomplishments:

Our group started by making a subset of the original dataset that only contains “face_with_mask” and “face_no_mask.” This subset of data was used for the first milestone, since we were focused on solely making the network identify these two categories. After the created the augmented data set, we read in the dataset and split up into a training set and testing set. The training set contained 75 percent of the data and the testing set contained the other 25 percent. Each data entry contains x and y coordinates which form a box around an individual’s face. We used this data to crop the faces from the complete images. We then resized the cropped images to be a standard size (224x224). We could then feed these standardized images to our neural network. We started by using the resnet34 pre trained neural network. Resnet is a deep convoluted neural network trained for facial image recognition, and has been proven to be more accurate and easy to train than its counterparts (Kaiming). After making our model, we trained the neural network to distinguish between the presence or absence of a mask using the selected dataset. After training for 5 epochs, with each epoch containing 430 entries, our average total loss of 0.034699 and 0.049615 for our test data. Lastly, we provided a means of visualizing the training dataset and the CNN guesses. We do this by displaying the full images with each of the assigned labels.

Design Decisions:

We chose to start by using the pre-trained neural network resnet34. This is a deep neural network, which has proven to be more effective than standard neural networks. By using this model as a starting point, we were able reach a low average training loss in fewer epochs. (Kaiming)

We decided to use the CrossEntropyLoss loss function since it was designated as most effective in classification problems. For the first milestone, this loss function may not be the best choice, but this will future-proof our design as we add more classifications. (Pytorch.org)

Based on train and error testing, a learning rate of (number here based on results) seemed to produce the best results for the resnet34 neural network. As we add more classifications to the neural network, we will explore other pre-trained models, loss functions, and learning rates.

Results:

Results can be found in the attached lr0.005output1.txt and lr0.005output2.txt files.

The output from the first two runs with 5 epochs and a training data set containing 75% of our data ran with a test data average total loss of 0.034699 and 0.049615. This proves that our CNN is able to classify both mask wearing and non mask wearing images on the scale 224 by 224. Running with more epochs or a deeper resnet could yield better results for future runs. The training error showed a clear downward trend during the later epochs, proving that the model became more accurate with more training runs.

Questions/Going Forward:

- How should we approach the mask worn improperly problem?
    - Proposed solution: make a second CNN which identifies mask worn properly/improperly that only runs if the first CNN has identified that a mask is being worn
    - Alternative idea: add a third classification on the original CNN for "mask_worn_improperly"
        - How would we use this idea and also the Sigmoid function, that only accepts 1s and 0s?
- How should we leverage Cuda?
    - Training currently takes around 2 hours for 5 epochs
    - Running the model for more epochs could improve our accuracy, thus we would benefit from leveraging Cuda
- Should we explore other pre-trained facial recognition CNN models?
    - Resnet has deeper models such as resnet101 and resnet152
        - These also have the trade off of longer time for training

Works Cited

He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep Residual Learning for Image Recognition." 25 Dec. 2015. Web. 16 Oct. 2020.

Pytorch.org, "CrossEntropyLoss." Web. 16 Oct. 2020.