Group Proposal Project X

Problem Statement

The main aim of this project is to detect photoshopped photos from real ones. This research is part of a broader effort to improve digital image forensics; a field dedicated to detecting the authenticity of images. Earlier this year Adobe and UC Berkley published a paper in which they proposed a Neural Network architecture to detect photoshopped images. Previous research on image manipulation detection involved detecting changes in original images that were obtained from techniques like splicing, cloning, and removal, whereas Adobe's effort focuses on the Face Aware Liquify feature in Photoshop because it's popular for adjusting facial features, including making adjustments to facial expressions with user-friendly operations like "'increase nose width' and 'decrease eye distance'." They also mention that while GAN based methods to generate false images like DeepFakes have garnered a lot of attention, most fraudulent images that practitioners in digital image forensics deal with are relatively small manipulations created from tools like Adobe Photoshop. They used a special type of CNN (see Network) to solve the problem, and were able to achieve 90% accuracy on their test set, 50 original images and edits to them made by a professional artist. Although that is commendable, the small size of the test set and the fact that their accuracy was driven by their ability to predict original images (96% compared to 84% for photoshopped images) leaves room for improvement and further validation. Therefore, in our project we would like to validate the model that they published to github and use other networks / frameworks to attempt to perform better than their model on a dataset we selected.

Dataset

We aim to use a dataset from Kaggle that is created by the Computational Intelligence and Photography Lab, Yonsei University. Here is a link to the dataset. https://www.kaggle.com/ciplab/real-and-fake-face-detection. The dataset contains expert-generated high-quality photoshopped face images. The photoshopped images are composites of different faces, separated by eyes, nose, mouth, or whole face.

Dataset Size

Inside the parent directory, training_real/training_fake contains real/fake face photos, respectively. Is it a large enough dataset since there are 1000 real and 1000 photoshopped images. The photographs have a 2000Mb high quality images.

Network

Adobe trained a Dilated Residual Network variant (DRN-C-26) a type of CNN to recognize photoshopped images. After testing their network on our dataset, we plan to experiment with transfer learning using the following three.

1. VGG (e.g. VGG16 or VGG19).
2. GoogLeNet (e.g. InceptionV3).
3. Residual Network (e.g. ResNet50).

Framework

We will use pytorch python framework to implement the network to better understand and improve skills in pytorch.

Network Performance Metrics

We aim to utilize different performance measures like training loss, validation loss, confusion matrix, and , f1-score in addition to accuracy to determine the performance of network to get a clear idea of where our model(s) can be improved.

Schedule:

10/29-11/6: Begin preprocessing images for analysis (as noted previously they are high quality, so will need to conduct feature engineering) and conduct a literature review for all networks being tested
11/6-11/13: Finish image preprocessing and network literature review
11/13-11/20: Implement models
11/20-End of Semester: Finish Project.

References

- https://github.com/PeterWang512/FALdetector
- https://www.kaggle.com/ciplab/real-and-fake-face-detection?fbclid=IwAR3yvVXnC1ZZDEeoHLH_GdTL6XAL-21-C9D8-c-13OqSq2G-bWOVfbavdK4
- https://storage.googleapis.com/openimages/web/index.html
- https://arxiv.org/abs/1906.05856
- https://machinelearningmastery.com/how-to-use-transfer-learning-when-developing-convolutional-neural-network-models/