EECS 6322 Paper Reproduction Project Proposal

Paper: *CNN-generated images are surprisingly easy to spot... for now*

Wang, Sheng-Yu, et al. "CNN-generated images are surprisingly easy to spot... for now." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020.

https://openaccess.thecvf.com/content_CVPR_2020/html/Wang_CNN-Generated_Images_Are_Surprisingly_Easy_to_Spot..._for_Now_CVPR_2020_paper.html

Summary

In this paper, the threat of visual disinformation, spefically, the CNN-generated fake image is discussed. The author implemented and trained the CNN models and generated images by the models. The forensics models, which are able to detect fake images are trained on the CNN-generated images. During the evaluation, their forensics models shows a great amount of generalization to other CNN synthesis methods, which means the fake images generated by other methods can be detectable as well. They also introduce a new dataset and evaluation metrics to detect CNN-generated images, and explore the effects of pre-processing and post-processing on the detection.

Deep learning framework

They use PyTorch as their main Deep Learning framework.

Datasets required

The dataset is consist of CNN-generated fake images generated by 11 different CNN-based image generator models. And all the required datasets, including test dataset, training dataset, and validation dataset are all open-source and available to download.

1. Test Dataset: It contains images from 13 CNN-based synthesis algorithms, including the 12 testsets from the paper and images downloaded from whichfaceisreal.com.

   https://drive.google.com/file/d/1z_fD3UKgWQyOTZIBbYSaQ-hz4AzUrLC1/view
2. Training Dataset: It contains images from either LSUN or generated by ProGAN.

   https://drive.google.com/file/d/1iVNBV0glknyTYGA9bCxT_d0CVTOgGcKh/view
3. Validation Dataset: It contains held-out ProGAN real and fake images.

   https://drive.google.com/file/d/1FU7xF8Wl_F8b0tgL0529qg2nZ_RpdVNL/view

List of experiments to be reproduced

To reproduce this paper, the following experiments will be expected to implemented:

1. The implementation, training and hyperparameters tuning of GAN-based image synthesis models. The models can generate fake images from the given image dataset.
2. The implementation, training of an "real-or-fake" binary classifier to distinguish between real or generated fake images, to evaluate the GAN-based image synthesis models.
3. The evaluation of robustness of binary classifier on detectability of synthetic images. This part will explore to what extend some post-processing techniques, including resizing and blurring, can affect the detectability of synthetic images.

Expected compute resources

As the size of training, testing and validation datasets are 89GB, the experiments would be run on Google Colab using free GPU resource. If more computing resources could be necessary, we will use NVIDIA Tesla P100 on Google Colab Pro or Compute Canada. https://ccdb.computecanada.ca/.