Application of General Adversarial Networks in Smoothing Handwriting

1 Master of Science in Data Science, Southern Methodist University,  
Dallas, TX 75275 USA

**Abstract.** – Dysgraphia is a mental disorder that causes poor writing skills. Application of Generative Adversarial Networks can be used to improve this skill.

1 Introduction

* **Problem Statement:** Because of the mental and motor disorder, people with Dysgraphia have poor writing skills. In addition to that, since so much focus is put on writing, people with dysgraphia also have a hard time expressing themselves on paper. However, so much development has happened in the world of machine learning, that there is a solution. The Neural Network, Generative Adversarial Network, general used for processing images can take bad handwriting and present it in a realistic handwritten text.

2 Literature Review

**Dysgraphia:** This is a mental and motor disorder which affects skills such as writing and spelling. Since people with Dysgraphia focus so much on writing, they tend to also find it difficult to express their ideas on paper. Dysgraphia is normally diagnosed early at the ages of five and six. The symptoms are:

* Bad handwriting
* Bad spelling
* Slow writing
* Poor formation of letters
* And poor spacing of letters effectively

**Generative Adversarial Networks (GANS)**: This is a Neural Network used for unsupervised machine learning. Developed by Ian Goodfellow in 2014, GANS can be simplified into two parts, the generator and the discriminator. The main idea is that the generator creates fake data, and the discriminator validates to see if its real or fake. The goal for the generator is to fool the discriminator. Since GANS are typically used in processing many images, it can be applied to improving bad handwriting.

**Hypothesis:** Having Dysgraphia does not mean the person isn’t clever, they just need extra help in handwriting skills. Consequently, the solution to challenges with having Dysgraphia can be solved by Generative Adversarial Networks. GANS will be able to take bad handwriting from a document and present it in a better realistic handwritten text. Solving this problem will help people with Dysgraphia focus more on expressing their ideas rather than how to write.

3 Methods

* Methods utilized for analysis
  + Interviews with Dr John Santerre and David Josephs
  + Decomposition of final desired product
* Initial methods we may use based on preliminary analysis
  + Generative Adversarial Network (GAN)
    - Two neural networks compete
    - Could merge samples
  + Style transfer
    - Famous example of applying artist style to photos
    - Would be used to merge samples
  + Auto encoder
    - Similar to principal components analysis (PCA)
    - More aggressive dimensional reduction
  + Singular value decomposition (SVD)
    - Linear algebra method
    - Sometimes helpful with image manipulation
  + Recurrent neural network (RNN)
    - Known applications in handwriting recognition
    - More difficult than GAN; will explore only as a backup option
  + Distributed computing
    - Used to minimize processing time
    - Explore Spark and Dask
    - Choose one to use
* Data sources that were used. Currently investigating:
  + MNIST – collection of handwritten digits in image format
  + UCI Machine Learning Repository - Letter Recognition Data Set
  + Chars74K dataset - Character Recognition in Natural Images
  + Data Augmentation – technique to expand existing datasets if needed

4 Results

* Demonstrate effectiveness of handwriting smoothing
* Identify performant models
* Produce a Python package other developers can use
* Release a command line utility any one can use to smooth handwriting
* Scope restricted to:
  + Printed handwriting
  + Numbers, upper, and lower case letters
  + Black and white samples
  + Texture of a standard ball-point ink pen
* Additional scope if time permits:
  + Punctuation
  + International characters from other languages
  + Color samples
  + Textures such as pencil, marker, paint, or crayon

5 Discussion

As explained in the introduction, dysgraphia is a learning disability which strictly affect the ability to write properly. We took an interest in finding a better way to help people with this condition communicate better when writing. In this project we would build an application that would take an image input of a writing and the application should correctly replace the writing with a better looking writing that is similar to the input but would be clearly spelled and written.

To tackle this, we considered 2 different approach. The first is GAN Generative Adversarial Networks which a generative deep learning neural network model that supposed to learn to generate a realistic new samples of the input. The second is the autoencoders which are also a generative deep learning models that would instead of generating new samples, would just reconstruct the input. We are going to explore GAN more because we believe it would make a better prediction model than the autoencoders. A GAN model has 2 main components, the generator and the discriminator. A discriminator is a classifier that tries to classify if a data is a real data or if the data was created by the generator. The generator learns to create a fake data and gets feedback from the discriminator.

Unfortunately, there no objective/loss function used in training a GAN model. There are no ways of checking the progress or efficiency of the model from a loss function. Models can only be evaluated by manually inspecting the generated image or using quantitative measures like inception score, boundary distortion. The evaluation process would be used at the same time in other to produce good results. There are other quantitative measures that could be used in evaluating a GAN model. We would explore the best option in our model building.

After the success of this Capstone project, we hope to make the application open source so that anyone that needs it would be able to have access to it. Finally, we do not think that there are ethics impact or violations that would be encountered for this project.

6 Conclusion

Due to the impacts of dysgraphia on children and adults, we are hoping to build an application that would help people communicate better. The application would create a better-looking handwriting from an original image of a person’s handwriting. We believe this would make life a lot easier for people with this learning disorder.

Acknowledgments. Jacquelyn Cheun, PhD. – Capstone Professor

References

1. Renzhuo Wan, Shuping Mei, Jun Wang, Min Liu and Fan Yang. Multivariate Temporal Convolutional Network: A Deep Neural Networks Approach for Multivariate Time Series Forecasting. Electronics 2019, 8(8), 876; <https://doi.org/10.3390/electronics8080876>
2. Shengyu Zhao IIIS, Tsinghua University and MIT Zhijian Liu MIT Ji Lin MIT Jun-Yan Zhu Adobe Research Song Han MIT. Differentiable Augmentation for Data-Efficient GAN Training. <https://arxiv.org/pdf/2006.10738.pdf>
3. Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler and Sepp Hochreiter. GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. <https://arxiv.org/pdf/1706.08500.pdf>
4. Phillip Isola Jun-Yan Zhu Tinghui Zhou Alexei A. Efros. Berkeley AI Research (BAIR) Laboratory, UC Berkeley. Image-to-Image Translation with Conditional Adversarial Networks. <https://arxiv.org/pdf/1611.07004.pdf>
5. Andrew Brock Heriot Watt University, Jeff Donahue DeepMind, Karen Simonyan DeepMind. LARGE SCALE GAN TRAINING FOR HIGH FIDELITY NATURAL IMAGE SYNTHESIS. <https://arxiv.org/pdf/1809.11096.pdf>
6. Yasin Yazıcı Nanyang Technological University (NTU), Chuan-Sheng Foo Institute for Infocomm Research, Stefan Winkler National University of Singapore (NUS), Kim-Hui Yap Nanyang Technological University (NTU), Georgios Piliouras Singapore University of Technology and Design, Vijay Chandrasekhar Institute for Infocomm Research. UNUSUAL EFFECTIVENESS OF AVERAGING IN GAN TRAINING. <https://arxiv.org/pdf/1806.04498.pdf>
7. Han Zhang, Zizhao Zhang, Augustus Odena, Honglak Lee Google Research. CONSISTENCY REGULARIZATION FOR GENERATIVE ADVERSARIAL NETWORKS. <https://arxiv.org/pdf/1910.12027.pdf>
8. Hao-Wen Dong · Yi-Hsuan Yang. May 2020. Towards a Deeper Understanding of Adversarial Losses under a Discriminative Adversarial Network Setting. <https://arxiv.org/pdf/1901.08753.pdf>
9. David Ha, Douglas Eck. A Neural Representation of Sketch Drawings. <https://arxiv.org/abs/1704.03477>
10. Adrian Rosebrock on August 24, 2020. Handwriting recognition with OpenCV, Keras, and TensorFlow. <https://www.pyimagesearch.com/2020/08/24/ocr-handwriting-recognition-with-opencv-keras-and-tensorflow/>
11. Omar Mohammed, Gerard Bailly, Damien Pellier, Univ. Grenoble-Alpes. Style Transfer and Extraction for the Handwritten Letters Using Deep Learning. <https://arxiv.org/pdf/1812.07103.pdf>
12. Bo Chang∗ Qiong Zhang∗ Shenyi Pan Lili Meng, University of British Columbia. Generating Handwritten Chinese Characters using CycleGAN. <https://arxiv.org/pdf/1801.08624.pdf>
13. Bharath Narasimhan University of Massachusetts Amherst. Calligraphy Style Transfer using Generative Adversarial Networks. <https://people.cs.umass.edu/~bnarasimhan/Cast.pdf>
14. Alex Graves Department of Computer Science University of Toronto. Generating Sequences With Recurrent Neural Networks. <https://arxiv.org/pdf/1308.0850>)
15. Piotr Teterwak Aaron Sarna Dilip Krishnan Aaron Maschinot
16. David Belanger, Ce Liu, William T. Freeman Google Research. Boundless: Generative Adversarial Networks for Image Extension. <https://arxiv.org/pdf/1908.07007.pdf>

Appendix:

Leave here if needed for additional information.