

On the Use of K-Competitive Networks for Writing Style Transfer

Luca Leeser
Cornell University
ll698@cornell.edu

Ryan Butler
Cornell University
rjb392@cornell.edu

Yuji Akimoto
Cornell University
ya242@cornell.edu

Cameron Ibrahim
Cornell University
cai29@cornell.edu

ABSTRACT

Writing style is a key component of the quality of any on-line review - a detailed and informative review is only useful insofar as it is able to retain a reader's attention. In this paper, we explore the use of k -competitive layers in convolutional neural networks to classify the author of a given piece of text. We then apply the techniques used in neural style transfer for images to transfer the writing style of one author onto the content written by another. Given sufficient examples of well-written reviews, our techniques can be used to globally improve the quality of reviews on Yelp.

1. INTRODUCTION

Neural style transfer [2] has proven very successful at transferring the artistic style of an image onto the content of another. Due to the layered structure of convolutional neural networks, it was observed that later layers in the network produce feature representations of the content of an image, while the correlations between these features give a good representation of artistic style. More specifically, given a content image and style image, the algorithm treats the input to the network as a variable and simultaneously minimizes the ℓ_2 norm between the input and content image's feature representations, and the input and style image's style representations.

Figure 1: Cornell's Gates Hall stylized in the form of Vincent Van Gogh's *Starry Night*.



Modifications that optimize computational time [4], preserve the texture of the style image [3], and preserve facial structures in the content image [5], among others, have further advanced the art of neural style transfer for images.

However, applying similar techniques to text has proven difficult for a number of reasons. Firstly, textual data lives in an extremely sparse space, and therefore it is difficult to iteratively produce better stylized text while avoiding phrases that make no sense. Secondly, style and content are much more closely linked in text than they are in images - the same writer may write two passages of text in a completely different tone, yet maintain the same underlying style in a manner that is very subtle.

In this paper, we explore the use of k -competitive convolutional layers, a generalization of k -competitive layers designed for autoencoders in [1], which in turn are based on k -sparse layers as defined in [10]. These layers are integrated into a word-level convolutional neural network [6], and we perform a qualitative analysis on the quality of content and style representations generated by these networks.

2. RELATED WORK

3. PRELIMINARY ANALYSIS

The majority of implementations of style transfer for images use the VGG-19 [12] architecture to generate content and style representations of an image. This network was trained to classify between 1000 images of the ImageNet challenge dataset [11] and therefore is able to generate informative feature representations of a very large variety of objects. Given the resources required to train a similarly robust classifier for text, we first explore the tradeoffs of using an underlying network that is trained to discriminate between a much smaller number of classes by training a VGG network with the CIFAR-10 dataset [7] (here, we use pre-trained weights provided by [8]). Although this modified network is far less generalizable in terms of classification, we found that the content representations provided by later layers of the network were still informative enough to reconstruct the original image, even for images containing objects not included in the 10 classes of CIFAR-10.

4. MODEL

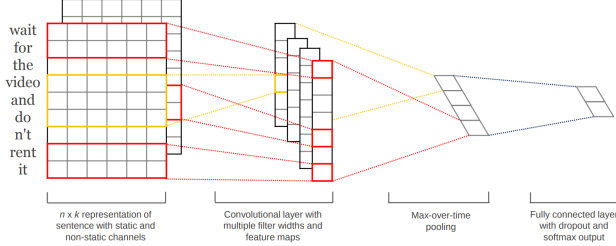
4.1 Convolutional Neural Networks for Text Classification

Although recurrent neural networks have proven adept at language recognition tasks, evaluating the overall content and style of written text is a task that is less dependent

on sequential information, and a hidden state is likely too small to encode enough information for our purposes. We believe that a convolutional neural network allows us to take advantage of the success of style transfer for images, in contrast to aforementioned approaches that treat style transfer as a machine translation problem.

As the baseline for our comparisons we use a word-level convolutional neural network as designed by Yoon Kim in [6].

Figure 2: Structure of the baseline convolutional neural network model.



4.2 K-Competitive Convolutional Layers

It has been observed that encouraging sparsity when learning representations acts as a capable regularizer that improves performance on classification tasks. In [10], Makhzani and Frey designed the k -sparse autoencoder and demonstrated its effectiveness as a preprocessing step on image data. In the feedforward phase, sparsity is used as the sole non-linearity: after computing the activations $z = W^T x + b$, only the largest k entries of z are kept and the rest are set to 0.

Since textual data often has high-dimensionality, sparsity, and power-law distributions [1], autoencoders for text have often learned trivial features such as proper nouns specific to a certain passage. Chen and Zaki made a modification to k -sparse layers by introducing competition in the place of truncation: after computing the activations $z = \tanh(W^T x + b)$, only the k entries with the largest absolute value are kept non-zero. These k most “competitive” neurons are then augmented (with an amplification constant) with the **energy** of the “losers” that were made inactive, where the energy of a subset of neurons H is defined as $\sum_{h_i \in H} |z_i|$.

Here, we propose a generalization of k -competitive layers to convolutional layers.

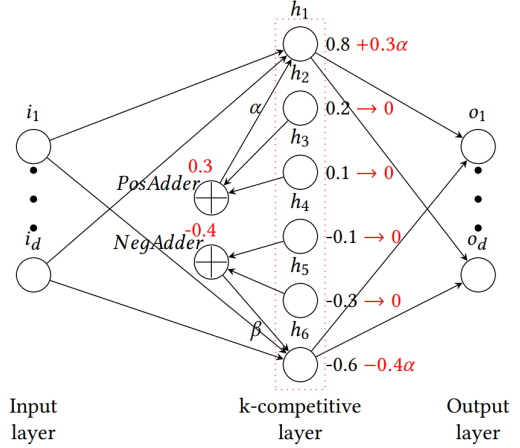
5. EXPERIMENTS

5.1 Datasets

We use the most prolific reviewers on Yelp¹ to generate a large amounts of text for various styles of writing. Although the reviews for a frequent user of Yelp may span a number of years, we make the assumption that a user’s writing style remains constant over that period of time. This may not strictly be the case as language, particularly of the colloquial variety that is often found online, evolves over time and certain words serve as clear indicators of certain time periods. However, dealing with a time-varying vocabulary

¹<https://www.yelp.com/dataset/challenge>

Figure 3: A fully connected k -competitive layer.



set is an extension that we have not yet considered. After identifying the most frequent reviewers, we tokenize all of their reviews using NLTK [9] to generate a corpus.

5.2 Hyperparameters and Training

5.3 Comparison of Architectures

6. RESULTS

7. CONCLUSION

8. FUTURE EXTENSIONS

9. ACKNOWLEDGMENTS

The authors would like to thank Professor Thorsten Joachims at the department of Computer Science at Cornell University for his guidance, and their fellow members of the Cornell Data Science NLP Research Team for their support.

10. REFERENCES

- [1] Y. Chen and M. J. Zaki. Kate: K-competitive autoencoder for text. *KDD*, 2017.
- [2] L. A. Gatys, A. S. Ecker, and M. Bethge. A neural algorithm of artistic style. *CoRR*, abs/1508.06576, 2015.
- [3] L. A. Gatys, A. S. Ecker, and M. Bethge. Texture synthesis and the controlled generation of natural stimuli using convolutional neural networks. *CoRR*, abs/1505.07376, 2015.
- [4] J. Johnson, A. Alahi, and F. Li. Perceptual losses for real-time style transfer and super-resolution. *CoRR*, abs/1603.08155, 2016.
- [5] P. Kaur, H. Zhang, and K. J. Dana. Photo-realistic facial texture transfer. *CoRR*, abs/1706.04306, 2017.
- [6] Y. Kim. Convolutional neural networks for sentence classification. *CoRR*, abs/1408.5882, 2014.
- [7] A. Krizhevsky. Learning multiple layers of features from tiny images. Technical report, 2009.

- [8] S. Liu and W. Deng. Very deep convolutional neural network based image classification using small training sample size. 2015.
- [9] E. Loper and S. Bird. Nltk: The natural language toolkit. In *In Proceedings of the ACL Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics*. Philadelphia: Association for Computational Linguistics, 2002.
- [10] A. Makhzani and B. J. Frey. k-sparse autoencoders. *CoRR*, abs/1312.5663, 2013.
- [11] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, 115(3):211–252, 2015.
- [12] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014.

APPENDIX