Writing Style transfer

Deep Generative Adversarial Autoencoders

Ayush Singh, Ritu Palod
CSYE 7245 Fall ‘17
Northeastern University

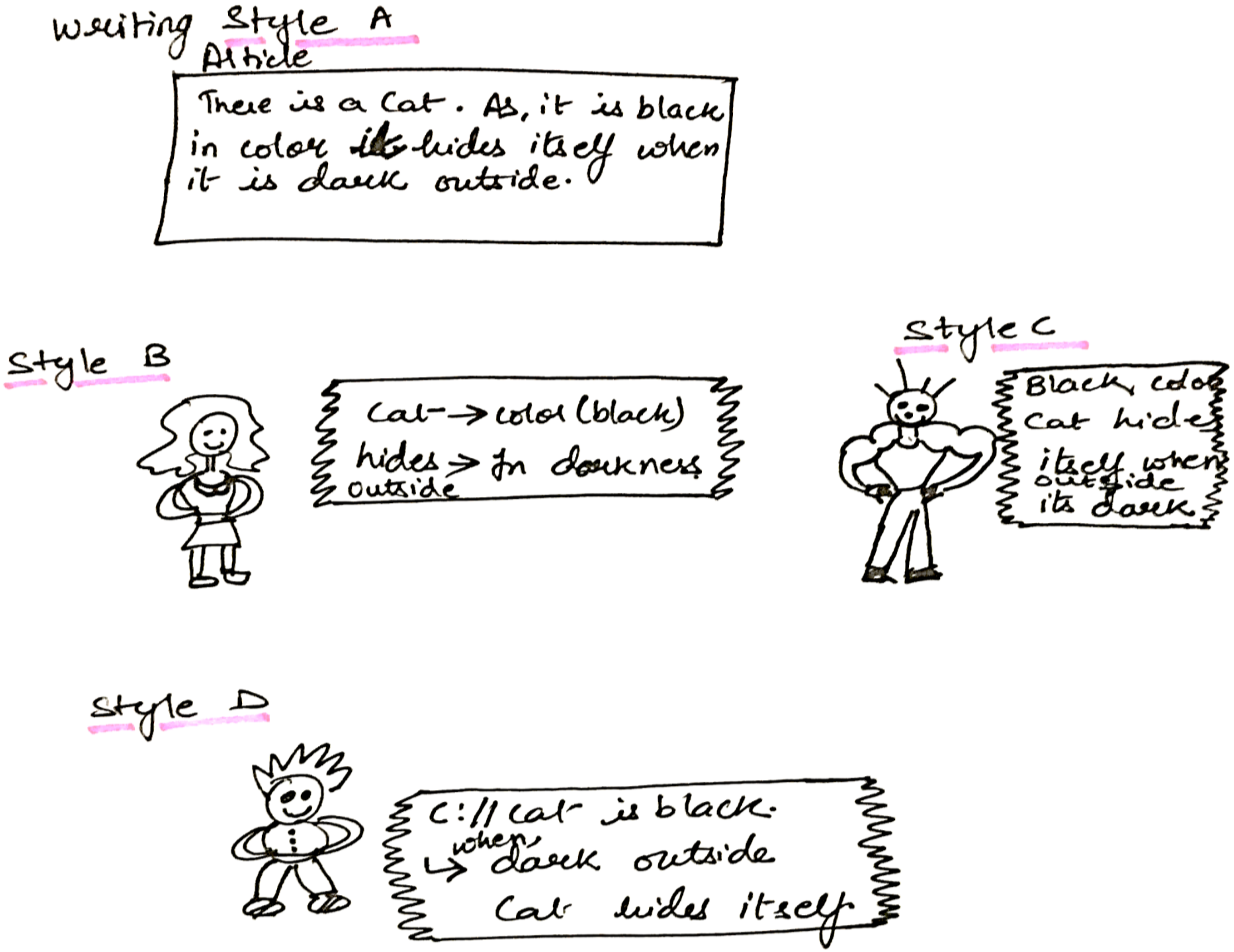
Progress Report

Overview of Presentation

- Introduction
- Related Work
- Intuition
- Approach
- Results
- Future Work
- References

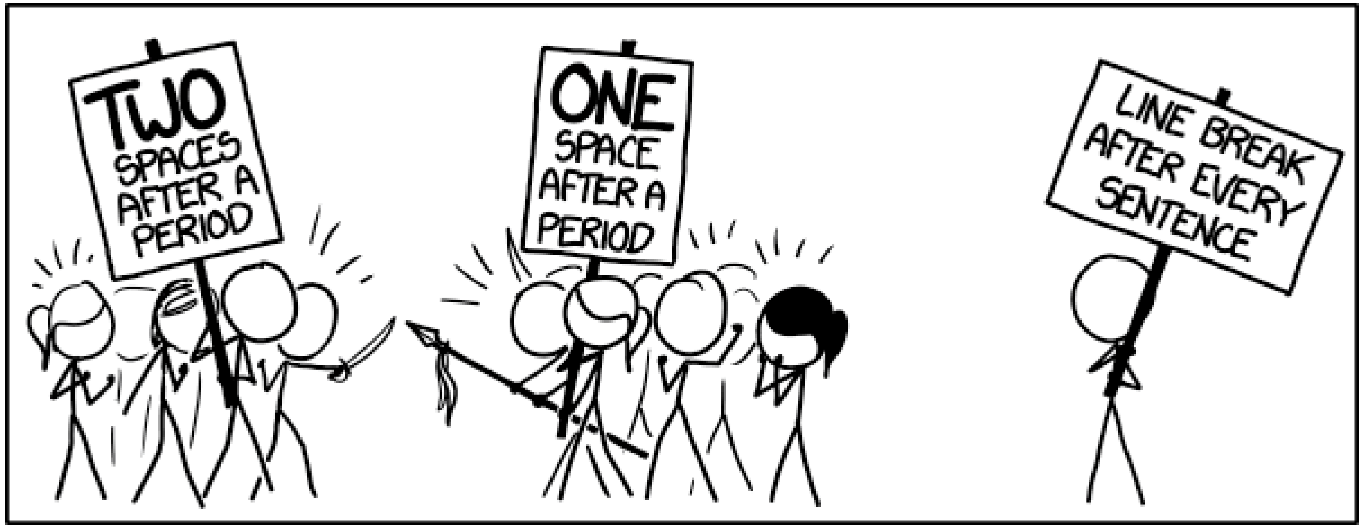
Introduction

- Different forms of writing a sentence leads to a style variation, which although enhances one’s creativity, is troublesome when dispersing information to a large audience.
- Everyone has a different style of reading and writing, apparently it all boil downs to the way their mind understand things.
- When there are multiple formats (internet era) easily available, one’s mind tends to read and understand it in its own way.



Problem

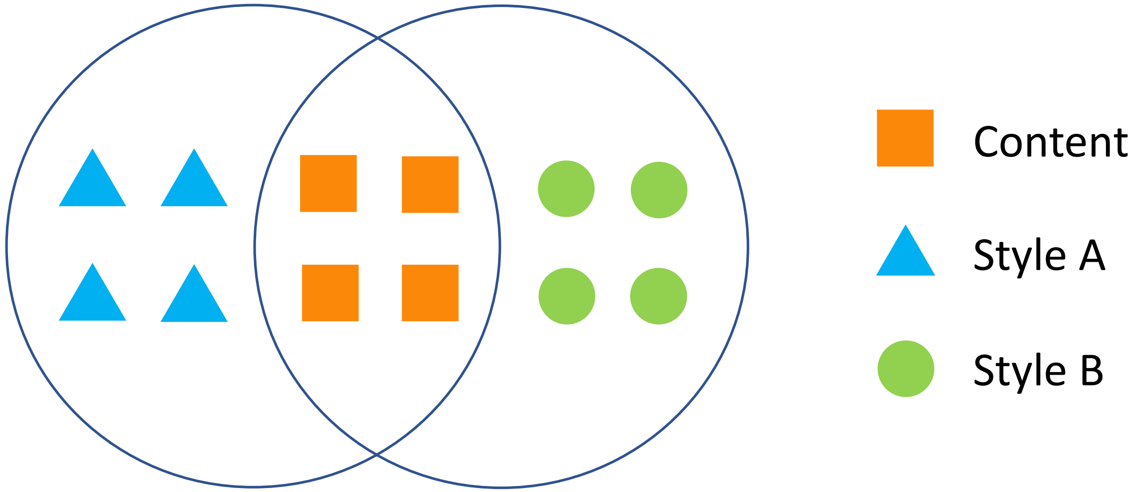
- Expressing in language is subjective
- People find it hard to understand text written by someone else
- What if we could always read text the way we understand it?



Goal

Achieve style transfer from Style A to Style B

- Learn style and content vector for Wikipedia
- Convert any sentence to Wikipedia style using these vectors
- Why Wikipedia? Because it is peer reviewed



Adversarial Shared-Private Model

Overview of Presentation

- Introduction
- Related Work
- Intuition
- Approach
- Results
- Future Work
- References

Image style transfer using convolutional neural networks (Gatys et al. 2016)

- Function to find how different two images are: content loss
- Function to find how similar two images are: style loss
- Minimize *content loss* and *style loss* as the *total variation loss*.

Image A

+

Image B

=

Image Style Transfer

Stylistic Transfer in Natural Language Generation Systems Using Recurrent Neural Networks (Kabbara et al. 2016)

1. For a given style transfer task between two styles A and B, collect relevant data for each styles.
2. Train model on each of the styles (separately).
3. During the testing phase, for a transfer from style A to style B, the system is fed texts having style A while the stylistic latent variables of the model are fixed to be those learned for style B (from previous step).

Style Transfer from Non-Parallel Text by Cross-Alignment (Shen et al. 2017)

- Two potential styles (e.g. positive and negative sentences)
- Discriminate between true style A, fake style A->B, and another for vice-versa

Sentiment transfer from negative to positive
i would recommend find another place . i would recommend this place again !
do not like it at all ! all in all, it 's great !
i regret not having the time to shop around . i have a great experience here .
average mexican food . authentic italian food .
Sentiment transfer from positive to negative
really good food that is fast and healthy . really bland and bad , and terrible .
you will notice that i have given this restaurant five stars . you should give this place zero stars .
definitely a place you can bring the family or just go for happy hour ! do not waste of your money, go somewhere else !
our waitress was very friendly and checked up on us a couple of times . our waitress was very rude and rushed with a couple of work .

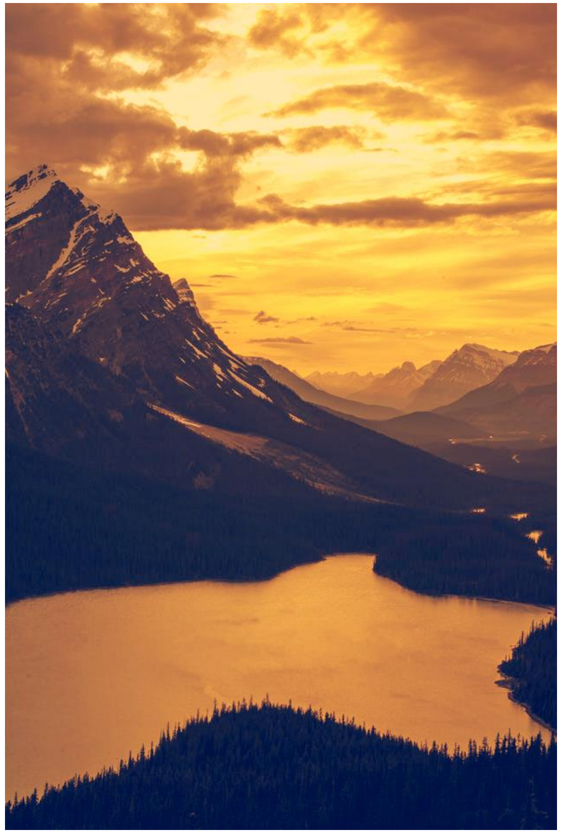
Samples from our cross-aligned auto-encoder. The first line is an input sentence, and the second line is the generated sentence after sentiment transfer.

Overview of Presentation

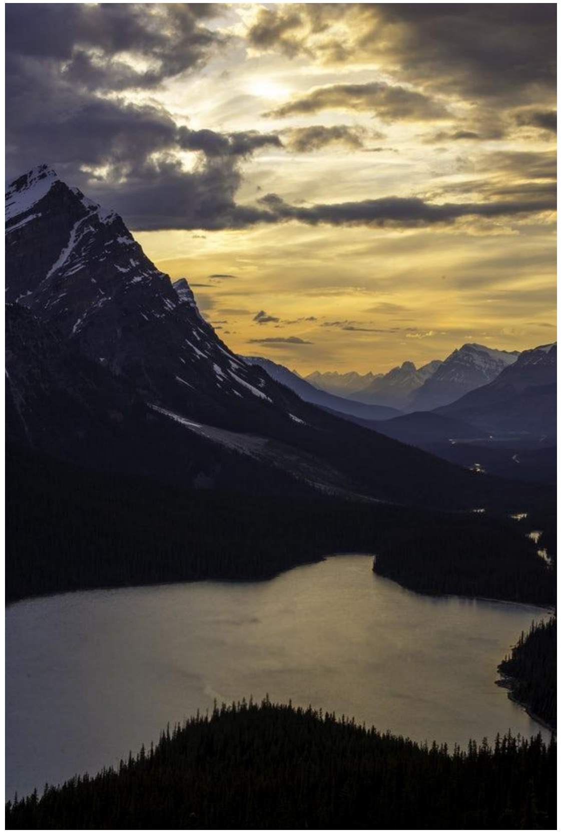
- Introduction
- Related Work
- Intuition
- Approach
- Results
- Future Work
- References

Image Vs Text

- Tolerance for image variation is lower
- Word order matters in text



There is a lake and mountain in the picture



And picture there is a lake in the mountain

Translation has parallel Data

- Millions of sentences of one language translated into another
- We don’t have that luxury in style transfer
  - o It is very difficult to acquire training data
  - o More difficult to obtain style pairs for data
- So this cannot be posed as another translation problem

Overview of Presentation

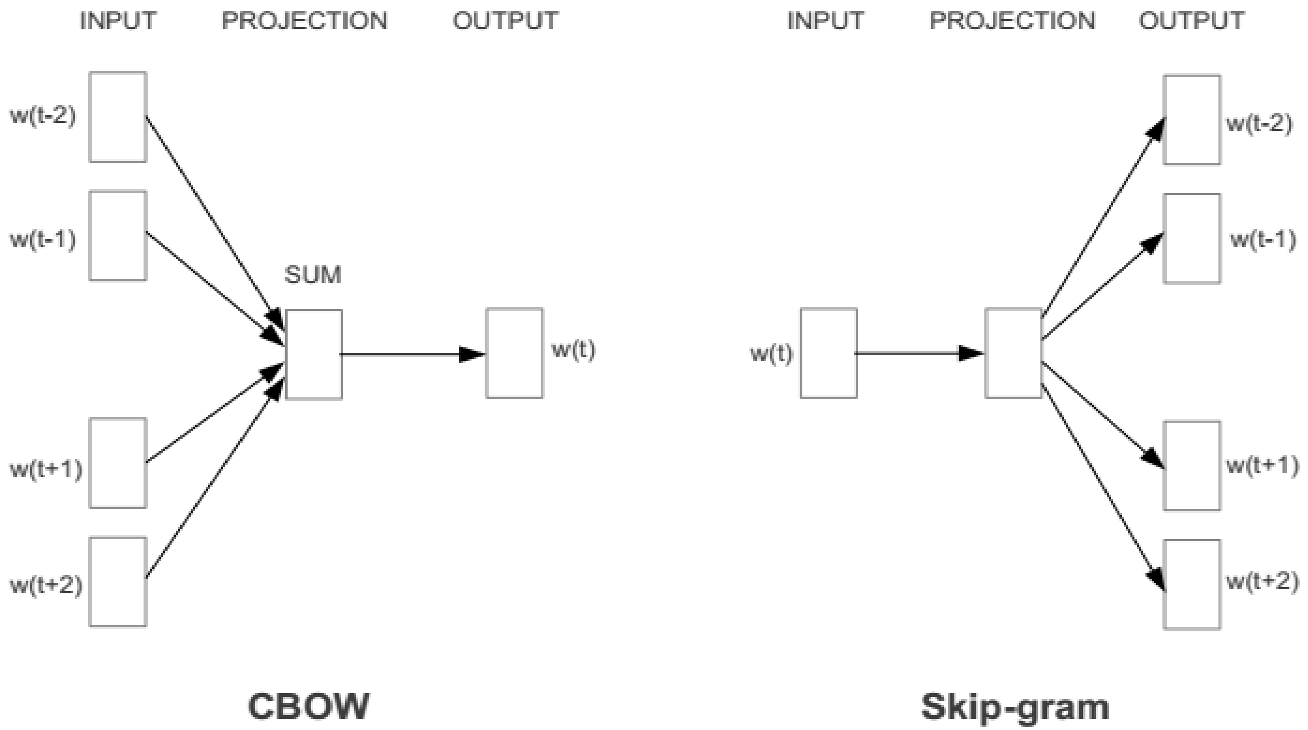
- Introduction
- Related Work
- Intuition
- Approach
- Results
- Future Work
- References

Approach

- Word Vectors
- Auto Encoders
- Sequence to Sequence
- Content based attention using Pointer Networks in Seq2Seq
- Generative Adversarial Networks (GANs)

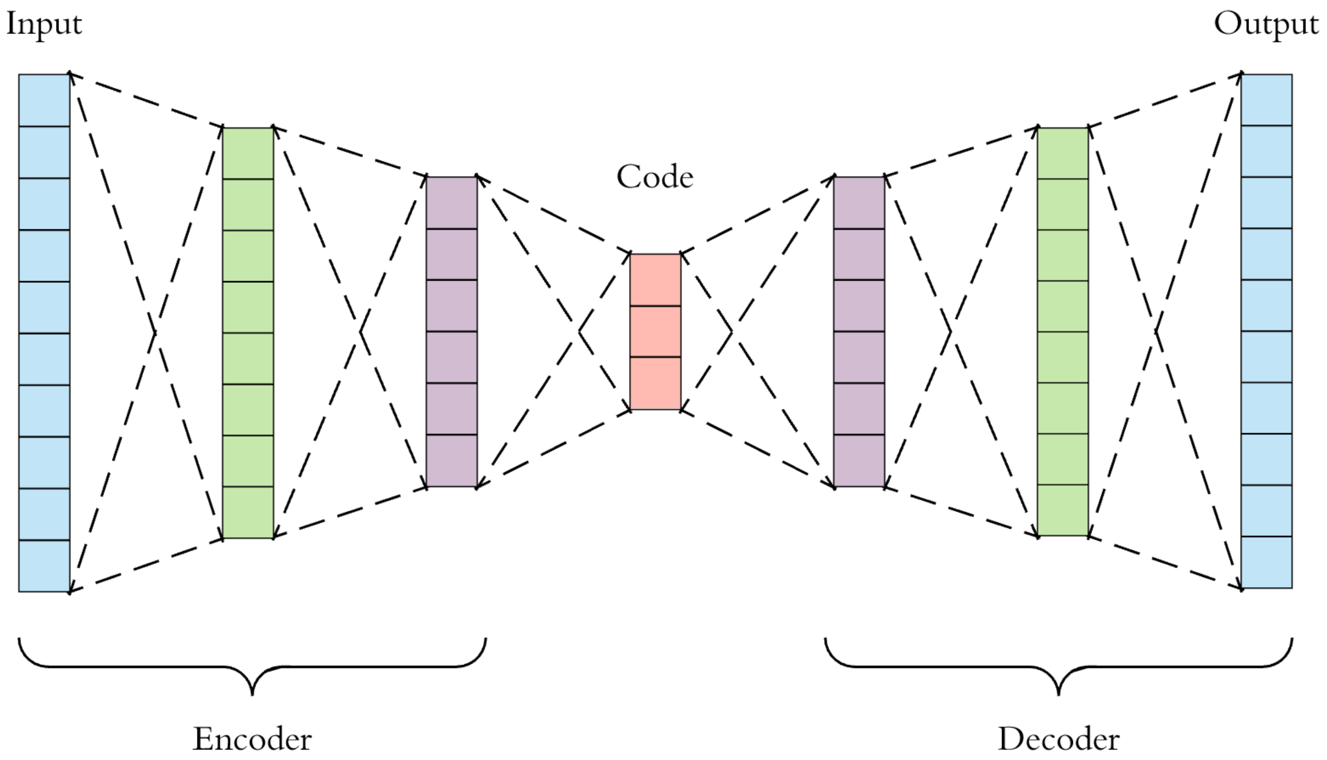
Word Vectors (Mikolov, Tomas, et al. 2013)

- Learn vector word representations using 1-Hidden layer perceptron
- Continuous Bag of Words: Predict center word from surrounding words
- Skip-gram: Predict surrounding word using center word



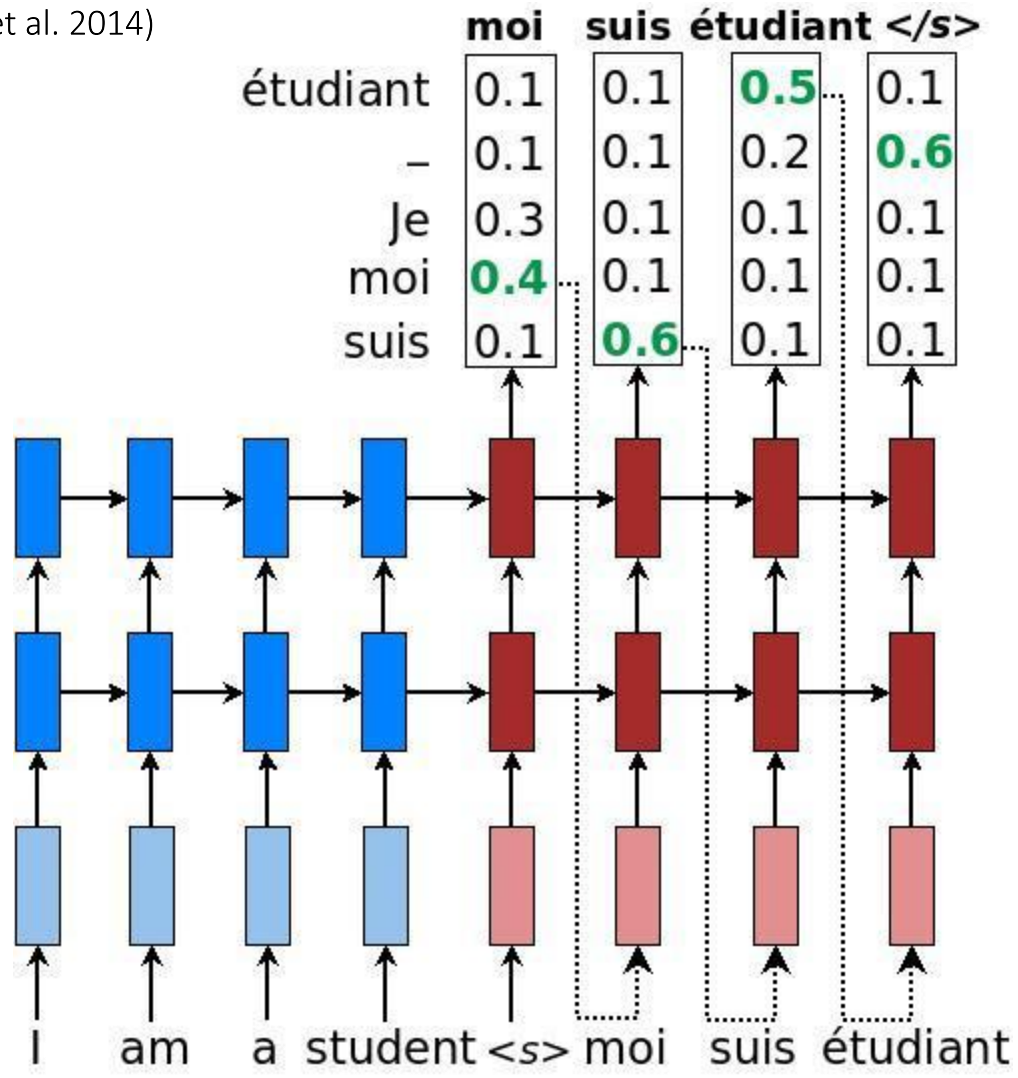
LSTM Autoencoder (Jiwei et al. 2015)

- Encoder-Decoder network where hidden layer is a compressed lossy representation of the original data



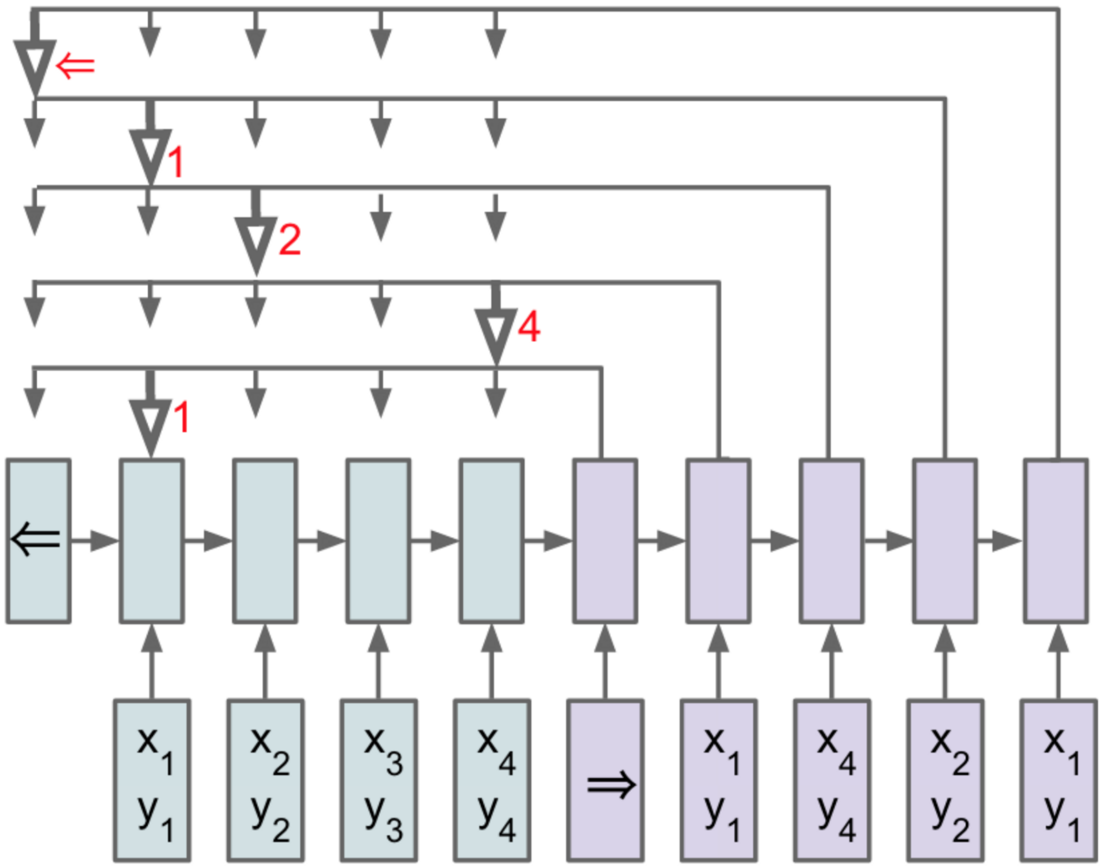
Sequence to Sequence (Sutskever et al. 2014)

- Learn language independent vector by encoding from one language
- Use this independent vector to decode into another language
- Use LSTM to encode long sequences as it prevents vanishing gradients
- Use Bidirectional LSTM to encode forward and backward token information



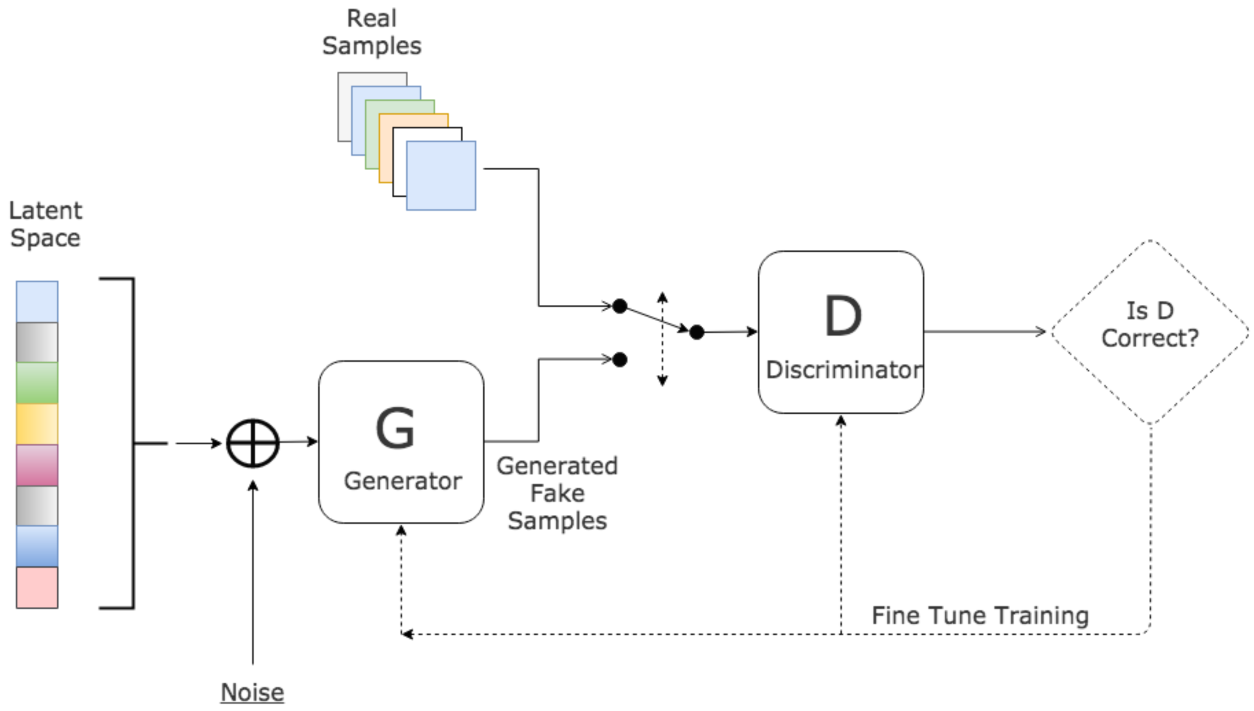
Content based attention using Pointer Networks in Sequence to Sequence (Vinyals et al. 2015)

- An encoding RNN converts the input sequence to a code (blue) that is fed to the generating network (purple)
- At each step, the generating network produces a vector that modulates a content-based attention mechanism over inputs
- The output of the attention mechanism is a softmax distribution with dictionary size equal to the length of the input.



Generative Adversarial Networks (Goodfellow et al. 2014)

- Generator to produce samples from random noise
- Discriminator learns to classify if sample is true or fake
- Both try to improve each other until convergence



Learning Style (Prior)

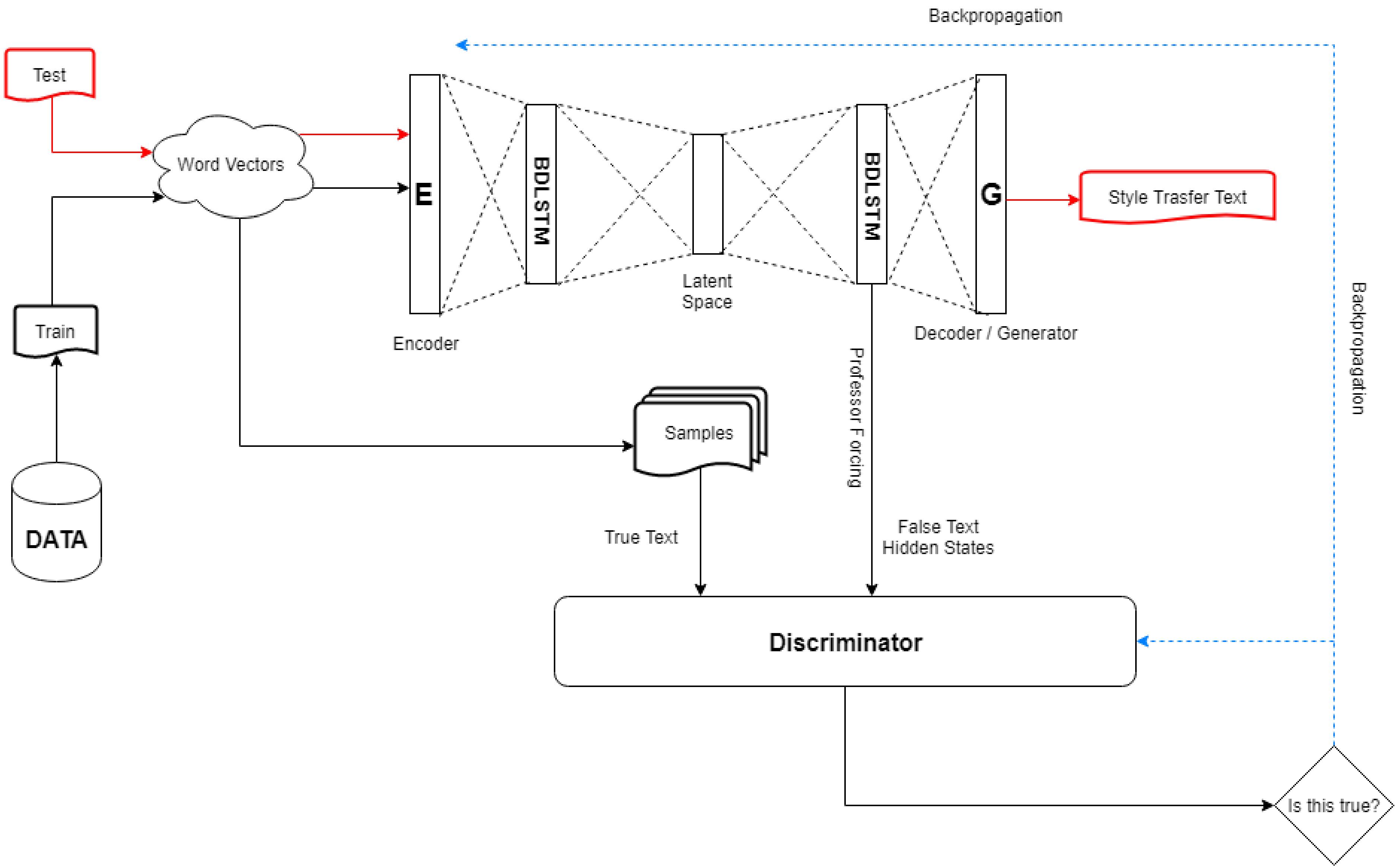
- Assume content prior to be same word vectors
- Use recurrent autoencoder to carry as much info as possible
- Update prior seeing new input data

Improving Prior learning using GANs

- Assign role of generator to decoder in autoencoder network
- Use discriminator along with professor forcing from decoder

Inference

- Remember our autoencoder is recurrent neural network Seq2Seq
- Feed the prior (hidden layer of autoencoder) to network’s decoder
- Decoder uses new text and prior to generate new style text



Deep Generative Adversarial Seq2Seq Autoencoder Network

Why is this approach better?

- Expected to even draw latent answers due to word embeddings
- Using Pointer with attention to tightly capture style
- Adversarial training allows non-parallel monolingual data
- Progress checking is easy using generative models

Overview of Presentation

- Introduction
- Related Work
- Intuition
- Approach
- Results
- Future Work
- References

Evaluation

- Soundness (generated texts being textually entailed with original version)
- Coherence (free of grammatical errors, proper word usage, etc.)
- Effectiveness (the generated texts actually match the desired style)
- BLEU Scoring using humans

Autoencoders

- Showing good signs of learning context vector
- For text: apply sigmoid over output before decoding to text
- Original

```
the quick brown fox jumped over the lazy dog from the quick tall fox
```

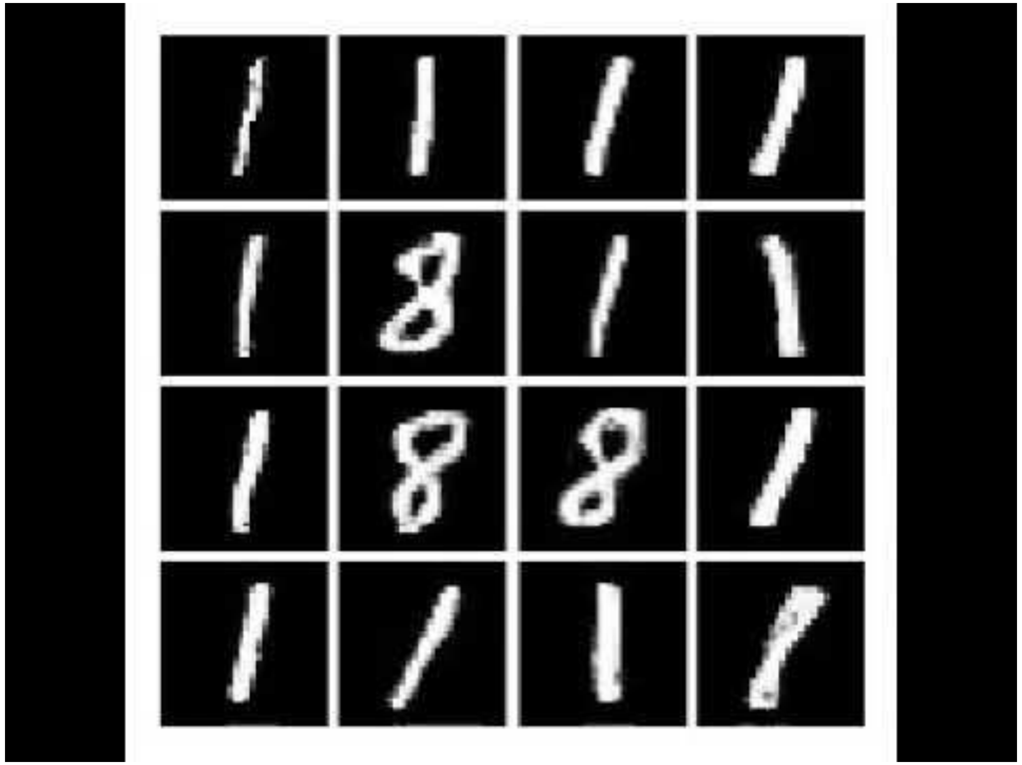
- Reconstruction

```
the quick fox tall over lazy dog jumped brown from
```

- 2 encoder and decoder layers each, try next with RNN biLSTM

GANs

- Mode collapse in 100k steps
- Back-propagating gradients on discreet text encoding
- Generating non sensical outputs since no context of Sequence
  . i ... that so have i'm all be are up
- love more it's we off 3 that's head down away sucks went summer face watch start boring sooo position re-ripped special squirrels 3rd standing it'll daily sannesias neighbor place uniform



https://www.youtube.com/watch?v=ktxhiKhWoEE

Overview of Presentation

- Introduction
- Related Work
- Intuition
- Approach
- Results
- Future Work
- References

Future Work

- Incorporate Seq2Seq models in Autoencoder and GANs
- Implement Pointer Network
- Use professor forcing technique

References

- Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. "Image style transfer using convolutional neural networks." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.
- Kabbara, Jad, and Jackie Chi Kit Cheung. "Stylistic Transfer in Natural Language Generation Systems Using Recurrent Neural Networks." *EMNLP 2016* (2016): 43.
- Shen, Tianxiao, et al. "Style Transfer from Non-Parallel Text by Cross-Alignment." *arXiv preprint arXiv:1705.09655* (2017).
- Alex M Lamb, Anirudh Goyal ALIAS PARTH GOYAL, Ying Zhang, Saizheng Zhang, Aaron C Courville, and Yoshua Bengio. Professor forcing: A new algorithm for training recurrent networks. In Advances In Neural Information Processing Systems, pages 4601–4609, 2016.
- Mikolov, Tomas, et al. "Efficient estimation of word representations in vector space." *arXiv preprint arXiv:1301.3781* (2013).
- Li, Jiwei, Minh-Thang Luong, and Dan Jurafsky. "A hierarchical neural autoencoder for paragraphs and documents." *arXiv preprint arXiv:1506.01057* (2015).
- Sutskever, Ilya, Oriol Vinyals, and Quoc V. Le. "Sequence to sequence learning with neural networks." *Advances in neural information processing systems*. 2014.
- Vinyals, Oriol, Meire Fortunato, and Navdeep Jaitly. "Pointer networks." *Advances in Neural Information Processing Systems*. 2015.
- Goodfellow, Ian, et al. "Generative adversarial nets." *Advances in neural information processing systems*. 2014.

Questions?