**Year** **–** 2019

**Idea –**

The main motivation of this task is to generate massive synthetic datasets of “historic” documents which can be used for the training of document analysis systems

Image-to-image translation problem, where the goal is to transform an image from a source domain (modern printed electronic document) to a target domain (historical handwritten document)

A GAN model for generateing an artificial historical document image that looks like a real historical document by transferring the “historical style” to the classical electronic document.

A new general-purpose generating framework that takes advantage of recent advancements in the design of Generative Adversarial Networks (GANs) and Neural Style Transfer Algorithms (NST).

**Method –**

Because of the dataset that the researchers use does not contain paired images between the source and target domains, they use a variant of GANs, called the cycleGAN, that uses the cycle consistency loss. The cycleGAN architecture performs a transformation of the images from the source to the target domains and vice-versa. The cycle consistency loss with the bidirectional mapping function coupled with the L1 distance loss increases the learning stability of the adversarial framework in an unpaired image setting.

For the Neural Style Transfer, the researchers choose the VGG-19 CNN implementation to minimize the content loss and the style functions conjointly.

The first stage produces a modern style document that can be customized with several parameters. It is achieved with a Latex framework. The second stage uses deep neural networks to transform the modern style printed document to a historical handwritten document, by train the network to learn a mapping function between the source domain (modern document) and the target domain (historical handwritten document). With two different approaches (GAN and NST) for the second stage, we demonstrate that it is feasible to transform the domain of an input document image from modern printed to historical handwritten.

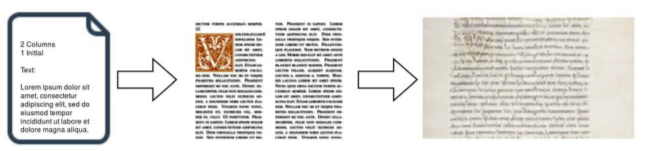


Fig. 1. The first step is to create modern electronic printed documents from a Latex specification document. The second step involves using a deep neural network to learn a mapping function to transform the modern printed document to a historical handwritten document.

NST Training Procedure

We use the VGG-19 based NST model in two different settings – using ImageNet weights and using weights from a pre-trained model. When using the model with the ImageNet weights, only the last layers of the network are reinitialized, and then the NST procedure is applied to the images.

In the case of the pre-training scenario, we first train the VGG-19 on our PDD dataset for a classification task. The network is trained for 25 epochs with a batch size of 4, learning rate of 0.001 and momentum of 0.9. The weights of this model are then used for the NST procedure.

**Results –**

The synthetic images generated by the cycleGAN appear significantly better than those generated with NST.

Regarding the semantic content (font shape, readability of words and letters, marginal annotations), the researchers notice many similarities between the target domain samples and the synthetic samples.

The overall style content of the target domain (background color, texture, paper degradation, initials style) is well expressed.

However, in a structural content point of view (column-mode, number and presence of initials, textual artifacts), the initials are not well detected and expressed. The two column-mode is not at all expressed.

When considering the synthetic documents produced with the NST, the structural content is better preserved. However, the style is mixed and standardized over the entire synthetic document, leading to the presence of a lot of colored artifacts. Also, the font does not change as compared to the synthetic images generated by the GAN. We can see these results in Figure 4.



Fig. 4. Examples of images generated by the cycleGAN and the NST after training on the Complete Document images. The first and second columns contains samples from the Target Domain and Source Domain respectively. The third column contains samples generated by the cycleGAN trained from scratch. The samples generated by the NST model (pre-trained on the PDD dataset) can be seen in the fourth column. Every sample contains a zoomed-in view to see the quality of the generated pages.

We believe that our promising results show that this area of research requires further investigation.

This research demonstrates the ability of GANs to generate synthetic documents that preserve the global characteristics of medieval manuscripts. However there remains scope for significant generative improvements at lower levels for characters and words.