

# **M.Sc. in Computer Science (Intelligent Systems)**

## **Dissertation Project Proposal**

Please describe the research topic on which you propose to work under the following headings:

**Student Name:**

Xu Zheng

**Student Email Address:**

xzheng@tcd.ie

**Research Area:**

Image Data Augmentation with Artistic Style

**Research Question:**

My research questions are listed below

1. Can style transfer be used as a data augmentation strategy for other image-based deep learning algorithms?
2. How useful style transfer can be compared and combined with the other traditional approaches for data augmentation?
3. Which styles and levels of abstraction work best?

**Proposed Supervisor:**

Prof. Aljosa Smolic <smolica@scss.tcd.ie>

## Research Summary:

I plan to use the image style transfer by CNN as a data augmentation strategy for image-based deep learning algorithms, such as image recognition, image classification and image segmentation.

The success of training deep learning algorithms has depended heavily on a large amount of annotated data. For many applications, gathering this kind of data can be very time-consuming or difficult. Hence, datasets are usually enhanced by data augmentation, i.e. applying random transformations on the data to enlarge the set. Often, only relatively simple transformations are applied e.g. random cropping or mirroring an image. In style transfer, the goal is to apply the style of an image to another image without changing its content. It is also possible to choose the level of abstraction when applying the style.

Since the content of an image should remain the same after application of new style, it seems natural to use style transfer as a data augmentation strategy for image-based deep learning algorithms. In this research project, we will evaluate how useful style transfer can be as a data augmentation strategy.

## Research Motivation:

Deep learning is one of the most important technologies in the field of artificial intelligence, which brought about great transformation in many industries, such as self-drive cars, intelligent robots and face recognition. Having worked as a software engineer for more than two years, my understanding of and interest in artificial intelligence, machine learning and deep learning grew drastically, and I want to know more about these technologies.

Computer vision is a very interesting and promising field. Therefore, I want to take it as my first step to gain a better understanding of these technologies. By doing this research work, I will explore many image-based algorithms in the AI industry, such as image recognition and image segmentation. What is more, I will also learn how to use many deep learning platforms, such as Tensorflow and Pytorch, which are very popular tools in the current AI industry, and it would be beneficial for my future work.

## Research Aims:

My research aim is to explore how image style transfer can be used as a data augmentation strategy for image-based deep learning algorithms.

Following are the research objectives:

1. Master how to implement a CNN model for style transfer.
2. Understand different image-based deep learning algorithms.
3. Explore how to train CNN with various platforms, such as Tensorflow and Pytorch.
4. Compare and combine style transfer with the more traditional data augmentation approaches.
5. Evaluate the results and continuously improve the model.
6. Analyze final results and check the research questions.

### **Potential benefits of the study for the field:**

There are two main potential benefits:

1. Deep learning techniques are quite data-intensive, making data collection and annotation an important and tedious task. If style transfer could be used as a data augmentation technique, it could ease the problem of data collection to some degree.
2. It may also lead to new innovative ideas on data augmentation thus improving the performance of image-based deep learning algorithms.

### **Proposed methodology/implementation approach:**

The proposed approach will be:

1. Literature Review
  - a) Understand the mechanism behind style transfer algorithm in CNN.
  - b) Understand the approach of how to implement image based tasks (recognition, classification) using CNN.
  - c) Get familiar with existing data augmentation techniques, such as zooming, cropping and re-scale.
2. Baseline Selection
  - a) Select proper image tasks and algorithms.
  - b) Select proper evaluation metrics.
3. Testing
  - a) Compare and evaluate different style transfer options in combination and against existing data augmentation techniques.
4. Analyzing the Results
  - a) Analyze all the results from many different configurations to see how useful style transfer can be and what options of style transfer work the best.

### **Background of your research area which puts your research into context:**

There have been lots of new CNN architectures proposed in recent years, such as AlexNet, VGGNet and ResNet. They are using many kinds of traditional data augmentation technologies during training, such as cropping, re-scaling and label-preserving transformations (He et al., 2016, Krizhevsky et al., 2012, Simonyan and Zisserman, 2014).

The recent advance of deep CNN has made it possible for computer vision systems to extract high-level semantic information from natural images. Based on it, a new neural algorithm of artistic style transfer was introduced, which allows us to apply different artistic styles on an image and also keep the image semantic content explicit at the same time (Gatys et al., 2015). Therefore, it is possible to use such style transfer technique as a data augmentation strategy for image-based deep learning algorithms.

In this research project, style transfer will be used as a data augmentation strategy on the image-based algorithms, and there has been no similar research on this specific topic. Based on current computer vision tasks, we will evaluate how well the style transfer works as a data augmentation strategy.

## Ethical Issues:

This research project involves no ethics application.

## Evaluation criteria (How will the results of your work be evaluated?)

To evaluate my work, I will choose

1. Image classification task as my baseline evaluation algorithm.
2. Use Top-1 error rate and Top-5 error rate as the metrics.
3. Apply style transfer as a data augmentation strategy with many different configurations on the original CNN classification model and evaluate if the error rate decrease.
4. If the error rate decreases in some configurations then we can prove the style transfer is a promising way to do image data augmentation.

Apart from it, I may try few other possible evaluation criteria. I will choose more image classification tasks and use the corresponding metrics to evaluate my work to see if I can have more interesting findings, such as computing precision and recall accuracy on object detection task.

## Source of literature:

Deep Learning Information

- <http://www.deeplearningbook.org/>
- <http://cs231n.stanford.edu/>
- [Christopher Olah's blog](#)
- [Andrei Karpathy's blog](#)
- [Wild ML](#)

Deep Learning Frameworks

- <http://caffe.berkeleyvision.org/>
- <https://www.tensorflow.org/>
- <http://pytorch.org/>
- <https://keras.io/>
- <http://lasagne.readthedocs.io/en/latest/index.html> + <http://github.com/dnouri/nolearn>

Style Transfer

- <https://arxiv.org/abs/1508.06576>
- <https://magenta.tensorflow.org/2016/11/01/multistyle-pastiche-generator>
  - <https://arxiv.org/abs/1610.07629>
  - <https://research.googleblog.com/2016/10/supercharging-style-transfer.html>
- <https://arxiv.org/pdf/1603.08155.pdf>

## Publication plan (What journals/conferences should be targeted or what organizations should be informed of your work?)

IMVIP (Irish Machine Vision and Image Processing Conference)

**Your own expertise and how well you are positioned to carry out the work:**

I have more than two years of software developing experience in a big R&D company in China. My core development area was in developing websites using Ruby on Rails. I also have a good knowledge of C and Python that can help me apply the CNN models quickly.

I also developed much interest in AI industry. As I find myself increasingly interested in AI industry in the process of taking various AI-related subjects, such as Machine Learning, Artificial Intelligence and Computer Vision, I believe I can complete this research project successfully where I can combine my skills and interests together.

**Prior work** (If your project is to be a continuation of a previous project, summarize the results of that project and say how your work will build on those results):

None

**Connections to funded/collaborative projects** (Will your project be connected to ongoing research, e.g., commissioned by a funding body? If so please outline the larger study, those involved in the work and your expected contribution).

None

**References:**

GATYS, L. A., ECKER, A. S. & BETHGE, M. 2015. A Neural Algorithm of Artistic Style. *arXiv:1508.06576*.

HE, K., ZHANG, X., REN, S. & SUN, J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. 770-778.

KRIZHEVSKY, A., SUTSKEVER, I. & HINTON, G. E. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 2012. 1097-1105.

SIMONYAN, K. & ZISSERMAN, A. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

**Table of research milestones** (Feel free to extend the table by adding further rows):

Project stage	Completed work	Date
Pre-Reading	1. CS213n online course. 2. TensorFlow tutorial.	29-Nov-17
Research question	1. Finalize the main research question.	1-Dec-17

Research Proposal	1. Copy of research proposal.	14-Dec-17
Literature review	1. Read style transfer papers. 2. Read image based tasks papers. 3. Get to know data augmentation technologies.	28-Feb-18
Model Selection	1. Choose baseline image tasks and baseline models. 2. Design style transfer model with CNN and try different configurations.	31-Mar-18
Training and Testing	1. Train and improve the architecture constantly to make the model stable. 2. Compare and combine the model with the more traditional approaches to get more results.	30-Apr-18
Analyze Data	1. Collect all the results and plot charts. 2. Analyze all the data to see how useful style transfer can be and see which configuration works the best.	31-May-18
Papers	1. Compose paper for IMVIP. 2. Compose dissertation paper.	31-July-18

### **Proposed Table of contents for your dissertation:**

Acknowledgments

Abstract

List of Figures

Chapter 1 Introduction

- 1.1 Overview
- 1.2 Motivation
- 1.3 Goals of the Study
- 1.4 Dissertation Structure

Chapter 2 State of the Art

- 2.1 Computer Vision Tasks
- 2.2 The Art of Data Augmentation

Chapter 3 Approach

- 3.1 Model Architecture
- 3.2 Configurations
- 3.3 Layers Selection and Analysis

Chapter 4 Implementation

- 4.1 Deep Learning Platform Selection
- 4.2 Training Strategy
- 4.3 Testing Strategy

Chapter 5 Evaluation

- 5.1 Evaluation Metrics
- 5.2 Single Model Evaluation
- 5.3 Combined Model Evaluation
- 5.4 Summary

Chapter 6 Conclusion

- 6.1 Project Overview
- 6.2 Limitations
- 6.3 Contributions
- 6.4 Future Work