WESTERN SYDNEY
UNIVERSITY

Generating Personalised Visual Art with Machine Learning

Sonali Marasini
19913230

A progress report submitted for INFO7017 Postgraduate Project B

in partial fulfilment of the requirements for the degree of

Master of Information and Communication Technology (Advanced)

Supervisor: Dr. Quang Vinh Nguyen

School of Computer, Data and Mathematical Sciences

Western Sydney University

April 2021

Contents

Introduction

Art is a universal language. It is the use of imagination and creativity. Unlike things civilizations stress over like medicine, agriculture, scientific breakthroughs, art is a form of human ingenuity and it is a stem of human possibility.

Painting is a form of art that many people enjoy. There have been many great painters who have had their own distinct style. Pablo Picasso was a Spanish painter who lived from the early 1900s to the 1960s. He was an extremely talented artist who is credited with creating "cubism". Cubism is when the artist paints the subject from lots of different angles all in the same picture. So, you see the front, the back, and the sides of the picture at the same time. In cubist artworks, objects are broken up into lots of geometric shapes and assembled. Cubism is a form of abstract painting. (Fry, 1988)

In cubism, every viewpoint of the object or the subject is painted. The artist not only paints what they see from their viewpoint but all the other hidden sides are painted as well. Suppose a person that is to be painted is sideways, the artist finds a way to paint the other side of the face that is not in the artist's line of sight. Cubism is interpreting a three-dimensional object on a flat piece of paper while keeping all the important details and ignoring all the unnecessary lines.

People love to take their photos, beautify those photos, and transform those photos into different styles. With the advances in photo editing technology, people have become more interested to transform their photos into a form of art. By using deep learning algorithms and image processing techniques, the goal of this project is to generate personalised cubism inspired art from portrait images. In this project, a self portrait image will be converted into a cubism inspired art using Python programming language.

From the earliest days, computers have been used to generate or draw images. In papers, books and articles, researchers were able to use computer generated images. This made the books more authentic and easy to understand as complicated concepts can be easily understood using an image. The representation of images and its accuracy entirely depends on the raw data and the representations that have been selected to create the images. (Peddie, 2013)

There have been several research conducted to generate different art styles. In Xiong and Zhang (2016), the authors have presented a research to generate Joan Miro's style of surrealism paintings automatically. The authors have used modelling approach to generate the specific art style. In the modeling process, the authors have used the Bezier function (Bertka, 2008) to draw the background curve lines on the canvas. Then, the graphical objects have been drawn using basic pictorials elements. Lastly, the objects are coloured using adaptive fill algorithm.

The mandala thangka paintings are a Tibetian Buddhist paintings which have very intricate details. These paintings are very time consuming to draw. Mandala art require high level of mastery skills. In Zhang et al. (2017), the authors have proposed a computer aided approach to generate mandala thangka paintings automatically. Firstly, the authors have constructed models of the patterns used in the mandala paintings. Based on the hierarchical structures observed from the hand drawn mandalas, the computer aided mandalas are generated using the initially constructed models.

Collomosse and Hall (2003) uses a set of several two-dimensional images taken from different angles as input. A non-photorealistic rendering (NPR) technique is used to produce a cubism inspired art as the output. Gooch and Gooch (2001) defines non-photorealistic rendering as a field in computer graphics that focuses on allowing a wide range of digital art expressive forms. It is exactly opposite to conventional computer graphics that concentrate on photorealism.

In Xing et al. (2018), the authors have proposed a portrait-aware artistic style transfer algorithm. This algorithm separates the foreground of the painting from the background and then a different transformation is applied to each of them to create a style of photograph.

The authors of Zhang and Yu (2016), have presented an experiment to generate art based on Wassily Kadinsky abstract painting style. The authors have been successfully in creating unique abstract paintings by randomising several parameters from the Kadinsky paintings. After setting up the parameters, the background is generated and then the styled patterns are drawn in randomised fashion. The authors have concluded that more styled patterns could be added in their approach to generate the art and make it more

detailed. This would make the paintings more accurate to the original Kadinsky paintings.

The authors of Heath and Ventura (2016) have argued that before a computer attempts to draw any image or painting, it must first learn to perceive the said image or painting. If the computational systems have perceptual skills, they tend to become more decentralised. This will, in turn, allow them to learn better models, analyse the input objects and produce intentional output objects.

Objective

The main objective of this project is to generate cubism inspired personalised art from a portrait by using machine learning techniques. This will allow people to enjoy abstract art which has been generated from their personal photograph. Moreover, it is an appreciation of the cubism style of art. This project would allow personalised cubism inspired art to be more accurate and make it more accessible to people for them to enjoy.

To be more specific, the main objectives of this research study have been listed below,

- To gain in-depth knowledge about deep learning and image processing techniques
- To have a stronger hold on Python programming language by executing and learning several machine learning algorithms
- To study if there is a way to generate personalised art by using deep learning methodologies in Python more efficiently

The objectives can be narrowed down to the following research questions which will be addressed through the literature reviews. The research study can be deemed successful if the following research questions shall be answered at the end,

- What are the different types of image processing techniques?
- How can the image processing techniques be enhanced to produce more accurate results?

- How can the enhancements be used to generate Picasso's cubism style art for personalised photo and make them accurate?
- What are the future prospects in this area?

Methodology

Background Removal

As the input, a portrait photograph is used. The background of the image is removed in the first step. The background of the image can be removed by classifying each pixel of the image to divide the background from the foreground. This task of labelling each pixel is known as semantic segmentation. Lian and Zhang (2019) states that two approaches could be used for this procedure, namely Fully Convolutional Network (FCN) and Conditional Random Fields as Recurrent Neural Networks (CRFASRNN). Both of these use deep neural networks for semantic segmentation of images. Fully Convolutional Networks are constructed from layers that are locally connected, such as convolution. Long et al. (2015) states that, in this sort of architecture, thick layer is not used. This decreases the number of parameters and the time to compute. CRFASRNN was introduced in Zheng et al. (2015), where the authors "formulated Conditional Random Fields with Gaussian pairwise potentials and mean-field approximate inference as Recurrent Neural Networks." These two techniques have been compared in Chen et al. (2014) and it has been proven that CRFASRNN is more structured and the image segmentation is more accurate as compared to Fully Convolutional Networks. This would mean that CRFASRNN provides a more accurate foreground by labelling the background pixels more precisely. Therefore, it would make sense to use CRFASRNN instead of Fully Convolutional Network for the process of background removal in this project.

Reconstruction

The next step would be to remodel the image. The left half of the portrait image is replaced by a profile face. Profile face refers to the side view of the input image which will be constructed from the input image.

First, the input image is rotated by 90 degrees to obtain a side view of the image. In Jackson et al. (2017), the authors have proposed a technique to achieve a reconstructed 3D image from an input 2D image using volumetric Convolutional Neural Network regression. The same technique can be applied to synthesise the left facing side view of the input image.

From the frontal portrait image, facial characteristics like eyes, ears and nose should be retained. The goal is to remove the left eye, the nose, left ear and mouth from the frontal image as these features will be preserved in the profile face instead. Now, to remove the above mentioned features from the frontal image, the 68-key points face detection technique is used as mentioned in King (2009).

Now that the left part of the input image has been removed and the profile face has been created, the next step is to splice these faces together and to fill the empty regions. These two images are joined together using the algorithm written in Lian and Zhang (2019). To fill the blank regions, continuous pixels are picked from the full frontal image.

Edge Detection

After reconstructing the input portrait image, the features of the image needs to be enhanced. This would ensure that the outline of the characters are more visible to reflect the cubist style. Edge detection is an image processing technique to determine the boundaries of the different objects within an image. This works by detecting discontinuities in the brightness of an image. According to Ziou et al. (1998), the edge detection technique includes, smoothing, differentiation and labeling. There are several edge detection techniques. These techniques have been studied and compared in Maini and Aggarwal (2009) as seen in table 1. Lian and Zhang (2019) have used Sobel operators to compute th edge points. However, as seen in table 1, this technique is sensitive to noise.

Table 1: Edge Detection Techniques

Operator	Advantages	Disadvantages
Classical	Simplicity, Detection of edges and their orientations	Sensitivity to noise, Inaccurate
Zero Crossing	Detection of edges and their orientations. Having fixed characteristics in all directions	Responding to some of the existing edges, Sensitivity to noise
Laplacian of Gaussian(LoG)	Finding the correct places of edges, Testing wider area around the pixel	Malfunctioning at the corners, curves and where the gray level intensity function varies. Not finding the orientation of edge because of using the Laplacian filter
Gaussian	Using probability for finding error rate, Localization and response. Improving signal to noise ratio, Better detection specially in noise conditions	Complex Computations, False zero crossing, Time consuming

Colouring

After the reconstructed image has been enhanced, the fourth step is for the background transformation by colouring. Four background layers are created which are scaled to the same size as the input photograph. Random background in generated styled similar to the background of a cubism painting.

Transformation

After the completion of above mentioned four processes, the input image is then transformed into cubism style using deep learning method. Gatys et al. (2015) have introduced "an artificial system based on Deep Neural Network that creates artistic images of high perceptual quality." The system uses neural representations to isolate and recombine arbitrary image content and style, providing a neural image algorithm for creative image creation.

Research Plan

Table 2: Timeline of the Research Plan

Task	Start Date	End Date
Planning for Research	01-Mar-2021	05-Mar-2021
Scope Identification	08-Mar-2021	12-Mar-2021
Development	15-Mar-2021	31-May-2021
Progress Report Preparation	09-Apr-2021	22-Apr-2021
Preliminary Results Discussion	12-Apr-2021	16-Apr-2021
Presentation Preparation	19-Apr-2021	23-Apr-2021
Progress Report Submission	23-Apr-2021	23-Apr-2021
Presentation	26-Apr-2021	26-Apr-2021
Final Report Preparation	03-May-2021	13-Jun-2021
Final Results Discussion	31-May-2021	04-Jun-2021
Final Report Submission	14-Jun-2021	14-Jun-2021

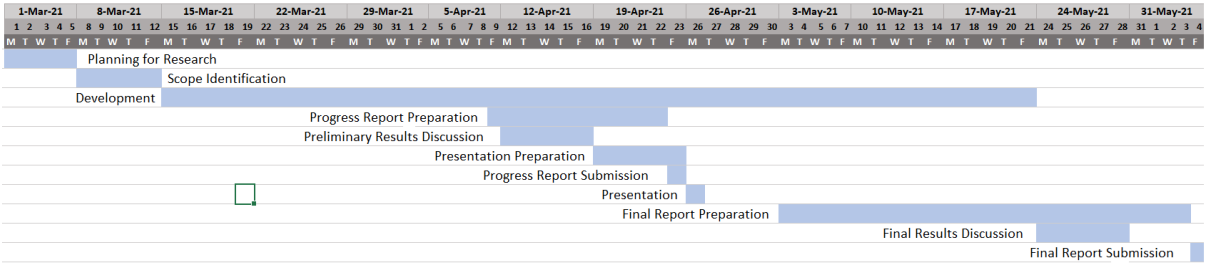

Figure 1: Gantt Chart of the Research Plan

References

Bertka, B. T. (2008). An introduction to bézier curves, b-splines, and tensor product surfaces with history and applications. *University of California Santa Cruz*.

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., and Yuille, A. L. (2014). Semantic image segmentation with deep convolutional nets and fully connected crfs. *arXiv preprint arXiv:1412.7062*.

Collomosse, J. P. and Hall, P. M. (2003). Cubist style rendering from photographs. *IEEE Transactions on Visualization and Computer Graphics*, 9(4):443–453.

Fry, E. F. (1988). Picasso, cubism, and reflexivity. *Art Journal*, 47(4):296–310.

Gatys, L. A., Ecker, A. S., and Bethge, M. (2015). A neural algorithm of artistic style. *arXiv preprint arXiv:1508.06576*.

Gooch, B. and Gooch, A. (2001). *Non-photorealistic rendering*. CRC Press.

Heath, D. and Ventura, D. (2016). Before a computer can draw, it must first learn to see. In *Proceedings of the 7th international conference on computational creativity*, pages 172–179.

Jackson, A. S., Bulat, A., Argyriou, V., and Tzimiropoulos, G. (2017). Large pose 3d face reconstruction from a single image via direct volumetric cnn regression. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1031–1039.

King, D. E. (2009). Dlib-ml: A machine learning toolkit. *The Journal of Machine Learning Research*, 10:1755–1758.

Lian, G. and Zhang, K. (2019). Transformation of portraits to picasso's cubism style. *The Visual Computer*, pages 1–9.

Long, J., Shelhamer, E., and Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440.

Maini, R. and Aggarwal, H. (2009). Study and comparison of various image edge detection techniques. *International journal of image processing (IJIP)*, 3(1):1–11.

Peddie, J. (2013). The history of visual magic in computers. *Springer, London. https://doi. org/10*, 1007:978–1.

Xing, Y., Li, J., Dai, T., Tang, Q., Niu, L., and Xia, S.-T. (2018). Portrait-aware artistic style transfer. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 2117–2121. IEEE.

Xiong, L. and Zhang, K. (2016). Generation of miro's surrealism. In *Proceedings of the 9th international symposium on visual information communication and interaction*, pages 130–137.

Zhang, J., Zhang, K., Peng, R., and Yu, J. (2017). Computer-aided generation of mandala thangka patterns. In *Proceedings of the 10th International Symposium on Visual Information Communication and Interaction*, pages 93–100.

Zhang, K. and Yu, J. (2016). Generation of kandinsky art. *Leonardo*, 49(1):48–54.

Zheng, S., Jayasumana, S., Romera-Paredes, B., Vineet, V., Su, Z., Du, D., Huang, C., and Torr, P. H. (2015). Conditional random fields as recurrent neural networks. In *Proceedings of the IEEE international conference on computer vision*, pages 1529–1537.

Ziou, D., Tabbone, S., et al. (1998). Edge detection techniques-an overview. *Pattern Recognition and Image Analysis C/C of Raspoznavaniye Obrazov I Analiz Izobrazhenii*, 8:537–559.