HYBRID LEARNING-BASED MODEL FOR EXAGGERATION STYLE OF FACIAL CARICATURE

SURIATI BINTI SADIMON

UNIVERSITI TEKNOLOGI MALAYSIA
HYBRID LEARNING-BASED MODEL FOR EXAGGERATION STYLE OF FACIAL CARICATURE

SURIATI BINTI SADIMON

A thesis is submitted in fulfilment of the requirements for the award of the degree of Doctor of Philosophy (Computer Science)

Faculty of Computing
Universiti Teknologi Malaysia

APRIL 2017
Dedicated to my beloved,

Late father, Sadimon bin Karmosari (al-fatihah)

Mother, Patimah binti Abd Hanan,

Husband, Mohd Razak bin Samingan

Princesses, Umairah and Uswah,

Princes, Muhammad Qayyim and Muhammad Qaid.
ACKNOWLEDGEMENTS

My foremost thanks and gratitude to Allah Almighty and Merciful for giving me strength and opportunities to complete my PhD study successfully. There is no power but with Allah. Next, I wish to express my sincerest appreciation to my supervisors, Professor Dr Habibollah bin Haron and Assoc. Prof. Dr Mohd Shahrizal bin Sunar for their support, guidance and advice throughout this study. My sincerest appreciation also extends to all my colleagues and other staff in Faculty of Computing for their assistance during my period of study. My thanks also to Ministry of Higher Education and Universiti Teknologi Malaysia for offering me the scholarship to pursue my PhD study. Lastly, I am very grateful to all my beloved family members for their constant prayers, encouragement and understanding throughout these challenging years.
Prediction of facial caricature based on exaggeration style of a particular artist is a significant task in computer generated caricature in order to produce an artistic facial caricature that is very similar to the real artist’s work without the need for skilled user (artist) input. The exaggeration style of an artist is difficult to be coded in algorithmic method. Fortunately, artificial neural network, which possesses self-learning and generalization ability, has shown great promise in addressing the problem of capturing and learning an artist’s style to predict a facial caricature. However, one of the main issues faced by this study is inconsistent artist style due to human factors and limited collection on image-caricature pair data. Thus, this study proposes facial caricature dataset preparation process to get good quality dataset which captures the artist’s exaggeration style and a hybrid model to generalize the inconsistent style so that a better, more accurate prediction can be obtained even using small amount of dataset. The proposed data preparation process involves facial features parameter extraction based on landmark-based geometric morphometric and modified data normalization method based on Procrustes superimposition method. The proposed hybrid model (BP-GANN) combines Backpropagation Neural Network (BPNN) and Genetic Algorithm Neural Network (GANN). The experimental result shows that the proposed hybrid BP-GANN model is outperform the traditional hybrid GA-BPNN model, individual BPNN model and individual GANN model. The modified Procrustes superimposition method also produces a better quality dataset than the original one.
ABSTRAK

# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DECLARATION</td>
<td>ii</td>
</tr>
<tr>
<td></td>
<td>DEDICATION</td>
<td>iii</td>
</tr>
<tr>
<td></td>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td></td>
<td>ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td></td>
<td>ABSTRAK</td>
<td>vi</td>
</tr>
<tr>
<td></td>
<td>TABLE OF CONTENTS</td>
<td>vii</td>
</tr>
<tr>
<td></td>
<td>LIST OF TABLES</td>
<td>xiii</td>
</tr>
<tr>
<td></td>
<td>LIST OF FIGURES</td>
<td>xv</td>
</tr>
<tr>
<td></td>
<td>LIST OF ABBREVIATIONS</td>
<td>xviii</td>
</tr>
<tr>
<td></td>
<td>LIST OF APPENDICES</td>
<td>xix</td>
</tr>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1</td>
<td>Overview</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>Background of Problem</td>
<td>2</td>
</tr>
<tr>
<td>1.3</td>
<td>Problem Statement</td>
<td>6</td>
</tr>
<tr>
<td>1.4</td>
<td>Objectives</td>
<td>7</td>
</tr>
<tr>
<td>1.5</td>
<td>Scope</td>
<td>7</td>
</tr>
<tr>
<td>1.6</td>
<td>Significance of the Study</td>
<td>8</td>
</tr>
<tr>
<td>1.7</td>
<td>Organization of the thesis</td>
<td>9</td>
</tr>
</tbody>
</table>
# LITERATURE REVIEW

## 2.1 Introduction

## 2.2 Computer Generated Caricature

### 2.2.1 Theories and Rules in the Art of Drawing Caricature

### 2.2.2 Process of Generating Caricature from Input Face Image

#### 2.2.2.1 Facial Feature Point Definition

#### 2.2.2.2 Facial Feature Extraction

#### 2.2.2.3 Distinctive Facial Features and Exaggeration

#### 2.2.2.4 Image Transformation

## 2.2.3 Different Approaches in Generating Facial Caricature

### 2.2.3.1 Interactive Approach

### 2.2.3.2 Regularity-based Approach

### 2.2.3.3 Learning-based Approach

### 2.2.3.4 Predefined Database of Caricature Illustration

## 2.3 Artificial Neural Network for Facial Caricature and Face-related Problem

### 2.3.1 Artificial Neural Network Model

### 2.3.2 Existing Neural Network Model for Facial Caricature and Face-related Problem

## 2.4 Evolutionary Artificial Neural Network

### 2.4.1 Genetic Algorithm Optimization

### 2.4.2 Artificial Neural Network Model Trained by Genetic Algorithm

### 2.4.3 Hybridization of Genetic Algorithm and Backpropagation Neural Network
3 RESEARCH METHODOLOGY

3.1 Introduction 53
3.2 Research Framework 53
3.3 Problem Definition 55
3.4 Data Preparation 56
3.5 Designing Backpropagation Neural Network (BPNN) Model 59
3.6 Designing Genetic Algorithm Neural Network (GANN) Model 62
   3.6.1 Problem coding 62
   3.6.2 Population 62
   3.6.3 Fitness function 63
   3.6.4 Evolutionary process 63
   3.6.5 Termination condition 64
3.7 Designing Hybridization of GANN and BPNN Model 64
3.8 Evaluation of Models 67
   3.8.1 Quantitative Error Measurement 67
   3.8.2 Statistical Tests 69
   3.8.3 Comparison of Models Performance 70
3.9 Hardware and Software Requirements 71
3.10 Summary 72

4 PROPOSED FACIAL CARICATURE DATA PREPARATION PROCESS 73

4.1 Introduction 73
4.2 Data Collection 73
4.3 Definition and Extraction of Facial Landmark Points 74
4.4 Normalization of Facial Landmarks Points
   4.4.1 Procrustes Superimposition Method
      4.4.1.1 Translation
      4.4.1.2 Isomorphic Scaling
      4.4.1.3 Rotation
   4.4.2 Modified Procrustes Superimposition Method
      4.4.2.1 Modified Translation
      4.4.2.2 Modified Isomorphic Scaling
      4.4.2.3 Modified Rotation
4.5 Generating Different Datasets and Average Face
4.6 Generating Input and Target Output
4.7 Summary

5 FINDING THE BEST PARAMETERS AND FEATURES FOR NEURAL NETWORK FACIAL CARICATURE PREDICTION MODEL
   5.1 Introduction
   5.2 Development of Backpropagation Neural Network (BPNN) Model
      5.2.1 Dividing the face dataset
      5.2.2 Selection of Optimal Parameters
      5.2.3 Training the network
      5.2.4 Experimental Setup
      5.2.5 Result and Analysis
         5.2.5.1 Quantitative Error Measurement
         5.2.5.2 Discussion
   5.3 Development of Genetic Algorithm Neural Network (GANN) Model
      5.3.1 Problem Encoding
5.3.2 Population Initialization 111
5.3.3 Fitness function 112
5.3.4 Chromosome Evolution 112
5.3.5 Termination Condition 115
5.3.6 Result of Testing Dataset 116
5.4 Summary 117

6 HYBRID NEURAL NETWORK MODEL FOR GENERATING FACIAL CARICATURE 119

6.1 Introduction 119
6.2 Development of Hybrid GA-BPNN Model 120
  6.2.1 Neural Networks Trained by Genetic Algorithm 120
  6.2.2 Fine Tuning using Backpropagation Algorithm 122
6.3 Development of Hybrid BP-GANN Model 124
  6.3.1 Several Backpropagation Neural Networks 125
  6.3.2 Genetic Algorithm Neural Networks 127
6.4 Result and Analysis 132
  6.4.1 Quantitative Error Measurement 132
  6.4.2 Statistical Test 133
  6.4.3 Comparison of Models Performance 135
  6.4.4 Discussion 138
6.5 Summary 141

7 CONCLUSIONS 144

8.1 Introduction 144
8.2 Thesis Summary 144
8.3 Research Contributions 148
8.4 Recommendation for Future Enhancements 149

REFERENCES 151

Appendices A - F 169-184
## LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Different approaches in generating caricature</td>
<td>23</td>
</tr>
<tr>
<td>2.2</td>
<td>Regularity-based approach</td>
<td>27</td>
</tr>
<tr>
<td>2.3</td>
<td>Learning-based approach</td>
<td>34</td>
</tr>
<tr>
<td>2.4</td>
<td>Previous works on optimization of neural network model using GA</td>
<td>48</td>
</tr>
<tr>
<td>2.5</td>
<td>Previous works on hybridization of GA and BP to train neural network model</td>
<td>50</td>
</tr>
<tr>
<td>4.1</td>
<td>Face landmark points description</td>
<td>76</td>
</tr>
<tr>
<td>5.1</td>
<td>Input samples for y-coordinate of the N2 dataset</td>
<td>96</td>
</tr>
<tr>
<td>5.2</td>
<td>Dividing the dataset based on 10-fold cross validation</td>
<td>97</td>
</tr>
<tr>
<td>5.3</td>
<td>Parameters values of Backpropagation Neural Network (BPNN)</td>
<td>104</td>
</tr>
<tr>
<td>5.4</td>
<td>Experimental Results</td>
<td>105</td>
</tr>
<tr>
<td>5.5</td>
<td>Average NMSE for the combination of x- and y-coordinates</td>
<td>107</td>
</tr>
<tr>
<td>5.6</td>
<td>Parameters values of the GANN</td>
<td>110</td>
</tr>
<tr>
<td>5.7</td>
<td>Average MSE for different number of population and generation</td>
<td>111</td>
</tr>
<tr>
<td>5.8</td>
<td>MSE of the different selection and crossover mechanisms</td>
<td>114</td>
</tr>
<tr>
<td>5.9</td>
<td>Result of GANN model</td>
<td>117</td>
</tr>
<tr>
<td>6.1</td>
<td>Parameters values of GA-BPNN model</td>
<td>124</td>
</tr>
<tr>
<td>6.2</td>
<td>Parameter values for hybrid BP-GANN model</td>
<td>132</td>
</tr>
<tr>
<td>6.3</td>
<td>Results of hybrid GA-BPNN model and hybrid BP-GANN model</td>
<td>133</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>6.4</td>
<td>Paired sample t-test results for x-coordinate</td>
<td>134</td>
</tr>
<tr>
<td>6.5</td>
<td>Paired sample t-test results for y-coordinate</td>
<td>134</td>
</tr>
<tr>
<td>6.6</td>
<td>Comparison of quantitative errors for all models</td>
<td>136</td>
</tr>
<tr>
<td>6.7</td>
<td>Two sample t-test results for comparison between models for x-coordinate</td>
<td>137</td>
</tr>
<tr>
<td>6.8</td>
<td>Two sample t-test results for comparison between models for y-coordinate</td>
<td>138</td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>FIGURE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Process of generating caricature from face image</td>
<td>16</td>
</tr>
<tr>
<td>2.2</td>
<td>Standard face model (Du et al., 2015)</td>
<td>19</td>
</tr>
<tr>
<td>2.3</td>
<td>Golden ratio of an ideal face</td>
<td>20</td>
</tr>
<tr>
<td>2.4</td>
<td>(a) Photographic caricature (Li and Miao, 2013) (b) Sketch caricature (Tseng, 2007) (c) Hand-drawn like caricature (Chen et al., 2006) (d) Outline caricature (Dal et al., 2011)</td>
<td>22</td>
</tr>
<tr>
<td>2.5</td>
<td>(a) Original face image and its simple template (b) Caricature image and its simple template (Akleman et al., 2000)</td>
<td>24</td>
</tr>
<tr>
<td>2.6</td>
<td>(a) Input face image (b) Line representation of input face image (c) Caricature image (Brennan, 2007)</td>
<td>25</td>
</tr>
<tr>
<td>2.7</td>
<td>(a) Original face image (b) Resulting caricature after 1st iteration (c) Resulting caricature after 16th iteration (Tseng and Lien, 2012)</td>
<td>26</td>
</tr>
<tr>
<td>2.8</td>
<td>Original face image (photo) and the resulting caricature (Chen et al., 2011)</td>
<td>28</td>
</tr>
<tr>
<td>2.9</td>
<td>(a) Original face image (b) Hand-drawn caricature created by an artist (c) The resulting caricature (Lai et al. 2006)</td>
<td>30</td>
</tr>
<tr>
<td>2.10</td>
<td>Genetic algorithm procedure</td>
<td>44</td>
</tr>
<tr>
<td>2.11</td>
<td>Conversion of neural network weights and biases into chromosome</td>
<td>46</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3.1</td>
<td>Framework of the research</td>
<td>54</td>
</tr>
<tr>
<td>3.2</td>
<td>Data preparation process</td>
<td>57</td>
</tr>
<tr>
<td>3.3</td>
<td>Procedure in designing Backpropagation Neural Network (BPNN) model</td>
<td>60</td>
</tr>
<tr>
<td>3.4</td>
<td>Hybrid GA-BPNN model</td>
<td>65</td>
</tr>
<tr>
<td>3.5</td>
<td>Hybrid BP-GANN model</td>
<td>66</td>
</tr>
<tr>
<td>4.1</td>
<td>Some samples of original face (photo) and its corresponding caricature</td>
<td>75</td>
</tr>
<tr>
<td>4.2</td>
<td>Landmark points on the original face image</td>
<td>76</td>
</tr>
<tr>
<td>4.3</td>
<td>Landmark points for face contour</td>
<td>77</td>
</tr>
<tr>
<td>4.4</td>
<td>Corresponding points on the caricature face (Pritchett, 2010)</td>
<td>77</td>
</tr>
<tr>
<td>4.5</td>
<td>Landmarks points for several samples of original face image</td>
<td>78</td>
</tr>
<tr>
<td>4.6</td>
<td>Facial landmarks configurations after translation</td>
<td>80</td>
</tr>
<tr>
<td>4.7</td>
<td>Facial landmarks configurations after translation and scaling</td>
<td>82</td>
</tr>
<tr>
<td>4.8</td>
<td>Facial landmarks configurations after translation, scaling and rotation</td>
<td>84</td>
</tr>
<tr>
<td>4.9</td>
<td>Contour of original face and its caricature after normalization using Procrustes superimposition method</td>
<td>85</td>
</tr>
<tr>
<td>4.10</td>
<td>Contour of original face and its caricature after normalization using modified Procrustes superimposition</td>
<td>90</td>
</tr>
<tr>
<td>4.11</td>
<td>A sample of input and target output in dataset N2</td>
<td>92</td>
</tr>
<tr>
<td>5.1</td>
<td>Structure of the neural network model</td>
<td>98</td>
</tr>
<tr>
<td>5.2</td>
<td>Graph of number of hidden nodes versus average NMSE</td>
<td>106</td>
</tr>
<tr>
<td>6.1</td>
<td>Hybrid GA-BPNN Model</td>
<td>121</td>
</tr>
<tr>
<td>6.2</td>
<td>Assigning weights and biases value to the related arrays</td>
<td>123</td>
</tr>
<tr>
<td>6.3</td>
<td>Assigning the weights and biases values to the neural</td>
<td></td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>6.4</td>
<td>Hybrid BP-GANN Model</td>
<td>126</td>
</tr>
<tr>
<td>6.5</td>
<td>Function popinit( )</td>
<td>128</td>
</tr>
<tr>
<td>6.6</td>
<td>Function BPLV( )</td>
<td>129</td>
</tr>
<tr>
<td>6.7</td>
<td>Function createRandom( )</td>
<td>130</td>
</tr>
<tr>
<td>6.8</td>
<td>Graph for number of generations vs average MSE for hybrid BP-GANN model</td>
<td>131</td>
</tr>
<tr>
<td>6.9</td>
<td>Target output and predicted output derived from all models</td>
<td>135</td>
</tr>
<tr>
<td>6.10</td>
<td>The actual caricature and predicted caricature from BPNN and GANN models</td>
<td>139</td>
</tr>
<tr>
<td>6.11</td>
<td>Actual caricature and predicted caricature of GA-BPNN vs BPNN model and GA-BPNN vs GANN model</td>
<td>140</td>
</tr>
<tr>
<td>6.12</td>
<td>Actual caricature and predicted caricature for BP-GANN model and all other models.</td>
<td>142</td>
</tr>
</tbody>
</table>
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDFM</td>
<td>Exaggerating the difference from the mean</td>
</tr>
<tr>
<td>PLS</td>
<td>Partial Least Square</td>
</tr>
<tr>
<td>KNN</td>
<td>K-Nearest Neighbour</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>BP</td>
<td>Backpropagation</td>
</tr>
<tr>
<td>BPNN</td>
<td>Backpropagation Neural Network</td>
</tr>
<tr>
<td>GA</td>
<td>Genetic Algorithm</td>
</tr>
<tr>
<td>GANN</td>
<td>Genetic Algorithm Neural Network</td>
</tr>
<tr>
<td>CCNN</td>
<td>Cascade Correlation Neural Network</td>
</tr>
<tr>
<td>PCA</td>
<td>Principle Component Analysis</td>
</tr>
<tr>
<td>SVR</td>
<td>Support Vector Regression</td>
</tr>
<tr>
<td>ASM</td>
<td>Active Shape Model</td>
</tr>
<tr>
<td>AAM</td>
<td>Active Appearance Model</td>
</tr>
<tr>
<td>RBF</td>
<td>Radial Basis Function</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Squared Error</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Squared Error</td>
</tr>
<tr>
<td>MAE</td>
<td>Mean Absolute Error</td>
</tr>
<tr>
<td>NMSE</td>
<td>Normalized Mean Squared Error</td>
</tr>
<tr>
<td>GA-BPNN</td>
<td>Genetic Algorithm- Backpropagation Neural Network</td>
</tr>
<tr>
<td>BP-GANN</td>
<td>Backpropagation- Genetic Algorithm Neural Network</td>
</tr>
<tr>
<td>GAbp</td>
<td>Backpropagation as an operator of Genetic Algorithm</td>
</tr>
<tr>
<td>SRM</td>
<td>Self-Reference Model</td>
</tr>
<tr>
<td>CPU</td>
<td>Central Processing Unit</td>
</tr>
<tr>
<td>RAM</td>
<td>Random Access Memory</td>
</tr>
</tbody>
</table>
# LIST OF APPENDICES

<table>
<thead>
<tr>
<th>APPENDIX</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Basic Concept of Artificial Neural Network</td>
<td>169</td>
</tr>
<tr>
<td>B</td>
<td>Backpropagation Neural Network</td>
<td>172</td>
</tr>
<tr>
<td>C</td>
<td>Designing Neural Network Model</td>
<td>174</td>
</tr>
<tr>
<td>D</td>
<td>Basic concept of Genetic Algorithm</td>
<td>176</td>
</tr>
<tr>
<td>E</td>
<td>Example of Original Face and Facial Caricature Dataset</td>
<td>179</td>
</tr>
<tr>
<td>F</td>
<td>Publications</td>
<td>184</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Overview

Caricature is a pictorial description of a subject or a person by exaggerating the most prominent features in order to make it different from others and create an easily recognizable likeness. A caricature may have limited similarity to the original subject and incorrect proportion but the subject can still be recognized. Caricature has been extensively used in our daily life for the past few decades. It often appears in magazines and newspapers for various purposes. It is used for entertainment, social expression or political purposes using humor or sarcasm. It is used in greeting card software to create some zany creations as well. With the emergence of Internet and mobile technology, caricature of face images have been used in many internet and mobile applications for social communication and entertainment over the web (Zhou and Liu, 2009) or in mobile phone (Kim, 2011) such as in games, live video chatting, forum and instant messenger services. A user can protect their identity and real image from other users for security purposes but still allows his/her basic facial gestures to be recognizable. Moreover, caricature is used as an avatar in virtual community (Liu et al., 2008; Dutta et al., 2012). Caricature is also more recognizable than a veridical portrait, which makes it very beneficial for face recognition (Smitaveja et al., 2009; Burton et al., 2015).

Caricaturing process involves two basic steps: observation and exaggeration. Caricaturists need to observe and identify distinct features in an individual’s face that
makes a person recognizable. The distinctive facial feature is the feature that is larger, smaller, sharper or rounder than a “mean face”. A “mean face” is an average face that human being encounter in everyday life and record in their brain (Rhodes et al., 1997). Caricaturists compare one’s face with this mean face to extract the distinct features and draw caricatures by reducing the unimportant details and exaggerating the distinctive facial features. How the caricaturists exaggerate the features depends on their style of drawing. A caricaturist has an amazing ability to capture the distinguished facial feature of someone’s face and has an inborn talent of drawing caricature, which is embedded in his/her subconscious mind. Unfortunately, this talent does not exist in all people and how the caricaturist draws the caricature is difficult to be explained. Therefore, how to generate caricature from input face image has become an interesting subject of research.

1.2 Background of Problem

Computer generated caricature is developed to assist users in producing caricature automatically or semi-automatically. It is derived from rapid advances in computer graphics and computer vision, and introduced as a part of non-photorealistic rendering technologies as well. Researchers in computer generated caricature try to convert the process of drawing caricature done by the caricaturist into formula and algorithm that can be executed by a computer. Different approaches have been utilized by previous works to generate facial caricature from face image. These approaches differ in terms of how the distinctive facial features are determined and how those features are exaggerated. Akleman (1997) came up with a very simple algorithm which utilizes interactive morphing tool to generate caricatures. He used trial and error method to find the distinctive features to be exaggerated interactively. Akleman et al. (2000) and Akleman (2005) further came up with a new deformation technique that uses simplicial complex to generate caricature. It can intuitively and interactively produce extreme caricatures. However, these works require skilled user such as professional caricaturist which limits their applicability and also add more loads on users (Nakasu et al., 2009; Lee and Byun, 2013). For ordinary, non-artist
users, it is difficult to identify the distinctive features, takes a long time in doing the trial and error process to get the desired result, and are prone to generate unrecognizable caricature.

Brennan (2007), the first person to attempt developing a caricature generator, formalized caricature as a process of exaggerating the differences between the features of a subject face and average face. The rate of exaggeration was defined interactively in order to produce line drawing caricature. Yamaguchi et al. (2003) and Tokuda et al. (2007) also used the notion of “exaggerating the difference from the mean” (EDFM) to produce caricature. They proposed an interactive system (PICASSO), which takes the model, caricaturist and gallery into consideration in order to determine the scalar of exaggeration rate. Chen et al. (2009) employed the handcraft rules of a particular caricaturist in his work. Du et al. (2015) established effective exaggeration rule based on the difference between the face image and standard face model. The exaggeration rates in those works were defined interactively by the user. Although those works have specific rules in determining the distinctive features, the exaggeration rate of that features needs to be defined interactively and intuitively by the user. These works still require skilled user input to control the facial feature points and to define the exaggeration rate in order to generate an interesting caricature and not a weird one. Some other works (Tseng and Lien, 2012; Yu and Zhang, 2013; Tseng and Lien, 2007; Kamimura and Chen, 2009) define the exaggeration rates empirically using eigenvector but they were unable to produce caricature like a real artist’s work.

In the process of generating artistic caricature using a computer, the style of the artist needs to be considered since the artist drawing style plays an important role in producing artistic caricatures (Lai et al., 2006; Sao, 2009). The caricatures of the same person painted by different artists will look different since every artist has his or her own drawing style to describe the unique features with different exaggeration rate. Yanushkevich and Shmerko (2007) also stated that if the art style of caricaturists can be understood, various benefits are expected in the application of caricature especially in face identification, recognition and matching technique. Unfortunately, previous works mentioned above did not consider the artist’s style of
exaggeration in their works. Very few works have attempted to observe and learn from the artist’s products. The challenge is that not all similar facial features will be exaggerated in the same style by a particular artist. It depends on many factors which include the distinctiveness of the features observed by the artist and intangible rules of drawing caricature in the mind of the artist that is hard to be explained explicitly and difficult to be coded in algorithmic way.

Learning-based approach has been proved to be a very promising way to deal with this problem. Liu et al. (2006) proposed a mapping learning approach to generate facial caricature. They employed Principle Component Analysis (PCA) to obtain the principle component of the facial features and used Support Vector Regression (SVR) to predict the caricature for the input face image. Liu et al. (2009) further came up with semi-supervised manifold regularization learning. However, both of the works only learnt general style of the artist by using hand drawn caricatures that are created by many artists over the world. Liang et al. (2002), Lai et al. (2006), Shet et al. (2005) and Yang and Lai (2010) learnt an individual artist style and used caricatures that are drawn by a particular artist. Liang et al. (2002) proposed a caricature generating system based on an example using Partial Least Square (PLS). This work used a linear method to map the original face image to its corresponding facial caricature. The exaggeration direction and selected facial features determined by this work were limited and the distinctive facial features selected by the artist may cover different prototypes. Lai et al. (2006), Tun (2009) and Shet et al. (2005) believe that generating facial caricature involves non-linear exaggerations. They proposed a neural network based caricature generation. Yang and Lai (2010) proposed a learning based system which uses K-Nearest Neighbour (KNN) regression to learn the relationships between the shape of the original image photo and the caricature. Although the caricatures produced by these works were claimed as successful results, there are still much room for improvements since the caricature was not exactly the same as the artist’s drawing. Additionally, there was no attempt to improve the accuracy of the model because no quantitative or statistical evaluation was made on these studies. Only a simple subjective evaluation from human perspective was performed which can only be accurately evaluated by an expert. A quantitative evaluation is required to assess the performance of such
methods and it can serve as a reference or benchmark to measure the performance and improvement of the learning based method used in generating a caricature.

Thus, the method of generating an artistic caricature that is similar to the one created by an artist without the needs of skilled user input still remain as an open research problem. In order to obtain the desired result, a study in capturing and imitating the style of a particular artist in drawing caricature should be conducted. Among the most crucial problems to be addressed by such study are inconsistency of the artist style and limited data collection of face image-caricature pairs (Lai et al., 2006; Yu and Zhang, 2013). Sometimes, the way of exaggeration done by an artist varies at different time periods or different conditions even on the same faces due to human factor. Thus, a method that can generalize the style inconsistency needs to be established in order to optimize the results. Besides, it is difficult to collect a huge number of face image pairs and facial caricature created by only one artist. This data limitation generally leads to inaccurate training results and causes unreliable prediction of the caricature if the new input face image is far from all the face images in the data collection. In addition, human faces have high similarities with each other and its caricature (Xu and Biederman, 2013). Hence, the most significant parameters needed to describe the data and a new approach to model the data should be determined so that an accurate result still can be obtained even on small amount of data collection.

Fortunately, accurate predictions can still be produced using Artificial Neural Network (ANN) on limited dataset because these models essentially depend on good quality dataset (Langer et al., 2006). As mentioned in Yu et al. (2006) and Yu et al. (2010), the data preparation process is very important since an effective data preparation process can produce significant information and good quality data, and this will result in the increase in the generalization ability of the prediction model. In addition, most of the previous works in learning based approach as mentioned above only used individual machine learning method to predict facial caricature. In other fields of study related to face image, hybrid methods have been successfully proved to produce a better result than individual methods such as Lin et al. (2011), Bhaiya and Pali (2012) and Melin et al. (2012). This is because hybridization techniques can
complement the strength of both methods and compensates each other’s weaknesses. Therefore, this finding motivates this study to use a good quality dataset and hybridization technique in developing the prediction model to improve the generalization performance of the model.

According to the aforementioned problem, due to the inadequacy of previous works and possible strategies that can be used to improve the results, this research will model the exaggeration style of facial caricature that will be used to predict facial caricature of the original face image in order to produce caricature that is almost as close as possible with the real artist’s drawing.

1.3 Problem Statement

It is shown in the previous section that the involvement of an artist’s exaggeration style in the caricature generation process will be able to produce an artistic caricature, but it is not an easy task to come out with such algorithm. Recent advancement in learning-based approach is used to deal with this problem, which is capable in capturing and learning the exaggeration style. However, inconsistency of the artist style due to human factor and limited data collection of face image-caricature pairs (Yu and Zhang, 2013; Li et al., 2016) may lead to inaccurate results and can generate inaccurate caricature that is different from the real artist products. This thesis proposes data preparation process that can produce good quality dataset and a hybrid technique in training the dataset to tackle this problem. This is due to the fact that good quality dataset which contains the most significant parameters to describe the exaggeration style of facial caricature can increase the generalization ability even using small dataset (Yu et al., 2010). High similarity of human faces (Xu and Biederman, 2013) and inevitable error caused by normalization process (Ni et al., 2008) are the challenges in the facial caricature dataset preparation process. Thus, a modified data normalization method is proposed based on Procrustes superimposition method which is simple to use (less complex), and facial features parameter is proposed based on geometric morphometric which is able to distinguish
subtle differences in facial features. Moreover, the hybrid technique which involved backpropagation (BP) and genetic algorithm (GA) could improve the prediction accuracy by taking the advantage of local and global searching ability to optimize the neural networks parameters, although such technique has never been explored before in this field. Thus, the main research question will be:

*How to model exaggeration style of a particular artist for better prediction accuracy of facial caricature of the original face image?*

### 1.4 Objectives

The objectives of this study are defined as follows:

i. To propose parameters of facial features and modified data normalization based on Procrustes superimposition method for better facial caricature dataset quality.

ii. To develop a hybrid neural network model that captures exaggeration style for prediction of facial caricature.

### 1.5 Scope

This study is carried out under the following scopes:

i. The pair of face image (photo) and its caricature is in 2D frontal view without accessories (such as a hat) or items that can obstruct the facial features (such as a finger), and both images must have similar pose and expression.

ii. The caricatures are drawn by only one artist that is John Pritchett.

iii. The number of sample is 32 pairs of face image and its corresponding caricature.
iv. This study only considers the exaggeration style of the artist. Other style of the artist such as materials used, types of brush stroke, colours used, or type of line drawings are not considered.

v. This study only focuses on the face contour or face shape due to the time limitation for completing the study. The same proposed methods and process can be extended to other facial features such as eyes, nose, and mouth.

vi. The facial landmark points are extracted manually to ensure the reliability of the dataset.

vii. Landmark-based approach is used for numerical representation of the facial features of the original face and its caricature.

viii. Neural network is chosen as the non-linear model and GA for the optimization algorithm.

1.6 Significance of the Study

This study is a significant endeavour in enhancing the artistic effect of the current computer generated caricature. If this study can successfully capture and quantify the exaggeration style of a particular artist and reliably predict facial caricature from a given input face image using the proposed facial caricature prediction model, the resulting caricature which is similar or very close to the artist’s works can be produced by integrating this result into the process of generating caricature which involves other areas of knowledge such as image processing, computer graphics and non-photo realistic rendering technology. The generated caricature will not only be interesting but also leave a deep impression and memory on the viewer and shows the style of a particular artist. There are various uses and benefits of caricature especially in face recognition field (Yanushkevich and Shmerko, 2007). This study can also be seen as a path to preserve a caricature artist’s style because his or her style can still be produced even if the artist is no longer available. Apart from that, this study also provides a way to evaluate a facial caricature prediction model quantitatively which has not been done before. Quantitative evaluation can also be used as a reference or guide for further
improvement of the facial caricature if learning based approach in generating caricature is used along with the qualitative evaluation.

1.7 Organization of the thesis

This thesis has seven chapters and is organized as follows:

i. Chapter 1: Introduction
   This chapter describes the problem background, specific problems to be tackled, objectives, scope and significance of the study.

ii. Chapter 2: Literature Review
   This chapter reviews the main subjects of interest which include the basic concepts, theory and process of generating caricature, existing approach in computer generated caricature, related algorithms primarily based on Artificial Neural Network (ANN), related methods which are Backpropagation Neural Network (BPNN), GA and Hybridization of both techniques.

iii. Chapter 3: Methodology
   This chapter presents the framework of the research which includes data preparation, design of proposed models, model evaluation and the hardware and software requirements.

iv. Chapter 4: Proposed Facial Caricature Data Preparation Process
   This chapter explains the process of data preparation which involves data collection, definition and extraction of facial landmark points, data normalization, different datasets generation, average face, input and target output.
v. Chapter 5: Finding the Best Parameters and Features for Neural Network Facial Caricature Prediction Model
   This chapter describes the development of Backpropagation Neural Network (BPNN) model which involves the selection of optimal parameters, dividing the face datasets and experimental setup. The development of Genetic Algorithm Neural Network (GANN) model which includes the selection of optimal parameters and trial experiment is also explained.

vi. Chapter 6: Hybrid Neural Network Model for Generating Facial Caricature
   This chapter discusses the implementation of the hybridization of Backpropagation Neural Network (BPNN) and Genetic Algorithm Neural Network (GANN) in two ways: GA-BPNN model and BP-GANN model.

vii. Chapter 7: Conclusions
   This chapter provides the conclusions which include research contributions and future enhancements.
REFERENCES


Network Based on Improved Adaptive Genetic Algorithm in Bridge 

propagation neural network prediction and finite-element model simulation to 

Fujiwara, T., Tominaga, M., Murakami, K., & Koshimizu, H. (2000). Web- 
PICASSO: Internet implementation of facial caricature system PICASSO. In 
*Proc. of 3rd International Conference on Advances in Multimodal Interfaces* 
(pp. 151–159). Springer-verlag, Berlin.

portrait caricature generation with facial components analysis. *IEEE 
International Conference on Intelligent Computing and Intelligent Systems.*

algorithms. In P. Siarry & Z. Michalewicz (Eds.), *Advances in metaheuristics 
for hard optimization* (pp. 199–221). Springer, Berlin Heidelberg.


algorithms in modeling of bimodal drug delivery. *International Journal of 
Pharmaceutics, vol 327, pp 126–138.*

realistic caricatures using a parametric facial appearance model. *Behavior 
Research Methods, 37, 170–181.*

Computing Model to Daily Average Temperature Analysis. *International 


Smitaveja, J., Sookhanaphibarn, K., & Lursinsap, C. (2009). Facial Metrical and Caricature-Pattern-Based Learning in Neural Network System for Face


