COMPUTERIZATION OF THE WARTEGG TEST IN HANDWRITING ANALYSIS

Lili Liu

A thesis
in
The Department
of
Gina Cody School of Engineering and Computer Science

Presented in Partial Fulfillment of the Requirements
For the Degree of Master of Computer Science
Concordia University
Montréal, Québec, Canada

December 2020

Concordia University
School of Graduate Studies

This is to certify that the thesis prepared

By: **Lili Liu**

Entitled: **Computerization of the Wartegg Test in Handwriting Analysis**

and submitted in partial fulfillment of the requirements for the degree of

Master of Computer Science

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining commitee:

______________________________ Chair

Dr. Charalambos [Charis] Poullis

______________________________ Examiner

Dr. Muna Khayyat

______________________________ Supervisor

Dr. Ching Y. Suen

Approved ______________________________

Chair of Department or Graduate Program Director

__________ 2020 ______ ______________________________

Dr. Mourad Debbabi, Interim Dean,

Faculty of Engineering and Computer Science

Abstract

Computerization of the Wartegg Test in Handwriting Analysis

Lili Liu

Wartegg Test (WT) is a drawing completion task designed to reflect the personality characteristics of testers. WT is a functional and usable psychological test which is not yet well known due to the language and region barriers among worldwide researchers. It is urgent to automate WT's psychological analysis process to let more people realize and use it, especially during the Covid-19 pandemic period, when people are asked to stay at home for a long time with limited access to in-person psychology consulting. The computerized WT allows people to attend psychology tests at home without taking the risk of virus infection. In this thesis, we proposed a system of computerizing the Wartegg test in handwriting analysis. We have implemented the image processing methods and machine learning algorithms on the Wartegg test to construct a personality analysis tool, which could mimic the evaluation process of psychologists automatically. We have extracted five features based on WT theory, they are space utilization, the numerical order of sequence, line curve ratio, animate or inanimate classification and category classification. After calculating all five features, we added them together and projected the summarized result to the Big five personality traits to get the final result. This system allows people to attend the Wartegg test at home and get the result immediately. The processing of the system has nothing to do with language and region, and only analyzes and obtains results based on drawings. Hopefully, this work could facilitate people to attend the Wartegg psychologist test with privacy security.

Acknowledgments

First and foremost, I would like to express my sincere thanks to my supervisor Dr. Ching Y. Suen of the Concordia University of Montreal, Canada, for his continuous guidance, encouragement and support leading up to complete this work. Dr. Suen has deep attainments in handwritten character processing and pattern recognition. He welcomes academic discussions at any time and spends a lot of time and energy to guide and inspire me to conduct more in-depth research and thinking. In terms of article writing and submission, Dr. Suen was able to modify and guide me in every detail. I learned a lot from him in doing research, as well as a social experience with people. I am very fortunate to have such an excellent and kind person as my supervisor, and I also enjoy my two years stay at the Center for Pattern Recognition and Machine Intelligence (CENPARMI).

My warmest thanks are extended also to Graziella Pettinati for her great efforts in providing training, valuable feedback and corrections in the field of psychology. Her suggestions and insights have greatly strengthened this work. I also want to thank all the friends, colleagues and staff of the CENPARMI for making this journey easier. Thanks to my committee members Dr. M. Khayyat and Dr. C. Poullis for their efforts and advice on my thesis.

I am very grateful to the collaborators Nam Tuan Ly and lab-mate Yunqi Xu of our papers. I am very fortunate to meet them and collaborate with them on experiments and papers. Thanks for the time of cooperation, all discussions and joint efforts. Thanks to my roommate and as same as the real family member Lu Liang, who has been by my side since I first arrived in Montreal as soon as I got off the plane, and spent the past two years exploring the mysterious Montreal with me. I thank her for the company, comfort and stability she provided in my life. Thanks to my friend Sidi Yang for helping each other in life and making progress together.

Finally, I would like to express my heartfelt thanks to my parents, who have been

confident in me and supported me throughout the journey. I am very grateful to them for their dedication and support. Their wise advice and support that guided me to go so far as to realize my dream. Finally, I would like to thank my younger brother and sister for taking care of our family. Sharing the duties for me so that I can study with peace of mind without worrying about our parents. Finally, thanks Xubo Lyu for his understanding, patience, support and academic suggestions and discussions.

I cherish and love this journey of studying in Montreal very much. This period is a precious memory of my life. Thanks for all the moments and people I met on this journey.

Contents

List of Figures **viii**

List of Tables **ix**

1 Introduction **1**

1.1 Wartegg Drawing Completion Test 1

1.1.1 The history of WDCT 1

1.1.2 Wartegg Test research in North America 4

1.1.3 Computer-aided Wartegg Test 5

1.2 Contributions of this Thesis 6

1.3 Outline of this Thesis 6

2 Related Work **8**

2.1 Wartegg Test research 8

2.2 Computerization of WT research 8

3 Methodology **11**

3.1 WT dataset pre-processing steps 11

3.2 WT feature extraction 14

3.2.1 Space Utilization 15

3.2.2 Animate and inanimate 16

3.2.3 The Numerical Order of Sequence 18

3.2.4 Line and Curve ratio 24

3.2.5 Categories 28

3.3 Big Five Personality evaluation method 38

3.3.1 BFP introduction 38

3.3.2 BFP criteria for Wartegg Test 39

4 Dataset and GUI. 41

4.1 WT dataset 41

4.1.1 WT dataset collection and environment setting 41

4.1.2 WT dataset description 42

4.2 Result from five features. 43

4.2.1 GUI Demo exhibition. 43

5 Conclusion and Future Work 45

5.1 WT Conclusion and discussion 45

5.2 Future Work 45

5.2.1 Feature extension to our system. 45

5.2.2 Multi labels sketch classification. 46

5.2.3 Color the sketches with generative models. 46

References 47

List of Figures

1 A sample of the Wartegg test form. 2
2 The literature review of Wartegg Test. 4
3 Flow chart of the whole system procedure 12
4 Split Images 14
5 Space Utilization results 17
6 Animate and Inanimate classification results 19
7 Four exemptions of order sequence results 23
8 Lines Process results 27
9 Diagram of the whole WT procedure 30
10 Model of inception v3 32
11 Pre-train and fine-tuning process of stage two 33
12 Quick Draw samples 34
13 TU-Berlin samples 34
14 Completed WT sample 41
15 GUI Demo 44

List of Tables

1 Lenet-5 model architecture 18
2 All combination of sequence's mix level 20
3 All possible combinations and CTD. 21
4 Concise rule of the CTD 22
5 First two last two combinations and evaluation principle for Emotional Stability 24
6 Evaluation Index in multi-categories classification experiment 36
7 Classification results of Q-T Sketch dataset 37
8 Extraversion and agreeableness 39
9 Emotional stability 39
10 Conscientiousness 40
11 Openness to experience 40
12 Big five personality result 40
13 Wartegg Test Dataset Information 43

Chapter 1

Introduction

1.1 Wartegg Drawing Completion Test

1.1.1 The history of WDCT

Have you ever attended a psychological test? Many people have attended psychological tests to predict their personality consciously or involuntarily. There is a psychological test that allows users to draw sketches and analyze their personality from their drawings: Wartegg test. Wartegg test (WT) is a drawing completion task that was designed to reflect and analyze the drawers' personality characteristics by observing the way testers finish a drawing with well-designed stimuli signs. WT was first proposed by Ehrig Wartegg in 1926 and was published in 1953. The Wartegg test questions form is designed as Fig. 1 shows. Eight semi-structured drawings are arranged in two rows, four boxes in each. Every box contains at least one of the little dots, lines or curves. The stimuli and their combination choice are specially designed based on psychology principles. Each box has its own structure with different stimuli[1] and meanings [2] of personality characters for psychologists to deal with.

- Box 1: A small round point located in the center of the box. It indicates the self-identifying, ego, the relationship with others.
- Box 2: A small curve located in top left corner parallels the dialogue direction, which is related to flexibility, movement in society, liveliness, emotional release (outward expression of emotions);

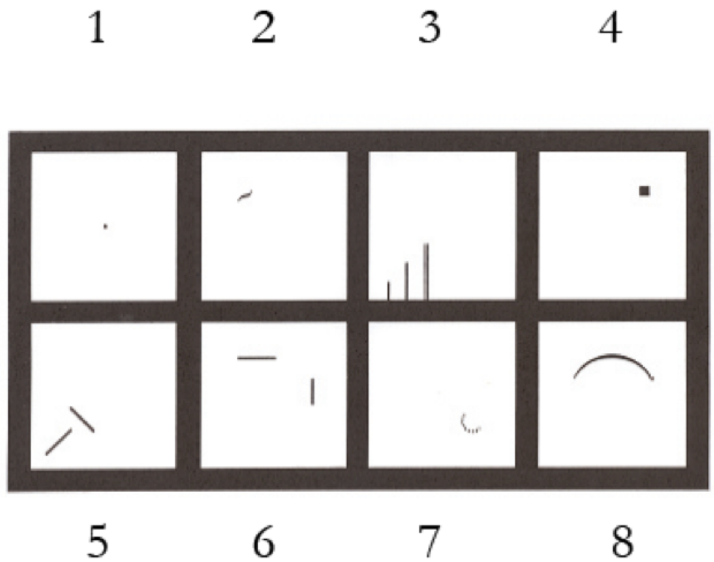

Figure 1: A sample of the Wartegg test form.

- Box 3: Three straight lines are parallel to each other, standing at the bottom of the box evenly, with increasing height from left to right. They related to attitude to the goals, money and feature, aspiration, eagerness, goal-directness, perseverance and conversely, a lack of orientation towards goals and achievement.

- Box 4: A solid black square placed at the top right of the box, which is related to the "obstacle" or problem, challenge, difficulty. It reflects how testers react when faced with challenges and problems.

- Box 5: Two perpendicular lines located at the bottom left corner of the box, with a small gap of the cross point. This is related to conflict, connection, aggression, disagreement and opposition. This initially reflects how the tester handles conflicts.

- Box 6: The two lines are perpendicular to each other but do not intersect. They are located in the upper middle of the box. These signs indicate testers' abilities to find out the potential relationship between signs and aggregate them together. This is related to kindness, consciousness, delicacy, cooperation, how expressed, repressed, sublimated or open to influence.

- Box 7: A series of points is located in the lower right corner of the box to form a semi arc, with an open opening to the upper right. They are related to protection. It shows where and how the subject feels protected and how he defends himself. The individual entries are indicated with a black dot, a so-called bullet.

- Box 8: A downward-facing curve which is related to the text in the entries may be of any length.

The tester was encouraged to develop these signs into meaningful drawings in their own way. During the administration procedure, the tester sits opposite to the administrator. The tester was given a WT paper form and a pencil of the HB #2 without an eraser and was told by "make a drawing in each box with a completed meaning, preferably the first which comes in your mind and abstract drawings and the drawing number. It's not necessary for you to follow the numerical order; work at your own pace and there are no time restrictions." To any questions from the tester, the administrator or psychologist will respond avoiding influencing the subject's response and reinforcing the belief that everything will be good because there are not right things or wrong things to do. The assumption of WT is based on the theory that one's picture can reflect his emotional status and personality characters, based on this assumption psychologists can analyze testers' personality. Some fields are connected with each other, those parts can be evaluated together. Boxes 1 and 8 are reflecting Self-Image. Self-esteem and security are connected and frequently dependent on whether or not the symbiotic union with the mother was successful. The emotion status can be classified as extrovertive and introvertive. Emotions are related to box 2, box 3 and box 7. Boxes 3 and 5, can indicate ability and productivity, work and achievement. Boxes 4 and 6 are a hint of tester‘s attitude towards the world.

The WT has been developed for at least 80 years and benefited from many aspects. The WT project was first developed in the 1920s and 1930s by the Austro German psychologist Ehrig Wartegg (1897-1983) [3]. During that time, the Wartegg test was practically unknown in English-speaking countries, while it was widely used in continental Europe and Latin America. In 1923 Wartegg adopted psychoanalytic ideas and made his first attempts as a therapist. According to the author, he was initially inspired to develop a drawing test from reading Richard Wilhelm's book, I Ching [4], on Chinese philosophy [5]. Another factor contributing to the development of Wartegg's test was the ascent of modern art in the 1920s [5]. Wartegg's academic

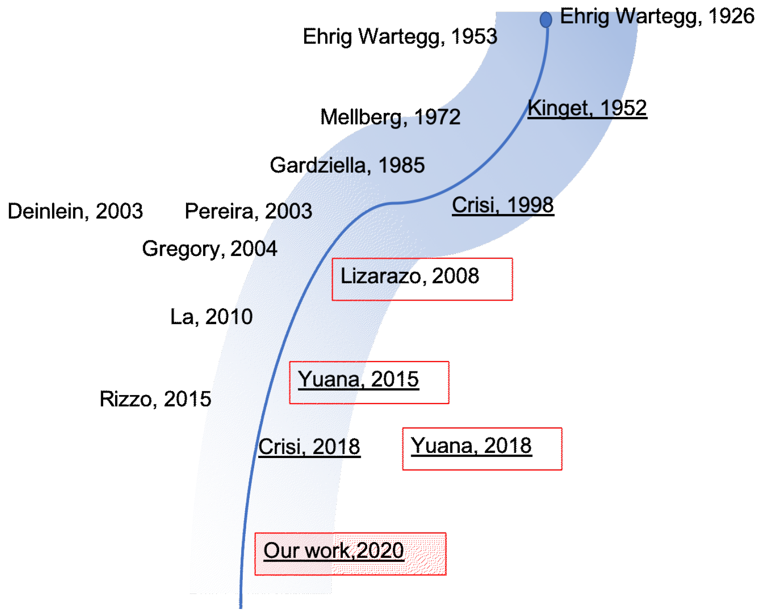

Figure 2: The literature review of Wartegg Test.

work on the WT at the University of Leipzig during the 1930s was based on the doctrine of Ganzheit psychology.

Wartegg test became famous in several countries including Finland, Brazil, Germany and so on. Several interpretation methods have been developed for the WT. The WDCT has been widely applied in sociability evaluation [6], personnel selection during recruitment process [7][8], vocational counselling [9], education counseling [10], [11], couple and family relationship therapy and inside condition diagnosis of patients [12], [13]. These applications prove its abilities, reliability and validity.

1.1.2 Wartegg Test research in North America

WT was widely spread in German-speaking countries and Latin America. However, it is not so popular in English-speaking countries [14]. In particular, in North America, the first WDCT manual was published by Kinget in 1952 [15]. Based on the work of Kinget, Italy researcher Crisi extended and validated the Wartegg test in 1998 and 2018 subsequently [16], [1]. It is reasonable for WDCT progress greatly in German-speaking countries since the creator is an Austro-German psychologist. Language and regional restriction could hamper the expansion of WDCT theory. The obstacles of WDCT mainly come from the academic exchanges between psychologists and the communication and analysis between psychologists and testers. Due to language barriers, there are fewer psychologists who speak English, so there are fewer opportunities for WDCT theory access affecting the spread of WT research and usage.

The good point is that WT is a graphical projecting psychological test. The test itself has very little language barriers. Unlike the text and audio, a sketch can work as a universal tool for communication. Its highly abstract and concise features express a great amount of information compared with other communication tools. Hand-drawn sketches were regarded as the representation of people's inside world. Especially for personality tests, drawing based tests could erase the barrier of language, gender, national origin and age, and easily reflect the preference and inside personality of the tester.

The WDCT proves its ability in a great range of applications and prospects dealing with different groups of testers. There is prospect for WT to involve children and disabled persons in psychology tests and contribute to the mental health carefulness of different groups. However, the traditional WT procedure is hosted by psychologists to predict tester's personality by analyzing their hand-drawn patterns on paper. This kind of in-person WDCT evaluation process is heavily restricted by the availability, language, accessibility and experience of the expert. Computerization of human-craft work is a typical solution in many fields.

1.1.3 Computer-aided Wartegg Test

In wartegg test area, only three works have tried to exploit the computerization analysis system implication. As shown in Fig. 2, the paper in the red rectangle indicates the progress of computerization analysis system implications. In 2008, Campos [17], the first try of organizing Wartegg test through a software, was created to analyze test results and also to provide the subject's personality. It transformed the WDCT form into an electronic version, which reduces the processing time by 60%. Fast Wartegg Analyzer Tool (FWAT) is another kind of web-based application designed for preparing the original image data for further exploration faster and easier, which was created by Yuanna in 2015. The speed of the evaluation process accelerated by 65% of time compared to manual operation [18]. In 2018, Yuana promoted an approach based on the FWAT to calculate the cosine similarity between the ideal characters and candidates' characters in order to select a perfect employee [19].

Hand-drawn pattern recognition is easy for humans because they possess prior knowledge from the real world, but it is difficult for the computer due to complex and abstract features. To overcome such issues, this thesis propose the computerization of the Wartegg test in handwriting analysis, a fully-automatic WDCT system based on Digital Image Processing (DIP) and Machine Learning techniques. This system extracts multi-modal features and analyzes them under the Big-Five traits automatically. This system can mitigate the heavy manual labour of psychologists and provide clients with flexible access. This thesis proposes the computerization of the Wartegg test in handwriting analysis.

1.2 Contributions of this Thesis

In this thesis, we propose to implement the image processing methods and machine learning algorithms on the Wartegg test to construct a personality analysis tool. The main contributions of this paper are: we collect a new dataset of finished Wartegg Test forms and label them by the content with a professional psychologist's help. We created a tool with a series of codes to evaluate the performance of selected features to predict personality with our new dataset. These contributions have been published in the Second International Conference on Pattern Recognition and Artificial Intelligence (ICPRAI 2020).

1.3 Outline of this Thesis

The content of this thesis is organized as follows:

Chapter 1 gives a general introduction of the Wartegg test. We describe the history of WT and the advantage of WT when compared with other format psychological tests. We moved our attention to the computerization Wartegg test, when analyzing the disadvantages of WT and proposed some new methods.

In Chapter 2, we researched the related work of computer-aided WT research and technologies used in the Image and Machine Learning method. The computer-aided WT research is more about general WT description. The Image and Machine

Learning methods are mainly focused on detailed hand drawing features from WT sketches.

In chapter 3, we integrate five selected features with different algorithms to achieve the personality test. These five features are categories of sketches, line and curve ratio, space utilization, animate an inanimate character and the numerical order of sequence. The evaluation of the five features is based on the Big five personality factors, which demonstrates the results of the experiment of BFF features and predicted personality.

In chapter 4, we illustrate the detailed information of the collected Wartegg test dataset and our GUI system of the Wartegg test.

In chapter 5, we discuss the conclusion and future work of this research.

Chapter 2

Related Work

2.1 Wartegg Test research

Wartegg Test was proposed for more than eight decades since 1939 and was first published in English for at least six decades since 1952 by Dr. Kinget. Following work hosted by Alessandro Crisi, which inspired by Kinget's theory, developed a functional and useful method and evolved several times in the last 30 years (separately in 1996 [20], 1998 [16], 2007 [21], 2016 [22], 2018 [1]). This thesis is partially based on Crisi's theory and Kinget's theory, the selection of theoretical knowledge is based on the realization of computerization.

2.2 Computerization of WT research

When comes to the computerization of Wartegg test, only a few of the researchers tried to change the WT evaluation into the computer-aided analysis or combined general or advanced computer technologies into WT evaluation procedure. In 2008, Campos [11] created the electronic version WDCT form and made it possible for a psychologist to organize it electronically. This is the first attempt of computerizing WT operations into automatically software applications, which helped to reduce the whole WT processing time by 60%. They created a YouTube video for the latter researcher to have an overall review of the software usage. In 2015, Yuana proposed the Fast Wartegg Analyzer Tool (FWAT), which was a kind of web-based Wartegg test process system. The FWAT optimized the psychologist's data acquisition operation by preparing graphs and ready to be evaluated. By preparing the original image data for further exploration faster and easier, the speed of the evaluation process accelerated by 65% compared to manual operation [18]. Taking advantage of computer technology PHP and MySQL, FWAT allows clients to access web-based software through the internet from any place in the world. Moreover, FWAT could proceed and compare many WT samples at the same time, which is more powerful and effective than the traditional method. According to the paper, [18] makes the WT analyzes procedure faster and efficient when compared with the manual system, however, all these changes were based on the format of the original paper Wt form instead of content and logical analysis of WT with intelligence. The author pointed this problem out at the end of this paper and advised to improve FWAT by adding an expert system. Pattern recognition and fuzzy logic are another two provincial solutions for WT automatic problem.

In 2018, Yuana continued his work of 2015 [18] and improved it with Cosine-Similarity for data similarly filtering for WT personality selection result. He evaluated each Wartegg test character of selected testers and then calculated the cosine-similarity value of selected several testers. The calculated value will be compared with the expected character value of WT and became a key component for filtering candidates while recruiting. This work introduced a complex computer technology, cosine-similarity, into the WT application, which improved the accuracy and efficiency of recruitment selection exploration.

Crisi has developed the Crisi Wartegg Test system, Italian scoring online software [21] and English-language scoring online software [23]. The online scoring service was created aiming to improve the liability, precision and accuracy when evaluating WT consultation. Working like a retrieval system, the online software could return a most related scoring guideline to clinicians or let certified expert-created new guidelines according to the client's submitted query form. Formally, the service changes traditional administration procedure more electronically by simplifying the procedure with the database query.

In 2020, Lili Liu tried to combine WT with artificial intelligence and image processing technology to automatically analyze the WT results to provide psychologists with numerical and quantitative analysis reference results. They proposed the Computer-aided Wartegg Drawing Completion Test framework (CA-WDCT), a fully-automatic WDCT system based on Digital Image Processing (DIP) and Machine Learning techniques. The CA-WDCT system extracts the specific features and analysis them following the big five traits rules automatically. This CA-WDCT feature extraction and analysis method can mitigate the heavy manual labour of psychologists and benefit clients with easier access. It helps institutes or individuals to have a better knowledge of intrinsic personality characters. This is the first attempt to introduce the intelligent algorithm into Wartegg test analysis [24].

In 2020, Nam Tuan Ly proposed to use object detection models to help with scoring the emotional quality of WT. The object in this project was the completed drawings in each box. The whole method consisted of two steps, he used the object detection model YOLOv3 to detect square frames from WT pages firstly and then detect objects in the WT square frame [25].

In 2020, Lili Liu developed a technique that allows computers to recognize hand-drawn sketches from the Wartegg test. In particular, she transferred and fine-tuned a deep learning model that is well-trained on a huge natural image dataset (Imagenet) on the limited sketch images. The method achieved the state-of-the-art result of 94% of top-1 accuracy and 98% of top-5 accuracy of open source sketch dataset and 63% top-5 Wartegg test results. Since the classification of natural images has been developed greatly, the transfer learning idea could help by taking advantage of applying prior research into handwriting sketches of WT.

The research in English-speaking countries is rarely far away from the blossom, let alone the computer-related WT research, there is still a long way to go in computerization the Wartegg test in handwriting analysis.

Chapter 3

Methodology

The computerized Wartegg test process is described in four steps as shown in Fig. 3 First of all, we need to isolate the eight single boxes of the Wartegg test form. Then, five predefined features are extracted automatically from each box by Computer-aided Wartegg Drawing Completion Test. Once the features are extracted from drawings, character analysis can be processed for a better understanding of hand-drawings and corresponding personalities based on the Big Five personality traits. Finally, the software system collects all these functions and generates personality evaluation.

Feature extraction is the most important part of WT system. It makes fully-automatic analysis possible without human intervention. Moreover, the extracted features come from the key clues for personality evaluation. Once we get the single box of drawings, the data for further steps are well prepared.

3.1 WT dataset pre-processing steps

When the questionnaire forms were collected and scanned into images, we can start to design the system and analyze the features from WT. As we know, the WT form is consisted of eight drawings and the personality analysis is based on each drawing. So the first step is to split eight single drawings from the scanned questionnaire form.

We tried a variety of built-in image processing methods for image cutting, such as OpenCV's rectangular contour detection, template matching, region of interest detection, using different threshold methods to cut contour images and so on. Some algorithms are effective for part of images, but it is difficult to cut arbitrary image

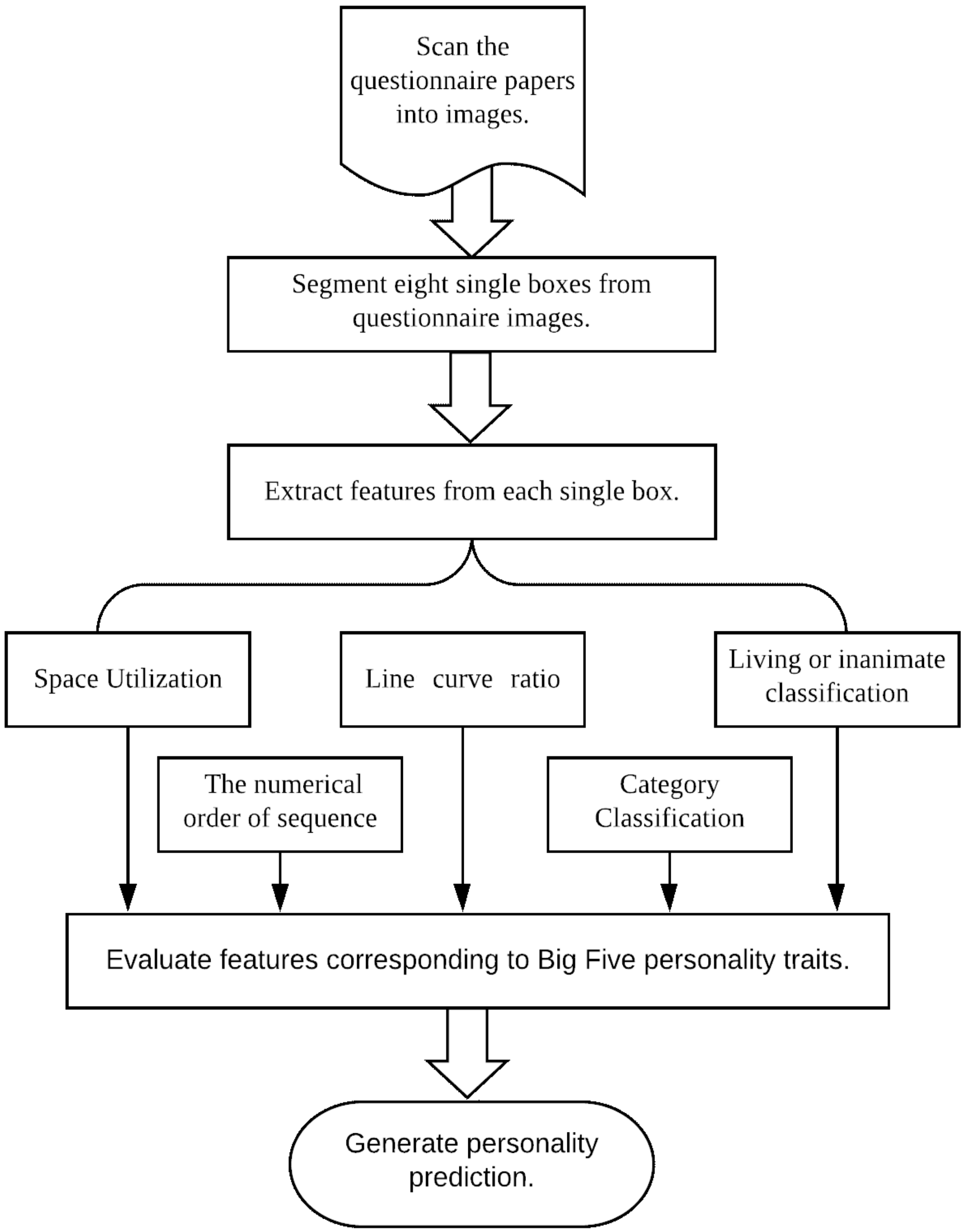

Figure 3: Flow chart of the whole system procedure

perfectly. The main influencing factors are the following four points: 1. WT sketches are arbitrarily connected with the black border, which destroys the characteristic of the rectangular structure. 2. The WT form is slightly rotated when pasted to questionnaire, which causes difficulty to framework detection. 3. The pixel color of the border of the WT form is uneven and with holes inside, which affects the cutting and test results. 4. Due to the characteristics of the experiment, the single box after cutting requires high accuracy and cleanness without black residual pixels. These four points caused the hardship for build-in algorithms to achieve the functions meeting the requirements of WT experiment.

In order to extract eight individual boxes from the questionnaire (completed sketches). Therefore, we propose a Watershed Scissors algorithm, which is specifically designed for WT form segmentation. It uses a modified template matching method to extract the outer contour of the black rectangle in the WT form. This black rectangular outline has obvious features and hardly changes with the affine transformation and rotation of the picture. The image distortion caused by the rotation generated can be reversely corrected by calculating the angle and roll it back. In the cropped table, the statistical pixel distribution is scanned by column and row to find the boundary and the watershed of the figure. Then, "scissors" started to work, cutting the aligned watershed into eight separate boxes like paper-cutting scissors and storing them in their original numerical order. Due to the problem of arbitrary connection between hand-drawn image strokes and form boundaries, the existing methods are difficult to cut out hand-drawn images clearly and losslessly. However, the watershed scissors algorithm we proposed can be 100% accurate when the image has rotation and chaotic strokes rate perfectly cut out target images. After that, we use this model to predict each single WT image and get the classification result. This algorithm has also been used in paper [24]. The split result shows in Fig. 4. The algorithm firstly cutting the WT form with black border out of the questionnaire and the cutting out each single drawing from the WT form. When we cut out eight individual single drawings from scanned questionnaire forms, we can extract five features from them and perform numerical analysis.

Algorithm 1 Watershed Scissor algorithm

Input Image of scanned Wartegg Test Questionnaire Form.

Step 1: Finding contours from gray input images.

Step 2: Matching the pre-defined black border rectangle template and filter target shape from matched out results.

Step 3: Cutting the form (target shape) out and compute its rotation angle.

Step 4: Form rotation correction and Warp Affine.

Step 5: Compute the pixel distribution of the form both horizontally and vertically.

Step 6: Trimming black borders at the bottom of the pixel distribution watershed.

Step 7: Sort and filter sub-images and get target drawings.

Result: Eight single drawings from each questionnaire form.

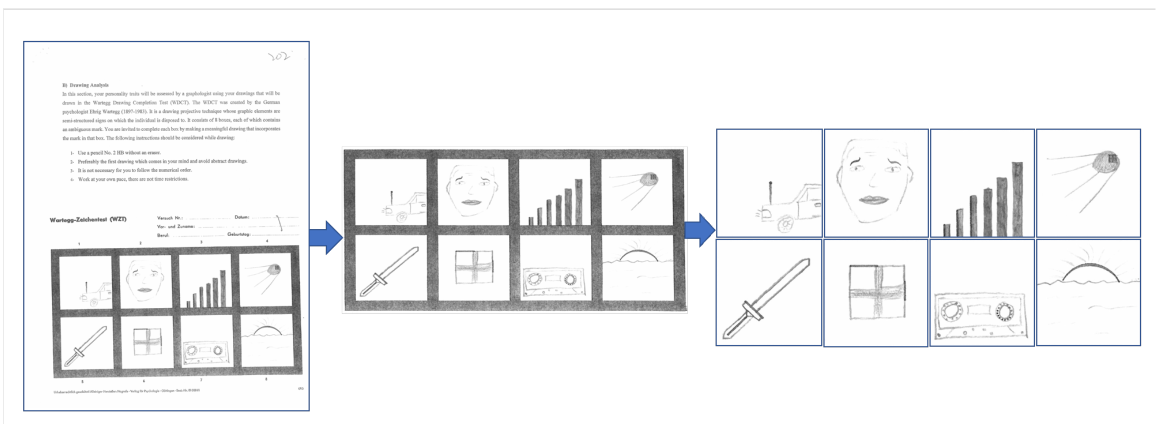

Figure 4: Split Images

3.2 WT feature extraction

How to analyze personality from hand-drawn images? This is the primary issue of psychological analysis. WT is based on the user's hand drawing. Features are important hints for WT analysis based on the characteristic information for personality analysis. The WT analysis process involves a mixture of multiple painting features and analyzes the personality features that are mainly for the user based on these features. In this thesis, five features are selected: space utilization, animate and inanimate, the numerical order of sequence, the ratio of straight lines and curved lines and category classification. The following paragraphs analyze each feature of each section.

3.2.1 Space Utilization

Space Utilization definition and description of WT.

Space Utilization aims to capture the tester's space and balance manipulate abilities. The space utilization feature is derived from the analysis of the execution of drawings [15]. This theory was extended by Kinget from Wartegg's original Wartegg test administration principle and proposed for further exploring the execution of drawings instead of focusing solely on content as many other evaluate systems do. The feeling of space is an important feature of tester's interior world, it has a prodigious influence on our mood, behaviour and physical activity: in short, human psychology. The feeling of how to cover the space we choose for an interior or exterior area is driven by the psychology of space. Space psychology plays a major role in defining the look and feel of a place. The psychology of space is a subject in itself. Understanding the importance and usage of space could reflect on understanding one's personality. This feature values the coverage level of drawings, use of available space provided by each test box. The feature we have exploited is based on space utilization and drawings distribution.

The space utilization feature evaluates the coverage used in each box. The coverage rate is the percentage of the covered area over the whole box area. Basically, all strokes would be detected and find out the outer points, which are treated as the anchor of locating the coverage map. Specifically, we need to find at least four most outer points as the shape corners of coverage and then drawing a red coverage, which includes every pixel in the box drawn by testers. All these outer points consist of the key corners of the contour shape covering the whole sketch. The contour is detected by Canny method and then a percentage computed from the contour shape over the whole box area. For example, if you paint the whole box with your pencil penmanship, the space utilization of this box is 100%. Different people would prefer to manipulate the space differently, so the space utilization situation represents the testers' feeling of space, their choice to utilize the space and farther implicate the personality.

Result samples of Space Utilization.

Space Utilization aims to capture tester's abilities of space and balance manipulate. We compute the space utilization by using convex hull algorithm to extract the covered area of sketches , as shown in Fig. 5. For each single drawing, we change it into binary image and remove noises. Then the contours are extracted from the proceed images. While computing the space utilization, it is necessary for the system to treat all sings in one image as an object. So we combining all single contours together to generate one big contour. Then, applying convex hull algorithm to find out the red area, which covers every stroke in the image. By computing the ratio of red area and the image size, we can get the space utilization ratio of the image. Whether it is one object or multiple objects in each drawing, we guarantee that they will be covered by one red area instead of being covered by multiple small red areas. In the case of multiple small objects, there are multiple blank gaps between objects, but we assume that these areas are already used and treated them as an object.

3.2.2 Animate and inanimate

Animate and inanimate feature definition and description of WT.

The feature of animate and inanimate is also derived from Kinget's theory. It is designed for extending the content information. Animate and inanimate is describing the content of each box. The content object with life and related to life is Animate. Otherwise it is inanimate. For example, faces and sun can be treated as animate object and sword and car should be treated as inanimate object. In this part, we extracted the drawing features of their animate or inanimate category. A special convolution neural network Lenet-5 [26] is used for classifying animate and animate drawing objects. The first step is to enlarge the data with four different augmentations: Image Sharpening, Unsharp Mask, Adaptive Threshing and Binary Threshold. The ratio of training: validation: testing is $6:2:2$. Images are resized into 28×28 and normalization is applied. Lenet-5, a classical model of deep neural networks, has been commonly used in deep learning areas, especially in handwriting classification.

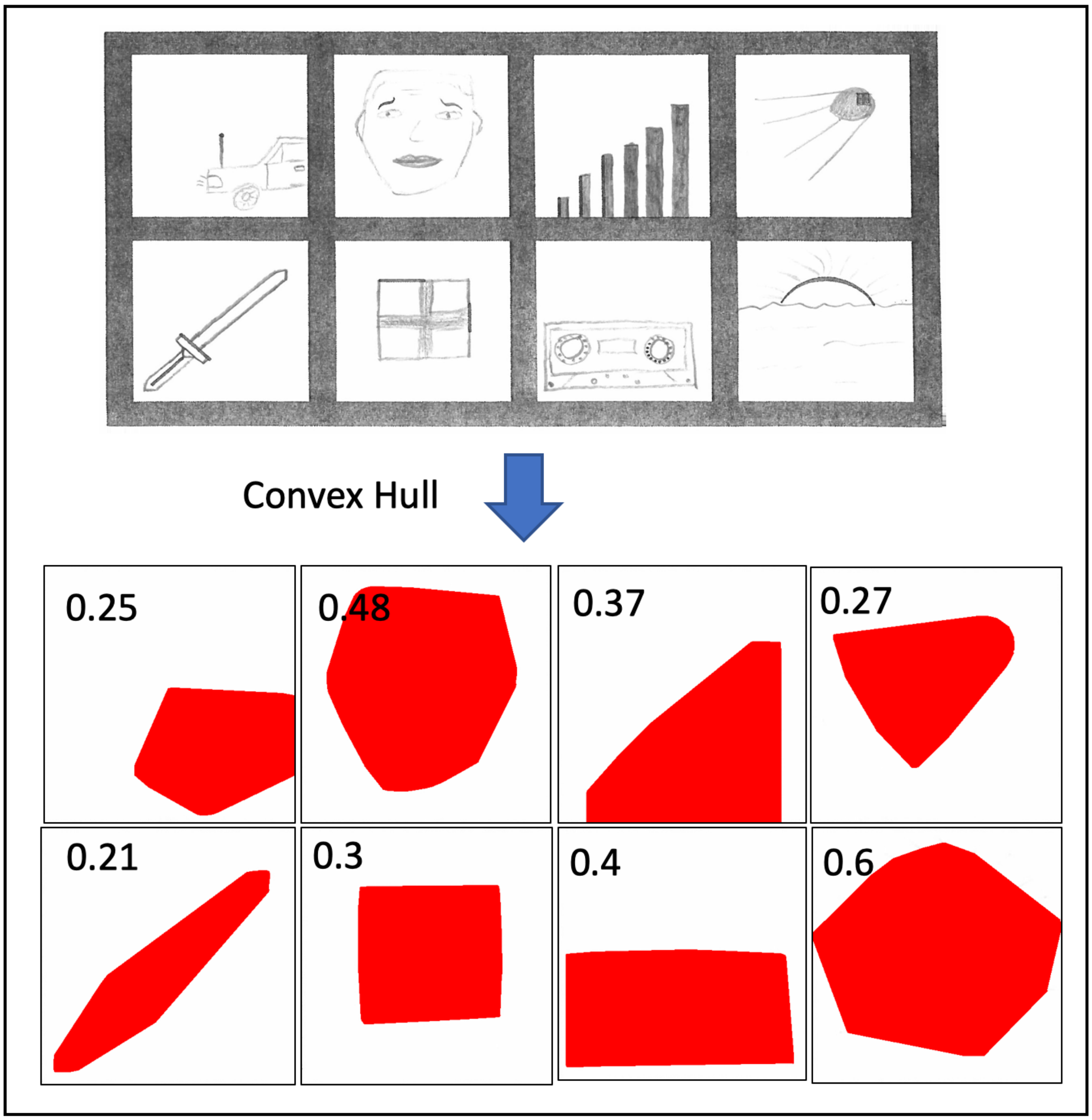

Figure 5: Space Utilization results

The model is a 6 layer convolution (Conv) neural network (CNN) with the architecture of Conv, Relu and Pooling repeated twice and following a fully-connected layer and Sigmoid function at the end. The model details are shown in Table 1. In order to improve the result, we used the Adam optimizer, Mini-batch and early stopping method while training the model. With this CNN model, we got the training accuracy of 98.46% and testing accuracy of 86.15%.

Table 1: Lenet-5 model architecture

Layer	Filter Shape	Activation function
Conv2D	(32, (7, 7))	relu
MaxPooling2D	(2, 2)	-
Conv2D	(64, (5, 5))	relu
MaxPooling2D	(2, 2)	-
Conv2D	(64, (3, 3))	relu
Flatten+Dense	(64)	relu
Dense	(2)	sigmoid

Animate and inanimate classification result.

The feature of animate or inanimate classification is a binary question. The famous neural network Lenet-5 was used to finish this function. We labeled all the single drawings as animate or inanimate and training the Lenet-5 model with these images. One of the WT sample was tested on this model, and its results are shown in Fig. 6. For the object with obvious creature feature, the function predicts it as animate, such as a face and an animal.

3.2.3 The Numerical Order of Sequence

Definition of the Numerical Order of Sequence

The number order of sequence is an important feature of Wartegg test personality analysis as several scoring systems are regarded. The order implicates tester's choice and preference. The choice of drawings and boxes, during which the client identifies

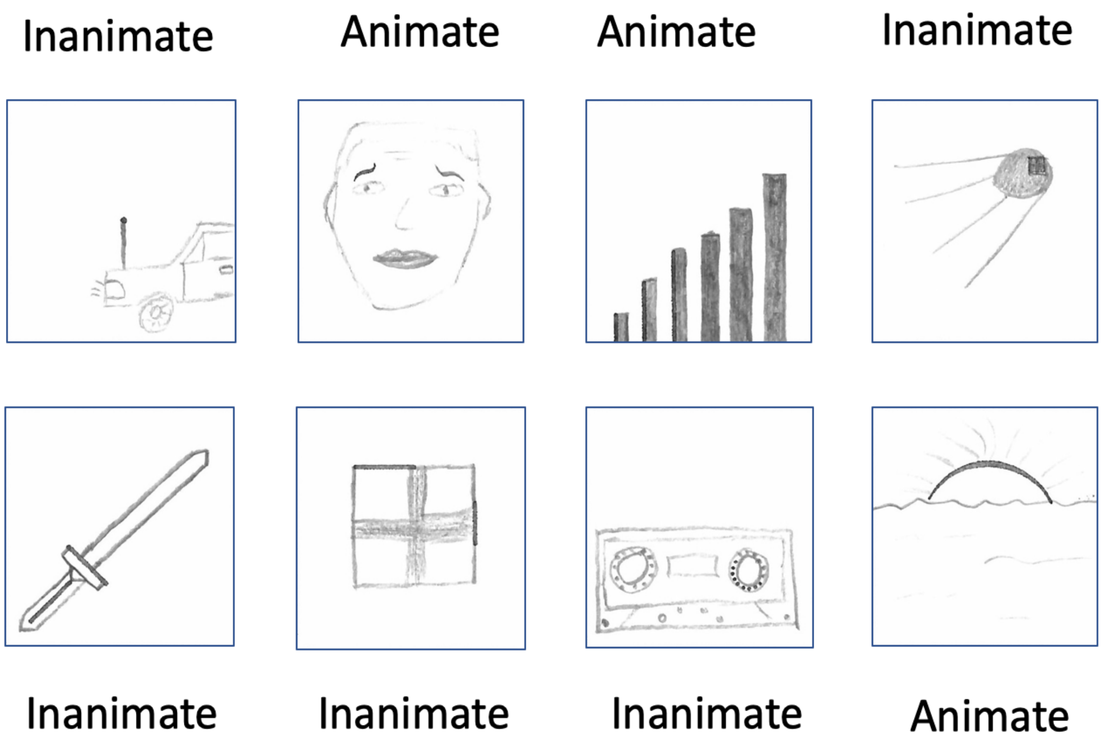

Figure 6: Animate and Inanimate classification results

preference, likes and dislikes related to their drawings and test stimuli. The appearance chances of boxes indicate tester's weakness. The earlier the box number appears in the sequence, the more the tester likes this image and the stimuli in the box are more likely to trigger the tester's drawing inspiration. For the same reason, the later the box number appears the more it proves that the tester does not like the stimuli and preferred to avoid them.

The order of sequence has three characters we need to analyze: the mix level of the sequence, the curve tendency and the first two and last two signs (F2L2) of the order sequence. Each box has its hypothesis regarding masculine boxes (referred to as technical-constructive boxes by Kinget) and feminine boxes (referred to as organic boxes by Kinget). Particular attention was paid to the order in which individuals completed these boxes, whether they demonstrated a preference for one group over the other. The order of sequence features is valued with the three characters. Firstly, the mix level of the sequence, which referred to the level of chaos, observes the inconsistency between two adjacent box choices. Mix level pays more attention to the changes (not the same) with adjacent characters, so it is enough to count the number of conversions between consecutive characters. By computing the Mix Level of all possible combinations, we find out the Change time range from 1 to 7. Now the question becomes how to define the mix level: High, average and low? We change the sequential order into CL sequence according to Kinget's theory. More specifically, replace box 1, box 2, box 7, box 8 with curve signs (C), replace box 3, box, 4, box 5, box 6 with line signs (L). For example, if you have an input sequence 12543768, when replaced them follow the rule we defined can get the new sequence CCLLLCLC. With this CL sequence, we can calculate the mix level, which is 4, since the second and third characters change for the first time, the fifth and sixth characters change the second time, the sixth and seventh characters change the third time and the last two characters change the fourth time. Under this rule we can figure out that the sequence with lowest mix level is 1, which changes only one time (CCCCLLLL), and the sequence with lowest mix level is 7, which changes every couple characters (CLCLCLCL or LCLCLCLC). All combination of sequence's mix level has been inducted and shown in Fig. 12. Notice, each underline represents a C or L sign within the condition that the amount of C or L is four.

Table 2: All combination of sequence's mix level

C.C.C.C.L.L.L.L	L.L.L.L.C.C.C.C.	Lowest, Change time:1
C.C.C.L._._._._.	L.L.L.C._._._._.	Low, Change time:2-3
C.C.L._._._._._.	L.L.C._._._._._.	High, Change time:3-5
C.L.C.L.C.L.C.L.	L.C.L.C.L.C.L.C.	Highest, Change time: 7

The order of sequence has three characters we need to analyze: the mix level of the sequence, the curve tendency and the first two and last two signs (F2L2) of the order sequence. The second character evaluation for the order of sequence is the Curve tendency degree (CTD). For Extraversion, the person is more likely to draw "curve" boxes at first and "line" stimulus last. When judging the curve tendency, we need to convert the number sequence into a sequence composed of C and L, and then calculate the curve tendency degree of the sequence. We define the Curve tendency degree to be equal to the sum of L's index number minus the sum of C's index number, $CTD = sum(L) - sum(C)$, where the object of summation is the order in which C or L appears in the sequence. For example, the CLLCLLCC sequence, $sum(C) = 1 + 4 + 7 + 8 = 20, sum(L) = 2 + 3 + 5 + 6 = 16$, the number is the index position of each C and L, thus the degree of curve tendency equals four $CTD = sum(L) - sum(C) = -4$. The curve tendency degree ranges from -16 to 16, where the curve tendency degree from -16 to -4 is low, the curve tendency degree from -4 to 4 is average, and the curve tendency degree from 4 to 16 is high.

Table 3: All possible combinations and CTD.

LOW
('llllcccc', -16), ('lllclccc', -14), ('lllcclcc', -12), ('llcllccc', -12), ('llclclcc', -10), ('lllccclc', -10), ('lclllccc', -10), ('lcllclcc', -8), ('cllllccc', -8), ('llccllcc', -8), ('llclcclc', -8), ('lllccccl', -8), ('lclcllcc', -6), ('llcclclc', -6), ('clllclcc', -6), ('llclcccl', -6), ('lcllcclc', -6), ('llcccllc', -4), ('clllcclc', -4), ('lcllcccl', -4), ('llcclccl', -4), ('lclclclc', -4), ('cllcllcc', -4), ('lcclllcc', -4),
AVE
('lclclccl', -2), ('lccllclc', -2), ('clllcccl', -2), ('llccclcl', -2), ('lclccllc', -2), ('cllclclc', -2), ('clclllcc', -2), ('llcccccll', 0), ('lcclcllc', 0), ('lclcclcl', 0), ('lccllccl', 0), ('cllclccl', 0), ('clcllclc', 0), ('cclllllcc', 0), ('cllccllc', 0), ('lcccllllc', 2), ('clcllccl', 2), ('lcclclcl', 2), ('cllcclcl', 2), ('cclllclc', 2), ('clclcllc', 2), ('lclcccll', 2),
HIGH
('ccllcllc', 4), ('lcccllcl', 4), ('clcclllc', 4), ('clclclcl', 4), ('lcclccll', 4), ('cllcccll', 4), ('cclllccl', 4), ('cclcllllc', 6), ('clccllcl', 6), ('lccclcll', 6), ('ccllclcl', 6), ('clclccll', 6), ('lcccclll', 8), ('cclcllcl', 8), ('ccllccll', 8), ('cccllllc', 8), ('clcclcll', 8), ('clccclll', 10), ('cclclcll', 10), ('cccllcl', 10), ('cclcclll', 12), ('cccllcll', 12), ('ccclclll', 14), ('ccccllll', 16)]
Difference range (-16 - 16)

As shown in Table 3, all of the possible combination groups and their corresponding CTD values are exhibited and they can also be summarized as a more concise rule of CTD as Table 4 shown. For a better understanding we compute an example here, if we have the sequence 8,7,6,3,2,1,5,4, replace box number as mentioned in last paragraph, then 1,2 equals CC and 7,8 equals LL, so the LC sequence is CCLLC-CLL. Computing the CTD value: $CTD = sum(C) = 1 + 2 + 3 + 4 = 10, sum(L) = 5 + 6 + 7 + 8 = 26, difference = sum(L) - sum(C) = 16$. So the CTD value is 16 and thus the Extraversion result is "High"

Table 4: Concise rule of the CTD

C.C.C.L._._._._	High
C.C.L._._._._._	
C.L._._._._._._	
L.C._._._._._._	Average.
L.L.C._._._._._	
L.L.L.C._._._._	
L.L.L.L.C.C.C.C.	Lowest.

The third character is the first two and last two signs (F2L2) of the order sequence. Emotional stability is the third feature which considers the consistency of first two and last two orders in the sequence. With the recommendation from the expert, we list all possible combinations of F2L2 and divide them into High, Average and Low levels as shown in Table 5.

The Numerical Order of Sequence result.

The order implicates tester's choice and preference. We evaluate three traits of the order sequence: Mix level, the curve tendency and first two last two position. For each traits we defined a mathematical function to change the psychological feeling into a numerical digit. Meanwhile these four orders are exempted because they are naturally existed and contain less information of these three traits. As shown in Fig. 7, the numerical order of sequence (NOS) of "12345678", "87654321", "12348765" and "87651234" are exempted. Any other order sequence excepted these four orders need to compute the traits we discussed above and get the feature property of the numerical order of sequence.

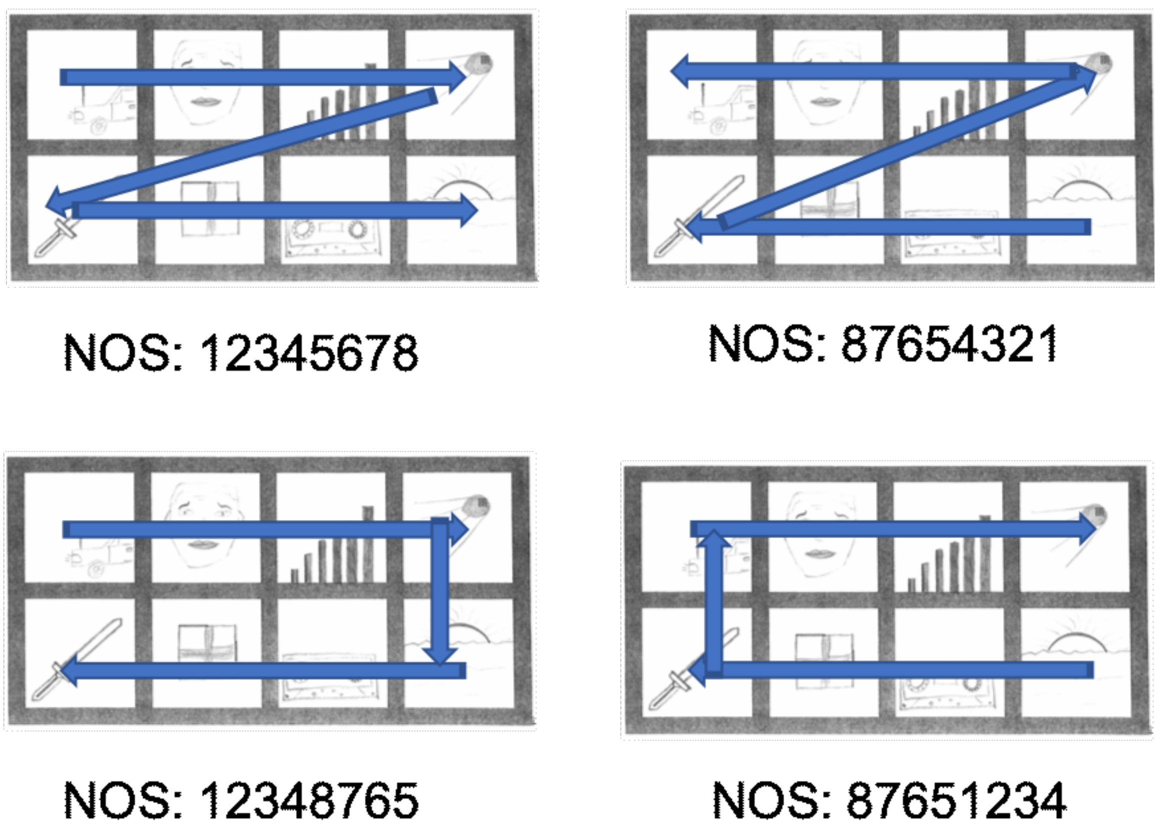

Figure 7: Four exemptions of order sequence results

Table 5: First two last two combinations and evaluation principle for Emotional Stability

Emotional stability	Order combination
HIGH	CC _ _ LL LL _ _ CC
AVE	CL _ _ CC LC _ _ CC LL _ _ CL LL _ _ LC CC _ _ CL CC _ _ LC CL _ _ LL LC _ _ LL
LOW	CC __ CC CL _ _ CL CL _ _ LC LC _ _ CL LC _ _ LC LL _ _ LL

3.2.4 Line and Curve ratio

Feature Definition and description.

The line-curve ratio feature is a kind of typical personality clue of drawings. There is at least one stimuli sign in each box, which was expected to be captured and extended by testers. A good response refers to keeping the same line or curve dominance with origin after extension. For example, the box one starts with a round dot at the center and its curve dominance, this means the whole completed drawing should keep in curve dominance as well to be a good response. The dominance is decided by the larger number of the straight and curve lines. To get the dominance of each box, we need to count the lines and curves. However, unlike the printed trace the handwriting sketches may not be so straight. Furthermore, when two lines connect, it is easier for the program to regard them as one object, which should actually be treated as two. The system was programmed to find out the cross corner of two or more lines and split all lines apart. Then trying to connect the crossed lines back with extra heuristic information or principal component analysis. This two-stage cut-and-connect operation splits the connected lines separately with understandable meanings [27]. The system will assign a number 1 or 0 to each drawing, 1 represents line dominate which means there are more lines than curves in this drawing. 0 represents curve dominate which means there are more curves than lines in this drawing.

Input Scanned WT single drawing image.

Step 1. Convert input image into gray image.

Step 2. Remove noise, dilate and erode.

Step 3. Stroke thinning, transforming all the strokes inside the sketch into one pixel wide lines.

Step 4. Separate all single lines from sketches.

while *Extract each separate stroke* **do**

  if *Only one adjacent pixel* **then**

    Single direction straight line: extend stroke with this pixel;

  else

    Branch: Multi-pixels exist in 8-pixel neighborhood.

  if *Stroke length < 12* **then**

    Compute the inner and outer angles of two crossed lines, the pixel with a smaller angular deviation should be selected.

   else

    Branch: stroke length greater than 12.

   if *The length of extended trace < Threshold (8 pixel)* **then**

     Discard as trail extension.

    else

     Compute the PCA value of each potential extend trace and product it with history principal component. The trace with maximum product result will be selected as extended direction.

    end

   end

  end

end

Step 5. Merging separate strokes corresponding to the possibility confidence:

$$\mathrm{E}\left(s_1, s_2\right) = abs(\cos(\Omega)) \times \frac{(\tau - \Delta)}{\tau}, \quad \text{where } \Delta < \tau$$

Step 6. Computing the curvature of each stroke, assign to line and curve group.
Step 7. Compare the number of line and curve, get the dominance result.

Result: Line curves dominance

Principal component analysis

The Principal component analysis we used was formed as $\operatorname{cov}(X,Y) = \frac{\sum_{i=1}^{m}\left(X_i-\bar{X}\right)\left(Y_i-\bar{Y}\right)}{m-1}$ and X,Y are the coordinates. The length we evaluated about the stroke is m equals 12. For the original stroke, given the principal component $\left(\vec{P_o}\right)$ and the future potentially extended stroke $\left(d\vec{P_e}\right)$ where d corresponds to the stroke future direction considered. The dot product of these two strokes are calculated as:

$$DP_d = \left(\vec{P_o} \bullet_d \vec{P_e}\right)$$

When calculated the multi future strokes which could be extended and then select the max one to extend:

$$\max_{1 \le d \le n} \left(\|DP_d\|\right)$$

Curvature Formula

When detected every single stroke, we evaluate them to get the property of line or curve.

We used curvature to determine if the stroke a line or curve. The curvature expression [28] is:

$$\kappa = \frac{\ddot{\vec{r}} \times \dot{\vec{r}}}{\left|\dot{\vec{r}}\right|^3} \tag{1}$$

Or:

$$\kappa = \frac{x''y' - x'y''}{\left((x')^2 + (y')^2\right)^{3/2}} \tag{2}$$

We use three points $(x_1,y_1),(x_2,y_2),(x_3,y_3)$ to calculate the curvature. With upper and lower limits of t_a and t_b, we can get the parametric equation:

$$\begin{cases} x_1 &= a_1 & -a_2t_a + a_3t_a^2 \\ x_2 &= a_1 & \\ x_3 &= a_1 & +a_2t_b + a_3t_b^2 \end{cases} \tag{3}$$

We could have:

$$\kappa = \frac{x''y' - x'y''}{((x')^2 + (y')^2)^{3/2}} = \frac{2(a_3b_2 - a_2b_3)}{(a_2^2 + b_2^2)^{3/2}} \tag{4}$$

Final result.

With the experiment of Line and curve ratio computing algorithm we described in Chapter 3, we get the result of each step shown in Fig. 8.

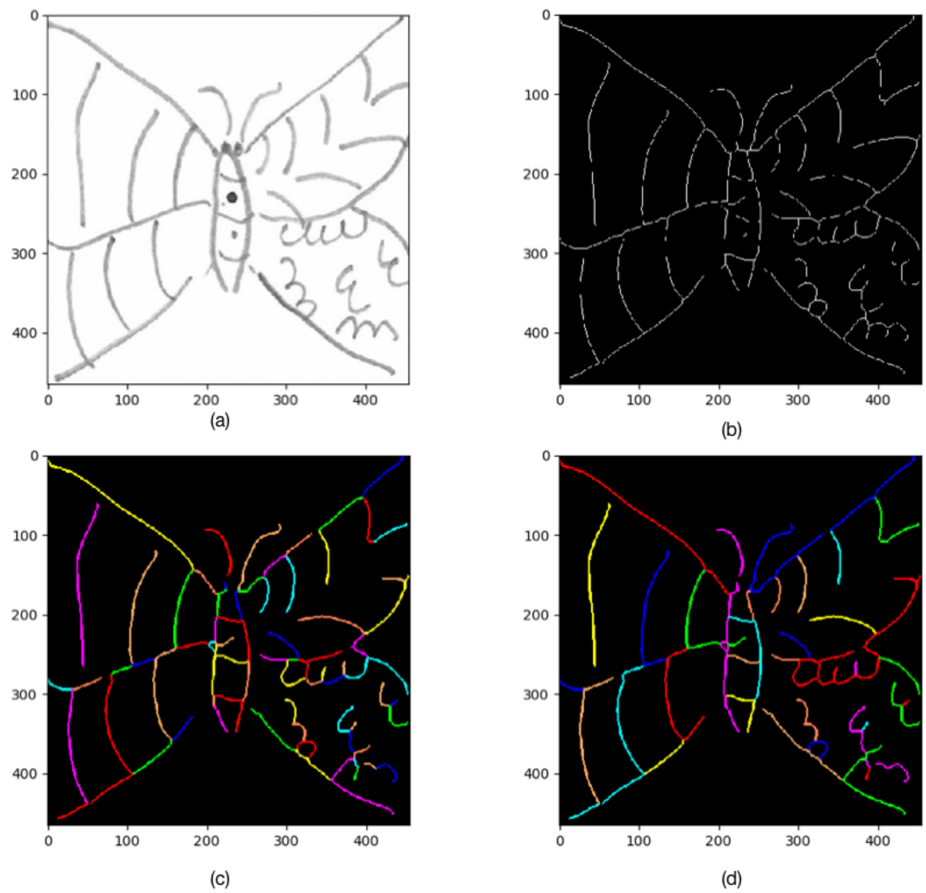

Figure 8: Lines Process results

- Picture (a): Original image to be evaluated. When segmented from WT form, the single images are changed into grey images and fed into lines and curves algorithm.
- Picture (b): Result after noise removal and thinning. An adaptive threshold method would be implemented on each image to remove the noise. Besides,

two reverse image processing operations Dilate and Erode would be applied in order to remove small artifacts and holes inside strokes. These operations are also designed for a better thinning result. Thinning is used to reduce each connected component in a binary image to a single-pixel wide skeleton [29].

- Picture (c): Separate every ambiguous line according to cross points. Scan and generate single strokes based on thinning skeleton.

- Picture (d): Group up lines. Merging lines together heuristically, which should be treated as one single line.

- Final result: Curves and lines dominate computing. This sample is curve dominate.

3.2.5 Categories

Current research and development status of Categories classification.

The categories is one of the most important features psychologist used in WT personality analysis. Category is a key element while evaluating tester's personality cause any further analysis related to content would be based on object's category. For a computer software, it is not easy to recognize the categories of a hand drawn sketches because of the highly abstract and great variety of sketch characters. Free-hand sketch classification itself is a single research topic. For sketch classification, a lot of research has been tried to improve the effect.

The most famous available datasets are Quick Draw [30] and TU-Berlin [31]. Original sketch benchmark proposed by [31] in 2012. [32] modified work uses SIFT, Gaussian Mixture Model (GMM) based fisher vector encoding for sketch recognition and fed into a SVM classifier. This approach enhances the recognition performance to near-human (73.1%) [31] accuracy. The major contribution has been presented in [30], a deep CNNs model namely Sketch-a-Net was introduced for sketch recognition and beats the human sketch recognition accuracy. For the first time, an effort has been made to specially design a deep convolution neural network (DCNN) architecture named sketch-DNN by [33]. Another research [34] extracts sketch features from two famous pre-trained CNNs, namely, AlexNet [35] and modified version of LeNet [36] and yield a little improvement in the recognition results.

More recently, a lot of deep learning methods for the free-hand sketch has been researched [37]. Especially for sketch recognition using transfer learning. Shaukat Hayat [38] used Deep CNN and transfer learning to classify Hand-Drawn Sketch. The author proposed to use the global average pooling of parameters from three pre-trained models as feature maps to produce the classification result, they got the result of 94.45%. In 2019, Mustafa Sert combined Principal Component Analysis (PCA) and selected the best performing layer features from the CNN-SVM recognition pipeline and got an accuracy of 72.5% on TU-Berlin dataset.

Transfer learning shows its abilities to be used on sketch classification job. It is a good method that can apply a neural network to learn from one dataset to another dataset after appropriate adjustment. Generally, the last layer of the network is closely related to the selected dataset and target task [39]. So, using the pre-trained model and sharing the available parameters, which also means pre-training and fine-tuning is a very effective method. Our method is also focused on transfer learning on sketch data classification.

Our method: Two-stages transferred inception-V3 for WT sketch classification.

In this thesis, we propose to use a two stages transfer learning for category classification of Wartegg Test sketches. The whole work flow is demonstrated in Fig. 9. As we can see from the flow chart, there are two-way process steps of the whole pipeline. The first step of the WT data process (left) is to isolate the WT form from the scanned questionnaire and then segment it into eight single boxes.

As we can see from the flow Fig. 9, there are two-way process steps of the whole pipeline. For the first step of the right, we need to obtain the pre-trained model, which is well designed and optimized on the ImageNet dataset. Then we fine-tune the pre-trained model by retraining and optimizing on the Sketch dataset and get the new model. The first step of the WT data process (left) is to isolate the WT form from the scanned questionnaire and then segment it into eight boxes, in order to extract eight individual boxes from the questionnaire (completed sketches). After that, we use this model to predict each single WT image and get the classification result. When we cut out the eight single drawings from the scanned questionnaire form, we can extract features from them and perform numerical analysis. For the

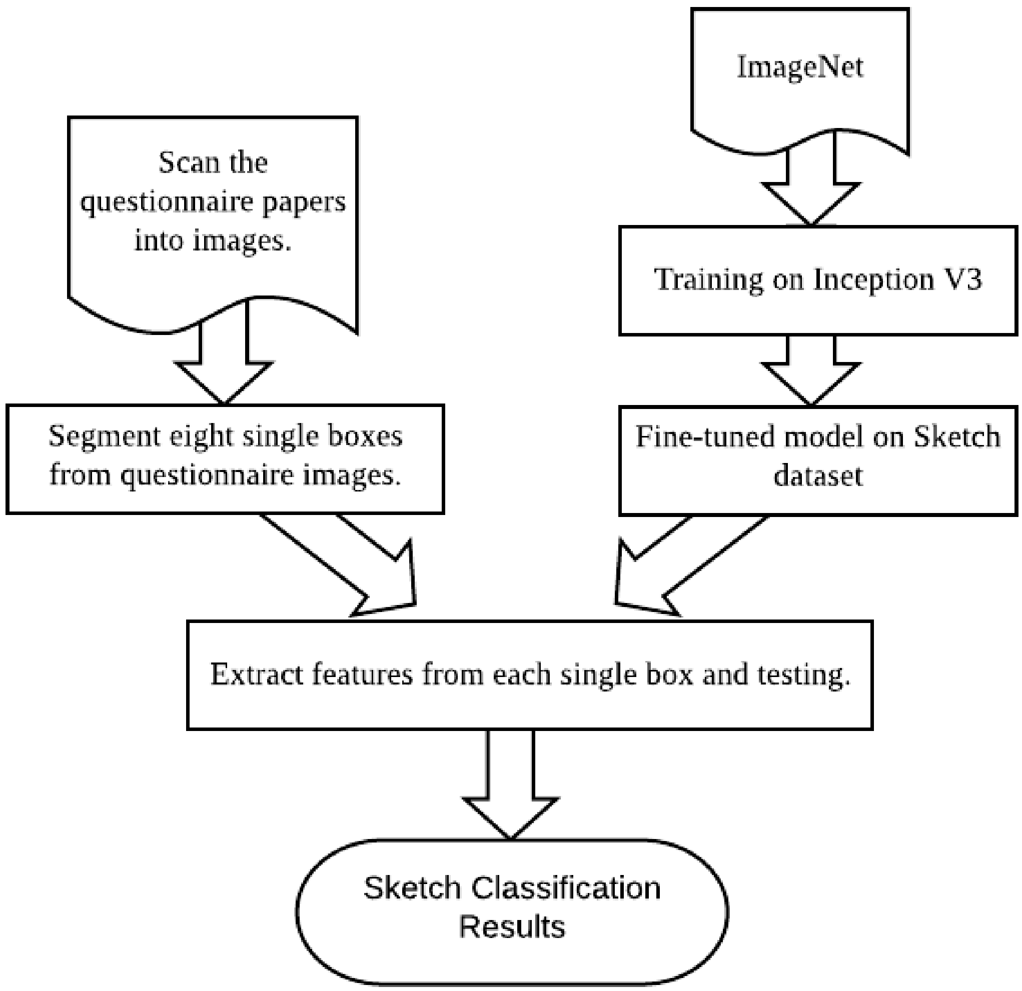

Figure 9: Diagram of the whole WT procedure

first step of the right, we need to obtain the pre-trained model, which is well designed and optimized on the ImageNet dataset. Then we fine-tuned the pre-trained model by retraining and optimizing on the Quick draw and TU-Berlin (Q-T) Sketch dataset in order to get the new model.

The pre-trained model we used is Inception-V3. We selected three models, which performed well on ImageNet datasets, VggNet-16[40], Resnet[41] and Inception-V3. Compared with these models, Inception contains fewer parameters and more effective computing efficiency. We also trained and tested these three models on our Q-T sketch dataset, the Inception-V3 earns highest score (91.89%) of sketch classification, while Resnet gets 77.1% and VggNet gets 78.93% as shown in Table. 7. Therefore, it was chosen as our pre-trained model for transfer learning. The Inception-v3 consists of seven basic elements: Convolution layer, Average Pooling layer (AvgPool), Max Pooling (Maxpool), Concat layer, Dropout, Fully connected layer and Softmax. Transfer learning is suitable for fixing the data shortage problem when training a neural network. Wartegg test happens to facing the hardship of collecting data under its strict experimental environment. Therefore, we use model migration to solve the problem of data shortage of the Wartegg test. The model we trained on ImagenNet performs fine-tuning on the collected and adjusted Q-T sketch dataset to obtain an effective model for sketch classification. After obtaining the model, we then migrate the model to WT data. Because there is too little WT data, which is less than one thousand valid samples, so it is unrealistic to perform model training or fine-tuning on it. Therefore, we use the fine-tuened model to be migrated method to achieve the model training, adaptation and get the final test result. All these steps consist our two-stages transfer learning based categories classification for Wartegg Test hand-drawn sketches.

The pre-trained model we used is Inception V3 as shown in Fig. 10. Since 2012, many models showing very high accuracy on ImageNet have appeared, like VggNet-16 [40]and Resnet [41]. Compared with these models, Inception contains fewer parameters and more effective computing efficiency. Inception-V3 performs best among these models when testing on sketches without any adjustment. The inception v3 consists of seven basic elements: Convolution layer, Average Pooling layer (AvgPool), Max Pooling (Maxpool), Concat layer, Dropout, Fully connected layer and Softmax.

There are three main advantages of Inception V3, The factorization, Auxiliary

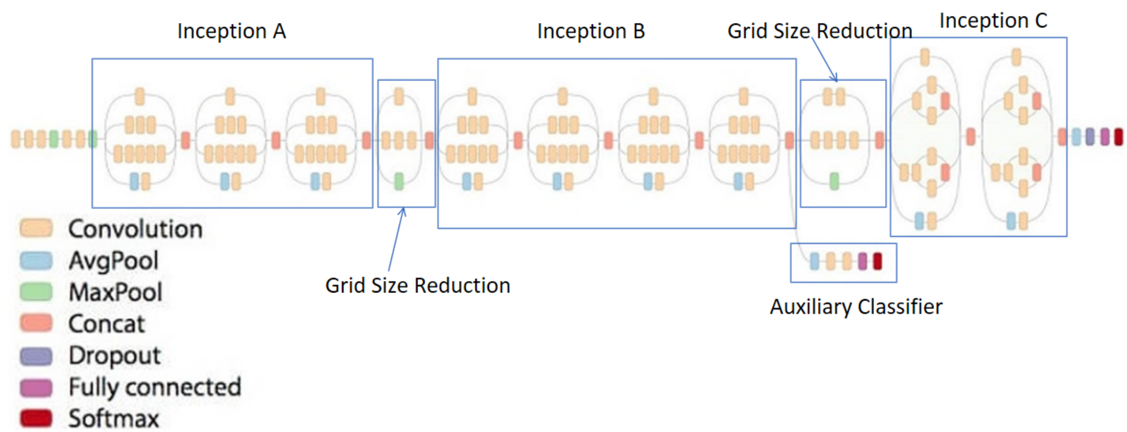

Figure 10: Model of inception v3

classifier and Grid size Reduction. In the first part of inception module (Inception A in Fig. 10), the inception V3 uses two 3 *3 kernels instead one 5*5 kernel which is used in GoogleLeNet, to do this the parameters can be reduced from 25 to 18. In the second inception module (Inception B in Fig. 10), the inception V3 divides a 7*7 kernel into one 1*7 and one 7*1 kernels, so for each divide, parameters will decrease from 49 to 14. In the last inception module (Inception C in Fig. 10), the inception v3 uses 1*3 and 3*1 kernel to replace a 3*3 kernel to reduce the number of parameters. By using factorization, the Inception V3 will use less parameters which can increase the calculation speed and change one convolution layer into two convolution layers that can increase the depth of the network.

Comparing with 2 Auxiliary classifier in Inception v1, the Inception V3 only has one Auxiliary Classifier which is located on the last layer of the second inception module aimed for regularization. The Inception V3 has two grid size reduction modules, by splicing the convolution layer which stride equals 2 and max pooling layer to achieve the purpose of overcoming the a representational bottleneck and reducing the amount of calculation.

Experiment

When process the feature categories, we need to firstly pre-processing the data for model training and testing. The training dataset we collected consisted of two parts: Quick draw [42] and TU-Burlin [43] dataset. This two datsets are the most suitable choice for us to extend our Wartegg test data and for model training, while they are

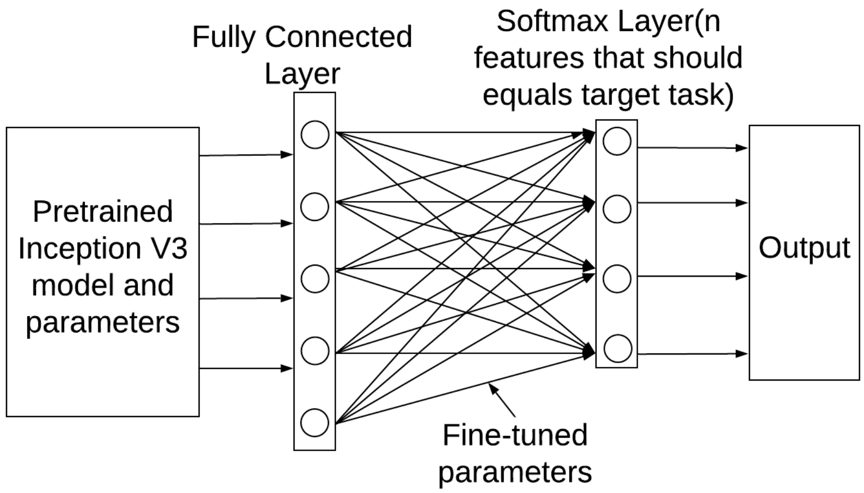

Figure 11: Pre-train and fine-tuning process of stage two

the most similar type of drawing styles with WT's.

Quick Draw is an online game developed by Google that challenges players to draw a picture of an object or idea and then uses a neural network artificial intelligence to guess what the drawings represent. This game collected a dataset with 50 million pictures of 345 categories drawn by players all over the world, the samples are shown in Fig. 12. The AI learns from each drawing, increasing its ability to guess correctly in the future [44], [45], [46]. The game allows player to draw a specific character within 20 seconds [45]. The concepts that it guesses can be simple, like 'car', or more complicated, like 'animals' [44]. Based on what they draw, the AI guesses what they are drawing. This game collected a dataset with 50 million pictures of 345 categories drawn by players all over the world.

The TU-Burlin dataset was first introduced by [43] in 2012. This dataset contains 250 categories and in total 20000 sketches, as shown in Fig. 13. The authors asked humans to sketch objects of a given category and gathered 20,000 unique sketches evenly distributed over 250 object categories. The categories exhaustively cover most objects that we commonly encounter in everyday life. The whole drawing procedure for players is 30 minutes. The organizers also provided with undo, redo, clear, and delete buttons for our stroke-based sketching canvas so that participants can easily familiarize themselves with the tool while drawing their first sketch. After finishing a sketch participants can move on and draw another sketch given a new category.

Figure 12: Quick Draw samples

They remove sketches that are clearly in the wrong category (e.g. an airplane in the teapot category), contain offensive content or otherwise do not follow our requirements (typically excessive context). However, the poorly pitched drawings remained. All drawings which contain context around objects, not easy to recognize, with text label and with large black area are not allowed and removed.

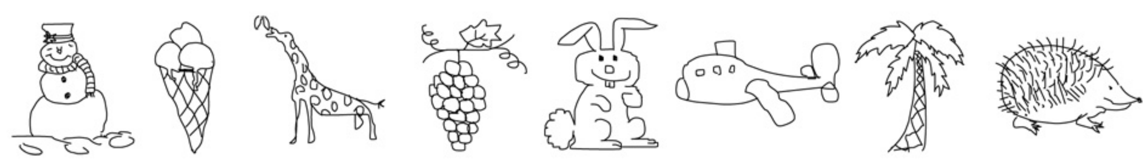

Figure 13: TU-Berlin samples

We combined Quick Draw and TU-Burlin datasets together to meet the training requirement of WT drawings classification. Since the WT drawings were collected without any limitation, the sketch results could be unexpected and they appear in a great number of categories. In this case, trying to cover as many as categories we could find is a solution to improve classification result and WT result. However, single dataset like Quick Draw or TU-Berlin is hard to meet the requirement of WT. Quick Draw has a large volume of drawing pictures while the content is quite simple and limited to 20 seconds to finish. These rules caused Quick Draw's drawings could only cover the simplest part of WT content. Same to TU-Berlin dataset, its finish process was extended to 30 minutes, which is much longer than Quick Draw's and similar to WT's. Besides time, the content of TU-Berlin finished in 30 minutes are also much similar to WT content. Though with such pros, TU-Berlin dataset provided players with an "eraser", which is in conflict with a key limitation of WT. All in all, each single dataset could hardly meet the requirement of our WT experiment, so combining different datasets and letting each of them fix their part of problems could be a reasonable solution.

When pre-processing these two datasets, we remove the duplicate categories and balance the image number of each category first. Second, for each category browsing each image and removing those just scrawl without meanings. Third, add the most frequently categories in WT which are not included in Quick Draw and TU-Berlin datasets. Fourth, change categories according to WT name rules.

We combined them together and selected a total of 281 categories for the training model. These categories are selected according to the appearance chance of Wartegg test response [1]. The most frequently responses are recorded in this book and provided us a strong reference for organizing our categories. Before training the model, we modify the training dataset to make it suitable for usage. First, we resize each single drawing image and centralize the sketch object inside the box frame and remove the noise from drawings. Finally, according to the category in our Wartegg dataset, the similar categories in the training dataset are merged. Then we created a new Sketch dataset with 281 categories, which is extended from TU-Berlin dataset and Quick Draw dataset. We collect these two existed sketch datasets, merged and modified them for our experiment and we named them as Q-T sketch dataset.

Evaluation Metrics.

Although inception V3 has many advantages, the significant problem encountered in our experiments is that the accuracy of Inception V3 is based on a large amount of labelled data, which is what we lack. To solve this problem, in our experiments, we adjusted the parameters of each layer of the Pre-trained inception V3 model which has already trained on Image-Net (an image dataset which has more than 14 million labelled images) by using the Q-T sketch dataset. And then we modified the last layer of this fine-tuned model from the Fully connected Layer to SoftMax Layer, we

Table 6: Evaluation Index in multi-categories classification experiment

Evaluation index	Function Definition
Recall	$\frac{TP}{TP+FN} \times 100\%$
Precision	$\frac{TP}{TP+FP} \times 100\%$
F1-score	$\frac{2 Precision Recall}{Precision+Recall} \times 100\%$

modify the features in SoftMax Layer which can fit with our output number and change the parameters between these two layers. To do this, we only need to retrain those parameters in this layer to make this model suitable for our target task, as shown in Fig.11. By doing these steps, we applied this new model to our Wartegg Test dataset.

WT test data adjustment

When dealing with the WT testing dataset, three steps have been applied. First is labeling the dataset. We counted the frequently occurring categories based on the Book of Crisi Wartegg System [1] and labeled our testing dataset. Second, due to the reason that Wartegg dataset was painted by real people who participated in the test, so some of the images are too abstract, meaningless, scrabble, and even hardly to recognize for human being, so we remove these kind of data. Finally, we put aside those images that have multiple content because they contain extremely different distribution with training dataset. By doing these three steps, we clustered the 281 sketch categories into 176 categories. So, for WT dataset classification the number of categories is 176 and even less.

Result and Discussion

In this experiment, we used Micro F1 to split the n-class evaluation into n two-class evaluations, add the TP, FP, RN corresponding to the n two-class evaluations, calculate the evaluation precision rate and recall rate, and the two accuracy rates and recall rates The F1 score calculated by the rate is Micro F1 [47]. Precision equals TP divided by the sum of TP and FP, which means the correct proportion among all positive in prediction. Also, Recall equals TP divided by the sum of TP and FN, which means the correct proportion in all real positive. TP means True Positive if

Table 7: Classification results of Q-T Sketch dataset

Methods	Datasets	Results
Human [48][49]	TUBerlin	73.1%
Resnet [41]	TUBerlin	77.1%
VGGnet [38]	TUBerlin	78.93%
SketchNet [50]	TUBerlin	80.42%
Inception-V3 [51]	TUBerlin	91.89%
Transfer Learning Inception-V3	TUBerlin+Quick Draw	**94% (Top5: 98%)**

label is Positive and predicted is positive. FP means False Positive, It means label is negative but the predicted result is positive. FN means False Negative, it means the label is positive but prediction is negative. The equations are shown in Table. 6

Our two stages transfer learning yields two classification results, on our Q-T sketch dataset and WT dataset respectively. For the middle result of Q-T sketches, we get 94% and 98% for top-1 and top-5 classification results. Comparing with the previous methods, our method boosts the effect by improving 2.02% accuracy and 6.11% for top-5 result. Details are shown in Table. 7.

For the WT data classification. The neural network classifier used in this experiment is also Inception-V3. We transferred and fine-tuned the training model of Inception-V3 on the Q-T sketch data to WT data. We tried several numbers of categories of Wt dataset. From the most frequent ten objects to 176 over 80% included the common response from the user. The classification accuracy of the model in 10 types of data is 85%, the classification accuracy in 17 types of data is 73%, and the classification accuracy in 176 types (All included) of data is 63% for top-5 results. The best result which balances the categories number and performance well is 176 with 63% accuracy. This result is obtained by our two stage transfer learning method and the previous step got the 98% for the top 5 categories.

3.3 Big Five Personality evaluation method

3.3.1 BFP introduction

The big five personality (BFP) traits were developed in the 1980s, it is also one of the best accepted and most commonly used data to describe the language of personality in academic psychology [52]. The big five comes from the statistical study of responses to personality items. Using a technique called the five-factor model [53], which can look at the responses of people to hundreds of personality items and ask the question "what is the best way to summarize an individual?". This has been done with many samples from all over the world and the general result is that, while there seem to be unlimited personality variables, five stand out from the pack in terms of explaining a lot of a person's answers to questions about their personality: Extraversion, Emotional Stability, Agreeableness, Conscientiousness and Openness to Experience. Extraversion is characterized by excitability, sociability, talkativeness and high amounts of emotional expressiveness. Emotional stability is characterized by happiness, cheerfulness and shows resilience. Agreeableness personality dimension includes attributes such as trust, altruism, kindness, affection and other pro-social behaviours. Conscientiousness, standard features of this dimension include high levels of thoughtfulness, good impulse control and goal-directed behaviours. Openness to experience, this trait features characteristics such as imagination and insight. People who are high in this trait also tend to have a broad range of interests. They are curious about the world and other people and eager to learn new things and enjoy new experiences. The big-five are not associated with any particular test, a variety of measures have been developed to measure them. This test uses the Big-Five Factor Markers from the International Personality Item Pool, developed by Goldberg in 1992 [54]. The test consists of fifty items that you must rate on how true they are about you on a five-point scale where 1 equals Disagree, 3 means Neutral and 5 equals Agree. It takes most people 3 to 8 minutes to complete. For each feature in the Wartegg test, we project their evaluation result to BFP traits, and we change the value range into 1 to 5 degrees also.

3.3.2 BFP criteria for Wartegg Test

In this section, we show the criteria for how the Wartegg test results project to big five personality traits as shown in Table 8-11. Table 12 shows the BFP result of the example we used in this thesis.

Table 8: Extraversion and agreeableness

1 - Extraversion, 3 - Agreeableness	High.	Average.	Low
Space utilization distribution.	> 3	2-3	< 2
Lines vs Curves.	< 2	2-3	> 3
Alive versus Inanimate.	> 3	2-3	< 2
Categories (+ versus -)	> 3	2-3	< 2
Numerical order of sequence: the person is preferred to draw curve boxes first then line's.			
For agreeableness, the person is more likely to prefer “curve” boxes first.			

Table 9: Emotional stability

2 - Emotional stability	High.	Average.	Low
Space utilization distribution.	$2-3$	> 3	< 2
Lines vs Curves.	$2-3$	> 3	< 2
Alive versus Inanimate.	$2-3$	> 3	< 2
Categories (+ versus -)	> 3	$2-3$	< 2
Numerical order of sequence: the person is more likely to mix “curve” and “line” boxes.			

Table 10: Conscientiousness

4- Conscientiousness	High.	Average.	Low
Space utilization distribution.	$2-3$	<2	>3
Lines vs Curves.	>3	$2-3$	<2
Alive versus Inanimate.	<2	$2-3$	>3
Categories (+ versus -)	>3	$2-3$	<2
Numerical order of sequence: the person is preferred to draw line boxes first then curve's.			

Table 11: Openness to experience

5 - Openness to experience	High.	Average.	Low
Space utilization distribution.	>3	$2-3$	<2
Lines vs Curves.	$2-3$	>3	<2
Alive versus Inanimate.	$2-3$	>3	<2
Categories (+ versus -)	>3	$2-3$	<2
Numerical order of sequence: the person is more likely to mix "curve" and "line" boxes.			

Table 12: Big five personality result

	Space Utilization	Lnines/Curvs.	Alive/Inanimate.	Categories	order
Score	0.36	0.625	0.375	0.9375	35648217
	1.8/5	3.125/5	1.875/5	4.6875	LLLLCCCC
EX	LOW	LOW	LOW	HIGH	Low
ES	LOW	LOW	LOW	HIGH	Low
AG	LOW	LOW	LOW	HIGH	Low
CO	AVE	HIGH	HIGH	HIGH	High
OP	LOW	AVE	LOW	HIGH	Low

Chapter 4

Dataset and GUI.

4.1 WT dataset

4.1.1 WT dataset collection and environment setting

The Wartegg Test data was designed to reflect the natural responses of testers to the existed stimuli of the boxes at their first glance. To achieve this function, the completion process of WT data should be strictly restricted to the rules. We collected 211 valid samples in person with the questionnaire survey and scanned them at the Center for Pattern Recognition and Machine Intelligence(CENPARMI) of Concordia University. Each sample consists of two questionnaire forms, the WT form and the describe table, one of the completed WT forms is shown in Fig. 14. In our work, we mostly focus on the WT form, the information of the table was organized as one of the features extraction system.

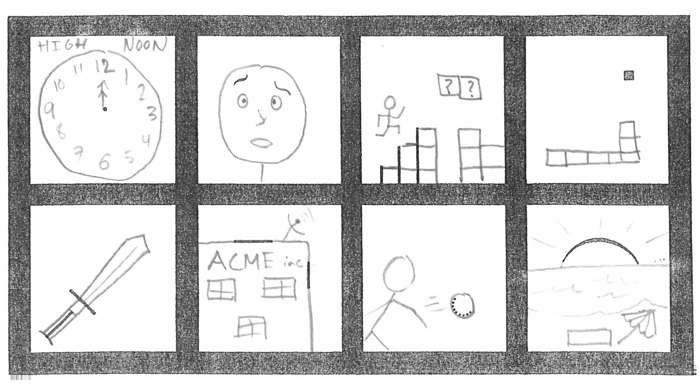

Figure 14: Completed WT sample

During the data collection stage, we need to set up a quiet environment to avoid testers being disturbed. The testers were encouraged to prepare enough time to participate in the questionnaire to ensure that they can calm down in emotion and avoid the rush to answer any question during their finishing period. In general, the completion time is about 15 minutes, while the time needed varies with different individuals. In order to provide enough time for testers, we set up three optional periods in a day for the attendant's flexible choosing, three-time periods are arranged separately in the morning, afternoon and evening, and the maximum time period is three hours. When the tester walks in our test environment, we provide them with a clean WT form and several #2 pencils, as well as multiple seats for the tester to choose arbitrarily. Besides, there is no eraser provided, because we expect to obtain the first response result of the tester, other than the result with careful modification and improvement.

The testers are encouraged to complete the form without any other limitation. They are told by "Making a drawing in each box with a completed meaning, preferably the first which comes in your mind and abstract drawings and the drawing number. It is not necessary for you to follow the numerical order; work at your own pace and there are no time restrictions." [24] when the tester needs any help from the administer, they would like to explain and help until the tester fully understands the questions without influencing their response. During the administration procedure, the client sits far away from the examiner to avoid causing pressure and near enough to provide help when needed. Wartegg test assumed that tester's drawings are able to project their inside world and personality characters. Each semi-structured sign is well designed to stimulate the tester's feelings. All of the completed drawing characters would be analyzed by a psychologist and then processed by our computer system.

4.1.2 WT dataset description

As illustrated in Fig. 14, each sample includes six factors: age, gender, country, degree, completed WT form and description table. The ages of the tester are divided into three ranges 18 to 35, 36 to 55 and older. The genders are collected in three categories, which are male, female and do not want to respond. The statistical results are demonstrated in Table. 13. The testers were distributed among 30 different majorities, the occupation type with the most participation is student. Since our survey

Table 13: Wartegg Test Dataset Information

Age		***Degree***		***Country***	
18∼35	181	Bachelor	44	Canadian	64
		College	2	Iranian	42
36∼55	29	Dipoma	4	Indian	35
		High school	41	Chinese	20
Elder	4	Master	80	Saudi	8
		P.h.D	19	Bangladeshi	5

is held in schools and most of the people who can be notified are students, the main occupations of the participants are students. Fortunately, the distribution of age and education in the student group is relatively good. Therefore, it is still in the students. It implicitly includes master of engineer and Ph.D. of researchers, as well as teaching assistants, so it can alleviate the single career in the student group. In addition to students, we also have a system of bookkeepers, dispatchers, fashion designers, housewives, nurses, self-employed persons, television and radio professionals and veterinary. The countries are distributed among 30 different names, the four countries with the most occurrences are Canada, Iran, India and China. Since this survey was conducted in Canada, this distribution basically corresponds to the Canadian population and nationality distribution.

4.2 Result from five features.

4.2.1 GUI Demo exhibition.

Based on the algorithms described in this thesis, we have also created software, which has implemented form segmentation and feature extraction for WT after scanning, and also includes image loading, preview display and save functions. The interface of the software is shown in Fig. 15. It contains the main interface, eight buttons (function buttons) and a message bar (result feedback). This software can cut the WT form into eight single images and extract five features fully automatically without human intervention. The pipeline of this software is extensible and we add further functions to it in the future.

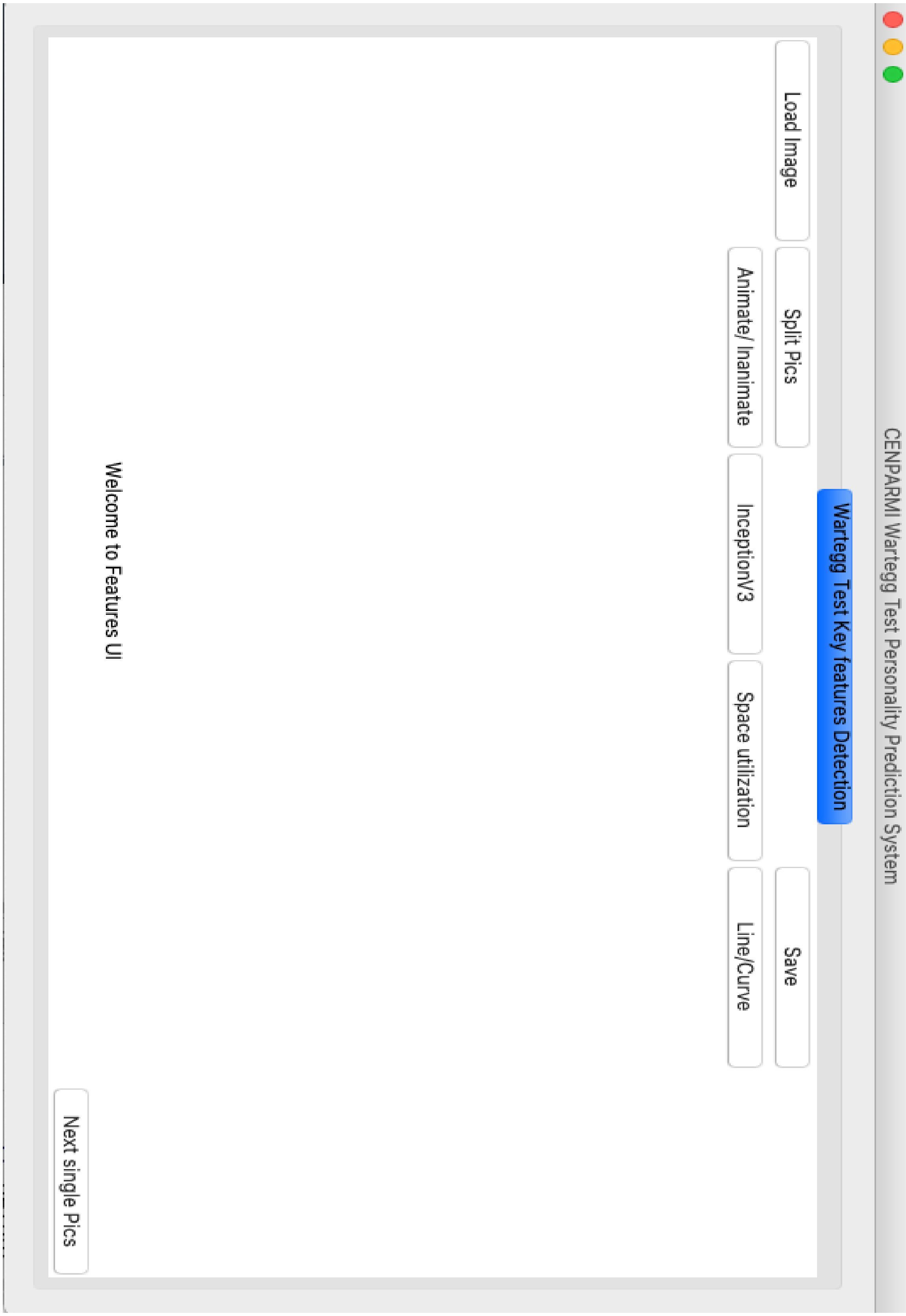

Figure 15: GUI Demo

Chapter 5

Conclusion and Future Work

5.1 WT Conclusion and discussion

In this thesis, we collected a new dataset of completed Wartegg Test forms and labelled their content with a professional psychologist's help. We evaluated the performance of existing architectures with our new WT dataset and Q-T sketch dataset. We proposed a WT system with the extraction of five features and analysis functions and project the scores of each feature to the big five personality traits. For each feature, we find out or create an effective algorithm to analyze the tester's personality automatically. By finishing these algorithms, we make it possible for testers to access the Wartegg test at any time. Hopefully, this work could help with patients and contribute to WT research and handwriting areas.

5.2 Future Work

5.2.1 Feature extension to our system.

This research has been designed based on five features, it is possible for a further extension with additional new features. When discussing with experts, we summarized several features that need to be studied and implemented in the future. Due to time constraints, constraints, only five of them have been selected in this research project. Based on this system, we can still add features: 1. Carefulness (based on detected lines): how carefully the tester drawing strokes 2. Shading: the percentage of shading within a drawing. 3. Positive or negative characters ratio. 4. Nice or ugly drawings ratio. 5. Surface and distribution (upgrades space utilized) 6. Meaningful or meaningless responses. 7. Direction of each drawing. 8. Outgoing or seclusive (reflect of curves) 9. Character meaning. I believe that after adding richer features, our system will become more solid and accurately reflect the personality characteristics of the tester, and the test results will be more accurate and convincing.

5.2.2 Multi labels sketch classification.

Speaking of content improvement, then for the future work, one important thing to be concerned about is classification or detection when multiple objectives exist in one drawing. In the current experiment, we can accurately process and classify single object drawings and clear images. This is a harder scenario for multi-class sketch classification. But in actual applications, many users will draw multi-class objects, even it is a messy picture, so in the next work, we should focus on designing a classifier algorithm or a detector algorithm, which can identify or detect multiple objects from the drawing.

5.2.3 Color the sketches with generative models.

The hardest part at the end is also the most interesting part. When we analyze the personality characteristics of the tester, we can give these pencils some colors to express their personality. The image sketch coloring is a very interesting and creative work. According to research, a survey shows that some generative models such as Pixel to Pixel Generative Adversarial Networks [55], SketchyGAN [56] and Image Colorization Using Generative Adversarial Networks, abbreviated as ICGAN [57] and etc. are all excellent works. This task is very challenging and groundbreaking because the training process of most current image coloring tasks is to provide additional information, such as Landscape coloring provides color image and sketch image pair for model training, cartoon coloring provides language description and sketch image for semantic analysis and then coloring. For WT data, we can choose to color the drawings according to the description or according to the analyzed psychological report. We are looking forward to the emergence of feasible coloring algorithms for the Wartegg test in the future.

Bibliography

[1] Alessandro Crisi and Jacob A Palm. *The Crisi Wartegg System (CWS): Manual for administration, scoring, and interpretation.* Routledge, Rome, Italy, 1st edition, 2018.

[2] Graziella Pettinati. The Wartegg test in a nut shell. `http://www.aspghandwriting.org/main/ASPGResources_Articles_Wartegg_Test.html`. [Online], Accessed: March 2015.

[3] Ehrig Wartegg. *Gestaltung und Charakter Ausdrucksdeutung zeichnerischer Gestaltung und Entwurf einer charakterologischen Typologie.* J.A. Barth, German, 1939. [Form and character. Interpretation of expression from drawings and an outline of a characterological typology.].

[4] Richard Wilhelm. *I Ging: das Buch der Wandlungen. 1.* Diederichs, 1924.

[5] E Wartegg. Zeichen der zeit. leben und forschung eines psychodiagnostikers. *Bernhardt, H. u. Lockot, R.(Hg.)*, pages 95–111, 2000.

[6] M Gardziella. Wartegg-piirustustesti. *Tr. It. The Wartegg Drawing Test. Psykologien Kustannus Oy, Jyväskylä*, 1985.

[7] Fabiana Marques Pereira, Ricardo Primi, and Claudia Cobêro. Validade de testes utilizados em seleção de pessoal segundo recrutadores. *Revista Psicologia-Teoria e Prática*, 5(2), 2003.

[8] Robert J Gregory. *Psychological testing: History, principles, and applications.* Pearson Higher Ed, New York City, 7th edition, 2004.

[9] W Deinlein and M Boss. Befragung zum stand der testanwendung in der allgemeinen berufs-studien-und laufbahnbereitung in der deutschen schweiz [the use

of tests in vocational counseling in german-speaking switzerland]. *Postgraduate thesis, NABB-6, Switzerlandhttp://www. panorama. ch/files (31.05. 2007)*, 2003.

[10] Krister Mellberg. The Wartegg drawing completion test as a predictor ofadjustment and success in industrial school. *Scandinavian Journal of Psychology*, 13(1):34–38, 1972.

[11] Oscar David Sánchez Forero. Calidad psicométrica del test Wartegg versión 16 campos, nuevas escalas y publicación de resultados para la aplicación e interpretación del test en la población colombiana. B.S. thesis, Catholic University of Colombia, Bogota, Colombia, 2013. [Psychometric quality of the Wartegg test version 16 fields, new scales and publication of results for the application and interpretation of the test in the Colombian population].

[12] CM La Spina, C Murasecchi, S Stinco, A Cassano, A Strippoli, N Trigila, M Di Salvatore, A Quaranta, E Straffi, and C Barone. The psychological projective test Wartegg-zeichentest: A contribution to the knowledge of patients with cancer. *Journal of Clinical Oncology*, 28(15_suppl):e19678–e19678, 2010.

[13] Amelia Rizzo, Lucia Della Villa, and Alessandro Crisi. Can the problematic internet use evolve in a pre-psychotic state? a single case study with the Wartegg. *Computers in Human Behavior*, 51:532–538, 2015.

[14] Eka Roivanen. A brief history of the Wartegg drawing test. *Gestalt Theory*, 31(1):55–71, 2009.

[15] G Marian Kinget. The drawing completion test. *NY: Grune & Stratton*, 1952. New York, USA.

[16] Alessandro Crisi. *Manuale del test di Wartegg: Norme per la raccolta, la siglatura e l'interpretazione*. Edizioni scientifiche Magi, Rome, Italy, 1998.

[17] JFS Lizarazo, JLS Rincon, and BMR Higuera. Wartegg automatic test qualifier (calificador de la prueba Wartegg 16 campos), 2008. Bogota, Colombia, Accession Number: 13-37-395. Install and use : https://youtu.be/LJArNfeuUsc.

[18] Rosihan Ari Yuana, Dewanto Harjunowibowo, and Nugroho Karyanta. Fwat (fast Wartegg analyzer tool) for personality identification. *Advanced Science Letters*, 21(10):3114–3117, 2015.

[19] Rosihan Ari Yuana, Dewanto Harjunowibowo, and Nugroho Karyanta. Fwat (fast Wartegg analyzer tool) for personality identification. *Advanced Science Letters*, 21(10):3114–3117, 2015.

[20] F Bianchi, A Crisi, and M Di Renzo. Il test di Wartegg nell'età evolutiva. un contributo psicodiagnostico allo studio clinico della balbuzie, della sordità e dei disturbi dell'apprendimento, 1996. [The Wartegg test in developmental age. A psychodiagnostic contribution to the clinical study of stuttering, deafness and learning disorders.] Rome, Italy.

[21] A Crisi. Manuale del test di Wartegg ii edizione [Wartegg test manual], 2007. Rome, Italy: Edizioni Scientifiche Ma. Gi.

[22] Alessandro Crisi and Francesco Dentale. The Wartegg drawing completion test: inter-rater agreement and criterion validity of three new scoring categories. *International Journal of Psychology and Psychological Therapy*, 16(1):83–90, 2016.

[23] Italian institute of waretgg. Italian institute of Wartegg, clinical assessment courses, counseling, training, Wartegg test teaching method. `"http://www.wartegg.com"`, 1999. [Online].

[24] Lili Liu, Graziella Pettinati, and Ching Y Suen. Computer-aided Wartegg drawing completion test. In *International Conference on Pattern Recognition and Artificial Intelligence*, pages 575–580. Springer, 2020.

[25] Nam Tuan Ly, Lili Liu, Ching Y Suen, and Masaki Nakagawa. Hand-drawn object detection for scoring Wartegg zeichen test. In *International Conference on Pattern Recognition and Artificial Intelligence*, pages 109–114. Springer, 2020.

[26] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[27] Pankaj Rajan and Tracy Hammond. From paper to machine: Extracting strokes from images for use in sketch recognition. In *SBM*, pages 41–48, Goslar, Germany, June 2008.

[28] Peijin Zhang. Numerical solution for the line curvature. "`https://github.com/Pjer-zhang/PJCurvature/blob/master/doc/doc_PJCurv.pdf`", July, 5, 2019. [Online].

[29] TY Zhang and Ching Y. Suen. A fast parallel algorithm for thinning digital patterns. *Communications of the ACM*, 27(3):236–239, 1984.

[30] David Ha and Douglas Eck. A neural representation of sketch drawings. *arXiv preprint arXiv:1704.03477*, 2017.

[31] Mathias Eitz, James Hays, and Marc Alexa. How do humans sketch objects? *ACM Transactions on graphics (TOG)*, 31(4):1–10, 2012.

[32] Ivan E Sutherland. Sketchpad a man-machine graphical communication system. *Simulation*, 2(5):R–3, 1964.

[33] Lumin Yang, Jiajie Zhuang, Hongbo Fu, Kun Zhou, and Youyi Zheng. Sketchgcn: Semantic sketch segmentation with graph convolutional networks. *arXiv preprint arXiv:2003.00678*, 2020.

[34] Peng Xu, Yongye Huang, Tongtong Yuan, Kaiyue Pang, Yi-Zhe Song, Tao Xiang, Timothy M Hospedales, Zhanyu Ma, and Jun Guo. Sketchmate: Deep hashing for million-scale human sketch retrieval. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8090–8098, 2018.

[35] Conghui Hu, Da Li, Yi-Zhe Song, Tao Xiang, and Timothy M Hospedales. Sketch-a-classifier: sketch-based photo classifier generation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 9136–9144, 2018.

[36] Jifei Song, Qian Yu, Yi-Zhe Song, Tao Xiang, and Timothy M Hospedales. Deep spatial-semantic attention for fine-grained sketch-based image retrieval. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 5551–5560, 2017.

[37] Peng Xu. Deep learning for free-hand sketch: A survey. *arXiv preprint arXiv:2001.02600*, 2020.

[38] Mustafa Sert and Emel Boyacı. Sketch recognition using transfer learning. *Multimedia Tools and Applications*, 78(12):17095–17112, 2019.

[39] Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. How transferable are features in deep neural networks? In *Advances in neural information processing systems*, pages 3320–3328, 2014.

[40] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[41] Jun-Yan He, Xiao Wu, Yu-Gang Jiang, Bo Zhao, and Qiang Peng. Sketch recognition with deep visual-sequential fusion model. In *Proceedings of the 25th ACM international conference on Multimedia*, pages 448–456, 2017.

[42] Jonas Jongejan, Henry Rowley, Takashi Kawashima, Jongmin Kim, and Nick Fox-Gieg. The quick, draw!-ai experiment. *Mountain View, CA, accessed Feb*, 17(2018):4, 2016.

[43] Mathias Eitz, James Hays, and Marc Alexa. How do humans sketch objects? *ACM Transactions on graphics (TOG)*, 31(4):1–10, 2012.

[44] Nicole Hennig. *Mobile learning trends: Accessibility, ecosystems, content creation*, volume 52. ALA TechSource, 2016.

[45] Ashwin Pajankar and Arush Kakkar. *Raspberry Pi By Example*. Packt Publishing Ltd, 2016. Birmingham, UK.

[46] Jillian Capewell. Let a computer guess what you're drawing in this high-tech pictionary game. `https://www.huffingtonpost.ca/entry/quick-draw-google-doodling_n_58331234e4b099512f83caca?ri18n=true`, 2016. [Online].

[47] Chuan Liu, Wenyong Wang, Meng Wang, Fengmao Lv, and Martin Konan. An efficient instance selection algorithm to reconstruct training set for support vector machine. *Knowledge-Based Systems*, 116:58–73, 2017.

[48] Emel Boyaci and Mustafa Sert. Feature-level fusion of deep convolutional neural networks for sketch recognition on smartphones. In *2017 IEEE International*

Conference on Consumer Electronics (ICCE), pages 466–467. IEEE, Las Vegas, 2017.

[49] Yonggang Qi, Yi-Zhe Song, Honggang Zhang, and Jun Liu. Sketch-based image retrieval via siamese convolutional neural network. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 2460–2464, 2016.

[50] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

[51] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. 2015. *arXiv preprint arXiv:1512.00567*, 2015.

[52] Murray R Barrick and Michael K Mount. The big five personality dimensions and job performance: a meta-analysis. *Personnel Psychology*, 44(1):1–26, 1991.

[53] Wikipedia contributors. Big five personality traits — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/w/index.php?title=Big_Five_personality_traits&oldid=984344054`, 2020. [Online], Accessed 27-October-2020.

[54] Lewis R Goldberg. The development of markers for the big-five factor structure. *Psychological assessment*, 4(1):26–42, 1992.

[55] P. Isola, J. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5967–5976, 2017.

[56] Wengling Chen and James Hays. Sketchygan: Towards diverse and realistic sketch to image synthesis. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pages 9416–9425, Salt Lake City, USA, June 2018. Springer.

[57] Kamyar Nazeri, Eric Ng, and Mehran Ebrahimi. Image colorization using generative adversarial networks. In *Proc International conference on articulated motion and deformable objects*, pages 85–94, Goslar, Germany, 2018. Springer.