CHAPTER 7

FACIAL GENDER RECOGNITION FOR VIRTUAL HELPDESK APPLICATIONS

Md Sah Hj. Salam and Ong Le Foo

7.1 INTRODUCTION

Computer-human face gender recognition is a technology that enables a computer to identify the gender of a human face. Typically, the system is equipped with a camera, which captures an image or video as the first step. The captured data is then pre-processed, and the face and gender are recognised using deep learning techniques.

Gender recognition refers to identifying an individual as male or female, although in rare cases, a person's gender may be both or neither. Human interaction is part of daily life, and recognising gender is usually straightforward for humans, who often adjust their behaviour according to the gender of the person they are interacting with. By incorporating gender recognition into computer systems, it is possible to provide tailored services or preferences, such as in marketing.

There are many potential applications of computer-based gender recognition, including security, advertising, and artificial intelligent (AI) virtual assistant services (Gharge et al., 2019). For example, people often approach help desks seeking assistance, but many help desks are now being replaced by intelligent virtual assistants. However, some intelligent

virtual assistants respond unnaturally, creating a gap in user expectations. While current smart assistants can listen and respond verbally, they cannot perceive visual feedback from users, leading to incomplete or impaired interaction and, consequently, reduced efficiency.

7.2 UNDERSTANDING THE BASICS

This section typically involves face detection, pre-processing feature extraction, and classification. While several classifiers exist, convolutional neural network (CNN) approaches provide the highest accuracy. An important application of gender recognition is in intelligent virtual assistants, sometimes called chatbots.

7.2.1 Gender Recognition

Gender recognition, as distinguished in Liew et al. (2016) which also referred to as a gender classifier, is a computer vision technique that detects human faces or bodies and classifies them as 'male' or 'female.' Numerous methods have been developed and studied for this purpose. The common approach to computer vision gender recognition consists of four phases: face detection, pre-processing, feature extraction, and classification.

Software libraries such as OpenCV can be used to perform face detection and image processing. A classifier is essential for gender recognition, and various methods are employed, including support vector machines (SVM), neural networks, and CNN. Among these, CNN-based methods have demonstrated the greatest potential for achieving high performance in the gender recognition process.

7.2.2 Intelligent Virtual Assistant

Intelligent Virtual Assistant, or sometimes also called 'Chatbot', is a software agent that can be used to listen to instructions, perform tasks, and provide services to every individual. Nowadays, communicating

with customers using a live chat interface is becoming increasingly popular and offering good real-time customer service. Many popular personal intelligence virtual assistants have been around for people to use, such as the smartphone intelligence virtual assistants, including Siri, Google Assistant, and others. AI virtual assistant has been reducing the cost of customer service by 30% and helping people to solve their difficulties automatically. The technologies in the AI virtual assistant include natural language processing, machine learning, and speech recognition to read the input of the user and perform tasks or give a response. As agreed by Maedche et al. (2019), an AI virtual assistant could be either text-based, speech-based, or both simultaneously.

7.2.3 Convolutional Neural Network

The CNN is a neural network that has one or more convolutional layers based on Albawi et al. (2017). It is mostly used in image processing, and it is also a suitable method for gender classification. A CNN consists of layers, including convolutional layers, a non-linearity layer, a pooling layer, and fully connected layers. CNN are unlike other methods because they can perform feature extraction and classification at the same time during the learning process. This feature has brought a significant strength over the other method, as it could reduce the need for pre-processing steps and allow CNN to be more easily trained compared to others because they have fewer parameters.

7.3 HOW IT WORKS IN PRACTICE

The development process has straightforward goals. Therefore, the waterfall method was selected and applied. The workflow is illustrated in Figure 7.1.