

Preface

Computers and digital cameras have become an essential part of our daily life. There have been many advancements in these two technologies. In terms of performance, both have become better than their previous generations and smaller. The size of these devices is comparable to that of a credit card. With all these advancements, there have been changes in the way interact with computers. Along with traditional input devices like a keyboard, mouse, touchpad, trackball, joystick, etc., speech and gestures are being used. The hand plays an important role in human-computer interaction (HCI). The hand gesture conveys certain information which a computer can understand and then produce an appropriate response. The hand can act as a touchless controller where the actions are intuitive and natural to the users. This can also have a wide range of applications for example in virtual and augmented reality, robotics, industrial automation, and many others. The major component of a computer-vision application is to understand the information conveyed through different hand poses to determine the position and orientation of the hand and its joints in the 3D space. The unique position of hand joints (also called hand keypoints) like fingertips, wrists, and knuckles poses a piece of unique information. However, to understand these pieces of information several challenges to overcome like occlusion, illumination variation, the scale of the hand, and skin color. This dissertation investigates the hurdles associated with hand keypoints detection from a monocular RGB image and presents algorithms based on deep learning techniques.

Chapter 1 presents the introduction of the dissertation. The types and processes of hand keypoints detection are described. The challenges that are present in the complete process are explored. As deep neural network architectures are used as the primary processing tool - a brief description of its working principle is given. Further, the motivation to take on this work and research problems are presented in this chapter.

Chapter 2 covers the complete detail of the process used for partial hand keypoints detection, basically fingertip detection from a monocular RGB image. Three different algorithms were used for the aforementioned task. The chapter covers the progress made through one process over another with the objective of achieving better fingertip detection accuracy.

Chapter 3 presents the details of methods used for complete hand keypoints detection. The details of the methodology used to make the process independent of hand localization are present. Along with that, a process to improve the hand keypoints detection is present when the scale of the hand is small.

Chapter 4 introduces a deep learning-based algorithm to perform hand keypoints detection directly from multiple hands. The chapter gives details of the process where the intermediate feature map is used. The process of training and inference is also described in this chapter.

Chapter 5 presents the summary and conclusions of this dissertation. Future research directions and challenges faced are also discussed in this chapter.

The salient outcomes of the research work have been reported in the international conferences and journals