HandTrack : A Library For Prototyping Real-time Hand Tracking Interfaces using Convolutional Neural Networks

Victor Dibia, *victor.dibia@gmail.com*

Abstract— The ability to track the human hand in real-time from monocular images has important implications for the design of unobtrusive, marker-free human-computer interaction paradigms. However, during real world interactions, the human hand is presented in multiple orientations, may be partially occluded from camera view, exbitit random motion and have large variations in skin tone making tracking difficult. Furthermore existing approaches to these challenges are limited (use of multiple sensors, dependence on manually engineered detection models) in terms of portability, ease of integration and scale. In this work, we introduce a HandTrack, an end-to-end trainable convolutional neural network model that learns to track the human hand using a single RGB image in realtime. HandTrack is fast, based on the SSD MobileNet architecture and can be easily integrated in designing user interactions (pointing, selections) across multiple platforms (web, mobile, desktop). It runs at over 21 fps on a consumer grade laptop (CPU) and outperforms existing approaches to object detection and tracking. We provide experimental results and code that demonstrates the qualitative and quantitative value of the model on a publicly available dataset. Finally, we provide an easy to use library (Python and Javascript) which application developers can import and leverage in rapidly prototyping gesture based interactions.

Index Terms—deep learning, hand tracking, deep, data visualization

1 INTRODUCTION

• *Author 1 is a Research Scientist focused on Machine Learning at Cloudera Fast Forward Labs. E-mail: victor.dibia@gmail.com*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx/

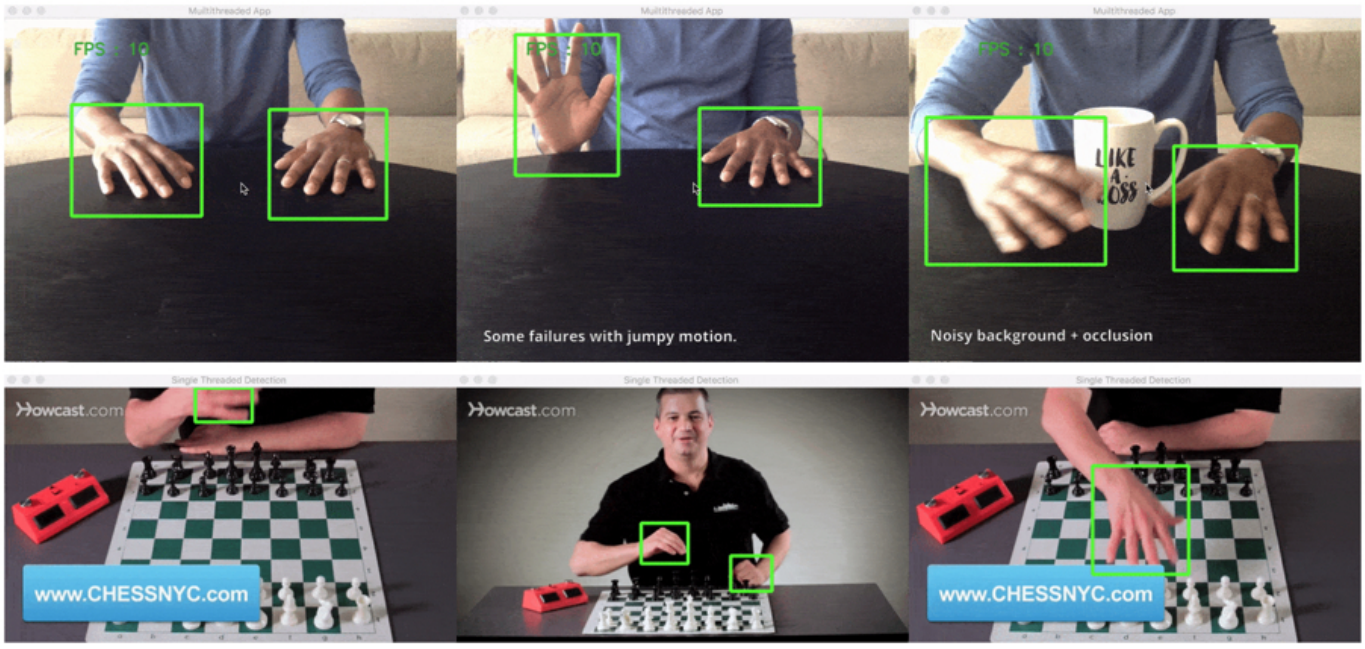

Fig. 1. Examples of bounding box tracking results from HandTrack with users with varied skin tones and orientations (a) HandTrack correctly tracks hands when applied to realtime video feed from a webcam. (b) HandTrack tracks hands when applied to a radomly selected chess tutorial video from Youtube.