

Real Time Detection of Hand Gestures for Interface Extension Using Convolutional Neural Networks

Satendra Varma
Indiana University
Bloomington, Indiana
satvarma@iu.edu

Shyam Narasimhan
shynaras@iu.edu

Scott Mathews
scomathe@iu.edu

Abstract

We present a method for real-time detection of hand gestures that employs a two phase image recognition pipeline, wherein hands are localized within the original input image, extracted, and fed into a CNN based classifier which outputs the detected gesture. Our results indicate that our method achieves a reasonable accuracy for real time detection on a predefined set of hand gestures.

1. Introduction

As the power of machine learning techniques continues to grow, the number of interfaces between machine and man continue to grow. Convolutional neural networks have demonstrated their efficacy in computer vision tasks. We employ a system based on convolutional neural networks for creating a vision based interface, which provides a new avenue for user interaction with their machine. Our goal with this project is to prototype a gesture based interface with computers, with the goal of being able to recognize a predefined set of hand gestures from a live video feed.

The space of voice assistants exploded recently with the advent of products from each major technology company aimed at providing a voice-based interface with computers. As desktop and phone processors continually grow more powerful, it is our belief that the next generation of interfaces with computers will utilize vision based interfaces.

Devices with built-in cameras would be able to utilize this interface by having an always-on camera, which performs actions on the system in response to certain visual inputs from the camera. One such visual input might be hand gestures. For example, a certain gesture might tell the camera to start taking commands, serving the same purpose as an activation phrase in current voice assistants. Following this “activation gesture” a series of additional gestures might be performed to indicate a command.

Another intriguing use case for the hand-gesture based

Method	Frobnability
Theirs	Frumpy
Yours	Frobbly
Ours	Makes one's heart Frob

Table 1. Results. Ours is better.

interface is with American Sign Language recognition. This interface could allow for text input using a purely visual interface, making use of the American Sign Language alphabet.

In order to facilitate these use cases, it is necessary for the algorithm powering the visual recognition to be lightweight enough to run in the background on low powered devices such as laptops and phones, while still providing enough performance to be useful.

Our goal in this project is to present a model for an efficient real time detection of hand gestures. Our results indicate that such interfaces are viable, given the proper performance minded setup.

2. Background and Related Work

There is

3. Methods

4. Results

5. Conclusion

6. References

List and number all bibliographical references in 9-point Times, single-spaced, at the end of your paper. When referenced in the text, enclose the citation number in square brackets, for example [1]. Where appropriate, include the name(s) of editors of referenced books.

6.1. Illustrations, graphs, and photographs

All graphics should be centered. Please ensure that any point you wish to make is resolvable in a printed copy of the paper. Resize fonts in figures to match the font in the body text, and choose line widths which render effectively in print. Many readers (and reviewers), even of an electronic copy, will choose to print your paper in order to read it. You cannot insist that they do otherwise, and therefore must not assume that they can zoom in to see tiny details on a graphic.

When placing figures in \LaTeX , it's almost always best to use `\includegraphics`, and to specify the figure width as a multiple of the line width as in the example below

```
\usepackage[dvips]{graphicx} ...  
\includegraphics[width=0.8\linewidth]  
    {myfile.eps}
```

6.2. Color

Please refer to the author guidelines on the CVPR 2018 web page for a discussion of the use of color in your document.

7. Final copy

You must include your signed IEEE copyright release form when you submit your finished paper. We **MUST** have this form before your paper can be published in the proceedings.

Please direct any questions to the production editor in charge of these proceedings at the IEEE Computer Society Press: Phone (714) 821-8380, or Fax (714) 761-1784.

References

- [1] Authors. The frobnicatable foo filter, 2014. Face and Gesture submission ID 324. Supplied as additional material `fg324.pdf`.