

Implementation of Deep High-Resolution Representation Learning for Visual Recognition

Shiyao Xie
2016xxxxxx

Zeyuan Yang
2017011577

Zifei Zhu
2017xxxxxx

Shengge Yang
2016xxxxxx

Abstract

High-Resolution representation is a popular topic in computer vision research field. After researching on related works in this field, we decided to implement a recent released paper at IEEE 2020, Deep High-Resolution Representation Learning for Visual Recognition (HRnet).[6] To completely comprehend the paper, we reproduced the code and ran it on an open source dataset. Moreover, the structure and outputs are compared with the source code.

1. Introduction

In many fields, high-resolution representation is important, such as human position detecting. How to extract key features from high-resolution images becomes a popular topic, also a challenging problem. Almost all previous state-of-art works share a similar method, which is applying a high-to-low resolution network and extracting high-resolution representation from low-resolution features. Instead, this work reshapes the entire network structure in two aspects. First, it parallels the convolution streams, Applying convolution on each resolution simultaneously. Also, this work adds multiresolution fusion modules. Every time initializing a new convolution stream, the information of each resolution is exchanged. By repeating this, this structure is more precise in spatial and get better results.

This paper is based on a similar paper released on CVPR 2019, Deep High-Resolution Representation Learning for Human Pose Estimation.[?] These two papers share a similar network structure. HRnet applied some finetuning on the previous network structure and promoted it to more tasks. Therefore, we mainly digged into this work and reproduce its code.

2. Related Works

3. Problem Definition

4. Approach

5. Technique Details

6. Experiments and Performance

7. Insights Analysis

8. Conclusion

References

- [1] FirstName Alpher. Frobnication. Journal of Foo, 12(1):234–778, 2002.
- [2] FirstName Alpher and FirstName Fotheringham-Smythe. Frobnication revisited. Journal of Foo, 13(1):234–778, 2003.
- [3] FirstName Alpher, FirstName Fotheringham-Smythe, and FirstName Gamow. Can a machine frobnicate? Journal of Foo, 14(1):234–778, 2004.
- [4] Authors. The frobnicatable foo filter, 2014. Face and Gesture submission ID 324. Supplied as additional material fg324.pdf.
- [5] Authors. Frobnication tutorial, 2014. Supplied as additional material tr.pdf.
- [6] Jingdong Wang, Ke Sun, Tianheng Cheng, Borui Jiang, Chaorui Deng, Yang Zhao, Dong Liu, Yadong Mu, Minghui Tan, Xinggang Wang, Wenyu Liu, and Bin Xiao. Deep high-resolution representation learning for visual recognition, 2019.