Fast-trained Image Inpainting for Irregular Holes

Naiqian Zheng (1800012904), Junwei Yang (1800013010), Dianze Li(1800012990)

Email：zhengnaiqian@pku.edu.cn, 1800013010@pku.edu.cn, aptx4869ldz@pku.edu.cn

Advisor：Jiaying Liu

PKU, Introduction to AI

2018-2019，Spring

Abstract

We use a Convolutional Neural Network of 14 Layers to encode and decode the picture and holes to get a reasonable prediction. We also use VGG19 pretrained model to extract features to improve the reasonability. Our loss function well balances effect and computing consumption. The network is fast-trained that it takes less time for convergence than most recent works and needs less computing resource.

Keywords： Image Inpainting, Partial Convolution

Introduction

Image Inpainting, the task of filling in holes in an image, is a hot spot with huge number of related works. We simplify a 16 layers network from Guilin et al.[1] and make a 14 layers network with a simplified loss function as well.

In the experiment, we use less than 1/10000 computing resource of work from Guilin et al.[1] and a data set of only 36000 images.

This is a open source project, you can download repository from https://github.com/NaturezzZ/Inpainting

Method

1 Dataset of Irregular Holes

To better predict irregular holes, we use a dataset of holes to match images randomly as input of network. So we get better predictions than work of Deepak et al.[2] in irregular holes inpainting.

2 Network Architecture

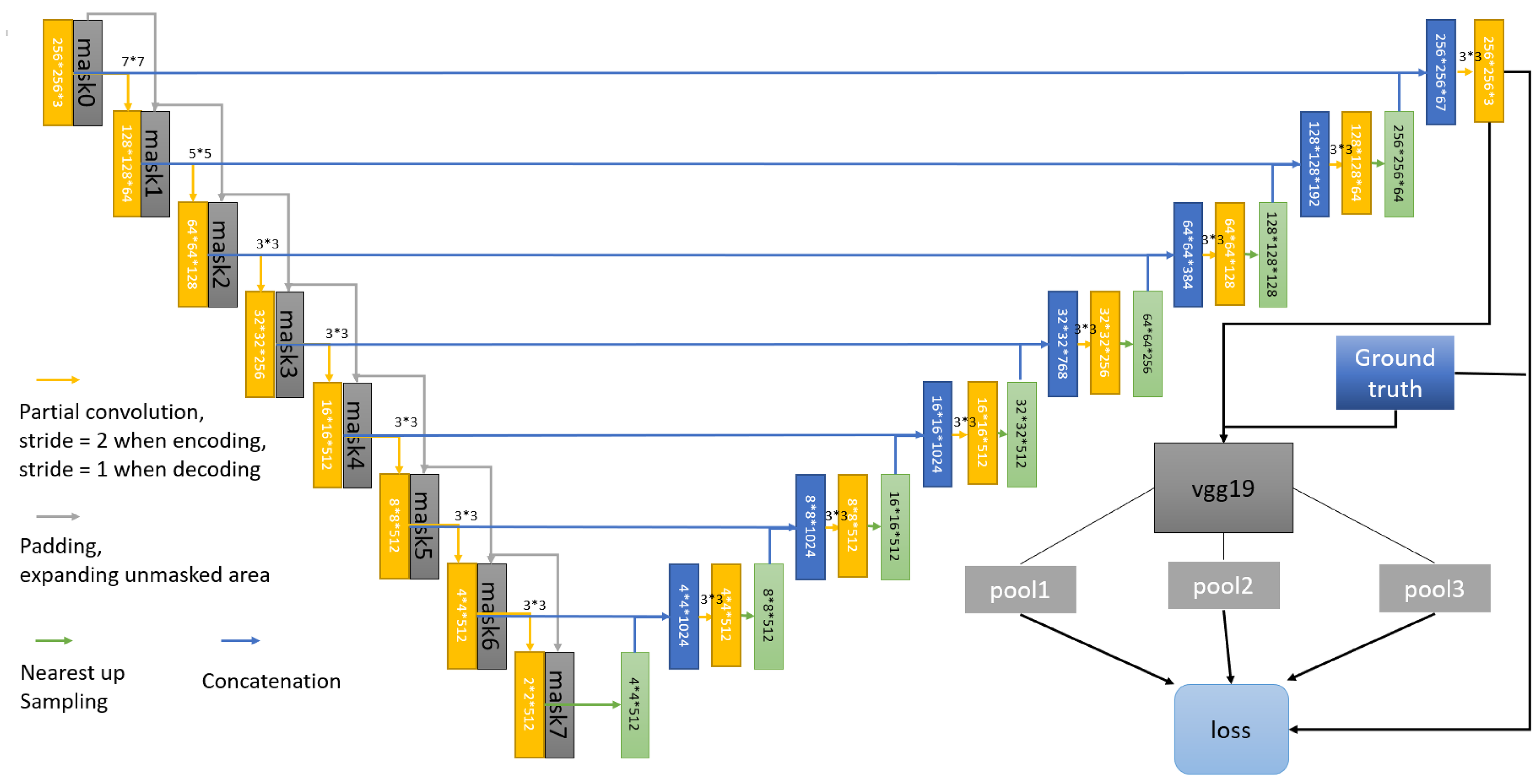
We use a 7 layers to encode the image into 2*2 and another 7 layers to decode it. After that, we use first 3 pooling layers of VGG19 to extract features.

3 Loss Function

We use per-pixel losses, perceptual losses, style losses and total variation

4 VGG19 pre-trained model

We add VGG19 model to get features from first 3 pooling layers and use them in loss function. As a result, our network has greater potential in understanding internal characters.



Experiments

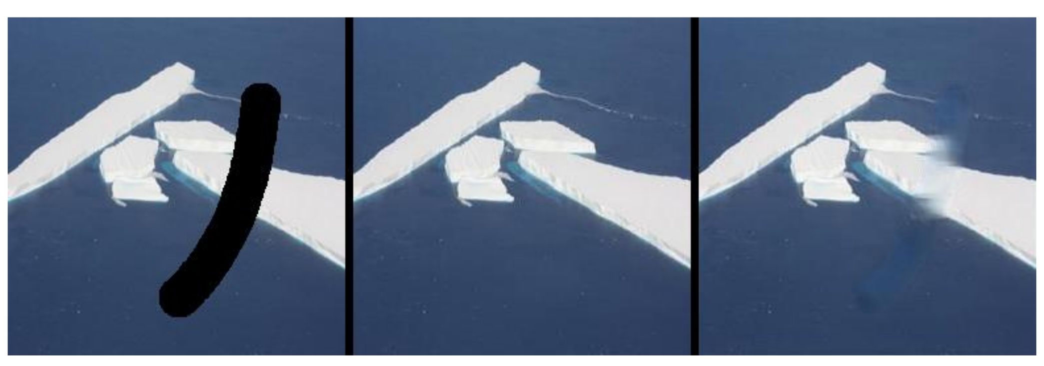


Fig.1

Fig.2

Fig.3

Fig.4

Fig.5

1 Training Procedure

We train on a single NVIDIA GTX 1060 GPU with a batch size of 6. The Initial Training takes us 16 hours and the fine-tuning takes us 24 hours.

2 Initial Training and Fine-Tuning

We firstly use a learning rate of 0.0002 for initial training, with test results shown in Fig.1 and Fig.2 (from left to right: masked image, ground truth and our prediction). Then we fine-tune using a learning rate of 0.00005 and freeze the layers in the encoder part of the network. This accelerates the convergence.

3 Result

Training result is shown as Fig.3, Fig.4 and Fig.5 (from left to right: masked image, ground truth and our prediction).

Summary

We develop work of Guilin et al.[1] in small holes by simplifying network and loss function in the initial training and we demonstrate that simplified network has good performance as well. We made a training dataset, finished training and put it into practical use with a user interface. The network will perform better in large holes inpainting if automatic mask update step (expanding from small to large) is taken. It has potential to perform better when supported with larger training dataset and more computing resource.

Reference

[1] Guilin Liu et al., M.: Image Inpainting for Irregular Holes Using Partial Convolutions. arXiv preprint arXiv: 1804.07723v2(2018)

[2] Deepak Pathak et al., M.: Context Encoders: Feature Learning by Inpainting. CVPR, 2016

PEKING UNIVERSITY 1898

AIPKU