

SymbolNet : First Step of Handwriting to Latex

Tao Cheng

East China Normal University

taocheng01@gmail.com

Jan, 2019

Overview

- 1 Motivation
- 2 Related works
- 3 Pipeline: preprocessing, recognition...

Motivation

- Latex is difficult to use for beginners
- Handwriting is easier and natural for us
- Important method to manage scientific documents

- Using end-to-end learning: based on encoder-decoder model^[1]

The diagram illustrates the mapping between LaTeX source code and its rendered output. The source code is shown at the top, and the rendered equations are shown below it on a grid background. Dashed lines connect specific tokens in the source code to their corresponding elements in the rendered equations.

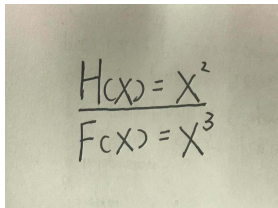
Source code: `Q = (b + 1/b)\rho, \quad \rho = \frac{1}{2} \sum_{\alpha>0} \alpha,`

Rendered equations: $Q = (b + 1/b)\rho,$ and $\rho = \frac{1}{2} \sum_{\alpha>0} \alpha,$

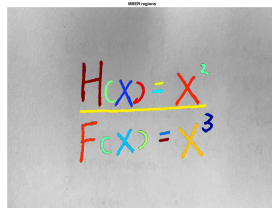
- Based on recognizing single character^[2]

Pipeline: preprocessing, recognition...

- Preprocess: Use Gaussian Blur algorithm to reduce noise in images and object detection algorithm(MSER) to detect every symbol in images.



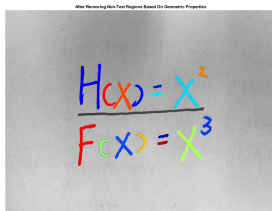
(a) Original Image



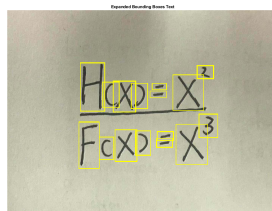
(b) Using MSER

Pipeline: preprocessing, recognition...

- Preprocess: Use Gaussian Blur algorithm to reduce noise in images and object detection algorithm(MSERER) to detect every symbol in images.



(a) Denoising



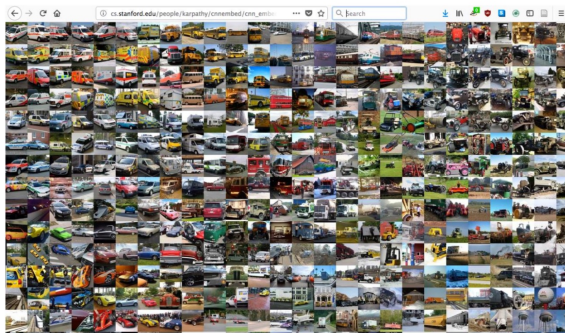
(b) Boxing Symbols

SymbolNet: recognize symbols in images

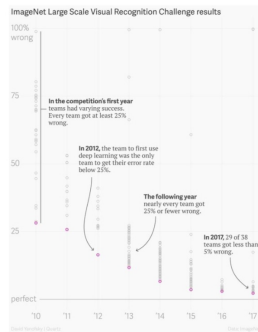
- AlexNet: proposed by Alex Krizhevsky, the state of art performance in ImageNet-2012 competition. milestone of deeplearning. Relu, training on multiple GPUs, dropout...
- SymbolNet: CNN(convolutional neural network) to recognize single symbol motivated by AlexNet^[3]

AlexNet and ImageNet

Dataset



[<http://cs.stanford.edu/people/karpathy/cnnembed/>, 30.11.2017]



[<https://qz.com/1034972/the-data-that-changed-the-direction-of-ai-research-and-possibly-the-world/>, 30.11.2017]

SymbolNet

- Dataset: Xai Nano's handwritten math symbols from Kaggle, 82 categories, 54,078 images for training, 13,505 images for testing

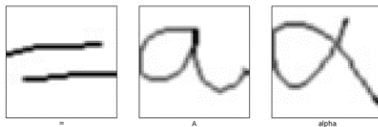


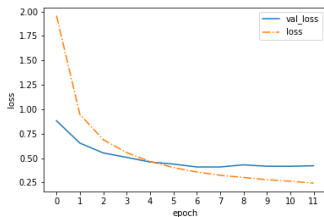
Figure: images after preprocessing

- SymbolNet Architecture

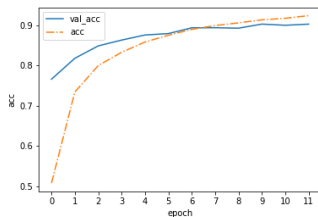
Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 43, 43, 32)	320
conv2d_2 (Conv2D)	(None, 41, 41, 64)	18496
max_pooling2d_1 (MaxPooling2)	(None, 20, 20, 64)	0
dropout_1 (Dropout)	(None, 20, 20, 64)	0
flatten_1 (Flatten)	(None, 25600)	0
dense_1 (Dense)	(None, 128)	3276928
dropout_2 (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 82)	10578
Total params: 3,306,322		
Trainable params: 3,306,322		
Non-trainable params: 0		
None		

Experiment Result

- Training Process



(a) Loss



(b) Accuracy

- Baseline Model: two dense layers with drop out

Experimental Result

Model	Accuracy	Time(min)/epoch	Model Size
Baseline	90.27%	7.5min	37.8MB
SymbolNet	80.49%	0.2min	3.10MB

Table: Model Comparison

- **TO DO:**

Build RNN(recurrent neural network) to take recognized symbols and their position as input and generate LATEX

Try residual net^[4] and end to end learning...

- 1 . Yuntian Deng, Anssi Kanervisto et al.(2017) Image-to-Markup Generation with Coarse-to-Fine Attention
- 2 . Xuan Yang, Jing Pu. (2015) MDig: Multi-digit Recognition using Convolutional Neural Network on Mobile
- 3 . Alex Krizhevsky, Ilya Sutskever.(2012)ImageNet Classification with Deep Convolutional Neural Networks
- 4 . Kaiming He, Xiangyu Zhang et al.(2015) Deep Residual Learning for Image Recognition

Thank You



Figure: Github