Deep Neural Networks for Recognizing Handwritten Mathematical Symbols

Domain Background

Given the importance of mathematics in all branches of science (physics, engineering, medicine, economics, etc.), the recognition of handwritten mathematical expressions has become a very important area of scientific research. However, recognition of mathematical symbols and expressions is a tough problem in the fields of computer vision and machine learning. Part of the reason is due to the variety in writing styles and enormous amount of mathematical symbols. As a student from mathematical department, diving in this problem is a meaningful thing to me. There are many related research about this problem such as

Xiaofang Xie On the Recognition of Handwritten Mathematical Symbols 2007

and

F. Álvaro, J.-A. Sanchez, J.-M. Benedi, Classification of On-line Mathematical Symbols with Hybrid Features and Recurrent Neural Networks, Proc. 12th ICDAR, Washington DC, USA, 25-28 Aug. 2013, pp.1012-1016 which is an application of RNN.

Besides , CHROME(*the Competition on Recognition of Online Handwritten Mathematical Expressions*) has also produced a lot of interesting result of this field.

Problem Statement

In this project, the mission for me is to let my trained model to recognize the single mathematical handwritten symbol. The possible solution to the problem could be SVM and CNN. I am going to use the accuracy to be the evaluation metrics, which is the same with deep learning project.

Datasets and Inputs

The dataset can be easily obtained from kaggle(here).

Dataset consists of (45x45) .jpg images. It includes basic Greek alphabet symbols like: $\alpha, \beta, \gamma, \mu, \sigma, \phi$ and θ ….

All math operators $\pm = \times \div$ …, set operators $\cup \cap \emptyset$ …. Basic pre-defined math functions like: $\log, \lim, \cos, \sin, \tan$. Math symbols like: $\sqrt{a}$ $\sum a$ $\int a$ and more.

This dataset is the whole dataset I need for training and testing. The main task here is to classify the mathematical symbols. The frequency table below of a sample of the data shows the distribution of those symbols

−	**33997**	!	**1300**	(	**14294**	)	**14355**	’	**1906**
[	778	]	780	{	376	}	377	+	25112
=	13104	0	6914	1	26520	2	26141	3	10909
4	7396	5	3545	6	3118	7	3765	8	3888
9	3737	A	12376	α	2546	b	8561	β	2025

cos	2986	lim	1675	√	8908	∀	45	≠	558

The table above shows that the number of pictures varies much and some of symbols do not have enough pictures for the algorithm to capture its characteristic, which is a problem that need to be concerned. Generally speaking, the data set is cleaning and large enough for the model to train.

Solution Statement

The solution that I is convolutional neural network. The most commonly used architecture had 3 convolutional layer each with 128, 64 and 32 filters respectively. These filters were of size 4 * 4, 3 * 3 and 2 * 2 respectively for each layer and had a stride of 2 * 2, 2 * 2 and 1 * 1 respectively for each layer. The activation function between each convolutional neural network should be ReLU.

Benchmark Model

After CHROME(the Competition on Recognition of Online Handwritten Mathematical Expressions), papers and research flourish. Xinyan et and al used a combination of Genetic algorithm and neural network to classify the mathematical symbols and got accuracies close to 90.6%. Another work in this area was done by Lu et and al who used a

convolutional Neural network to classify the symbols. They got accuracies close to 83%.

Evaluation Metrics

The main evaluation metrics that I am going to use is accuracy. However the training time and cross validation loss would also be considered. Besides, in order to evaluate the model thoroughly, F1-score would also be an option.

Project Design

The workflow:

1. Data cleaning and preprocessing
2. Model construction
3. Model evaluation and optimization
4. Future work