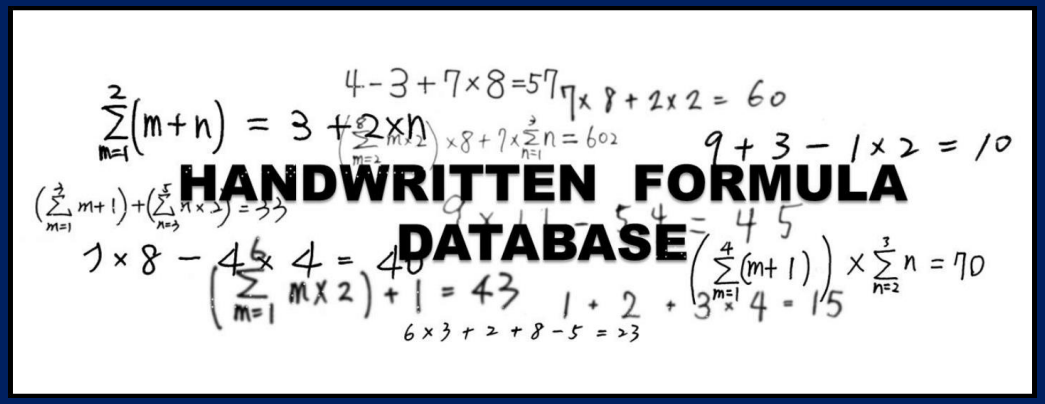
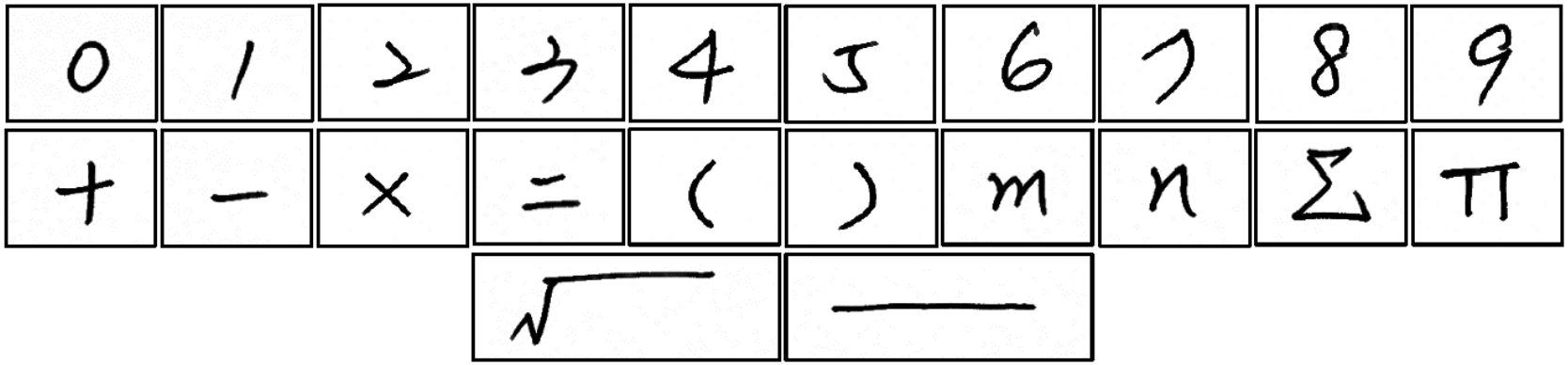
HANDWRITTEN FORMULA COMPETITION

0386032 – Jan Kristanto

TABLE OF CONTENTS

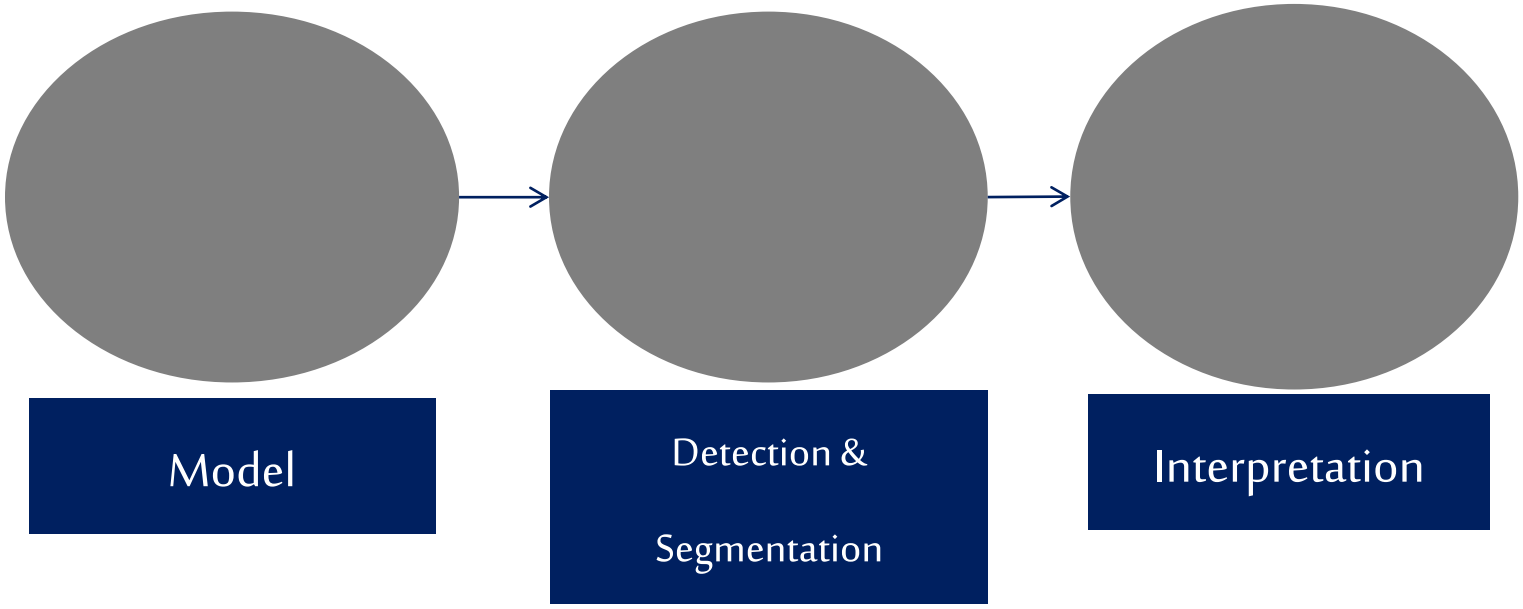


- There are 22 symbols
- Each symbol has 515 samples
- Each sample 100x100, except Fraction and Square Root (100x200)



WORK FLOW

Steps to solve the problem



Model : build classifier

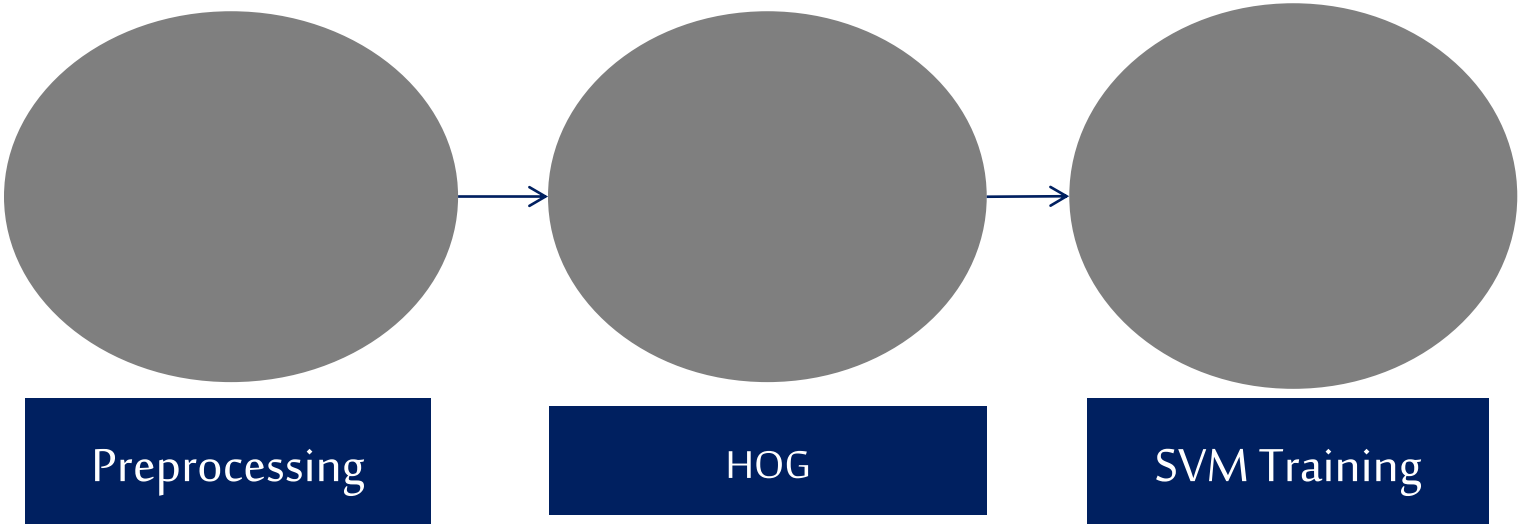
Detection Segmentation : Segment and detect the existing symbols

Interpretation : interpret the symbols into computable math formula

MODEL

Classifier To recognize the symbols

I am using SVM and HOG to build the classifier.

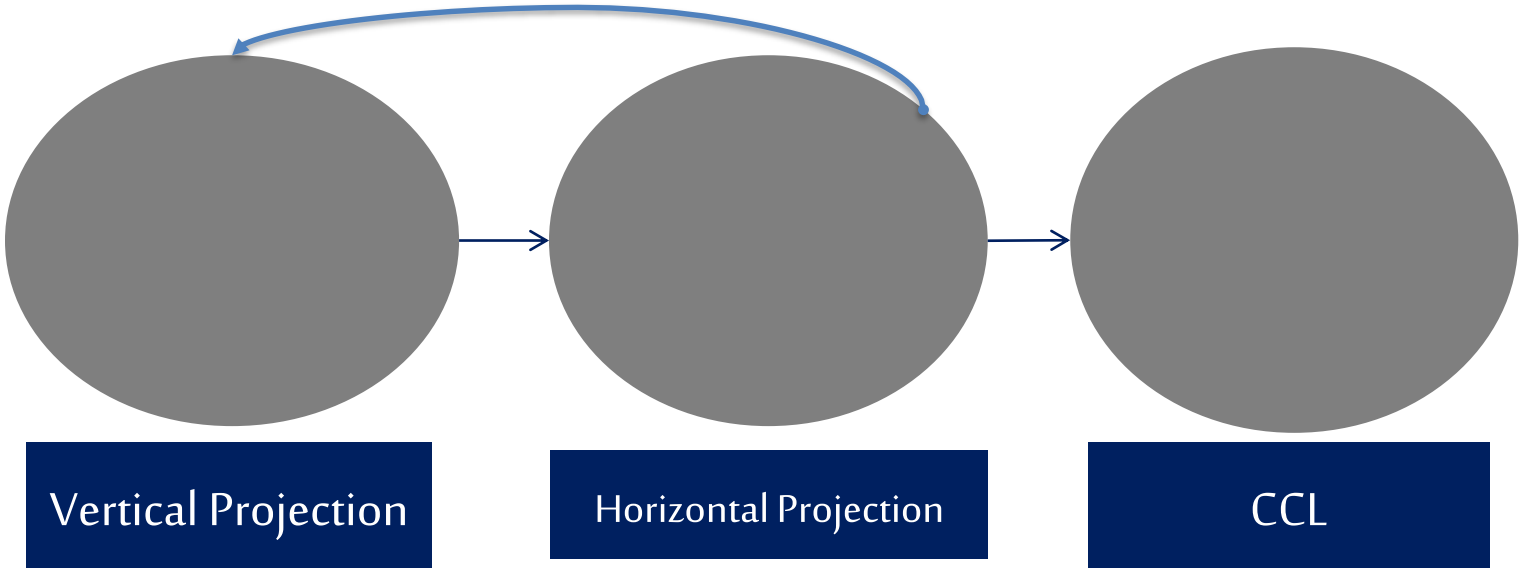


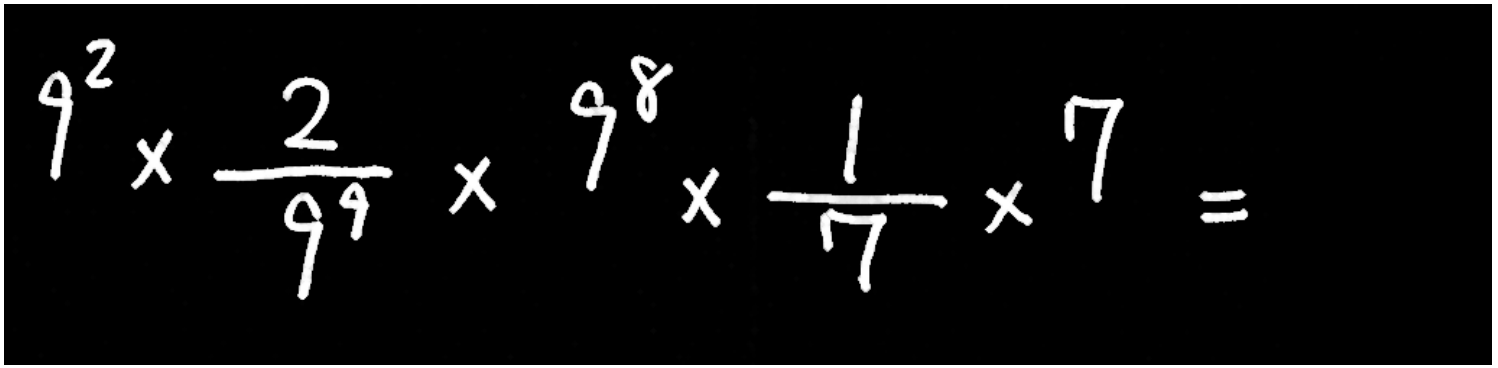
Based on this schema, I have a model classifier with 99.02% accuracy regards to Training data.

DETECTION & SEGMENTATION

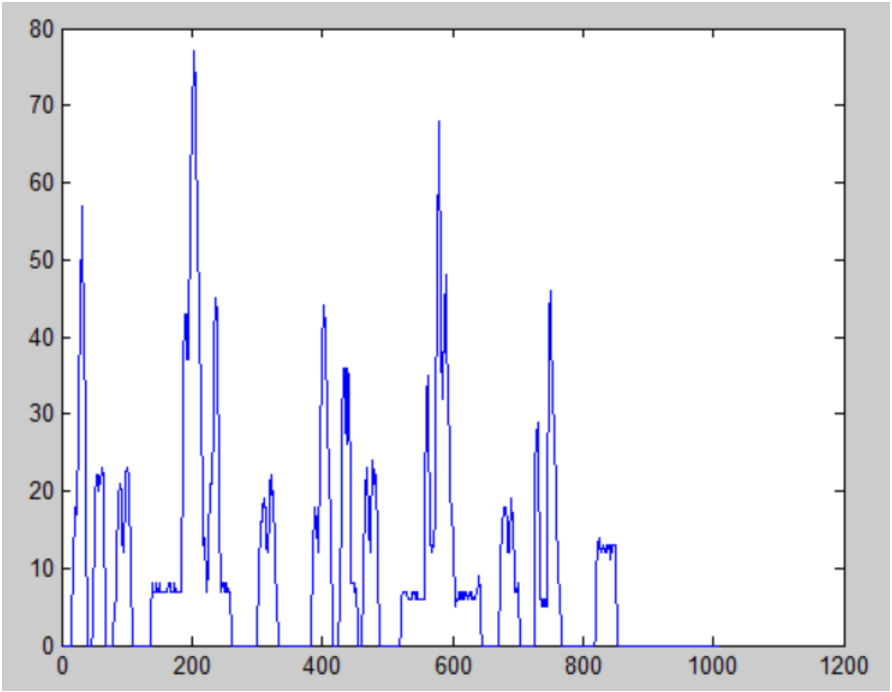
Segment and detect symbols

- Vertical & Horizontal Projection
- Component Connected Labelling (CCL)

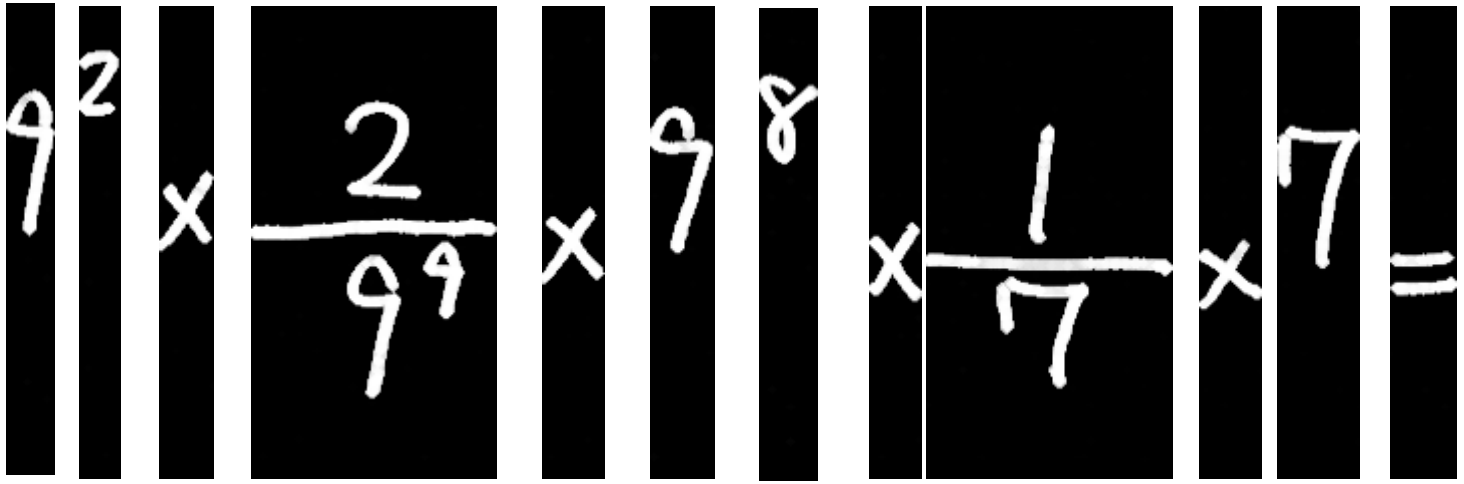


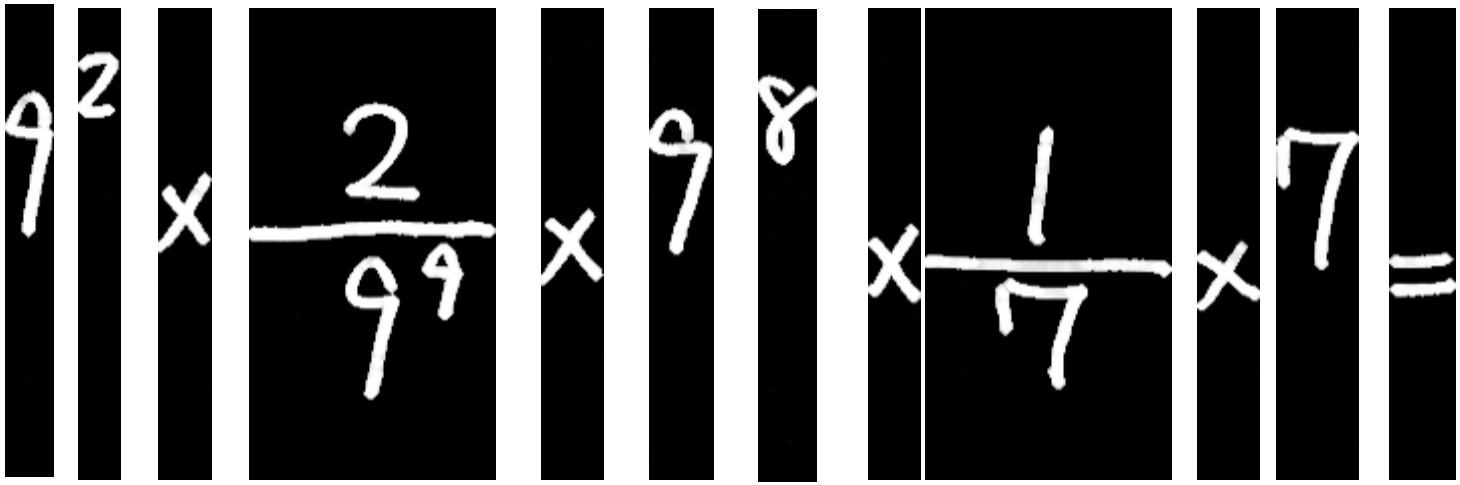


Vertical Projection

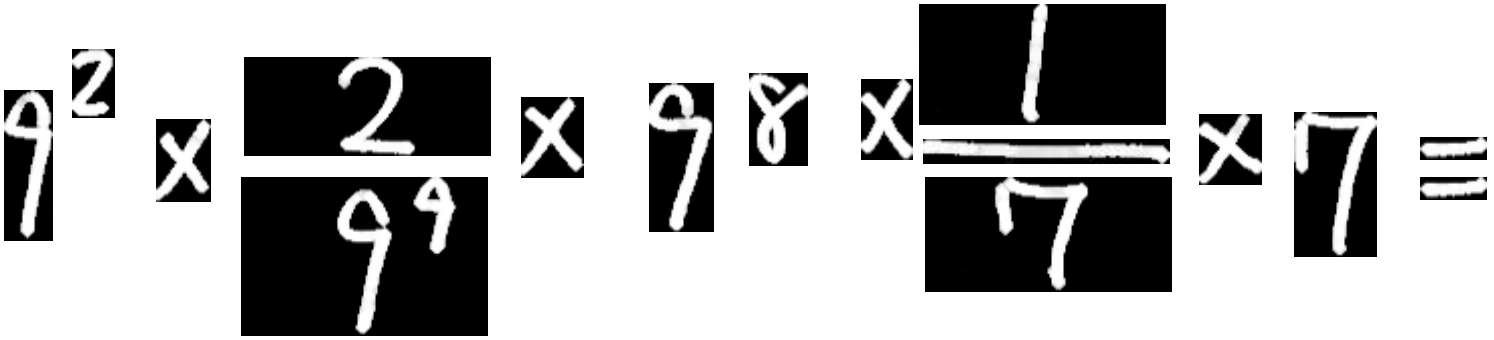


$$9^2 \times \frac{2}{9^9} \times 9^8 \times \frac{1}{7} \times 7 =$$





Horizontal Projection



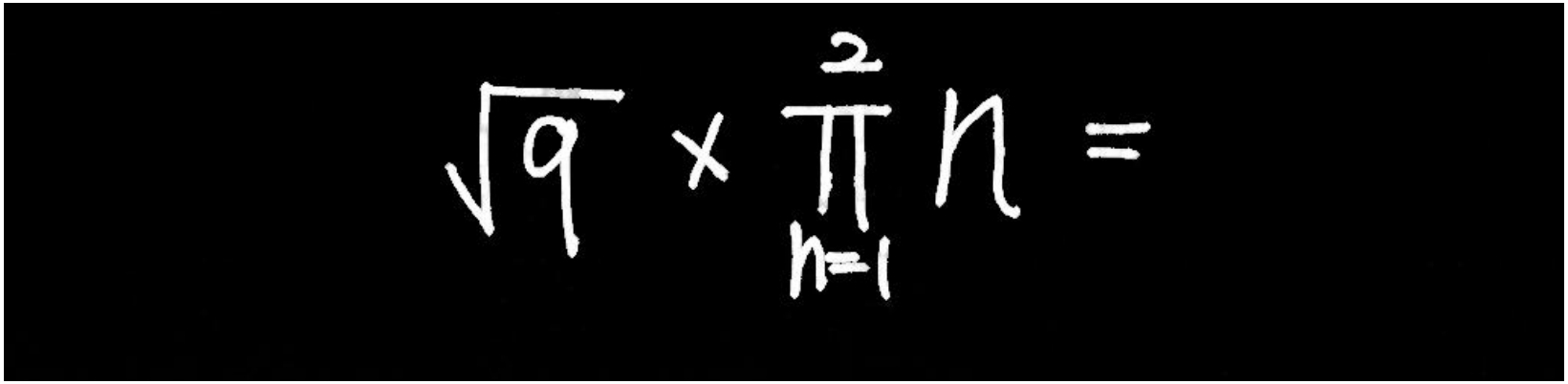
.....

.....

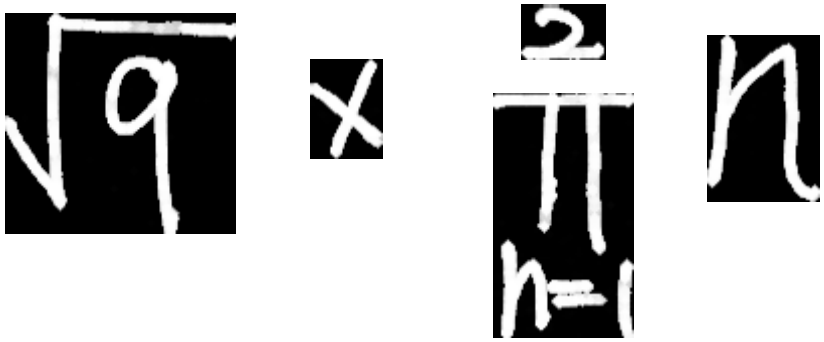
.....

.....

.....



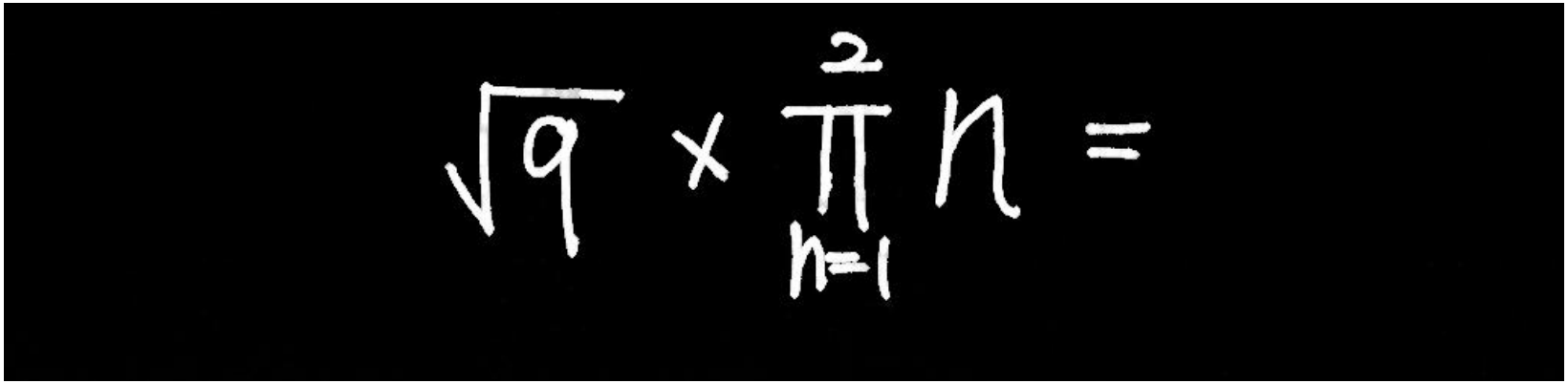
After Vertical & Horizontal



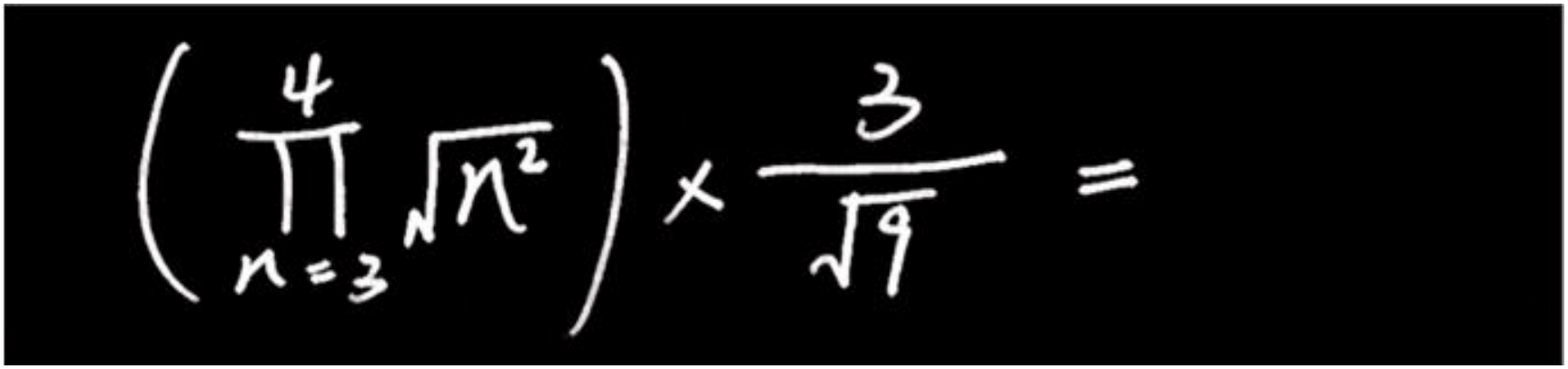
Component Connected Labelling

Fix Smallest Size

Add Padding (for to be square)



Resize to 100x100



1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16
17							

INTERPRETATION

How to transform and compute symbols that have been identified

- Interpret formula as string

- Simply just use eval matlab function.

Sigma and Product

1. Find sigma/product symbols
2. Find the digit before sigma/product symbol and recognize as digit ending.
3. Find the m or n after sigma/product symbols
4. Find the digit after m or n character
5. Find m or n character until non digit symbol.

INTERPRETATION

How to transform and compute symbols that have been identified

Square Root

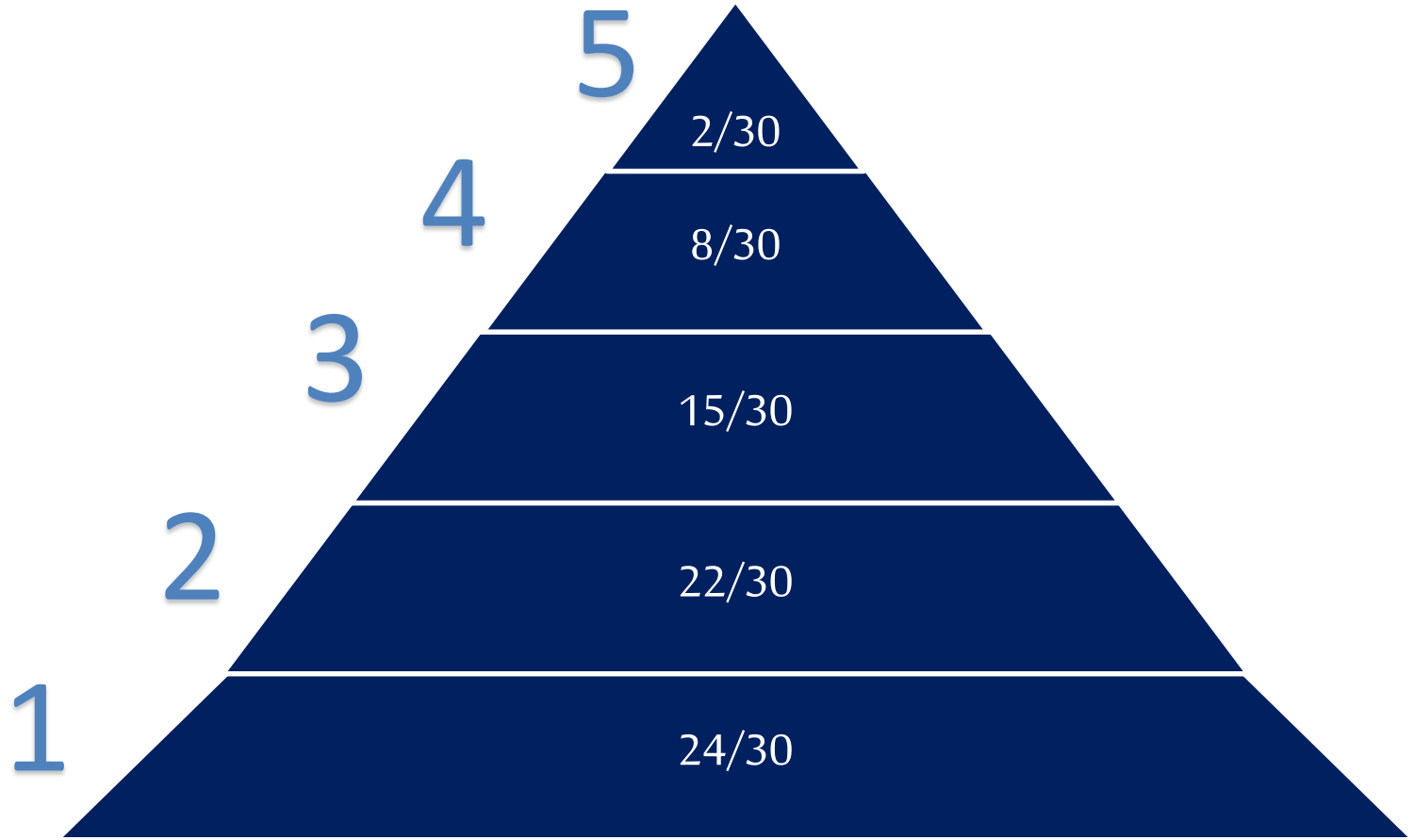
1. Find sqrt symbols
2. Find the small digit before sqrt symbol and
3. Find the m or n or digit after sqrt symbols until operator symbols

Power

1. Take a look the size of the digit and compare to the previous digit.

ACCURACY

Accuracy formula for each level



LIMITATION & FUTURE WORK

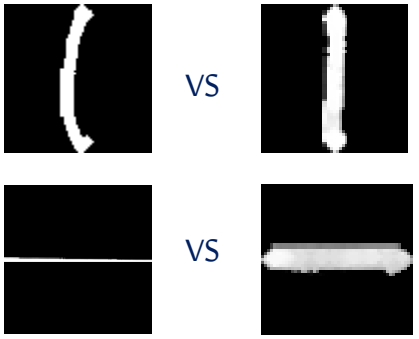
Drawback and proposed idea

1. The Classifier sometimes miss classify

   “(“ -> “1”

   “__” (fraction) -> “-” (subtract)

   Try other model than HOG and SVM



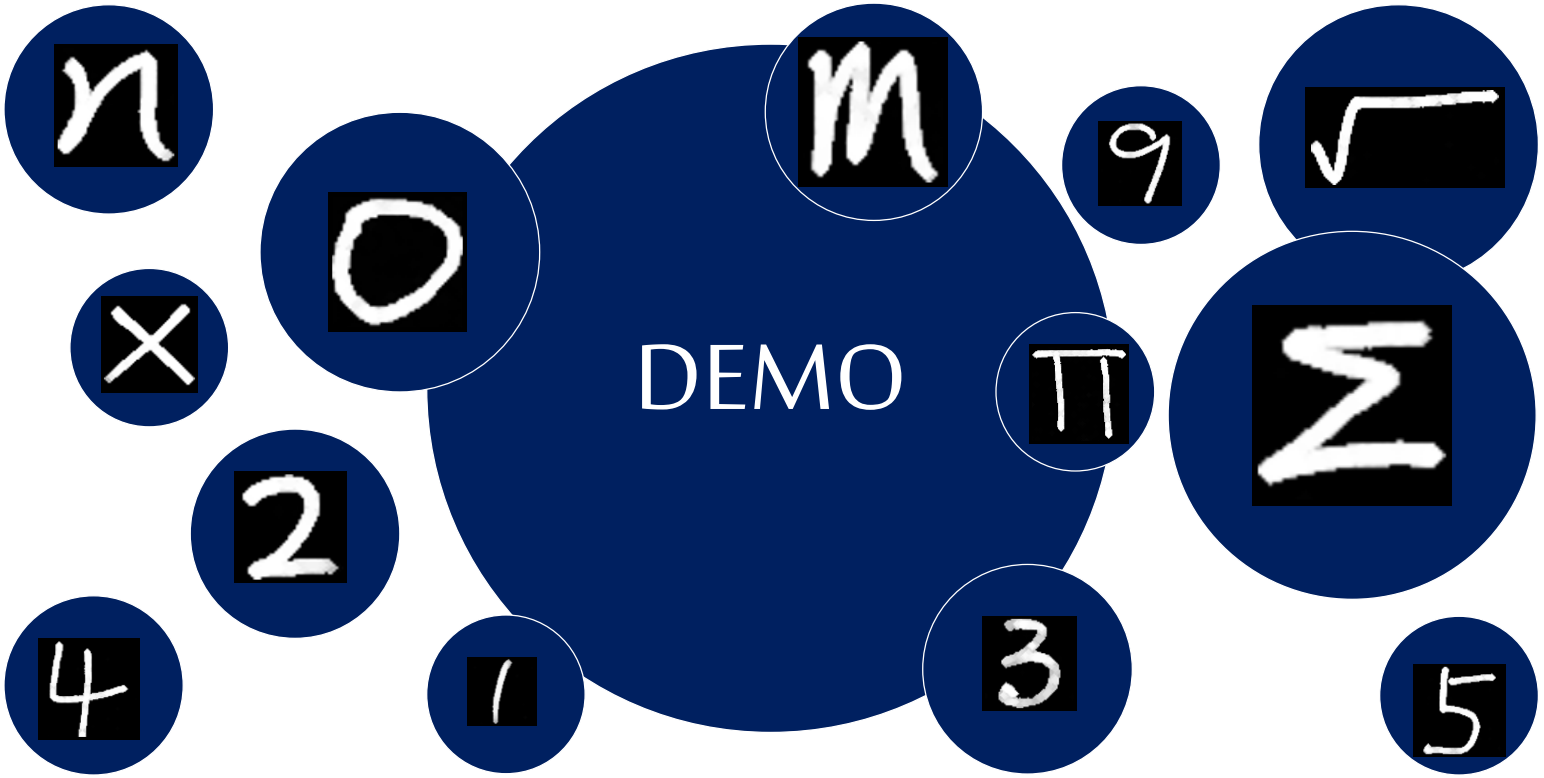
2. Fail to segment symbols properly.

   Usually because of the handwritten is too complicated.

   Try use window sliding after CCL

3. The formula interpreter fails to interpret complicated formula.

   Try to use standard format. MathML , Latex or at least Tree.


DEMO

$$8-6\times5\times1+2-4\times3+97=$$

$$8+3-5\times4=$$

$$2+3-1\times5=$$

$$(7+2)\times 9 =$$

$$((((((9+1)+2)+3)-5)-1) =$$

$$(12+34+5)\times 6 =$$

$$9^2 \times \frac{2}{99} \times 9^8 \times \frac{1}{7} \times 7 =$$

$$\frac{9}{3^2} \times 4^2 \times 3 =$$

$$\frac{1^3 + 2^2 + 3^1}{1 + 2 \times 3 - (3+2)} =$$

$$\left(\sum_{n=1}^{3} n^2 \times 2\right) =$$

$$\left(\sum_{m=1}^{2} m + 1\right) + \left(\sum_{n=3}^{5} n \times 2\right) =$$

$$\sum_{m=1}^{5} (m \times 2) - 30 + 3 =$$

$$\prod_{n=1}^{4} \sqrt{n^2} \times \frac{1}{\sqrt[5]{9}} \approx$$

$$\left(\prod_{m=1}^{3} m \right) \times \frac{1}{\sqrt{4}} =$$

$$\left(\sum_{n=0}^{3} \sum_{m=1}^{9} \frac{m^2}{\sqrt{m^2}} \times n \right) \times 4 - \prod_{n=5}^{6} (n+0)^2 \approx$$

THANKYOU

THANKS FOR WATCHING THIS PRESENTATION