Using Off-line Features and Synthetic Data for On-line Handwritten Math Symbol Recognition

Kenny Davila
Document and Pattern Recognition Lab
Rochester Institute of Technology
Rochester, New York, USA

dprl [DOCUMENT AND PATTERN RECOGNITION LAB]

ICFHR 2014 - Crete, Greece

R·I·T

September 3, 2014

Our Application

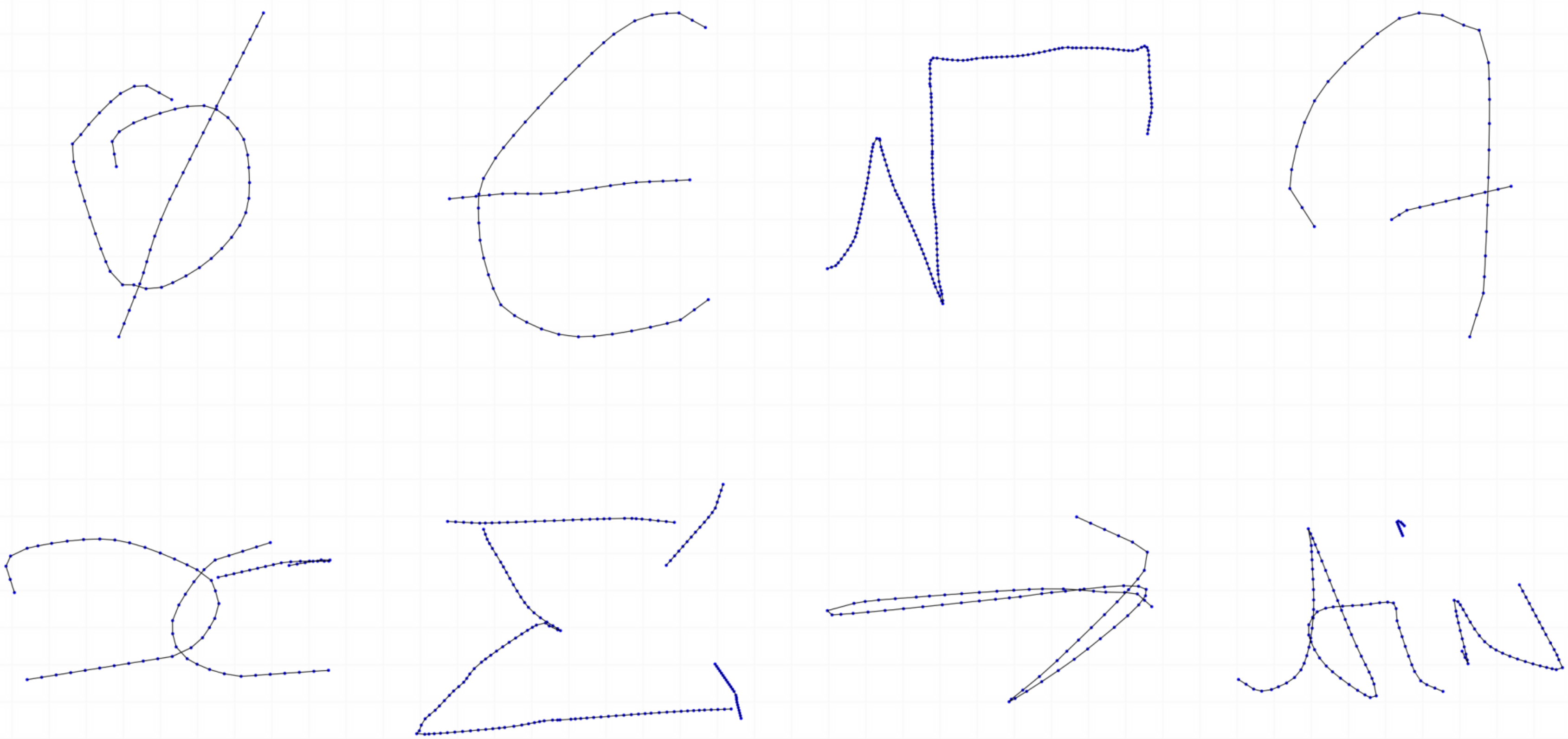
Format	• On-line
Source	• Handwritten
Domain	• Math (101 classes)
Scope	• Multiple writers

CROHME 2013 Classes

Group	Symbols	Count	Group	Symbols	Count
Digits	0-9	10	Arithmetic Operators	$+, -, \pm, \div, !, \times, /, \mathrm{sqrt}$	8
Letters	a-z,A-C,E-I,L-N, P, R-T,V,X,Y	44	Logical Operators	$\rightarrow, \|, \forall, \exists$	4
Greek Letters	$\alpha, \beta, \gamma, \lambda, \phi, \pi, \theta, \sigma, \mu, \Delta$	10	Set Operators	$\in$	1
Functions/ Relations	sin, cos, tan, lim, log, $=, \neq, <, \leq, >, \geq$	11	Operators with limits	$\sum, \int$	2
Fence Symbols	(,), {, }, [,]	6	Other Symbols	∞, *COMMA*, ., $\cdots$, PRIME	5

A total of 101 Classes

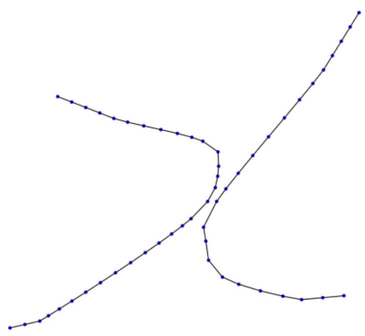
Math Symbol Examples



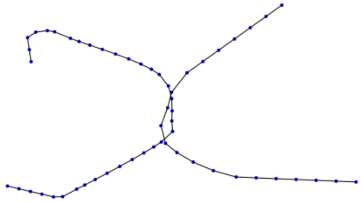
Extracted from CROHME 2013

Ambiguous classes

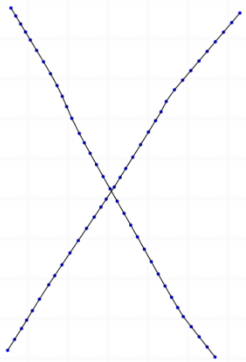
Class Group	Class Group	Class Group
X, x, times	C, c	S, s
1, | , (,) , comma	P, p	9, q
comma,), prime, /	V, v	O, 0



Uppercase X



Lowercase X



Times

Research Questions

- Can we define a good **set of features** for shape description which provide robust recognition results?
- Which **machine learning techniques** are best suited for our features?
- Can we improve recognition rates by **modifying the training dataset**?

Related Work: Handwritten Math Symbol Recognition

Method	Classifier	Features
Hu and Zanibbi [1]	HMM	NDTSE, curvature features
Alvaro et al. [2]	RNN	- Normalized Coordinates, first and second derivatives, curvature - Context Window with PCA
MacLean et al. [3]	Greedy DTW	Normalized Coordinates

Comprehensive Survey by Plamondon and Srihari (2000) [4]

[1] L. Hu and R. Zanibbi, "HMM-based recognition of online handwritten mathematical symbols using segmental k-means initialization and a modified pen-up/down feature," in ICDAR 2011
[2] F. Alvaro et al., "Classification of online mathematical symbols with hybrid features and recurrent neural networks," in ICDAR, 2013
[3] S. MacLean and G. Labahn, "Elastic matching in linear time and constant space," in DAS 2010
[4] R. Plamondon and S. Srihari, "Online and off-line handwriting recognition: a comprehensive survey." TPAMI, 2000

Proposed methodology

- Feature Extraction
- System Training
- Symbol Recognition

Features

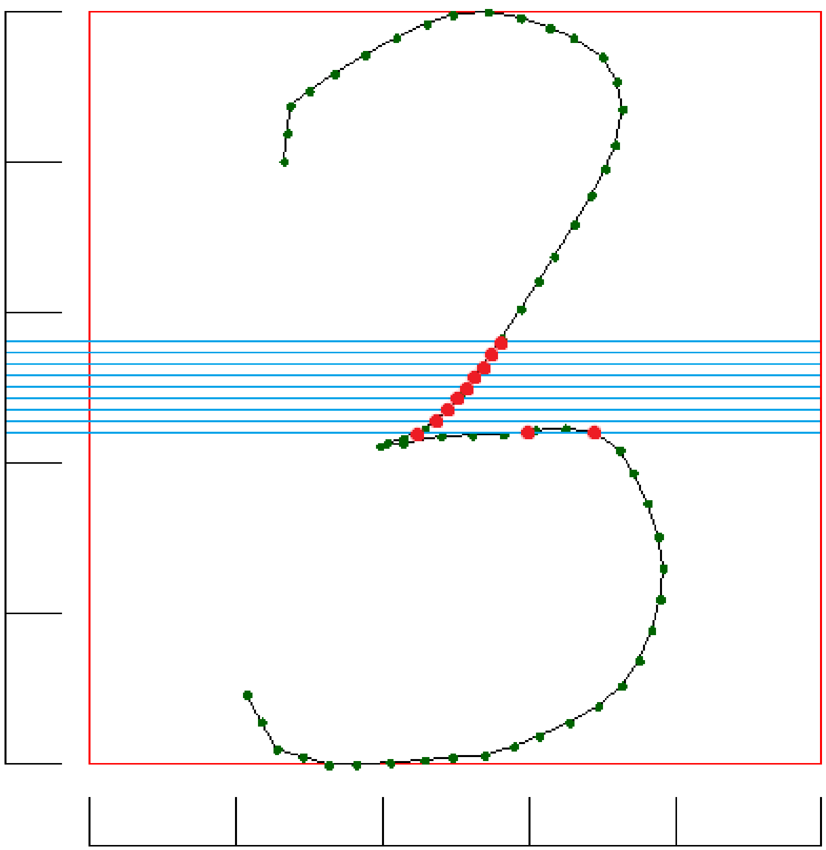
A total of 102 values in final vector

- Global Features (11)
  - global descriptors like aspect ratio, # of strokes, etc
- Crossings (30)
  - Intersections between strokes and lines at X, Y positions
- Fuzzy 2D Histograms (25)
  - 2D Histogram of points using fuzzy memberships
- Fuzzy Orientation Histograms (36)
  - Histogram of line segment angles with fuzzy memberships

Global features (11)

- Number of traces (1)
- Normalized Aspect Ratio (1)
- Center of mass (2)
- Covariance of X and Y coordinates (1)
- Per-trace average and total:
  - Angular Change (2)
  - Line Length (2)
  - Number of Sharp Points (2)

Crossings Features



Average Counts

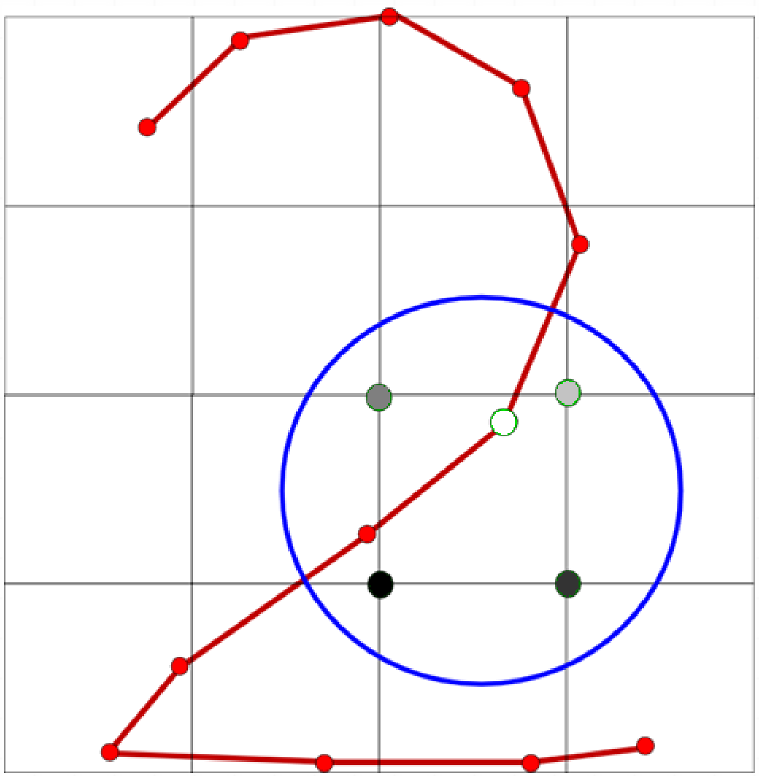
Horizontal	Vertical
2.0	0.0
1.0	1.8
1.2	4.0
1.0	3.8
1.6	0.0

Divide the symbol in regions:
5 Horizontal, 5 Vertical

Use 9 lines per region computing intersection: count, first and last

Compute averages per value per region for a total of 30 values

Fuzzy 2D Histogram



The symbol region is divided using a grid with 5x5 corners for 25 values

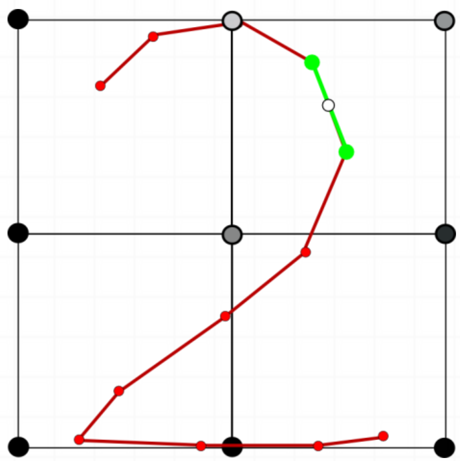
For point P compute the membership value over each corner C of the cell where P is located

$$P = (x_p, y_p)$$

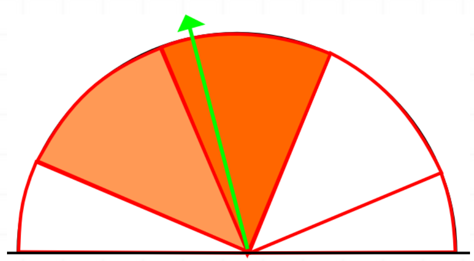
$$C = (x_c, y_c)$$

$$m_p = \frac{w - |x_p - x_c|}{w} \times \frac{h - |y_p - y_c|}{h}$$

Fuzzy Histogram of Orientations



Weights Per Distance



Weights Per Angle

Symbol is divided in cells with 3 x 3 corners with 4 angular bins per corner for 36 values in total

For each line segment we weight by:

1. Segment length
2. Distance to corners, same as Fuzzy 2D Histograms, and affects 4 sets of angular bins
3. Slope angle, it affects the 2 closest angular bins

Classifiers

- Four different methods applied
  - AdaBoost.M1 with C4.5 (Maximum 50 trees)
  - Random Forests (Maximum 50 trees)
  - SVM Linear Kernel
  - SVM RBF Kernel

- Parameters optimized using Grid Search

Symbol Recognition Experiments

- Each classifier was optimized using Grid search to find good parameter values

- We benchmarked the performance of our method using different classifiers

Math Symbol Recognition Benchmark

Method	Classifier	Top-1	Top-5
Hu et al.	HMM	82.9%	97.8%
Alvaro et al.	R-NN	89.4%	**99.3%**
MacLean et al.	Greedy DTW	85.8%	99.1%
Proposed Method	AdaBoost C4.5	88.4%	98.7%
	Random forests	87.9%	98.4%
	SVM Linear Kernel	88.6%	99.1%
	SVM RBF Kernel	**89.8%**	99.1%

Using a subset of MathBrush Dataset with 93 classes

Method	Classifier	Without Junk	With Junk
MyScript	MLP	91.04%	**85.54%**
Alvaro et al.	BLSTM-RNN	91.24% / 89.79%	84.14%
Proposed Method	SVM RBF Kernel	88.66%	83.61%

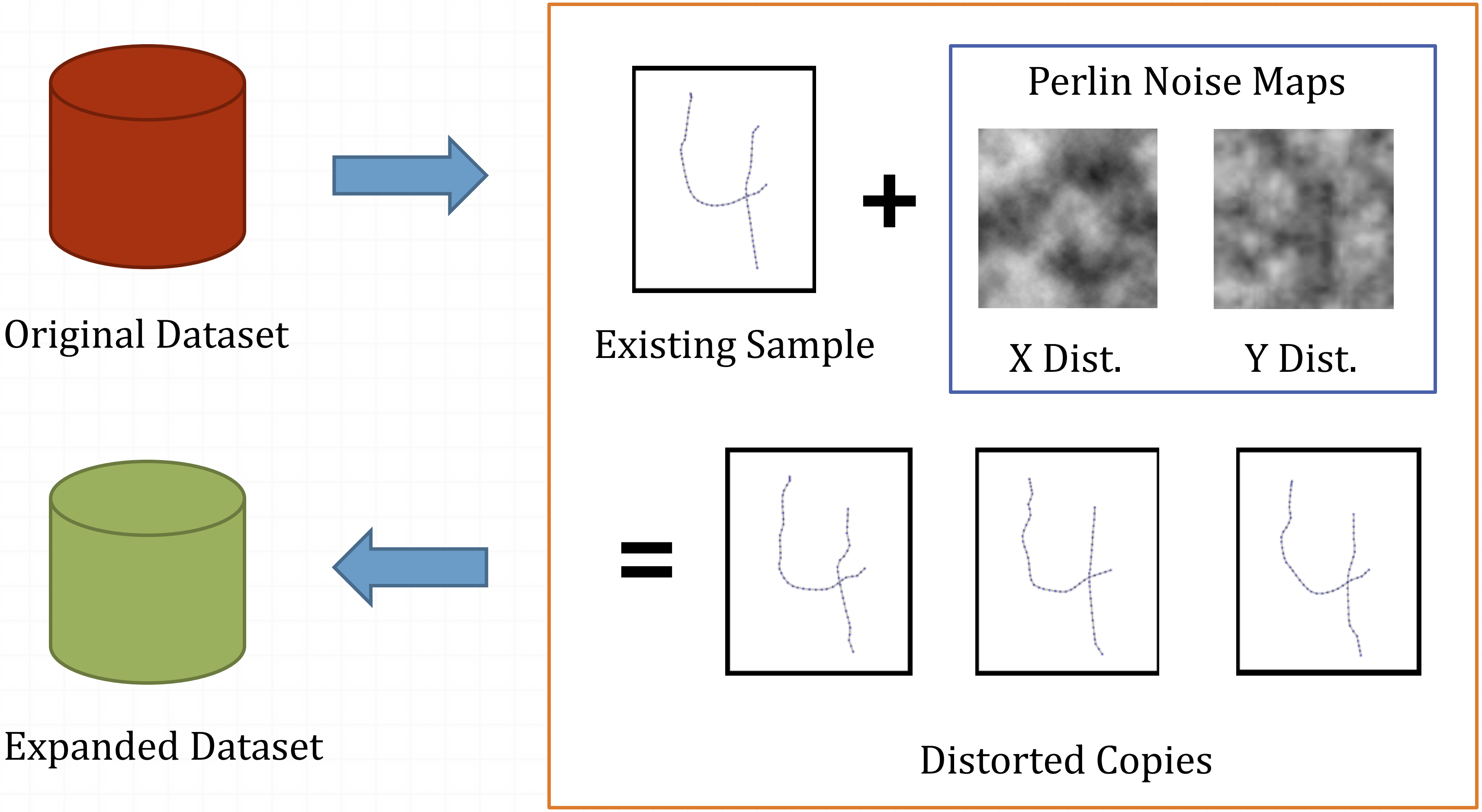
Using CROHME 2014 Dataset with 101 classes

Related work: Handwritten Data Generation

Method	Goal	Method
Simard et al. [1]	Synthetic Digit Images	- Elastic Distortion - Smooth Random Noise
Plamondon et al. [2]	Synthetic Strokes	- Training from Real Data - Kinetic Model
Sarkar et al.[3]	Style Identification	- K-means Clustering

[1] P. Simard et al, "Best practices for convolutional neural networks applied to visual document analysis." in ICDAR, 2003
[2] R. Plamondon et al, "Recent developments in the study of rapid human movements with the kinematic theory: Applications to handwriting and signature synthesis," in PR Letters, 2014.
[3] P. Sarkar, "Style consistent classification of isogenous patterns," TPAMI, 2005

Data generation

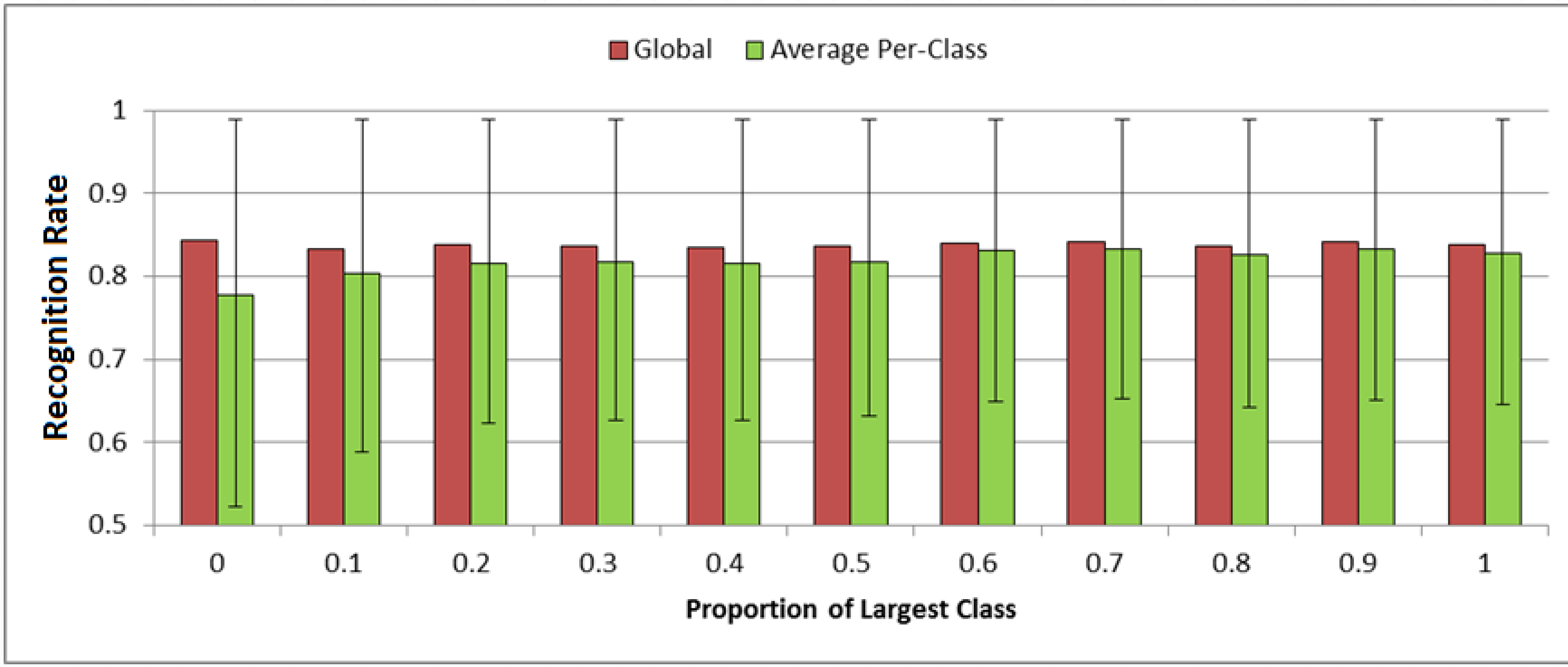


Data Generation Experiments

- First, we tuned up the data generation process itself to find out how much distortion is good for the system using a fixed amount of generated data
  - Perlin noise map sizes
  - Perlin noise map layers
  - Maximum displacements

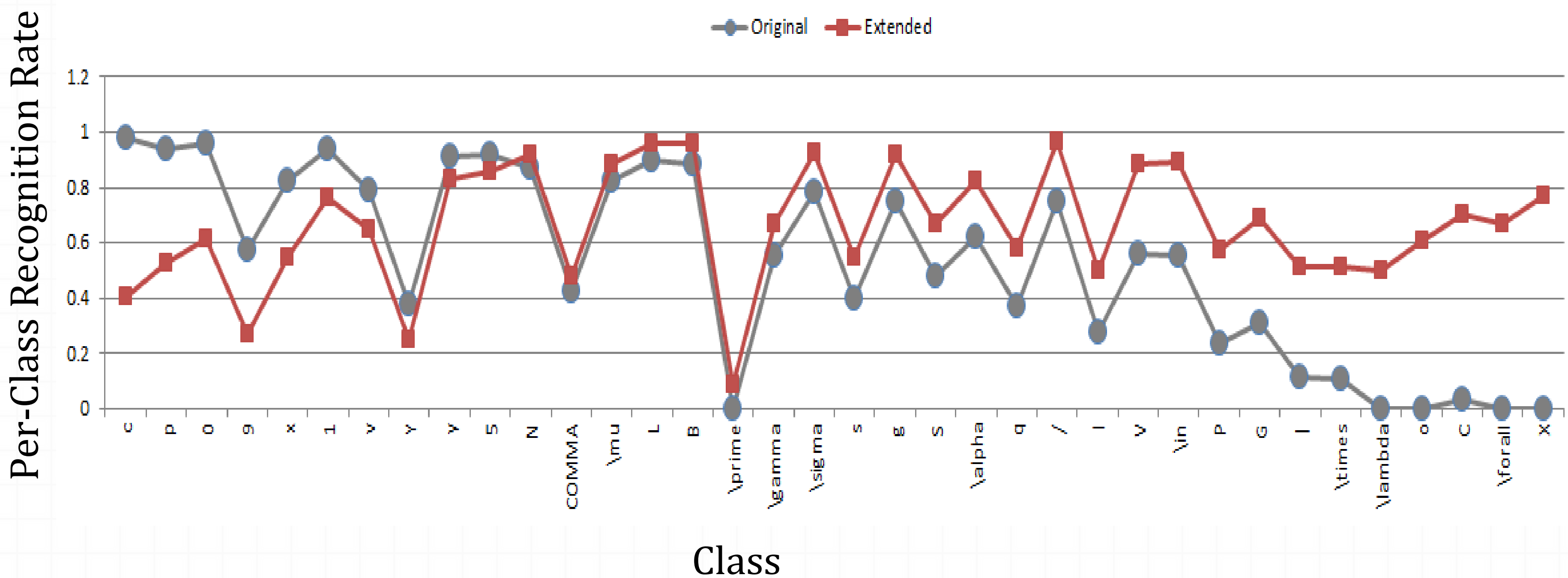
- Second, we tested how much data should be generated using our method using a fixed amount of distortion
  - Minimum number of samples per class relative to largest class

Recognition Rates For Different Amounts of Synthetic Data



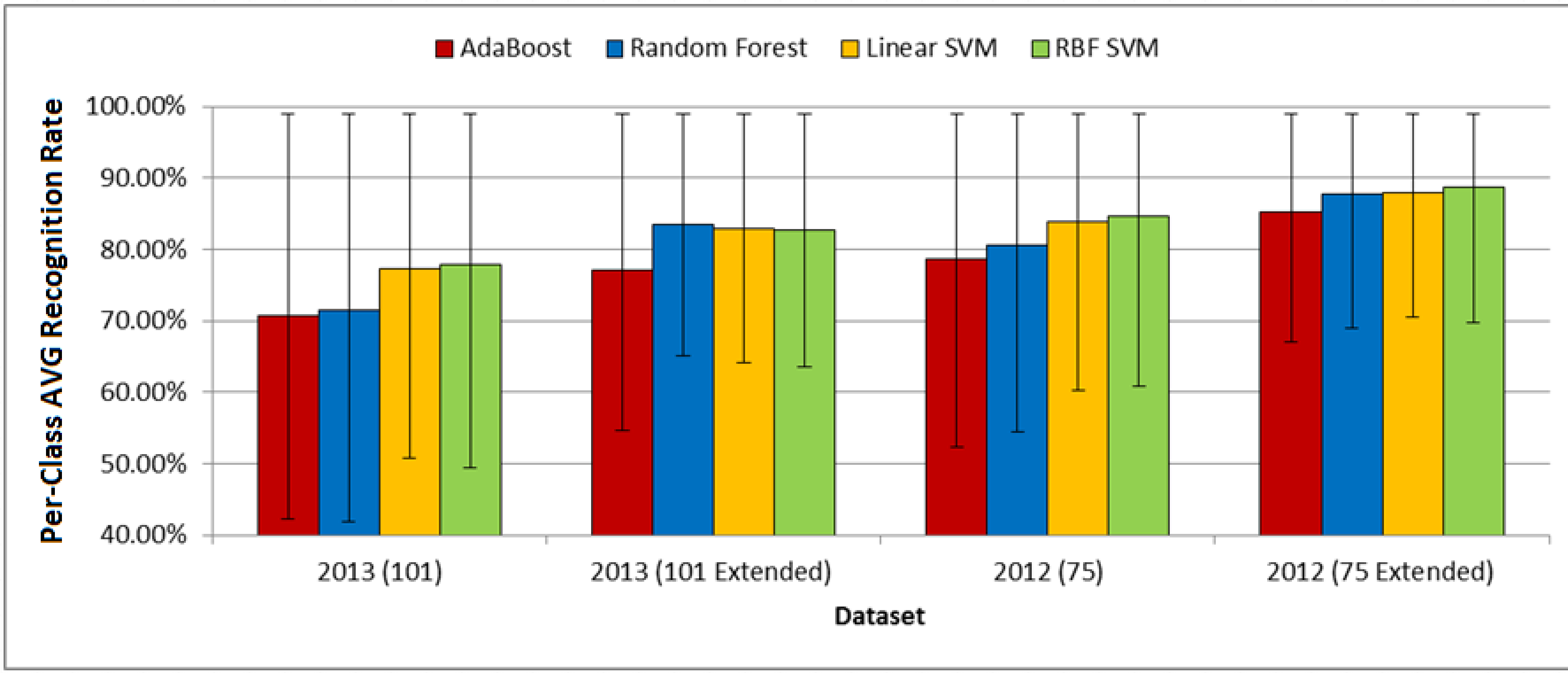
Using SVM with Linear Kernel over CROHME 2013

Data Generation Trade-Offs



Per-Class Recognition Rate Trade-off between
CROHME 2012 and CROHME 2012 Extended
Using SVM with Linear Kernel
(Only classes with more than 5% difference are shown)

Comparison of Learning Algorithms Average Per-Class Recognition Rate



Using CROHME 2012 and 2013 datasets

Discussion

- Data generation affected recognition rates with trade-offs:
  - Lower Global recognition rate
  - Higher Average Per-Class recognition rate

- Analysis of confusion matrix shows higher errors between **ambiguous classes**
  - Context is required to reduce errors

  - If we ignore these errors the new recognition rate is
    - 93.52% (vs 85.89%) for CROHME 2013 (101)
    - 96.36% (vs 94.49%) for CROHME 2012 (75)

Conclusions

- Can we define a good **set of features** for shape description which provide robust recognition results?

  Competitive recognition rates were achieved using a**daptations of off-line features.**

- Which **machine learning techniques** are best suited for our features?

  Best method depends on goals:
  - **SVM with RBF kernel** was best choice **for high recognition rate**
  - **Random Forests** was best choice **for speed**

- Can we improve recognition rates by **modifying the training dataset**?

  Trade-offs between ambiguous classes prevent data generation from achieving higher recognition rates

Future work

- Explore additional features
  - HBF49 by Delaye and Anquetil [1]

- Apply method on different datasets

- if context is available, use cascade classification
  - Group sets of ambiguous classes as a single class each
  - Use second classifier on each set of ambiguous classes with context features

[1] A. Delaye et al, "Hbf49 feature set: A first unified baseline for online symbol recognition," Pattern Recognition, 2013.

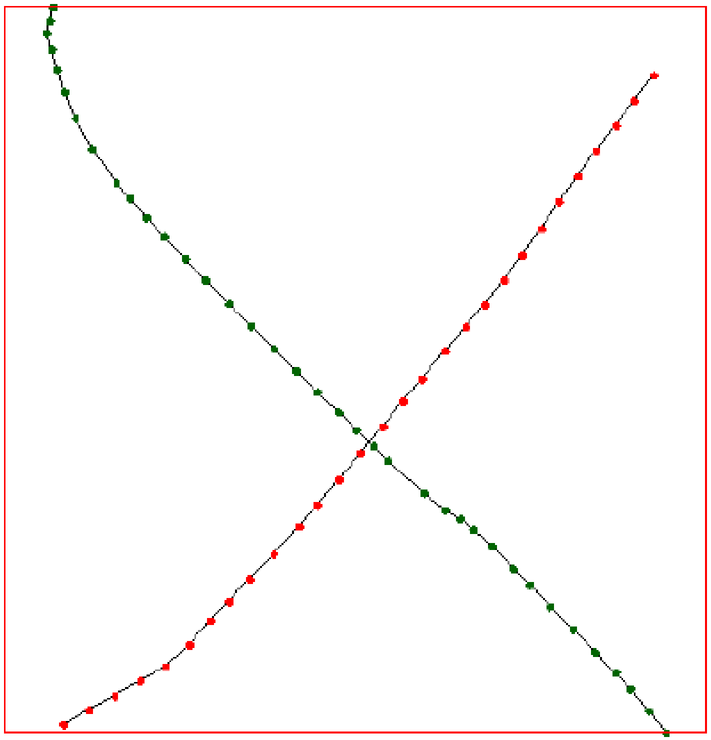
Questions?

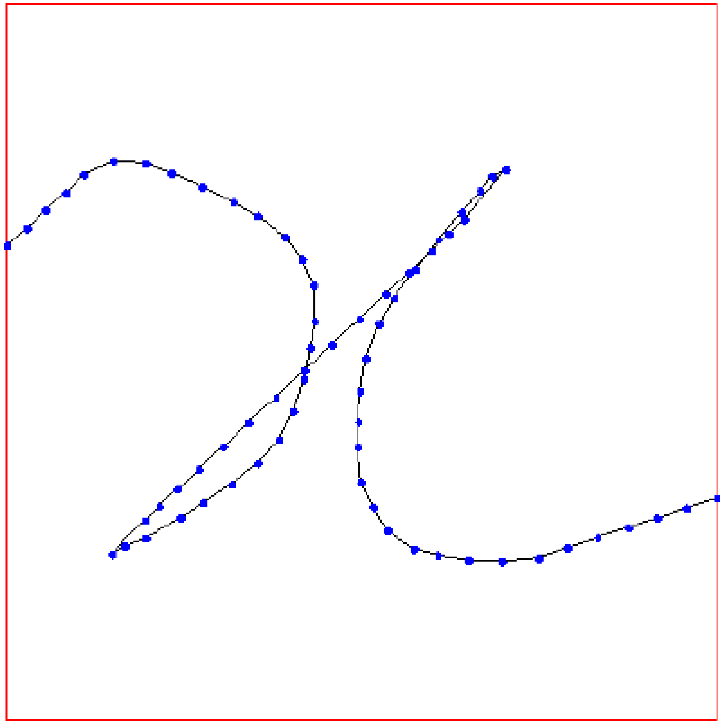
This material is based upon work supported by the National Science Foundation (USA) under Grant No. HCC- 1218801

dprl [DOCUMENT AND PATTERN RECOGNITION LAB]

R·I·T

Final shape vs. drawing process





- Two traces
- Small Angular Variation

- One trace
- Large Angular Variation

Data balancing strategy

- Balance class representation using

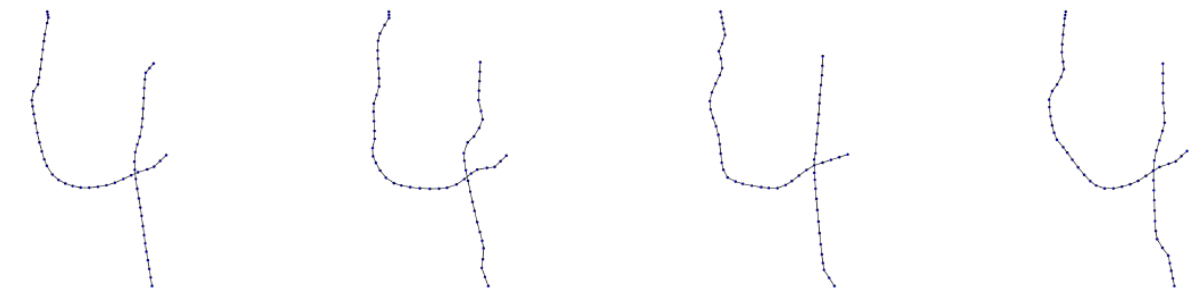
$$Min_{count} = T\,|C|$$

  - T is a parameter
  - C is the largest class

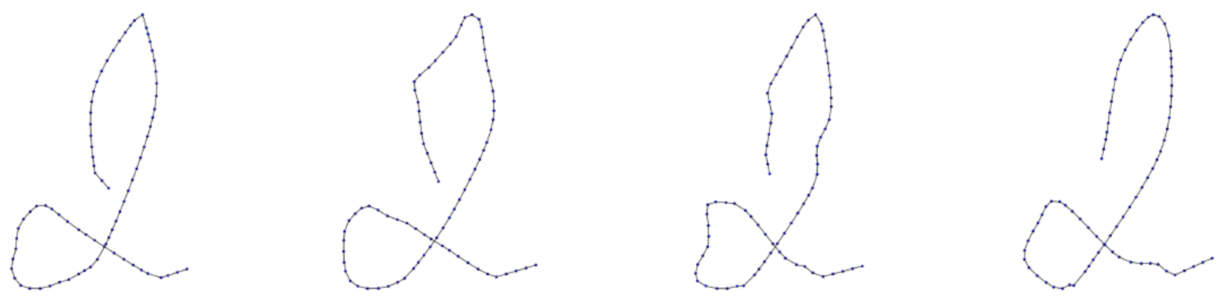
- The dataset is balanced if

$$T \geq 1.0$$

Data generation example



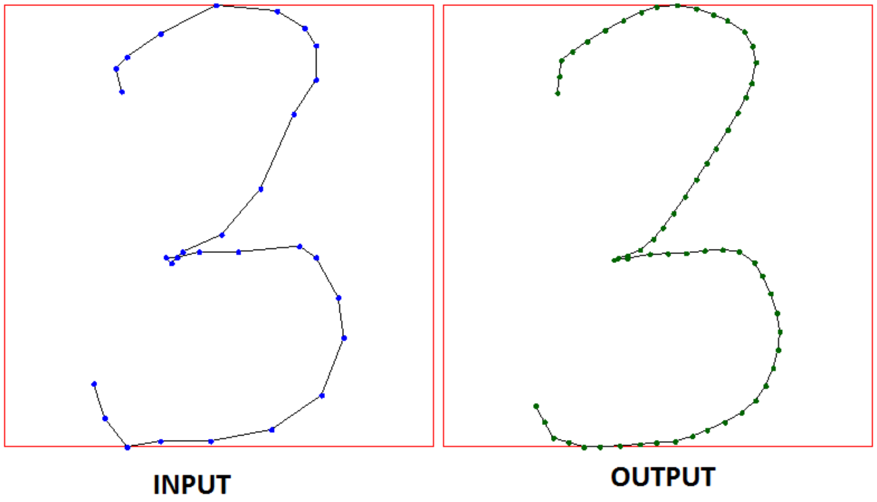
(a) Original (b) Copy 1 (c) Copy 2 (d) Copy 3



(e) Original (f) Copy 1 (g) Copy 2 (h) Copy 3

Data preprocessing

- Based on method by Huang et al. [1]
- Removal of noise by resampling traces

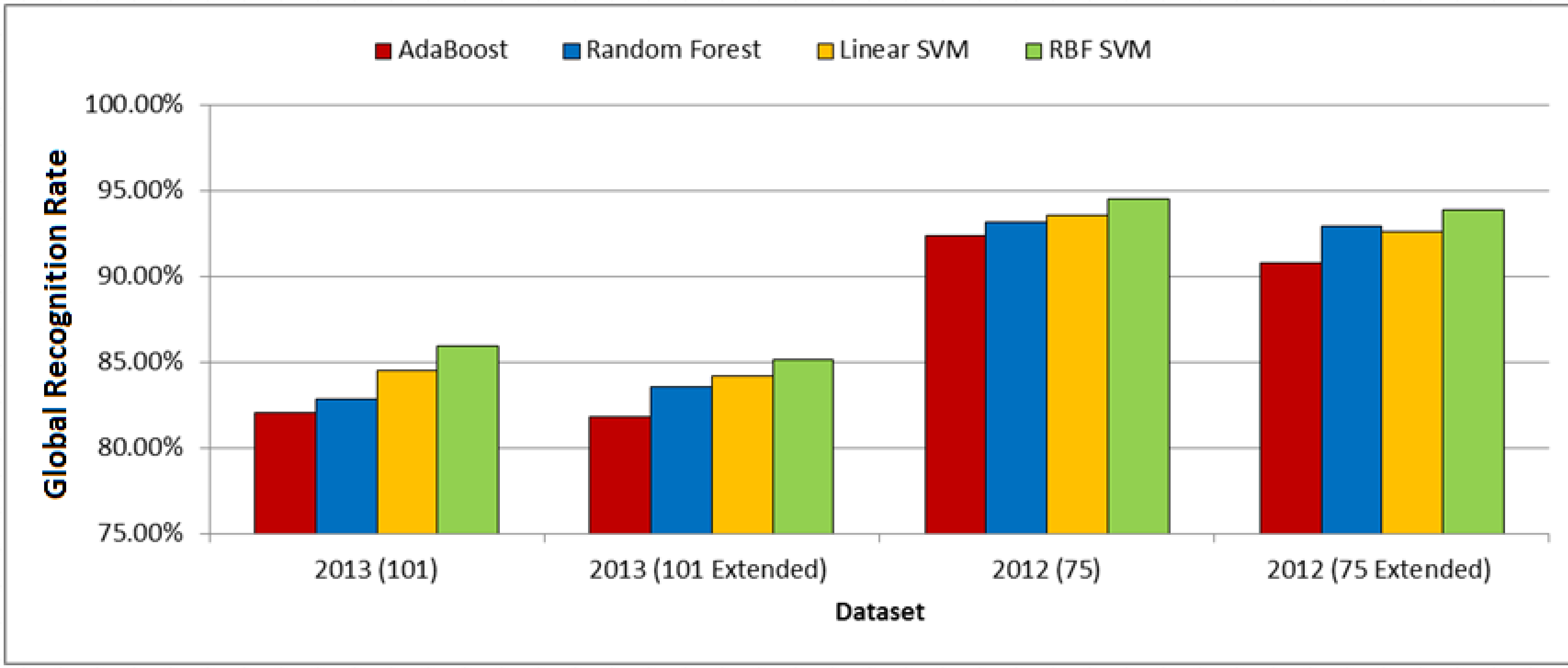


[1] B. Q. Huang, Y. Zhang, and M.-T. Kechadi, “Preprocessing techniques for online handwriting recognition,” in Intelligent Text Categorization and Clustering. 2009

Dataset information

	CROHME 2013	CROHME 2013 B	MathBrush
Classes	101	75	100
Folds	No	No	Yes
Training	68,598	65,544	22,305
Testing	6,082	6,336	2,531
Extended	451,637	291,292	-

Comparison of Learning Algorithms Global Recognition Rate



Using CROHME 2012 and 2013 datasets

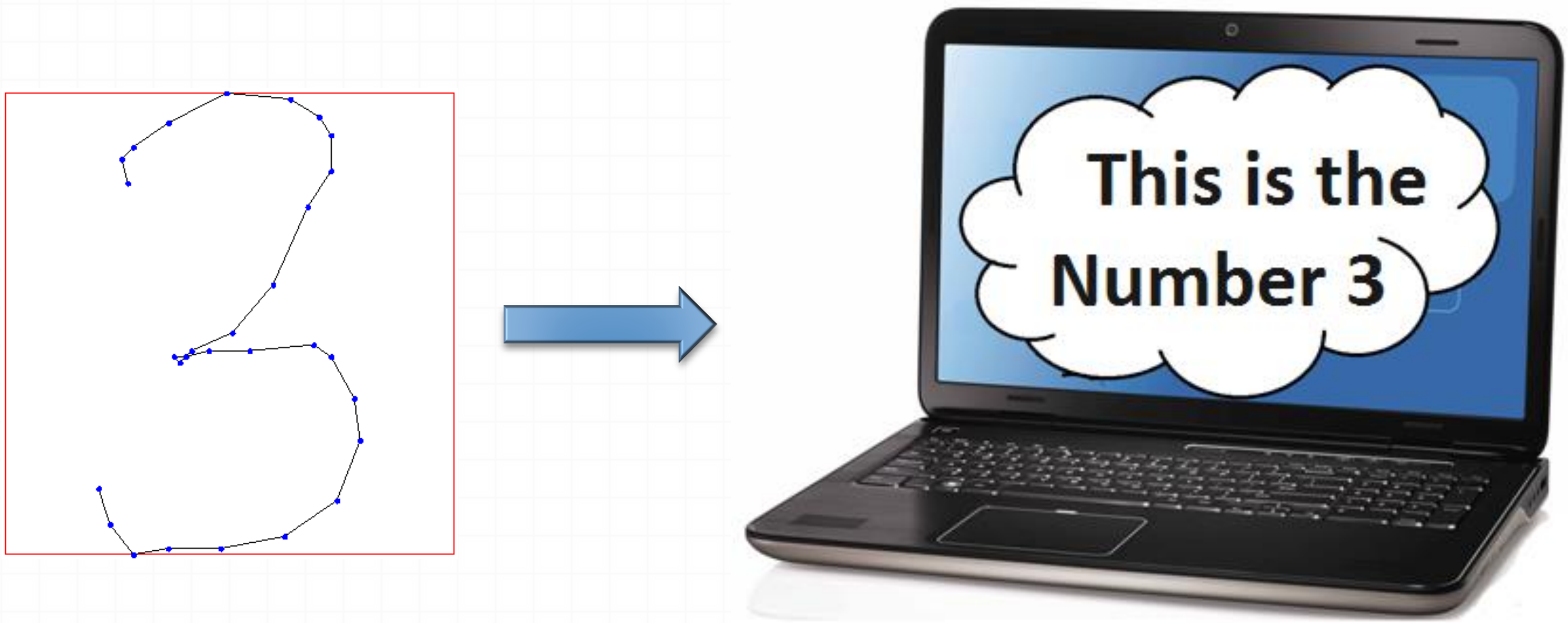
Ambiguous classes (CROHME 2013) #1

<table>
<tr><th>Class</th><th>Size</th><th>Similar</th></tr>
<tr><td>x</td><td>4,115</td><td>X, times</td></tr>
<tr><td>X</td><td>223</td><td>x, times</td></tr>
<tr><td>times</td><td>477</td><td>x, X</td></tr>
<tr><td>1</td><td>5,026</td><td>|,(,), comma</td></tr>
<tr><td>|</td><td>358</td><td>1, (,), prime</td></tr>
<tr><td>(</td><td>3,191</td><td>1,|</td></tr>
<tr><td>)</td><td>3,185</td><td>1,|, Comma</td></tr>
<tr><td>Comma</td><td>498</td><td>1,), prime, /</td></tr>
<tr><td>prime</td><td>51</td><td>|, Comma</td></tr>
<tr><td>/</td><td>157</td><td>Comma, Prime</td></tr>
</table>

Ambiguous classes (CROHME 2013) #2

Class	Size	Similar
c	754	C
C	206	c
p	453	P
P	85	p
v	230	V
V	85	v
s	191	S
S	94	s
q	208	9
9	583	q
O	90	0
0	1,438	o

The Problem



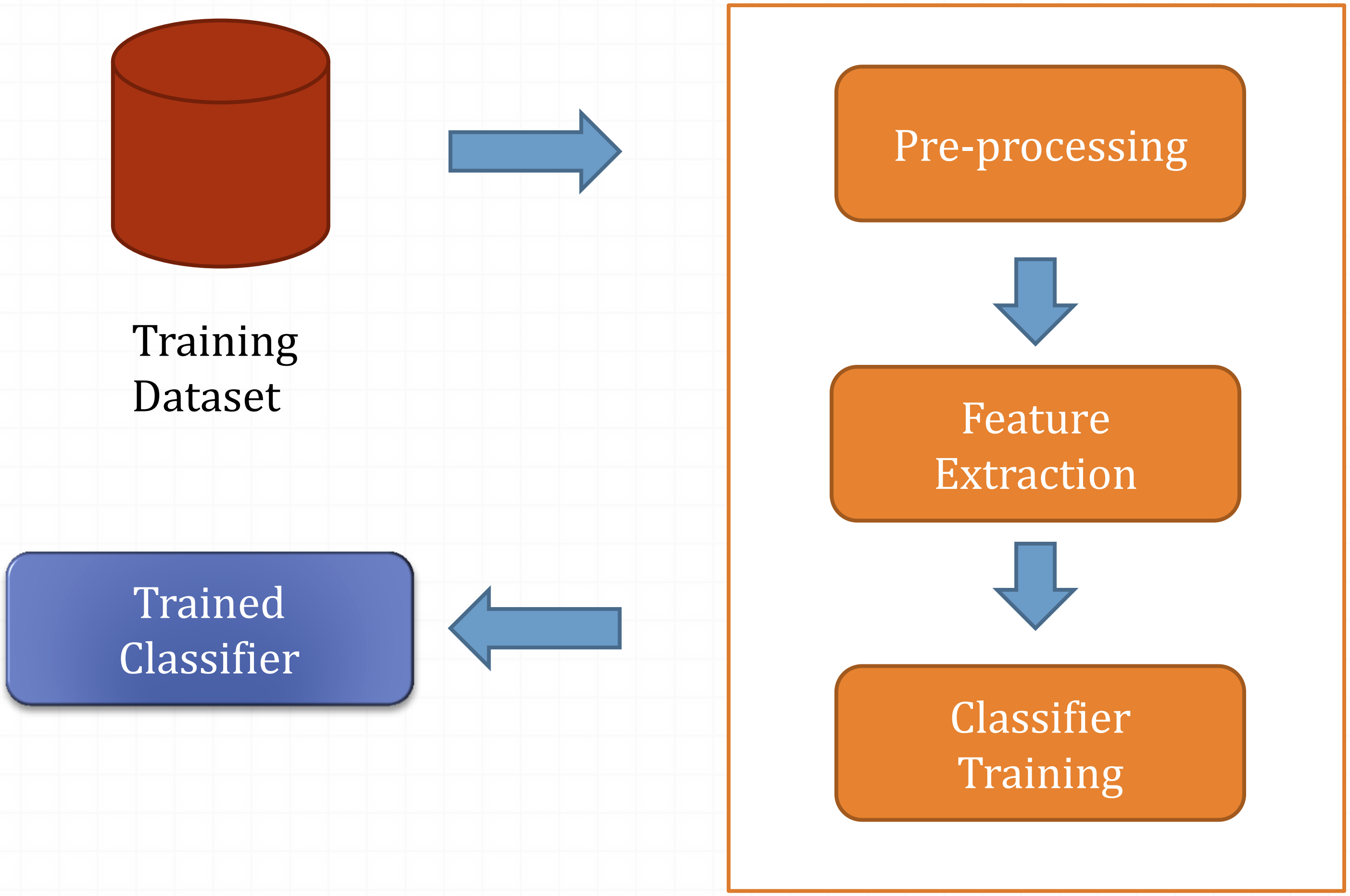
The Symbol Recognition Problem

Related Work

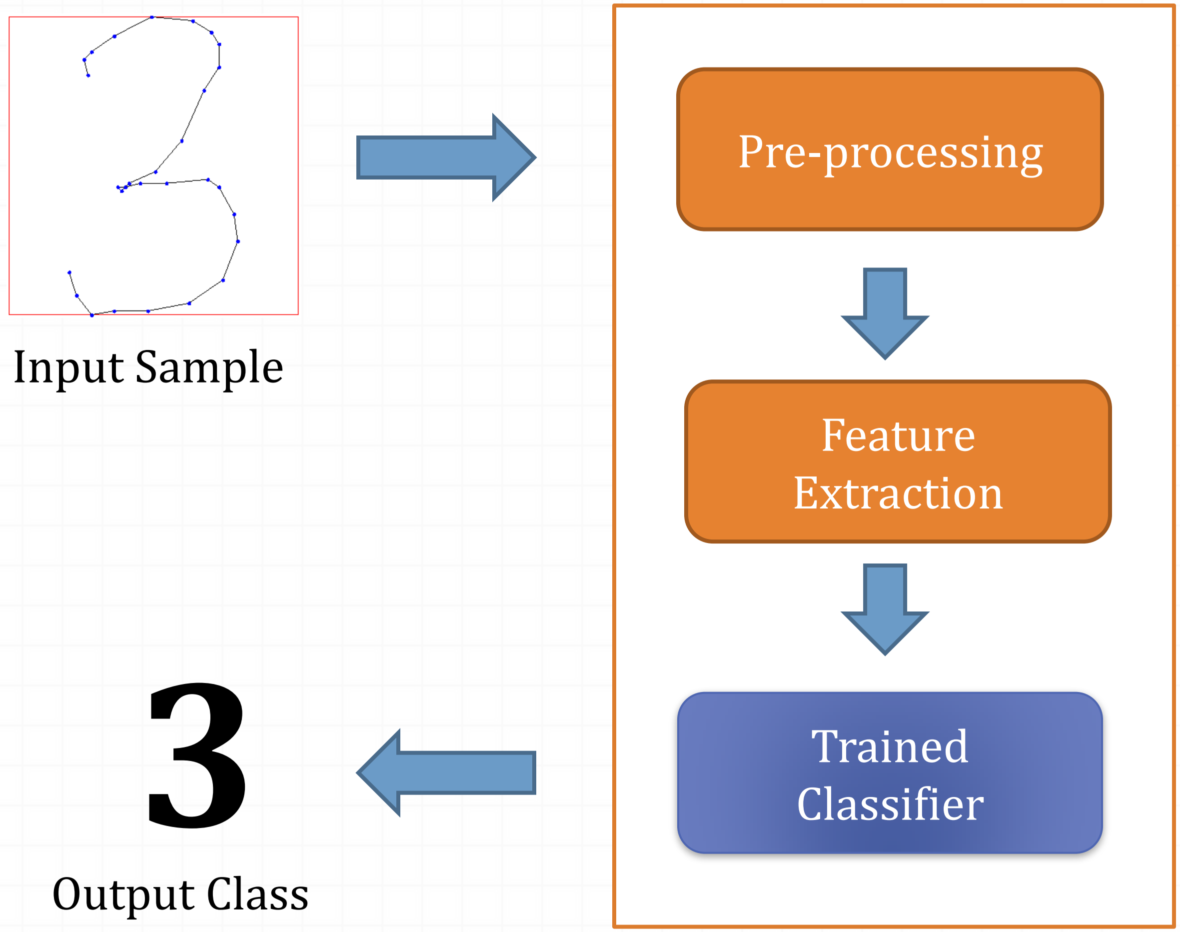
- Math Symbol Recognition

- Data Generation

System training



System recognition



Math Symbol Recognition Benchmark - CROHME 2014

Method	Classifier	Without Junk	With Junk
MyScript	MLP	91.04%	**85.54%**
Alvaro et al.	BLSTM-RNN	91.24%	84.14%
Proposed Method	SVM RBF Kernel	88.66%	83.61%