

# The art of breaking and designing captchas

**Elie Bursztein**



Session ID: HT02-402

Session Classification: xxxxxxxxxxxxxx

**RSACONFERENCE2012**

**Elie Bursztein**[View my profile page](#)**1,458**

TWEETS

**34**

FOLLOWING

**736**

FOLLOWERS

Compose new Tweet...

Who to follow - [Refresh](#) - [View all](#)**Christopher Sogholian** @csogholian

Followed by ashkan soltani and ot...

[Follow](#)**Dan Kaminsky** @dakami

Followed by Jeremiah Grossman a...

[Follow](#)Worldwide trends - [Change](#)[#SincelmBeingHonest](#)[#SoyDeLaGeneracion](#)[#ThingsThatILoveInLife](#)[Gina Lopez](#)[Mike Hussey](#)[Melón y Melames](#)[Visionz Of Home](#)[Despicable Me 2](#)[Dr. Sheldon Cooper](#)

twitter

© 2012 Twitter [About](#) [Help](#) [Terms](#) [Privacy](#)[Blog](#) [Status](#) [Apps](#) [Resources](#) [Jobs](#)[Advertisers](#) [Businesses](#) [Media](#) [Developers](#)

## Tweets

**Ross** @Hypn

26m

[@singe](#) [@lactichaze](#) and let's not forget the timthumb.php vuln that allowed code execution - all due to themes, nm WP's core code or devs.[←](#) [In reply to Dominic White](#)**Adam Baldwin** @adam\_baldwin

40m

Poll: If a site had all its user accounts, salts &amp; sha1 hashes downloaded would you consider that a breach? Would you require passwd resets?

**Leif Ryge** @wiretapped

2h

[#RIP](#) credibility of [@linode](#) — they still haven't said anything on their blog or twitter about being compromised. [#fail](#) [bitcoinmedia.com/compromised-li...](#)[↻](#) Retweeted by [Dominic White](#)**Marie Hattar** @MarieHattar

1h

New device may be useful for those long conferences... [bit.ly/Ay2AJ](#)**Adam Baldwin** @adam\_baldwin

2h

Every time I'm cracking passwords I'm sort of disappointed somebody didn't use the password zerocool.

**Celine Bursztein** @cealtea

4h

Photo: We finally have keys of our new home :) (Taken with Instagram at Maplewood Apartments) [tumblr.co/ZXk-ixHJnflm](#)**Jon Mitchell** @JonMwords

6h

Invalid Argument Ep. 1 - The App Store: (Banana) Republic? [rww.to/A4QZko](#) Guests: [@calebelston](#) [@drbarnard](#) [@mb](#) [@brianhax](#) [@macguitar](#)[↻](#) Retweeted by [ReadWriteWeb](#)**ReadWriteWeb** @RWW

5h

Daily Wrap: Siri's Imagined Physical Presence and more [rww.to/xvNPJK](#)**Crystal Beasley** @skinny

5h

I knew this sort of person existed, I just didn't ever fathom I'd become one of them. [pic.twitter.com/fX1TuTaM](#)[↻](#) Retweeted by [Priscilla Scala](#)

**Elie Bursztein**[View my profile page](#)**1,458**

TWEETS

**34**

FOLLOWING

**736**

FOLLOWERS

Compose new Tweet...

Who to follow - [Refresh](#) - [View all](#)**Christopher Sogholian** @csogholian

Followed by ashkan soltani and ot...

[Follow](#)**Dan Kaminsky** @dakami

Followed by Jeremiah Grossman a...

[Follow](#)Worldwide trends - [Change](#)[#SincelmBeingHonest](#)[#SoyDeLaGeneracion](#)[#ThingsThatILoveInLife](#)[Gina Lopez](#)[Mike Hussey](#)[Melón y Melames](#)[Visionz Of Home](#)[Despicable Me 2](#)[Dr. Sheldon Cooper](#)

twitter

© 2012 Twitter [About](#) [Help](#) [Terms](#) [Privacy](#)[Blog](#) [Status](#) [Apps](#) [Resources](#) [Jobs](#)[Advertisers](#) [Businesses](#) [Media](#) [Developers](#)

## Tweets

**Ross** @Hypn

26m

[@singe](#) [@lactichaze](#) and let's not forget the timthumb.php vuln that allowed code execution - all due to themes, nm WP's core code or devs.[←](#) [In reply to Dominic White](#)**Adam Baldwin** @adam\_baldwin

40m

Poll: If a site had all its user accounts, salts &amp; sha1 hashes downloaded would you consider that a breach? Would you require passwd resets?

**Leif Ryge** @wiretapped

2h

[#RIP](#) credibility of [@linode](#) — they still haven't said anything on their blog or twitter about being compromised. [#fail](#) [bitcoinmedia.com/compromised-il...](#)[↻](#) Retweeted by [Dominic White](#)**Marie Hattar** @MarieHattar

1h

New device may be useful for those long conferences... [bit.ly/Ay2AJ](#)**Adam Baldwin** @adam\_baldwin

2h

Every time I'm cracking passwords I'm sort of disappointed somebody didn't use the password zerocool.

**Celine Bursztein** @cealtea

4h

Photo: We finally have keys of our new home :) (Taken with Instagram at Maplewood Apartments) [tumblr.co/ZXk-ixHJnfLm](#)**Jon Mitchell** @JonMwords

6h

Invalid Argument Ep. 1 - The App Store: (Banana) Republic? [rww.to/A4QZko](#) Guests: [@calebelston](#) [@drbarnard](#) [@mb](#) [@brianhax](#) [@macguitar](#)[↻](#) Retweeted by [ReadWriteWeb](#)**ReadWriteWeb** @RWW

5h

Daily Wrap: Siri's Imagined Physical Presence and more [rww.to/xvNPJK](#)**Crystal Beasley** @skinny

5h

I knew this sort of person existed, I just didn't ever fathom I'd become one of them. [pic.twitter.com/fX1TuTaM](#)[↻](#) Retweeted by [Priscilla Scala](#)

# Twitter Follower Packages

Please Select One Of Our Targeted Follower Pages



## Silver Package

- ✓ **1000 Targeted Followers**
- ✓ Guaranteed REAL, Targeted People Interested In Your Business
- ✓ Added To Your Page Within 25 days
- ✓ Targeted To Your Business/Niche
- ✓ Select The Country/s Where You Want Your Followers From
- ✓ No Automatic Bots/Programs To Get Followers, We Proudly Target 100% Of Your Followers Manually

**\$49.99**

[Order Now >](#)



## Gold Package

- ✓ **5000 Targeted Followers**
- ✓ Guaranteed REAL, Targeted People Interested In Your Business
- ✓ Added To Your Page Within 40 Days
- ✓ Targeted To Your Business/Niche
- ✓ Select The Country/s Where You Want Your Followers From
- ✓ No Automatic Bots/Programs To Get Followers, We Proudly Target 100% Of Your Followers Manually

**\$139.99**



## Join the Conversation

Already on Twitter? [Sign in.](#)

Already use Twitter on your phone? [Finish signup now.](#)

Full name  ✔ ok

Your full name will appear on your public profile

Username  ✔ ok

Your public profile: [http://twitter.com/ eliedemo](http://twitter.com/eliedemo)

Password  ✔ Good

### Are you human?



Before we create your account, we need to make sure you're not a computer.

edisibil from

Type the words  
above

Can't read this?

[↻ Get two new words](#)

[🔊 Hear a set of words](#)

Powered by reCAPTCHA.

[Help](#)

Finish

Create my account

I want the inside scoop—please send me email updates!



Captcha Validation: \*

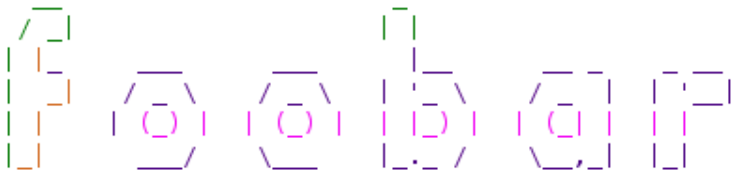


Captcha Validation: \*

Защита от автоматической регистрации

$$\lim_{x \rightarrow 0} \ln \left( 2 + \sqrt{\arctg x \cdot \sin \frac{1}{x}} \right)$$

Введите ответ



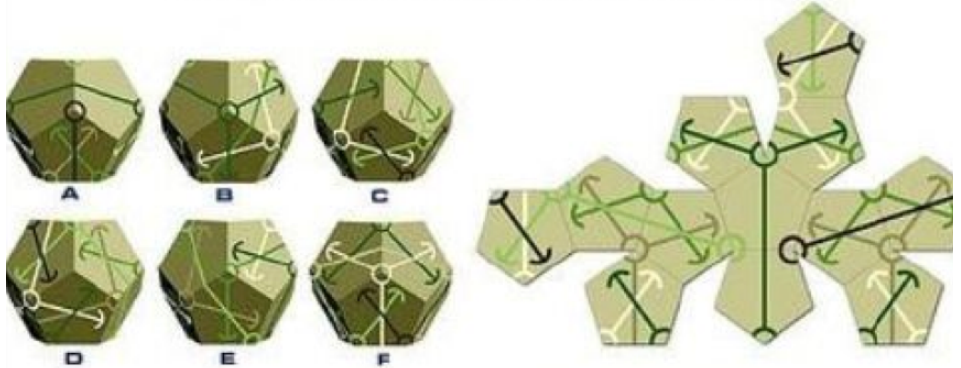
Captcha Validation: \*

Защита от автоматической регистрации

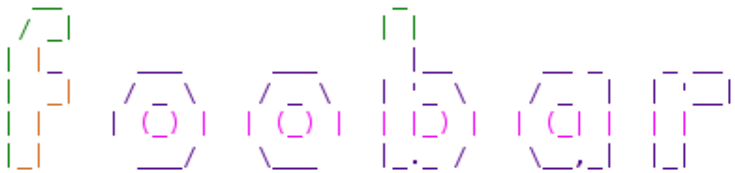
$$\lim_{x \rightarrow 0} \ln \left( 2 + \sqrt{\arctg x \cdot \sin \frac{1}{x}} \right)$$

Введите ответ

No premium user. Please enter the one that can NOT be created from the unfolded pattern. 29 seconds remain.



Download via Cogent #2



Captcha Validation: \*

Защита от автоматической регистрации

Введите ответ

$$\lim_{x \rightarrow 0} \ln \left( 2 + \sqrt{\arctg x \cdot \sin \frac{1}{x}} \right)$$


Please click on the images that show cats:



No premium user. Please enter the one that can NOT be created from the unfolded pattern. 29 seconds remain.





Captcha Validation: \*

Защита от автоматической регистрации

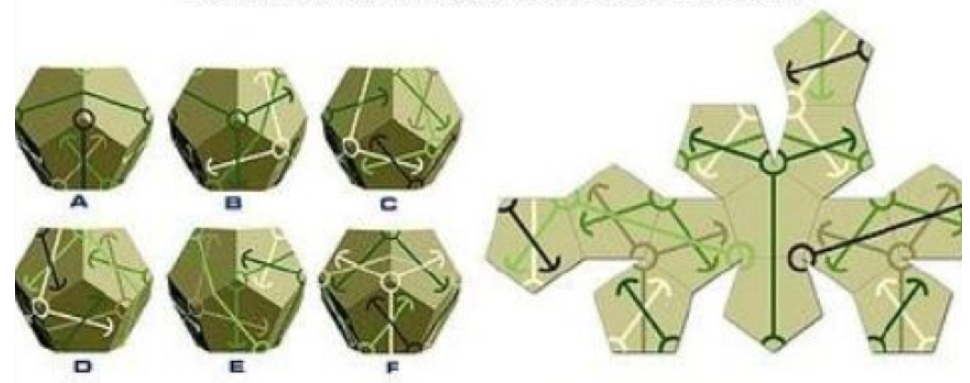
$$\lim_{x \rightarrow 0} \ln \left( 2 + \sqrt{\arctg x \cdot \sin \frac{1}{x}} \right)$$

Введите ответ

Please click on the images that show cats:



No premium user. Please enter the one that can NOT be created from the unfolded pattern. 29 seconds remain.



Download via Cogent #2

	 meet me	 meet me
	 meet me	
	 meet me	
<input type="button" value="switch to men"/>		

hotcaptcha by frozenbear

# World Most-Popular Captchas



ZKW4

[Megaupload]



[Reddit]



944 531

[eBay]



[CNN]

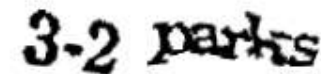


BAE3

[Baidu]



[Authorize]

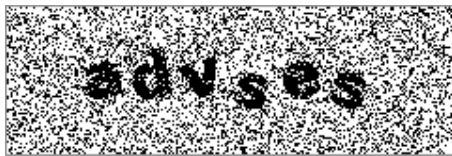


3-2 parks

[Recaptcha]



[Skyrock]



[Captcha.net]



[NIH]

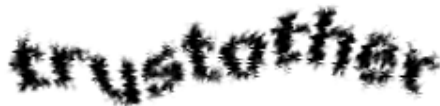


[Digg]



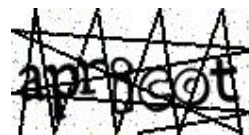
skytngomi

[Google]

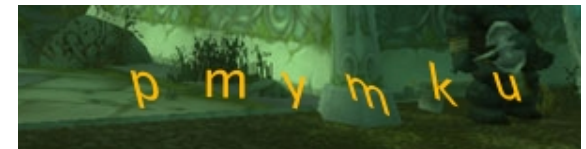


trustother

[Wikipedia]



[Slashdot]



[Blizzard]



# World Most-Popular Captchas



ZKW 4

[Megaupload]



[Reddit]



9A 51

[eBay]

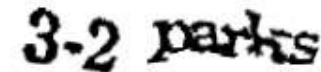


[CNN]



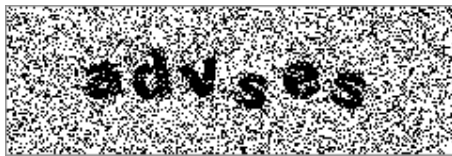
BAE3

[Baidu]



3-2 parks

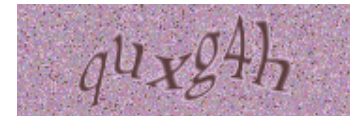
[Recaptcha]



[Captcha.net]



[Authorize]



[Skyrock]



[NIH]



[Digg]



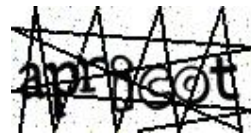
skyngomi

[Google]



trustother

[Wikipedia]



[Slashdot]



[Blizzard]

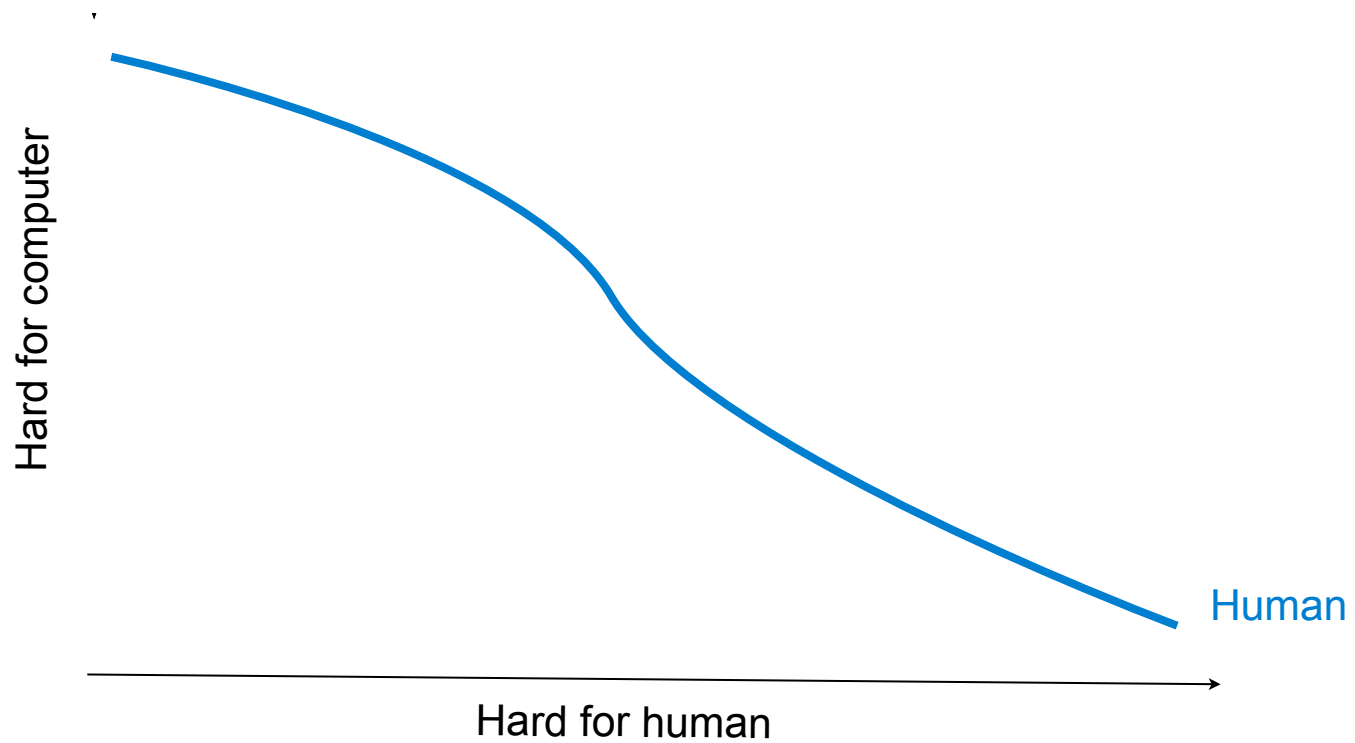
# Captcha Design Goal

Hard for computer

Hard for human

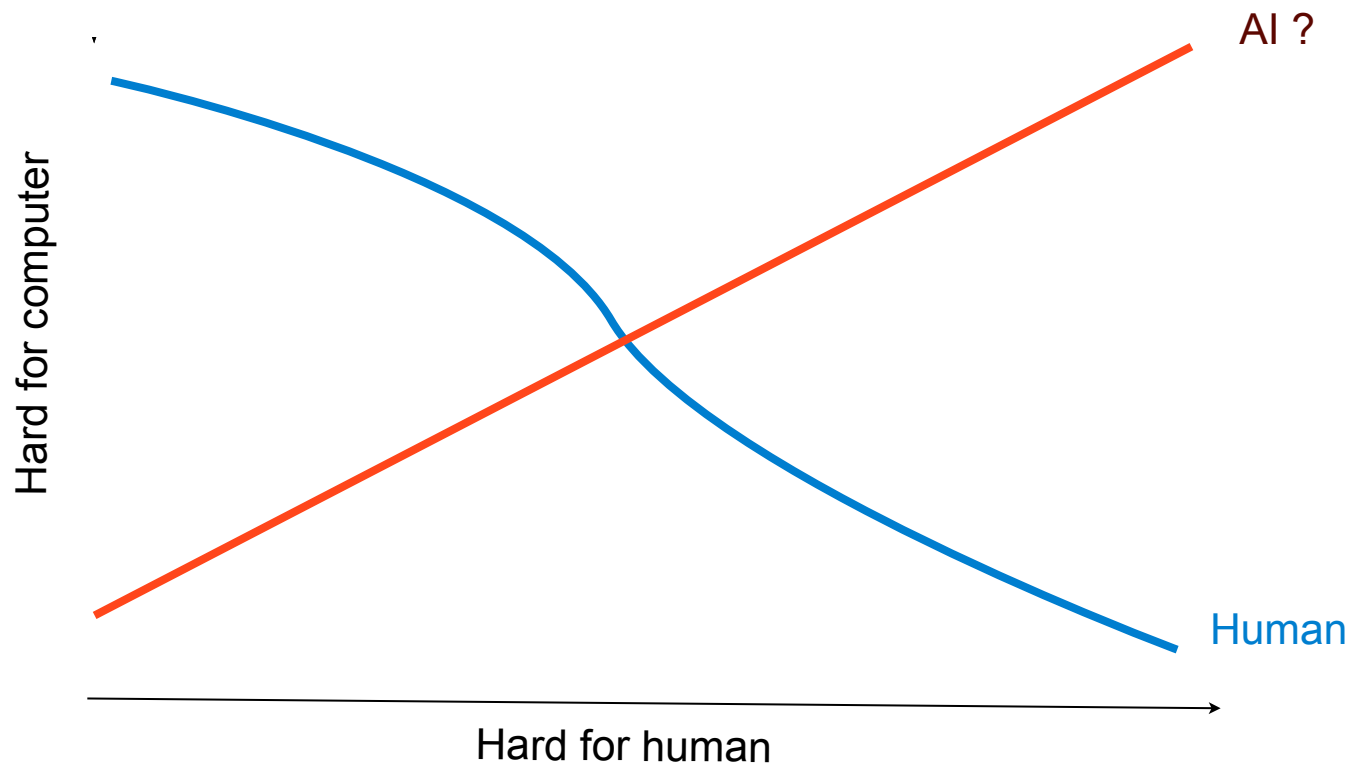


# Captcha Design Goal

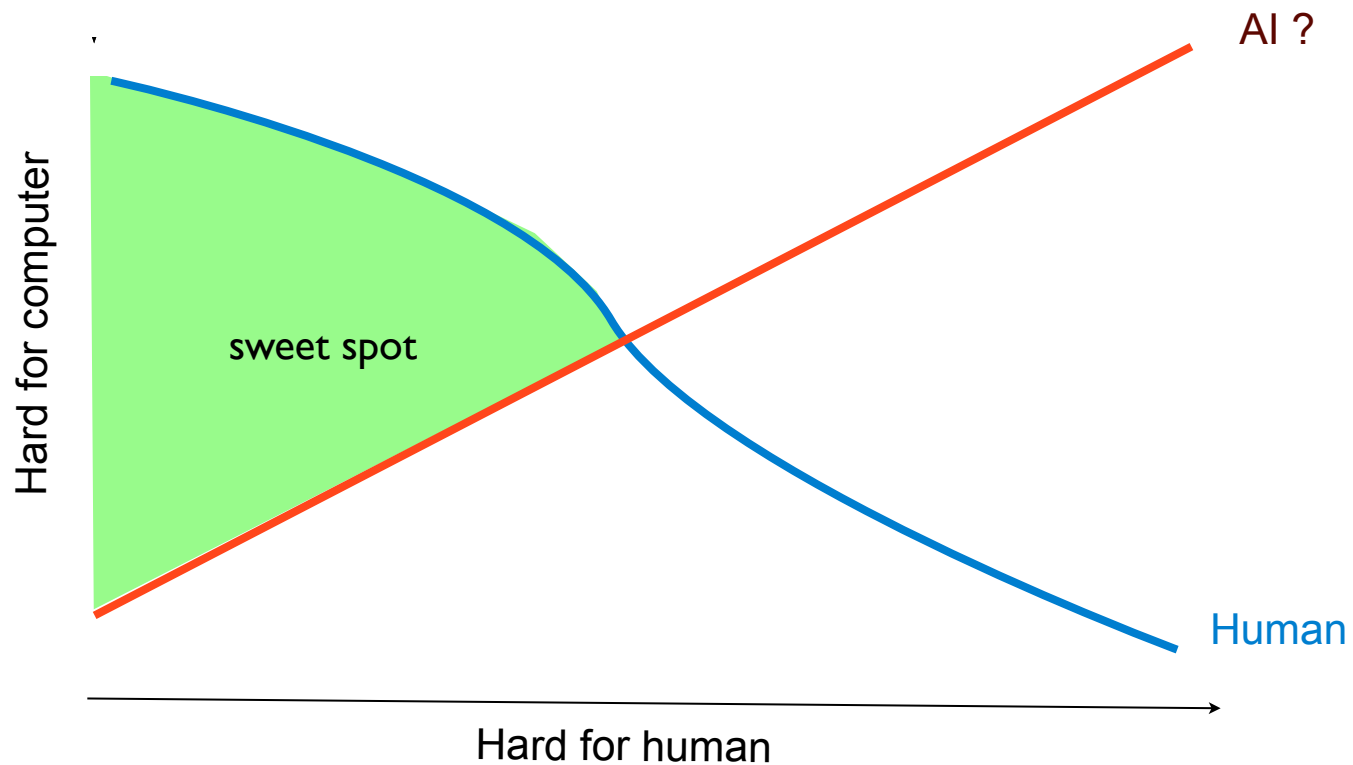




# Captcha Design Goal



# Captcha Design Goal





# How to **break** and **design** CAPTCHAs



Based on the **breaking 21** of the most **popular schemes** and **designing** the new Wikipedia captcha

# Outline





# Outline

- How to break text captcha



# Outline

- How to break text captcha
- How to make captchas easier for human



# Outline

- How to break text captcha
- How to make captchas easier for human
- How to break audio captcha



# Outline

- How to break text captcha
- How to make captchas easier for human
- How to break audio captcha
- How to break video captcha



# Evaluation metrics



Accuracy





# Evaluation metrics



Accuracy



Solving time



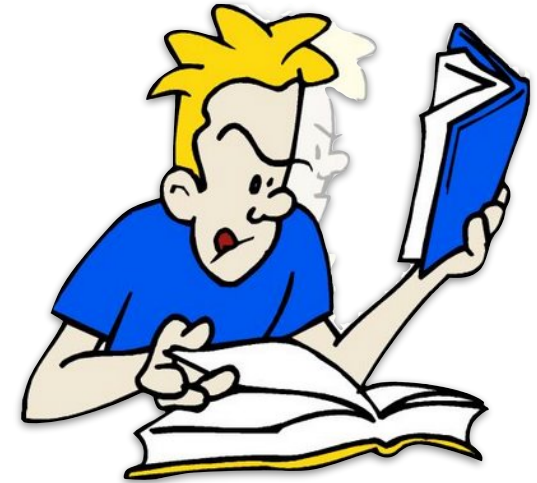
# Evaluation metrics



Accuracy



Solving time

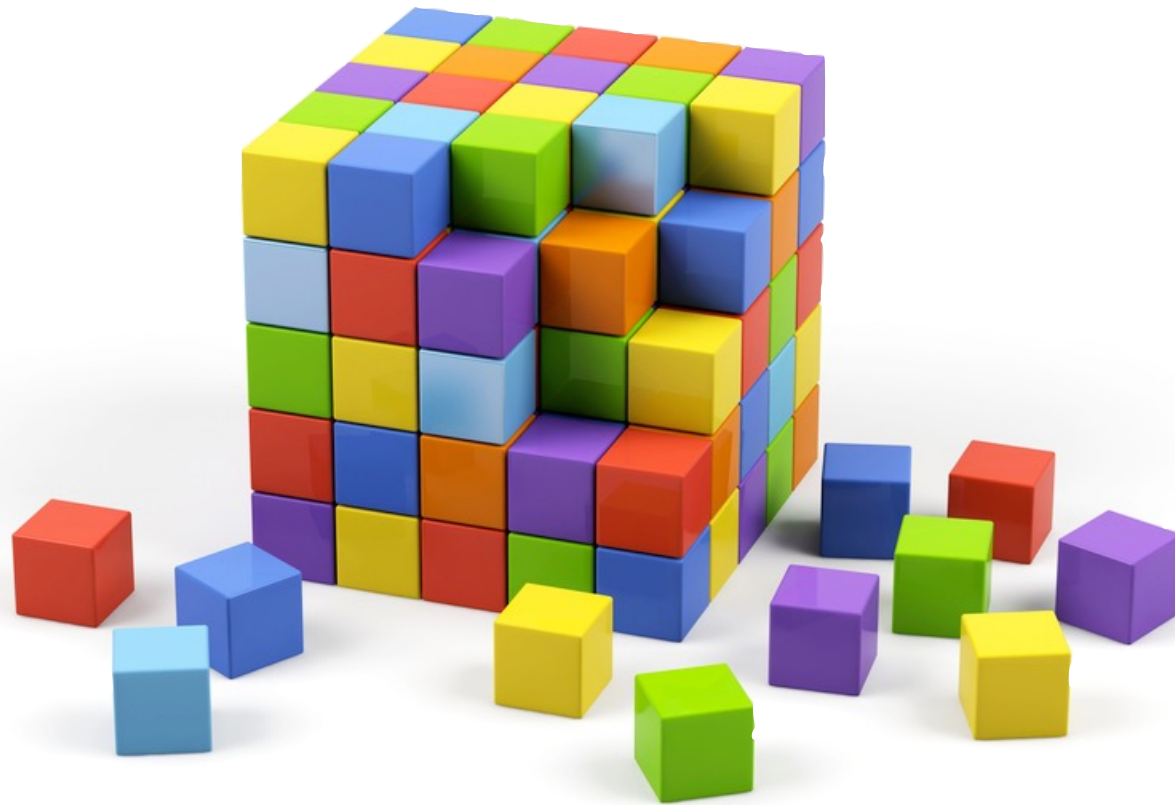


Learnability



# How to Break Text-Captchas





Think Lego



How to break a captcha: **example**







Pre-processing: **background removal**





Pre-processing: background removal

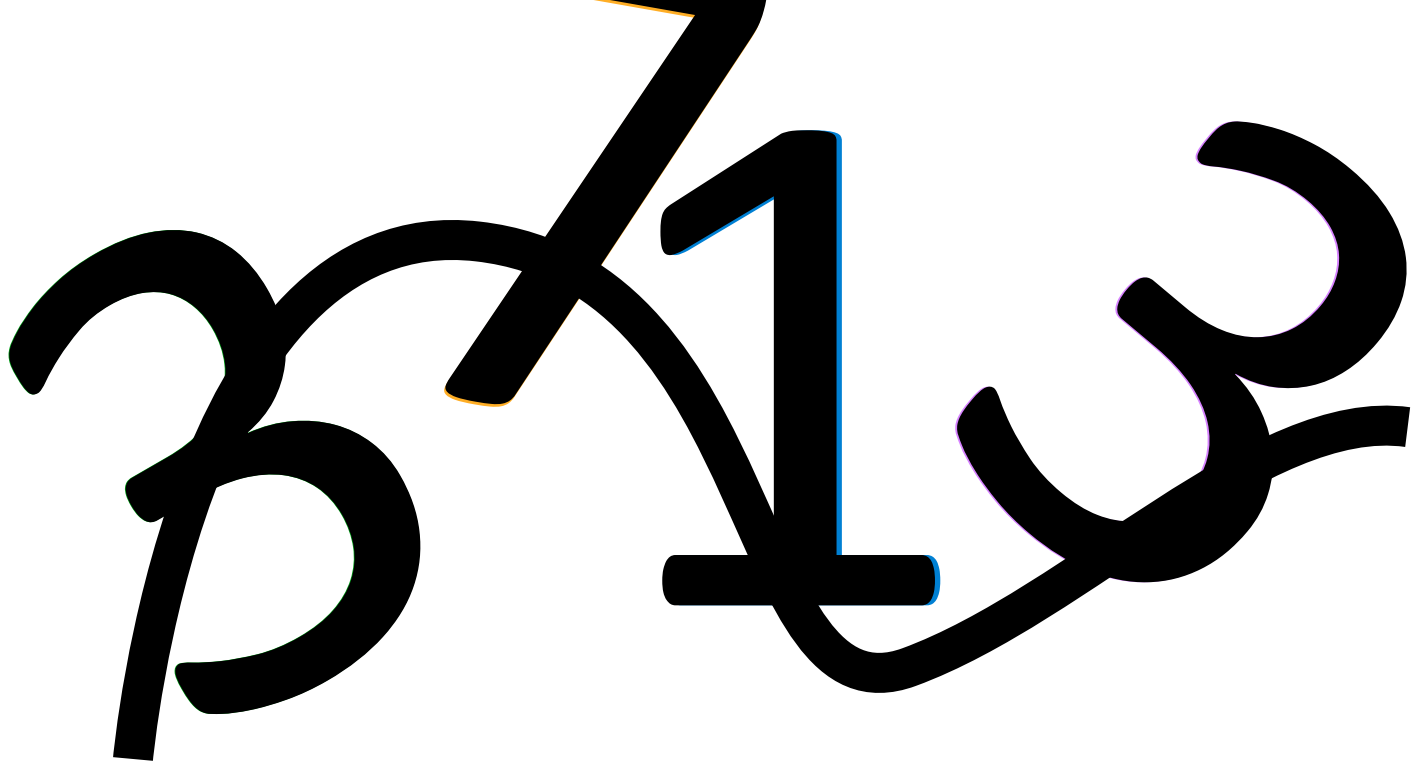




Pre-processing: **captcha binarization**

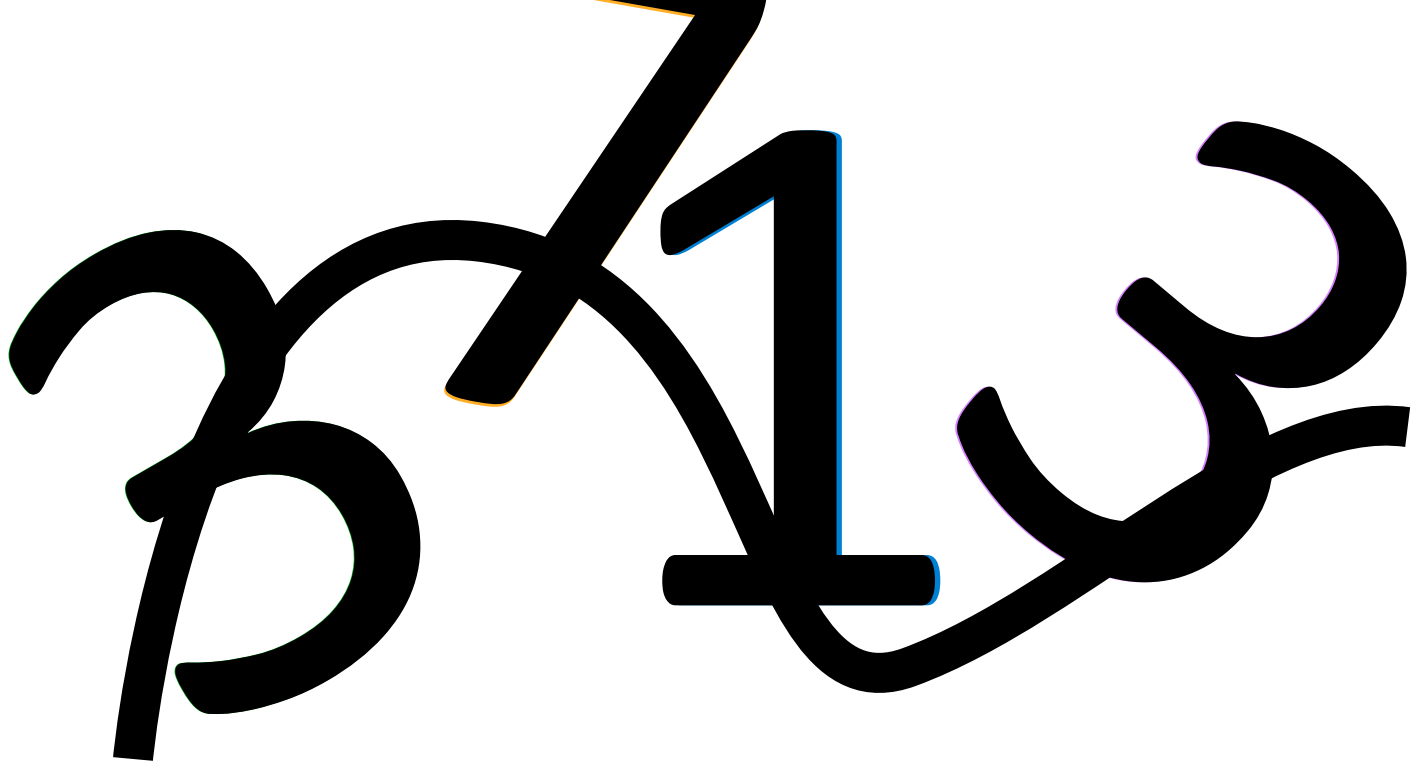






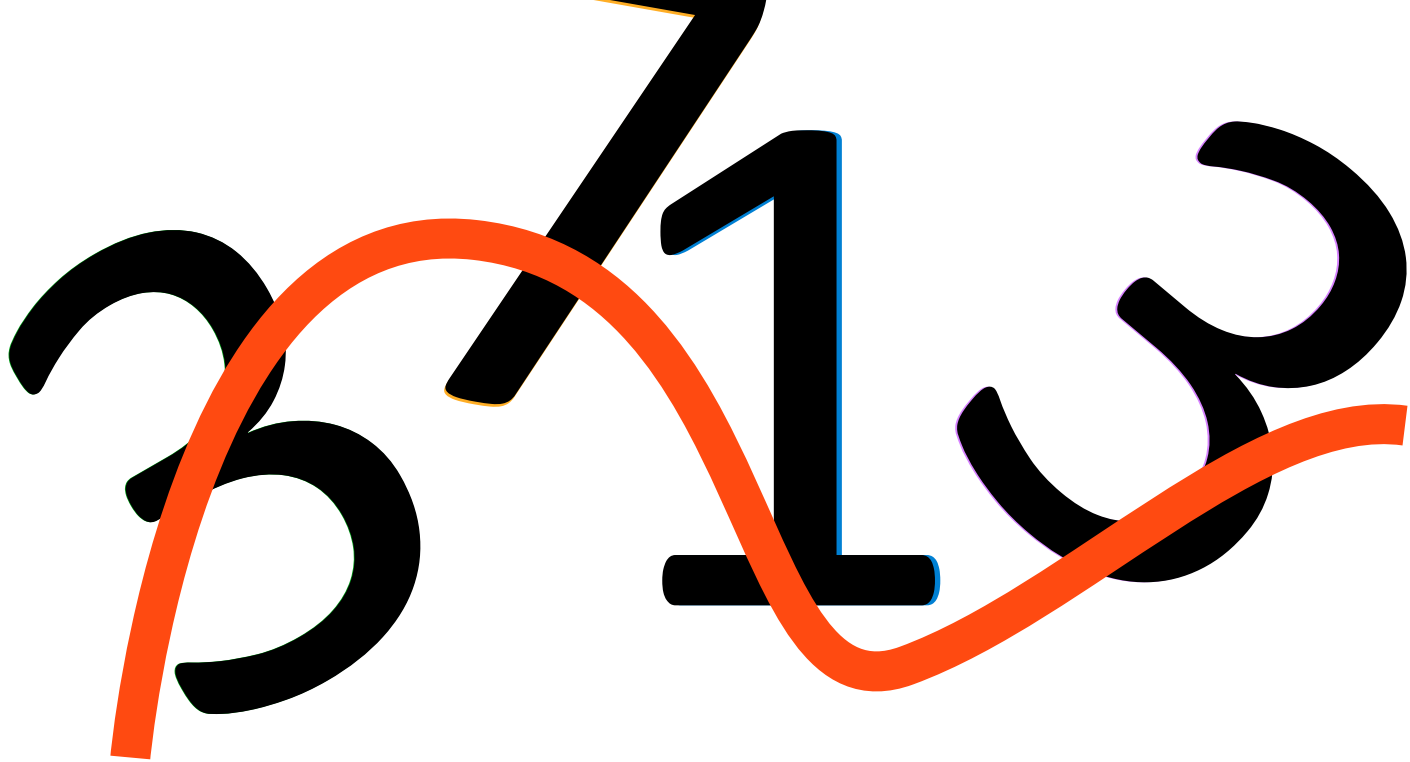
Pre-processing: **captcha binarization**





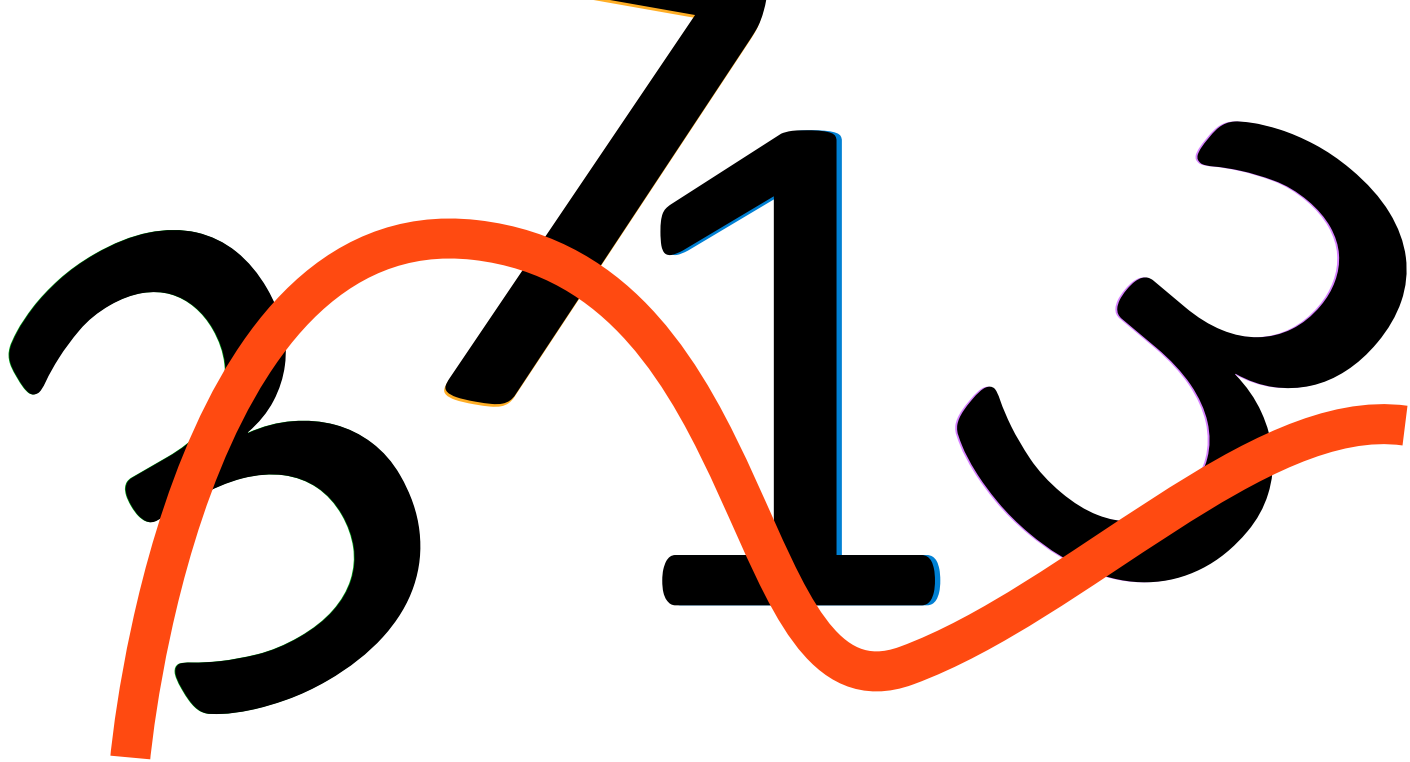
Pre-processing: **Line detection**





Pre-processing: **Line detection**





Pre-processing: Line removal



3 1 3

Pre-processing: Line removal



3 1 3

Segmentation: clustering algorithm





Segmentation: clustering algorithm





Segmentation: cluster separation







Segmentation: cluster separation





Post-segmentation: **inverting rotation**



3 7 1 3

Post-segmentation: **inverting rotation**



3

7

1

3

Recognition:



Recognition: 3 7 1 3

# Breaker 5 Stages Pipeline

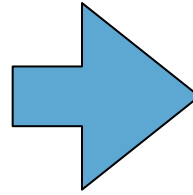


Slashdot captcha



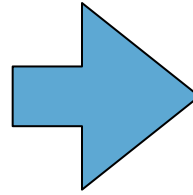
# Breaker 5 Stages Pipeline

Preprocessing



# Breaker 5 Stages Pipeline

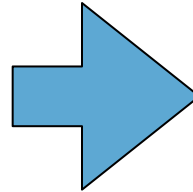
Preprocessing



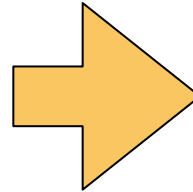


# Breaker 5 Stages Pipeline

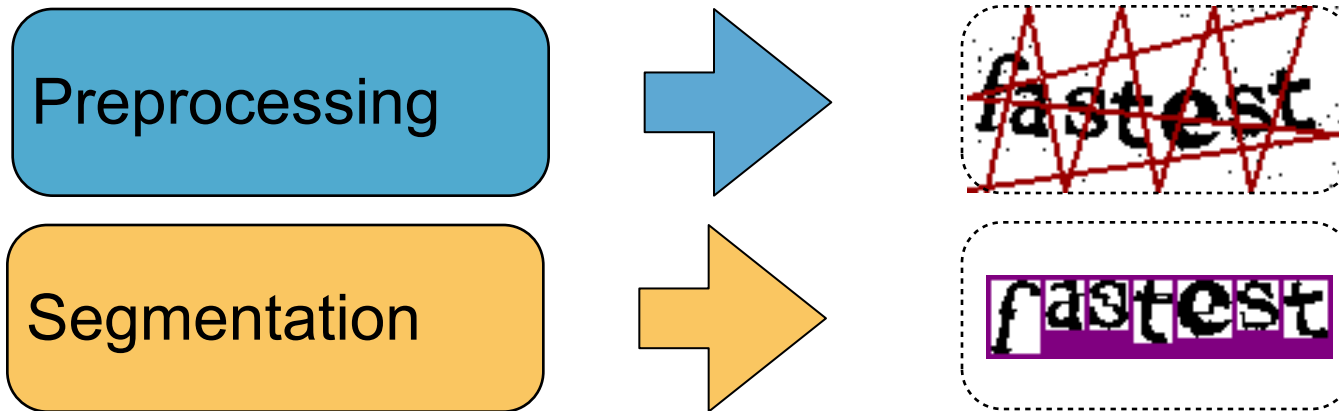
Preprocessing



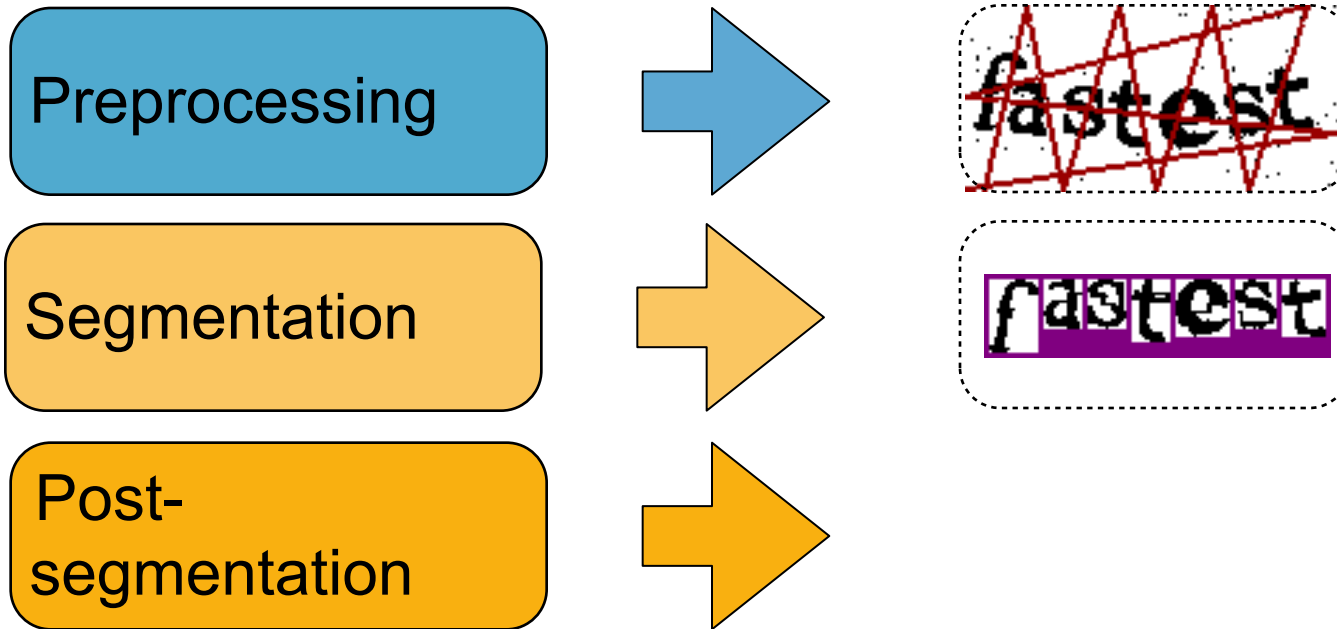
Segmentation



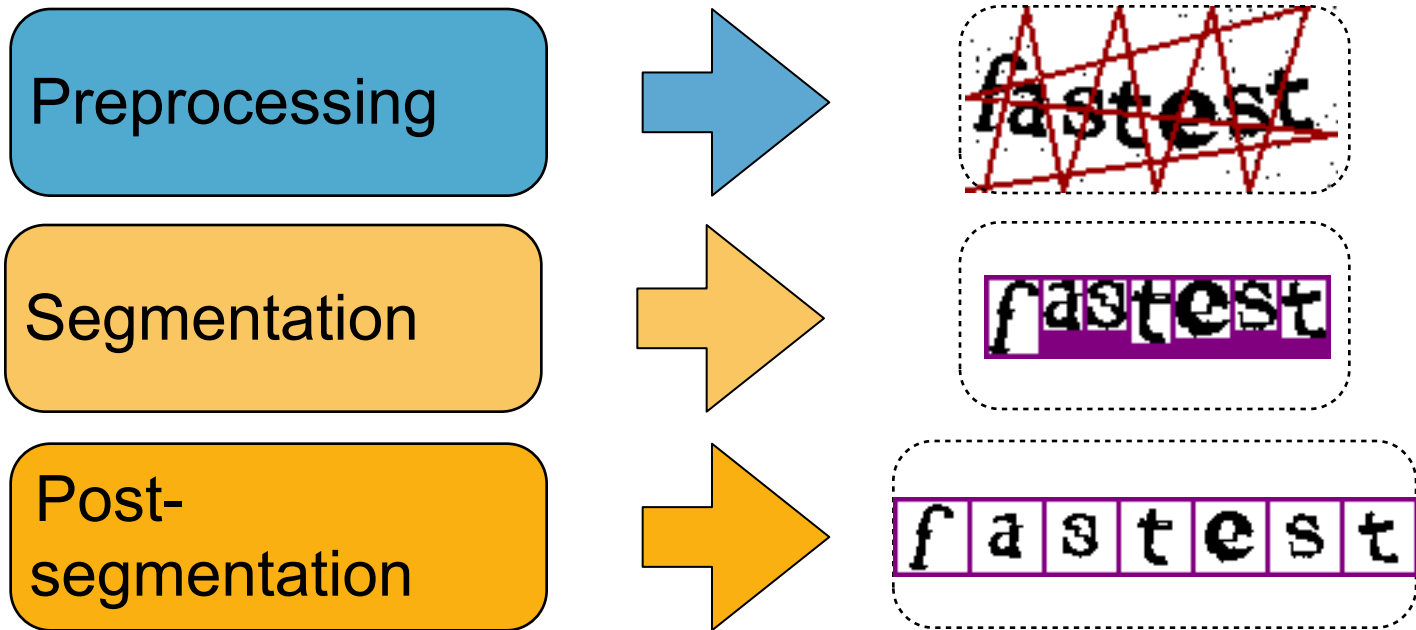
# Breaker 5 Stages Pipeline



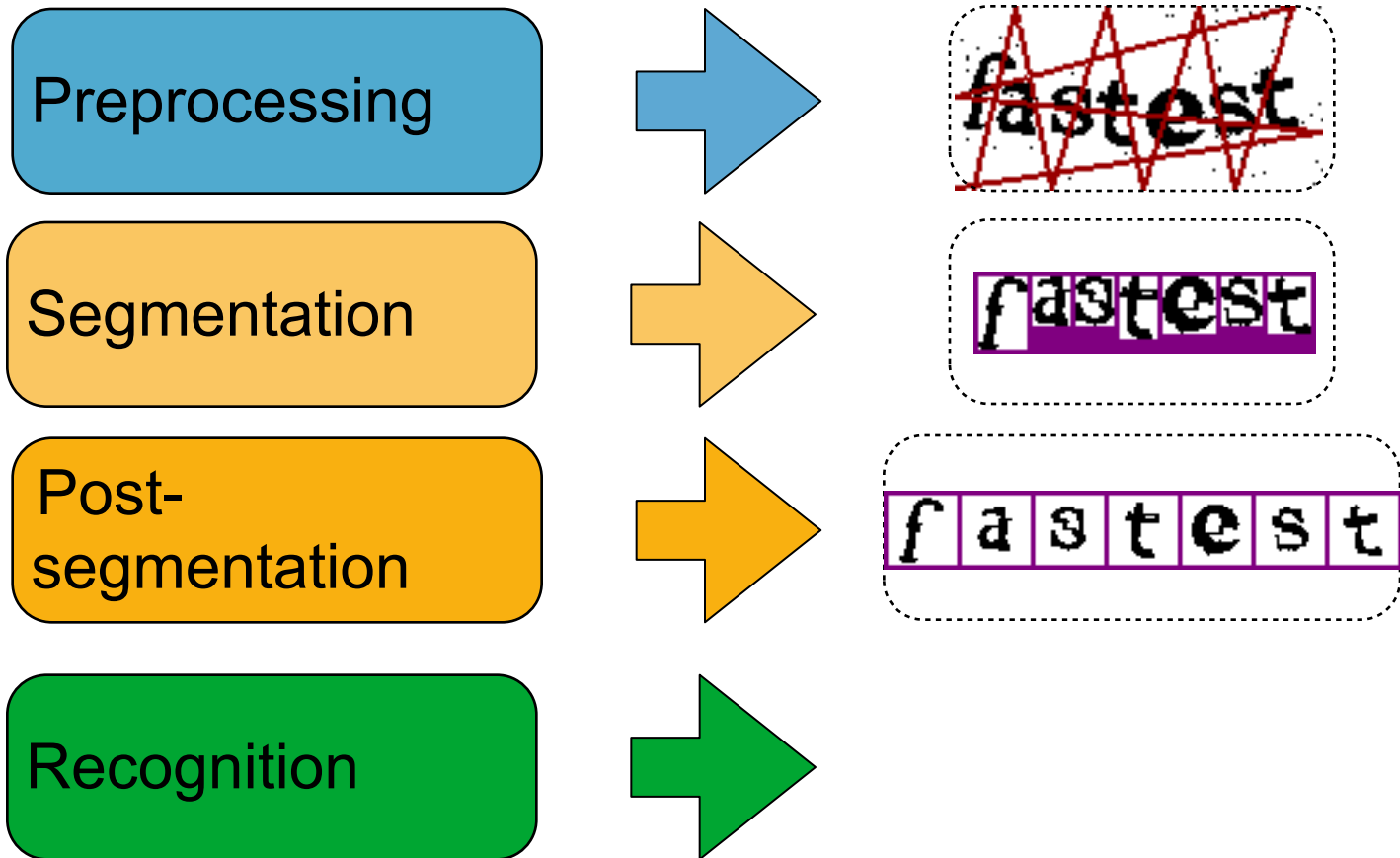
# Breaker 5 Stages Pipeline



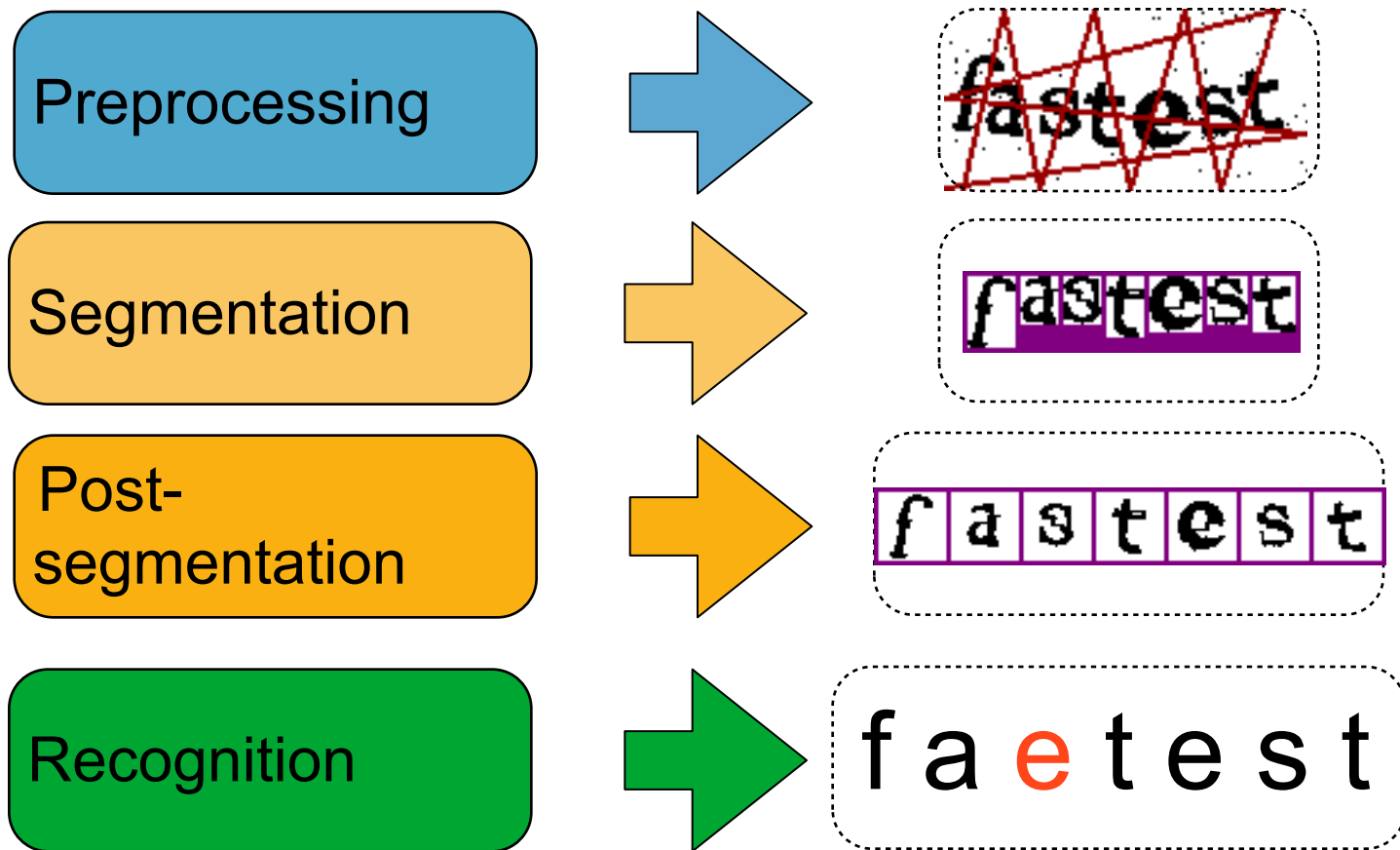
# Breaker 5 Stages Pipeline



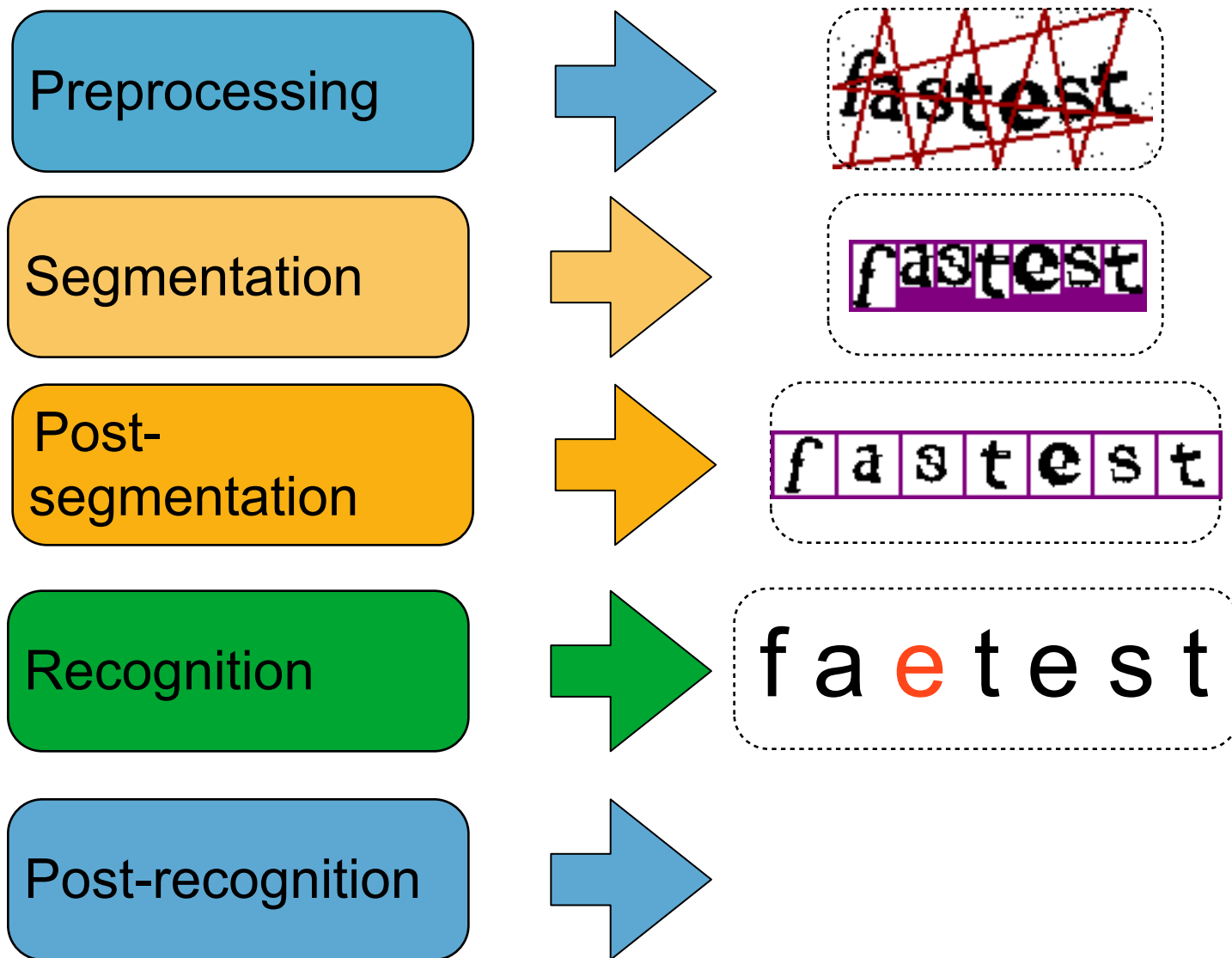
# Breaker 5 Stages Pipeline



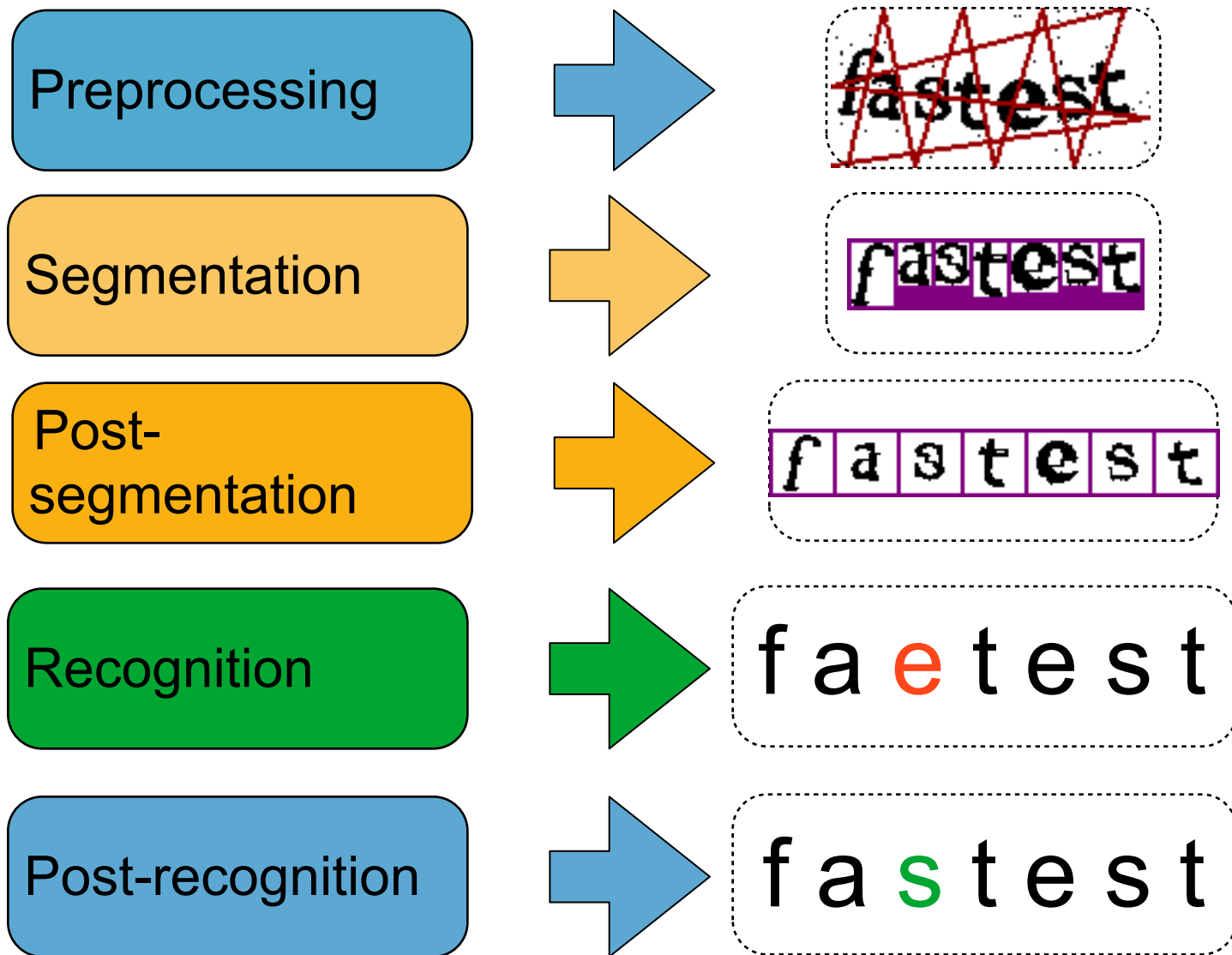
# Breaker 5 Stages Pipeline



# Breaker 5 Stages Pipeline



# Breaker 5 Stages Pipeline





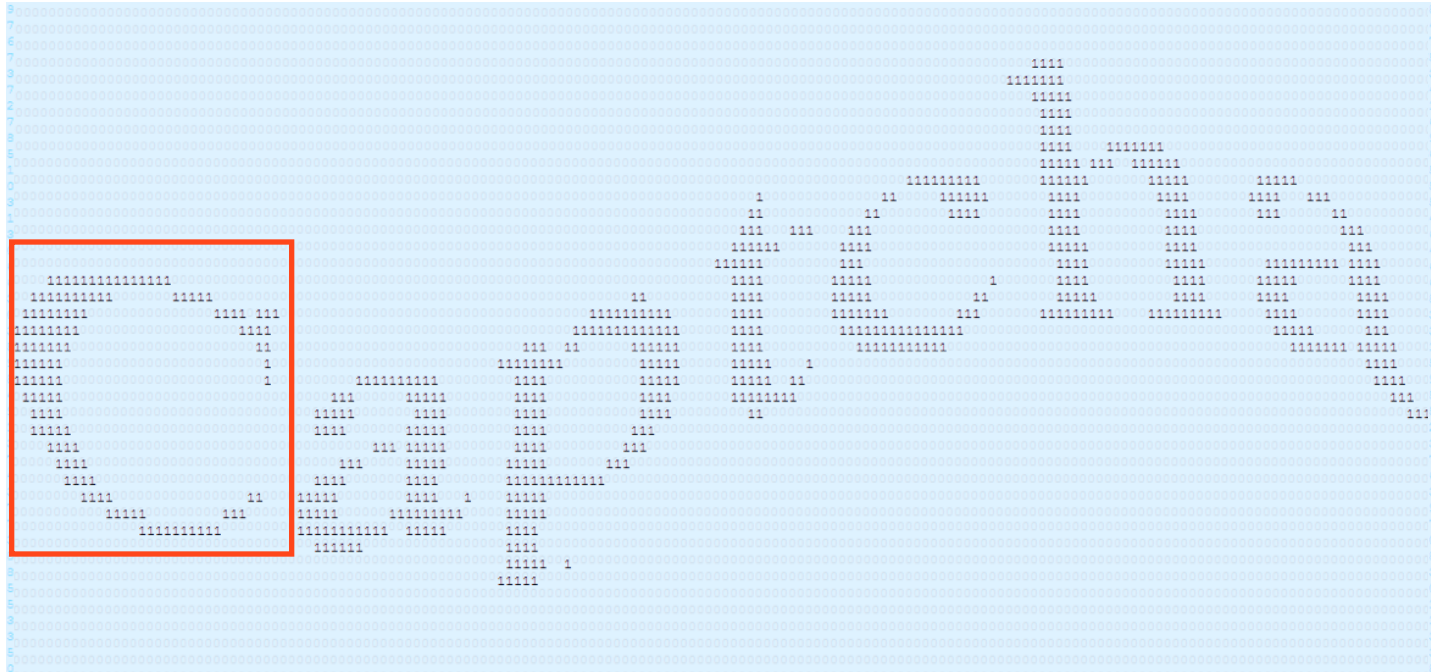
# CapTCHA

From the **image** to the **matrix representation**

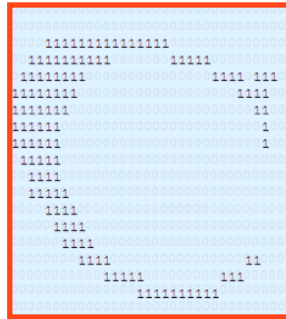


```
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
```

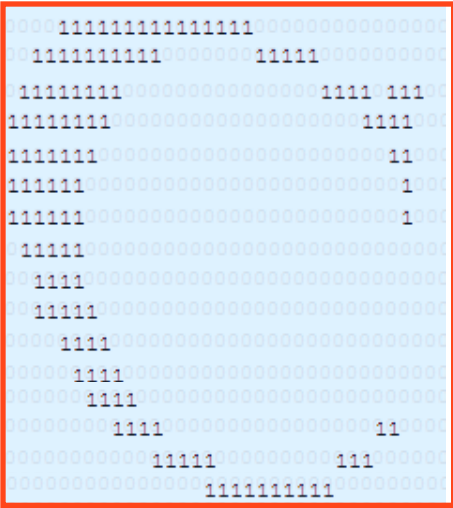
From the **image** to the **matrix representation**



From the **image** to the **matrix representation**



From the **image** to the **matrix representation**



From the **matrix representation** to the **vector representation**

```
001111111111000000001111110000000000
111111111000000000000000111101110000
11111111100000000000000000000011110000
11111111000000000000000000000000110000
1111111000000000000000000000000010000
1111110000000000000000000000000010000
11111000000000000000000000000000000000
00111100000000000000000000000000000000
00111100000000000000000000000000000000
00001111000000000000000000000000000000
00000111100000000000000000000000000000
00000011110000000000000000000000000000
000000001111000000000000000011000000
000000000011111000000000011100000000
000000000000111111111111000000000000
```

```
000011111111111111110000000000000000
```

From the **matrix representation** to the **vector representation**

```
11111111 0000000000000000 1111 111
11111111 0000000000000000 1111
11111111 0000000000000000 11
11111111 0000000000000000 1
11111111 0000000000000000 1
111111 000000000000000000000000
1111 000000000000000000000000
11111 000000000000000000000000
1111 000000000000000000000000
1111 000000000000000000000000
1111 000000000000000000000000
1111 000000000000000000000000
11111 00000000 111 000000
1111111111
```

```
1111111111111111 0000000000000000 1111111111 0000000011111 000000000000
```

From the **matrix representation** to the **vector representation**

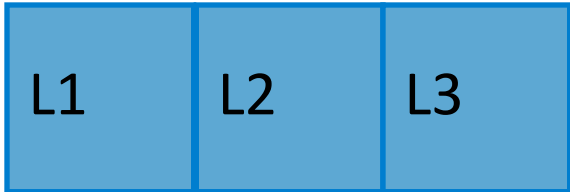


```
11111111 0000000000000000 1111 1111
11111111 0000000000000000 1111
11111111 0000000000000000 11
11111111 0000000000000000 1
11111111 0000000000000000 1
111111 000000000000000000000000
1111 000000000000000000000000
11111 000000000000000000000000
0000 1111 0000000000000000000000
00000 1111 0000000000000000000000
0000000 1111 0000000000000000 11
0000000000 11111 00000000 111
0000000000000000 1111111111
```



From the **matrix representation** to the **vector representation**

```
11111111 00000000000000000000 1111 0000
11111111 00000000000000000000 11 0000
11111111 00000000000000000000 1 0000
11111111 00000000000000000000 1 0000
 111111 0000000000000000000000000000
   1111 0000000000000000000000000000
    1111 0000000000000000000000000000
     111 0000000000000000000000000000
      11 0000000000000000000000000000
       1 0000000000000000000000000000
        1111 00000000 111 000000
         11111111 00000000
```



From the **matrix representation** to the **vector representation**

vector

From the **matrix representation** to the **vector representation**

Known vectors

Distance

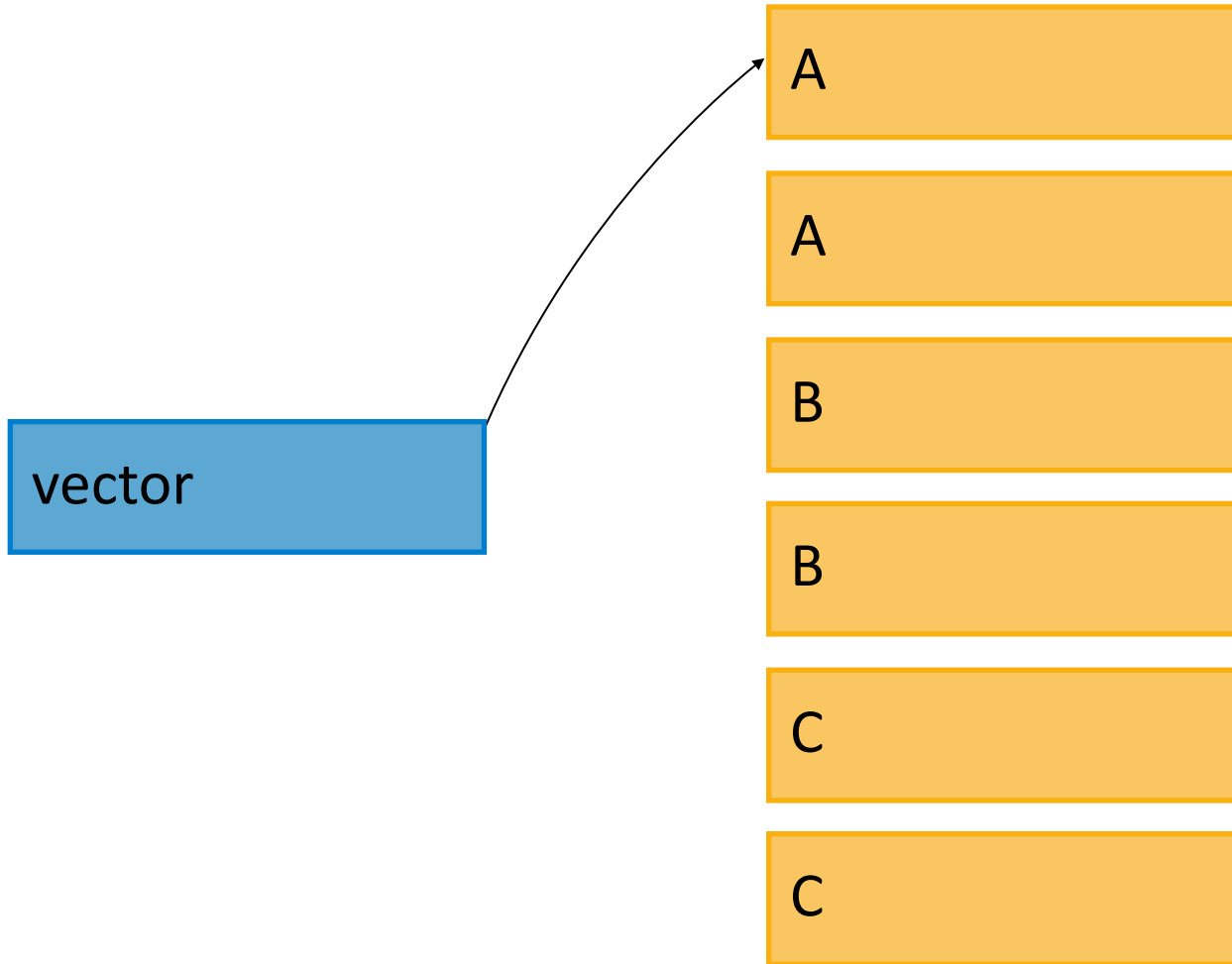
vector

- A
- A
- B
- B
- C
- C

From the **vector representation** to the **segment value**  
(classification)

Known vectors

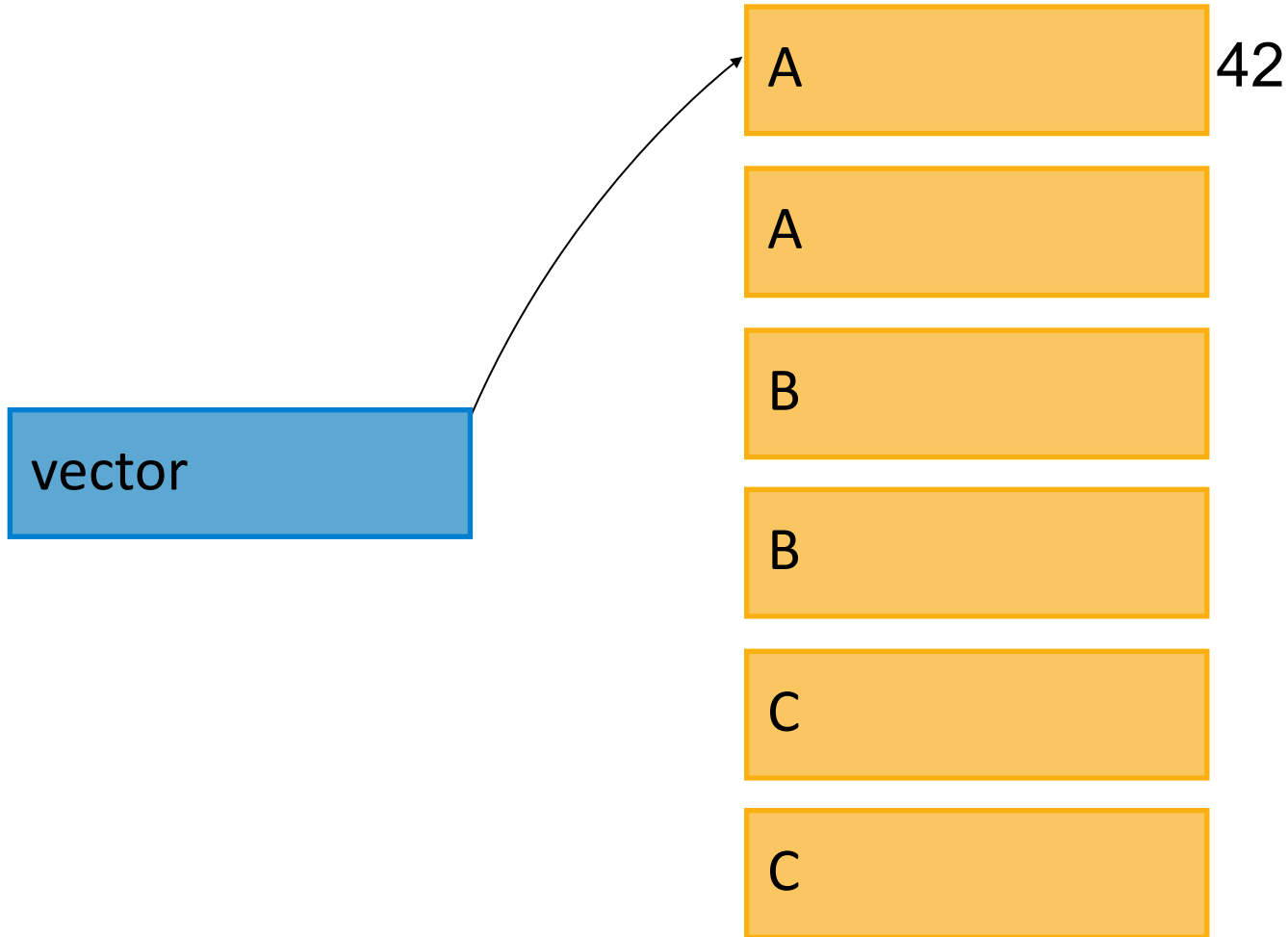
Distance



From the **vector representation** to the **segment value**  
(classification)

Known vectors

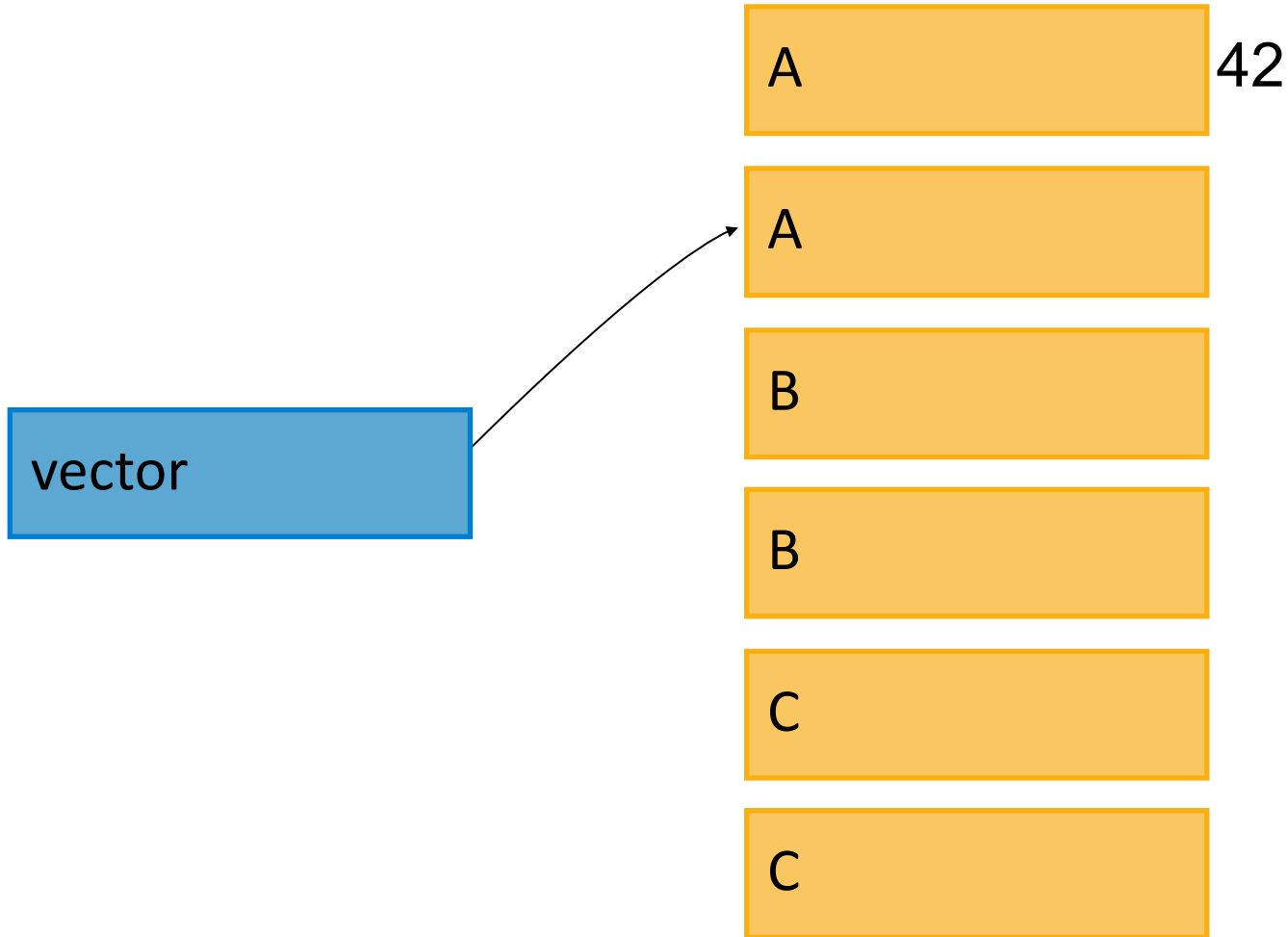
Distance



From the **vector representation** to the **segment value**  
(classification)

Known vectors

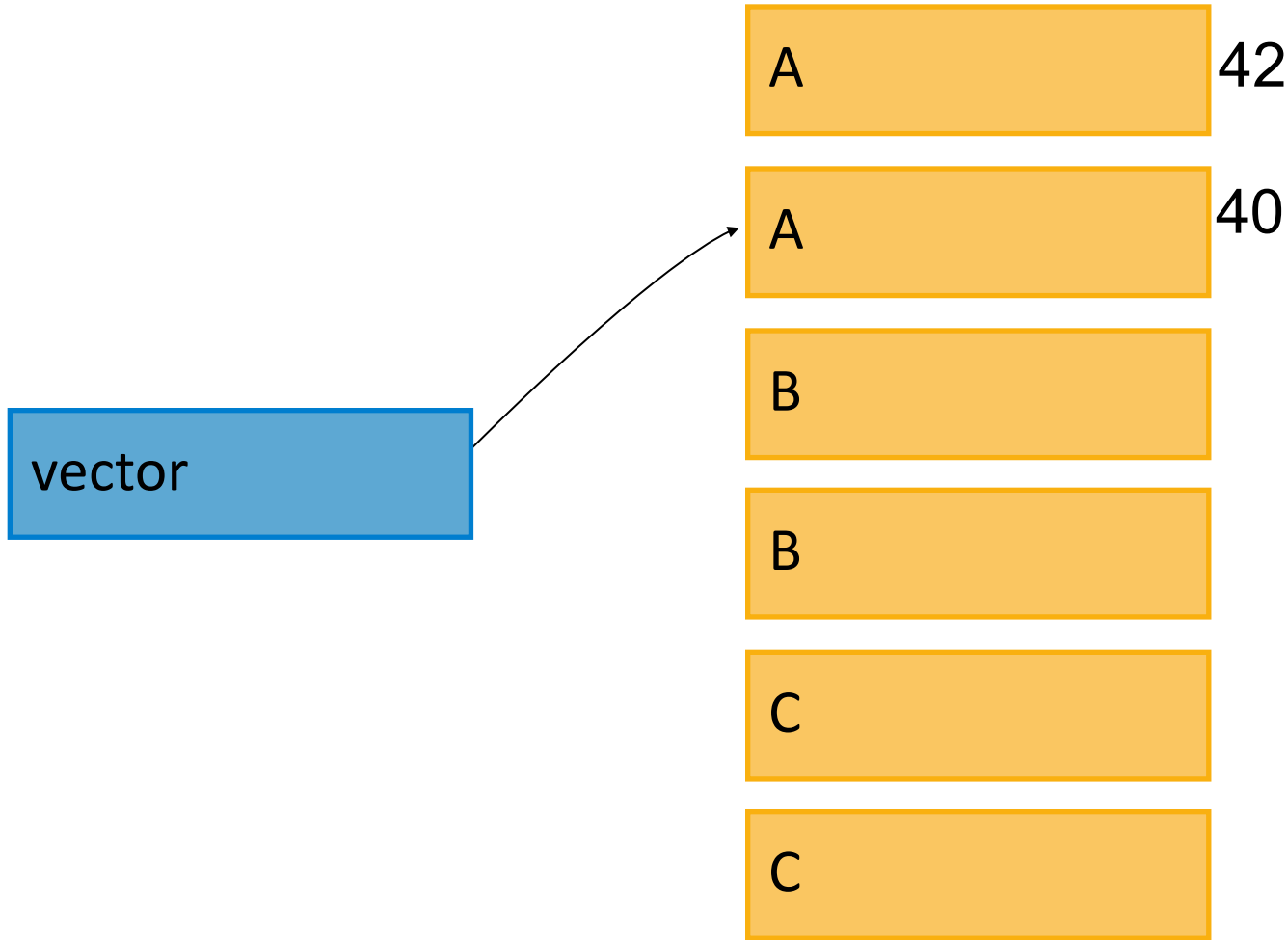
Distance



From the **vector representation** to the **segment value**  
(classification)

Known vectors

Distance



From the **vector representation** to the **segment value**  
(classification)



Known vectors

Distance

vector



A	42
A	40
B	
B	
C	
C	

From the **vector representation** to the **segment value**  
(classification)

Known vectors

Distance

vector



A	42
A	40
B	32
B	
C	
C	

From the **vector representation** to the **segment value**  
(classification)

vector

Known vectors

Distance

A	42
A	40
B	32
B	70
C	12
C	18

From the **vector representation** to the **segment value**  
(classification)

vector

Known vectors

Distance

A	42
A	40
B	32
B	70
C	12
C	18

From the **vector representation** to the **segment value**  
(classification)

# Breaker efficiency

$$\text{Solver accuracy} = \text{Coverage} * \text{Precision}^{\text{length}}$$

*Coverage*: Segmentation rate

*Precision*: Recognition rate



# Anti-recognition techniques



# Anti-recognition techniques

Blurring

3tr2bb



# Anti-recognition techniques

Blurring

3tr2bb

Distortion

0zt99n





# Anti-recognition techniques

Blurring

3tr2bb

Distortion

0zt99n

Rotation

o a v y s b



# Anti-recognition techniques

Blurring

3tr2bb

Distortion

0zt99n

Rotation

o a v y c b

Fonts

08G722



# Anti-recognition techniques

Blurring

3tr2bb

Distortion

0zt99n

Rotation

o a v y s b

Fonts

08G722

Charsets



# Anti-recognition techniques

Blurring

3tr2bb

Distortion

0zt99n

Rotation

o a v y c b

Fonts

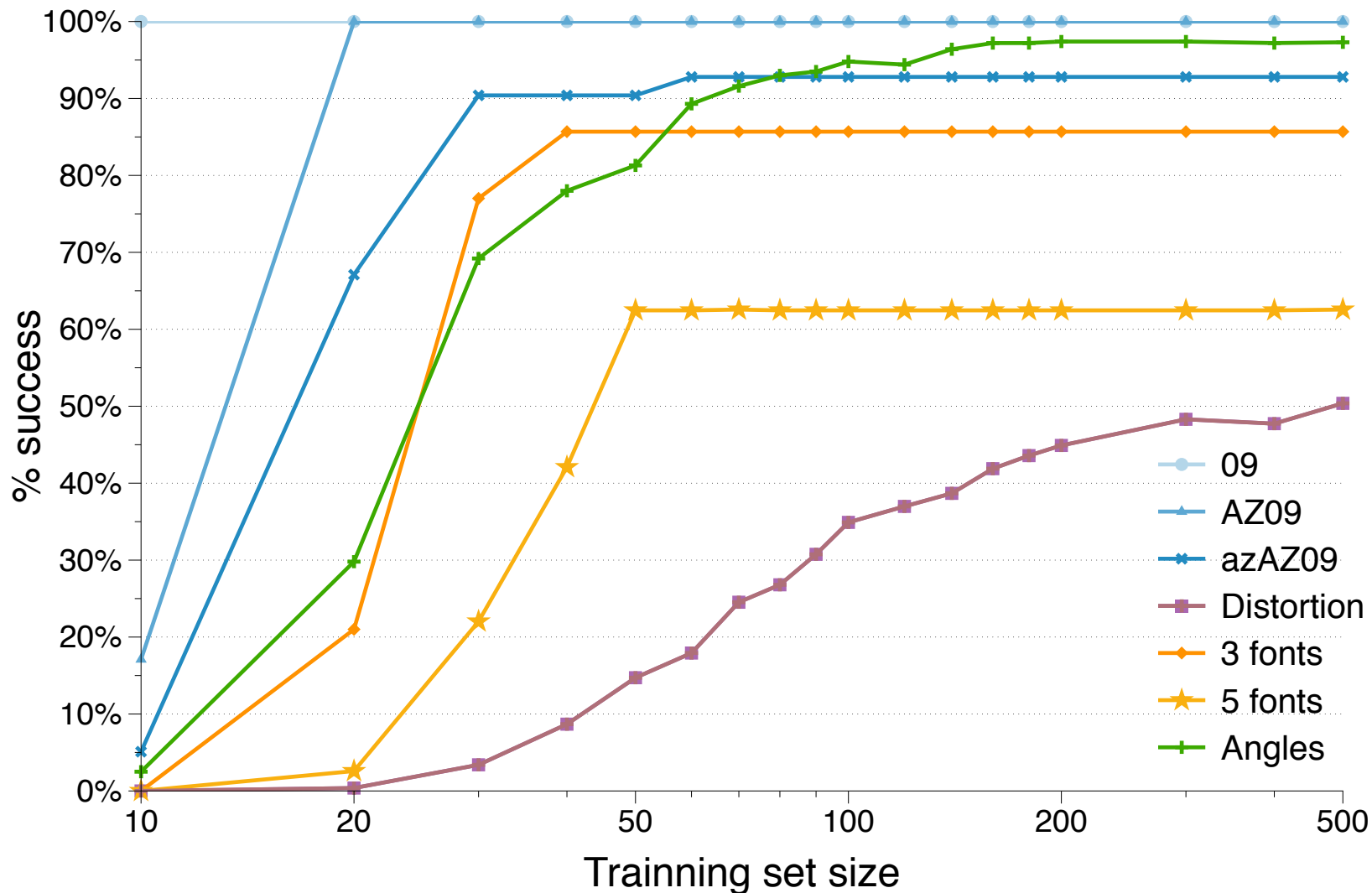
08G722

Charsets

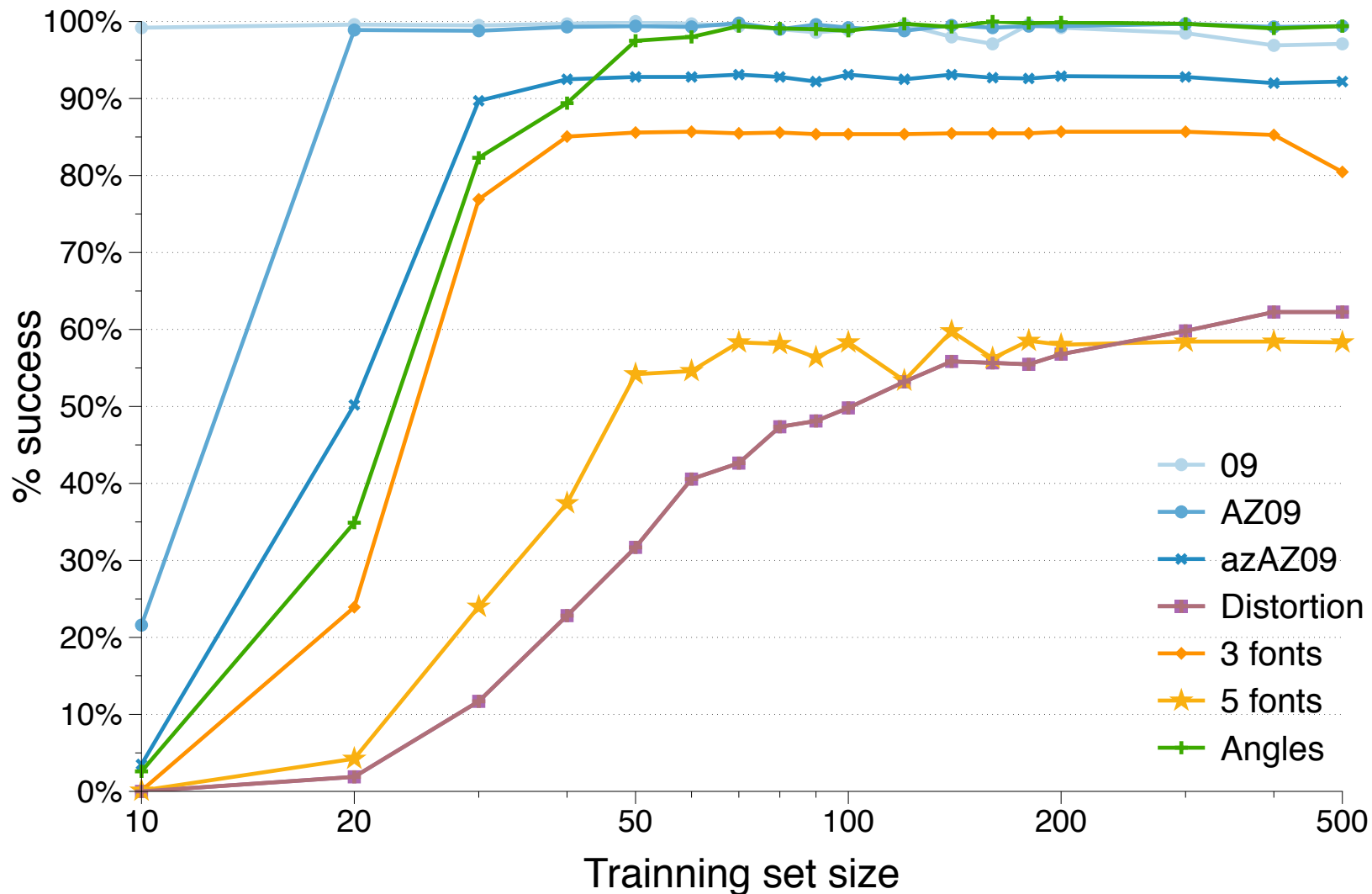
0123456789



# SVM learning rate



# KNN learning rate



# Anti-recognition taxonomy



# Anti-recognition taxonomy

## Background Confusion





# Anti-recognition taxonomy

Background Confusion



**3nc9z**



# Anti-recognition taxonomy

## Background Confusion

**3nc9z**

*quxg4h*



# Anti-recognition taxonomy

## Background Confusion



# Anti-recognition taxonomy

## Background Confusion



## Lines

# Anti-recognition taxonomy

## Background Confusion



## Lines

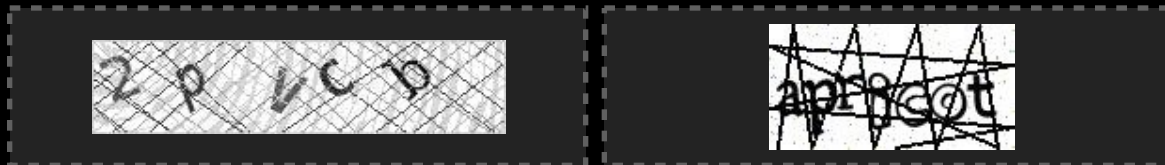


# Anti-recognition taxonomy

## Background Confusion



## Lines



# Anti-recognition taxonomy

## Background Confusion



3nc9z



quxg4h



p m y m k u


## Lines



z p v c p



apiscot



dramacharm

# Anti-recognition taxonomy

## Background Confusion



## Lines



## Collapsing





# Anti-recognition taxonomy

## Background Confusion



3nc9z



quxg4h



p m y m k u

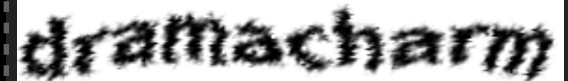
## Lines



z p v c p



apscot



dramacharm

## Collapsing



984505

# Anti-recognition taxonomy

## Background Confusion



3nc9z



quxg4h



p m y m k u

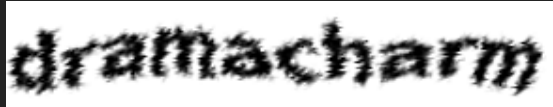
## Lines



z p v c p



apiscot



dramacharm

## Collapsing



984505



RAE3

# Anti-recognition taxonomy

## Background Confusion



3nc9z



quxg4h



p m y m k u

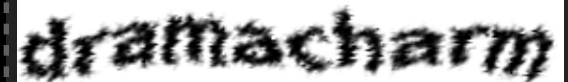
## Lines



z p v c p



apiscot



dramacharm

## Collapsing



984505

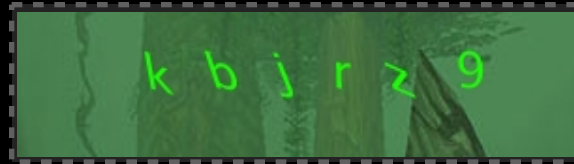


RAE3

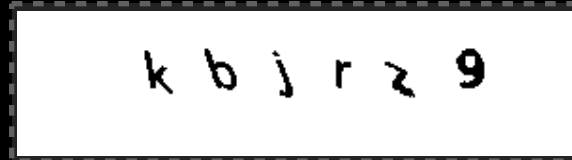


deactiesge

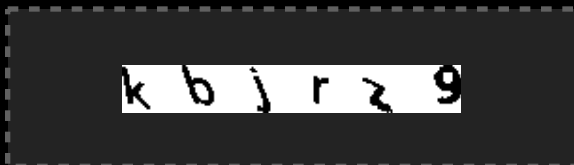
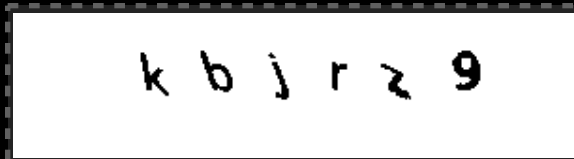
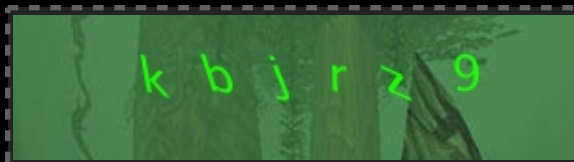
# Breaking World of Warcraft



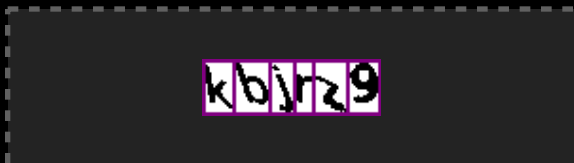
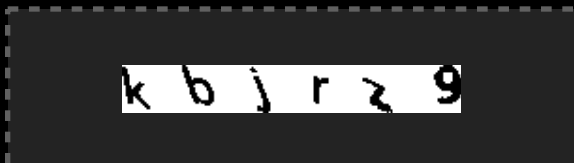
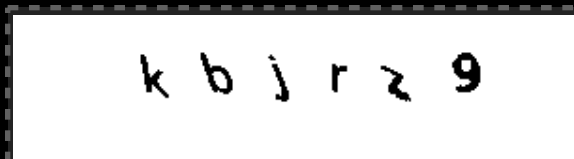
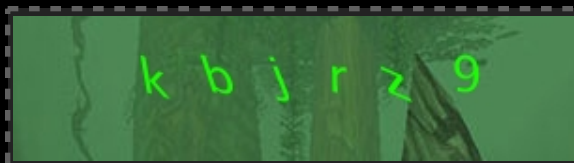
# Breaking World of Warcraft



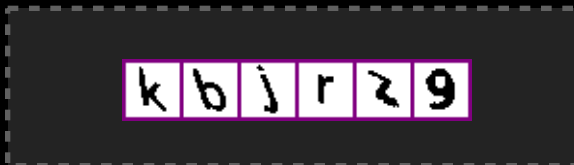
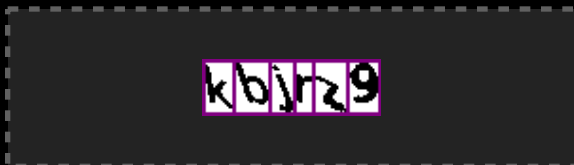
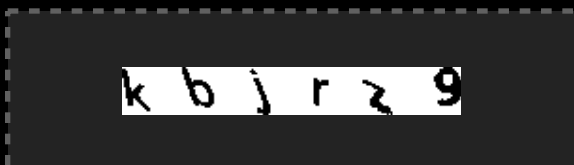
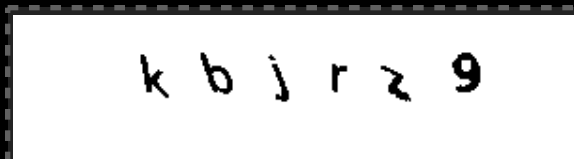
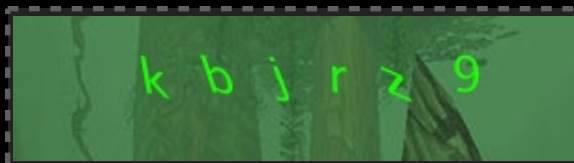
# Breaking World of Warcraft



# Breaking World of Warcraft



# Breaking World of Warcraft





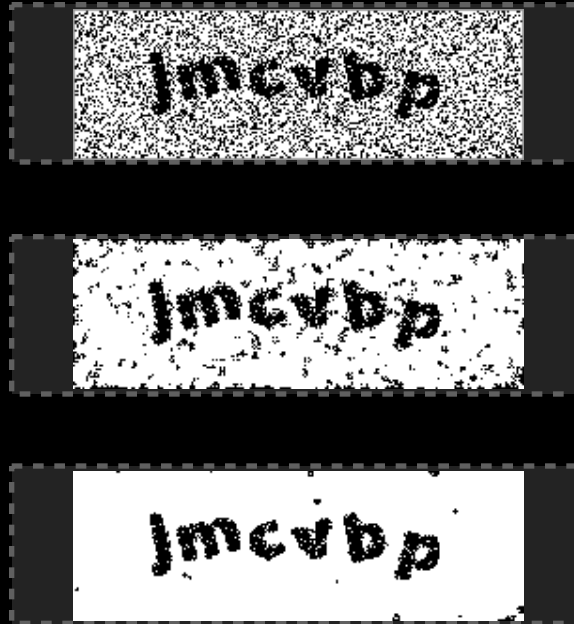
# Breaking Captcha.net



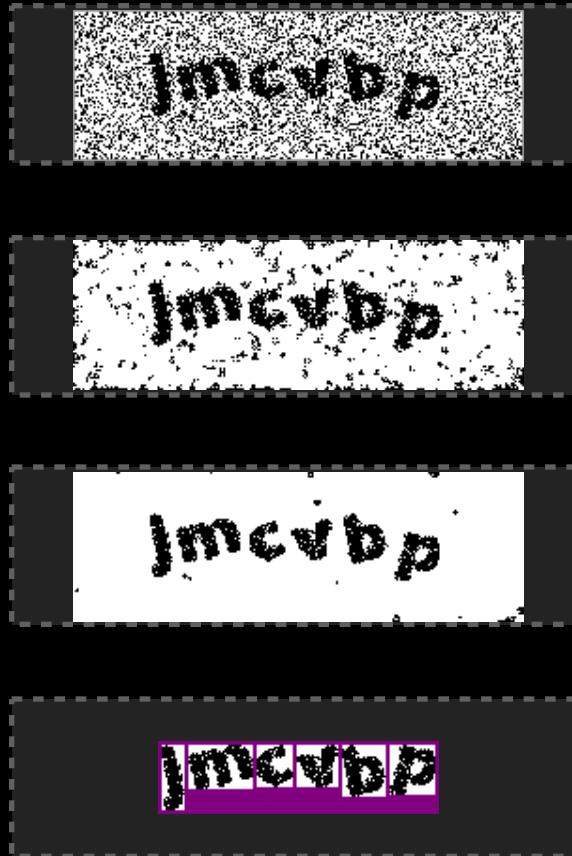
# Breaking Captcha.net



# Breaking Captcha.net



# Breaking Captcha.net



# Breaking Captcha.net



# Breaking Wikipedia

**dramacharm**



# Breaking Wikipedia

dramacharm

dramacharm



# Breaking Wikipedia

dramacharm

dramacharm

dramacharm





# Breaking Wikipedia

dramacharm

dramacharm

dramacharm

dramacharm



# Breaking Wikipedia

dramacharm

dramacharm

dramacharm

dramacharm

d r a m a c h a r m



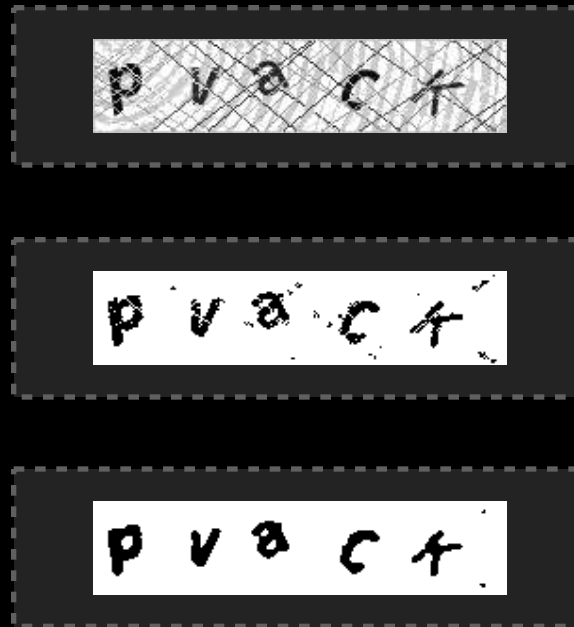
# Breaking Digg



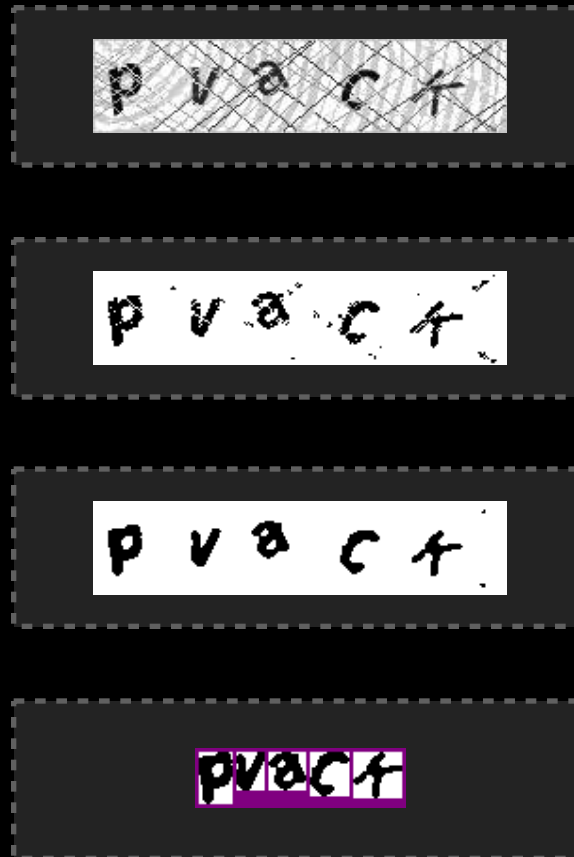
# Breaking Digg



# Breaking Digg



# Breaking Digg



# Breaking Digg

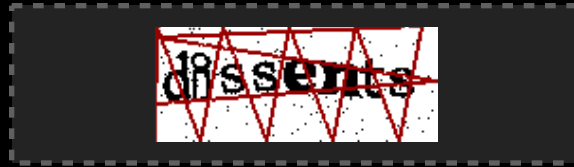


# Breaking Slashdot

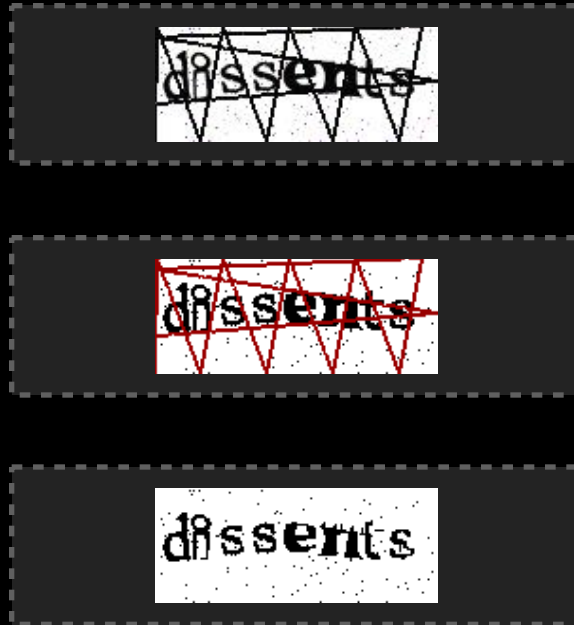




# Breaking Slashdot



# Breaking Slashdot



# Breaking Slashdot



dissents



dissents



dissents



dissents

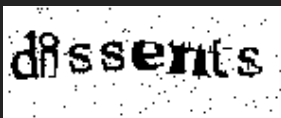
# Breaking Slashdot



dissents



dissents



dissents



dissents



d i s s e n t s

# Breaking eBay

944 531



# Breaking eBay

944 531

944 531



# Breaking eBay

944 531

944 531

944 531



# Breaking eBay

944 531

944 531

944 531

944 531





# Breaking eBay

944531

944531

944531

944531

9 4 4 5 3 1



# Failing to break eBay

584671

# Failing to break eBay

584671

584671



# Failing to break eBay

584671

584671

584671



# Failing to break eBay

584671

584671

584671

584671



# Failing to break eBay

584671

584671

584671

584671

5 8 4 6 7 1



# Breaking Baidu

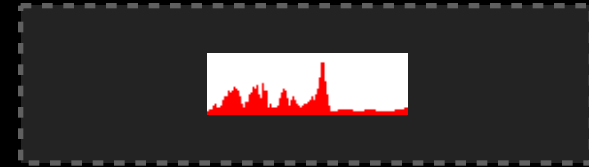


# Breaking Baidu

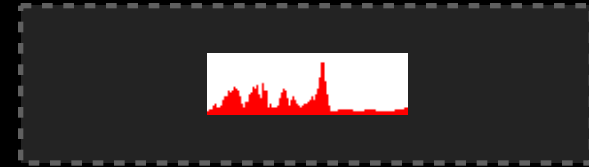




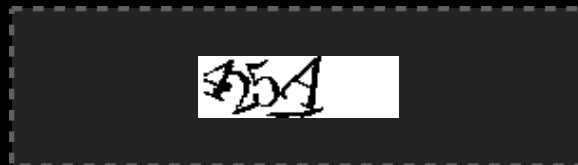
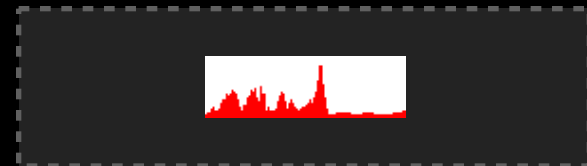
# Breaking Baidu



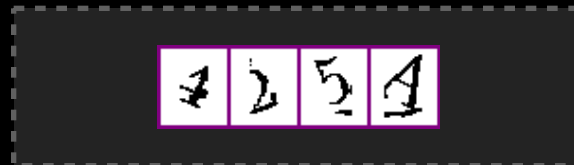
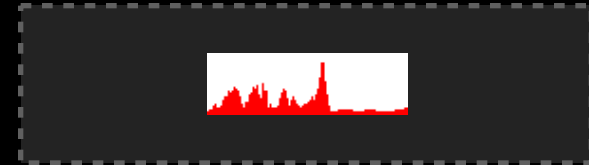
# Breaking Baidu



# Breaking Baidu



# Breaking Baidu

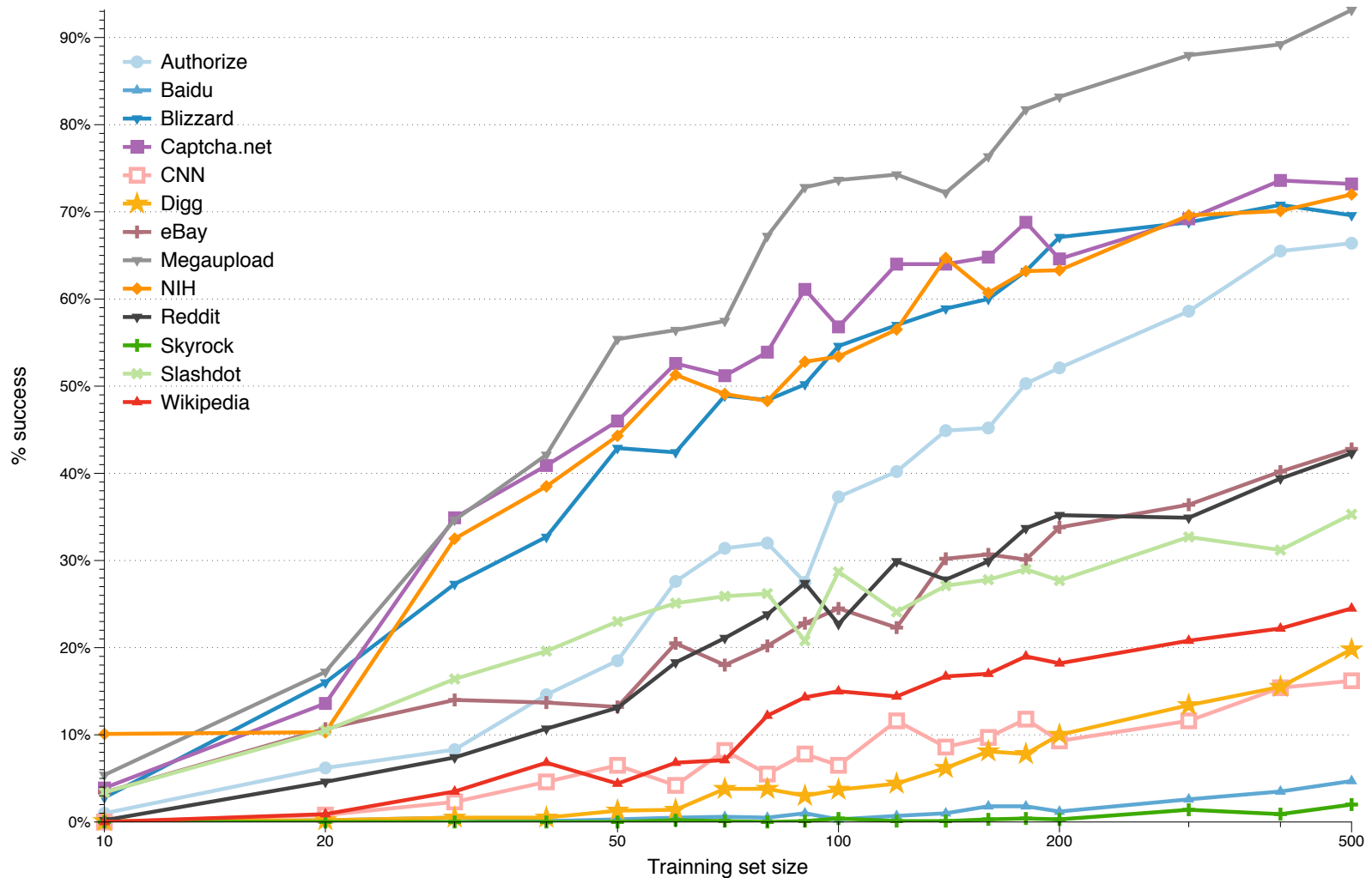


# Overall results

	Segmentation	Solving rate
Authorize	84%	66%
Baidu	98%	5%
Blizzard	75%	70%
Captcha.net	96%	73%
CNN	50%	16%
Digg	86%	20%
eBay	95%	43%
Google	0%	0%
MegaUpload	n/a	93%
NIH	87%	72%
Recaptcha	0%	0%
Reddit	71%	42%
Skyrock	30%	2%
Slashdot	52%	35%
Wikipedia	57%	25%



# Learning rate for real schemes



# Decaptcha main interface

The screenshot displays the DeCaptcha main interface, which is divided into several functional areas:

- Configuration File:** A tree view on the left showing the program structure, including Preprocessing, Segmentation, Reconstruction, and Recognition sub-modules.
- Results:** A central area featuring a pie chart titled "Efficiency breakdown" with the following data:

Category	Percentage
Success	66%
Preprocessing	11%
Segmentation	18%
Verification	5%
- Log:** A text area on the right providing a detailed record of operations, such as "AntiPattern-Preprocessing: Starting" and "Vanilla-Segmentation: Number of pixels found 18".
- States:** A vertical stack of image thumbnails on the far right, showing the original captcha and the corresponding processed results with bounding boxes.
- Images:** A list in the bottom-left quadrant showing the current batch of images and their assigned processing step (e.g., Segmentation, Preprocessing, Verification).
- Properties:** A section below the configuration file, currently empty.
- Control Panel:** A bottom bar containing sliders for "Train" (Size min, Size max, Step) and "Test" (Size), along with a "Done" button, a "Save images" checkbox, and "Train", "Test", and "Save" buttons.

# Apply design principles

- Core design principles
  - Randomize length
  - Randomize character size
  - Wave the captcha
- Use anti-recognition as a means of strengthening captcha security
- Don't use a complex charset
  - Bad for human (see our research on this)
  - Useless for security
- Use collapsing or lines



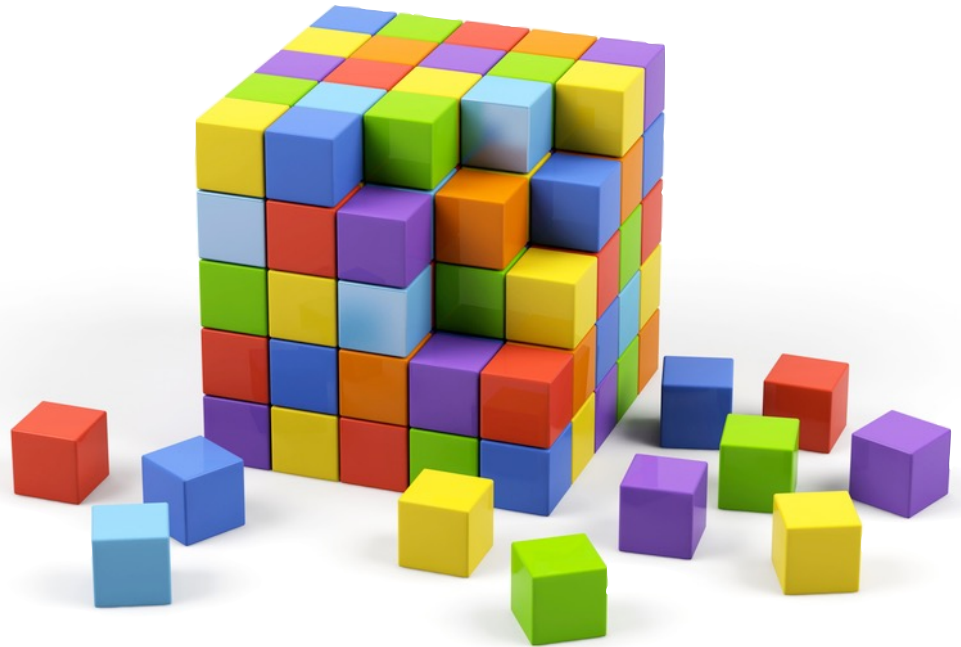




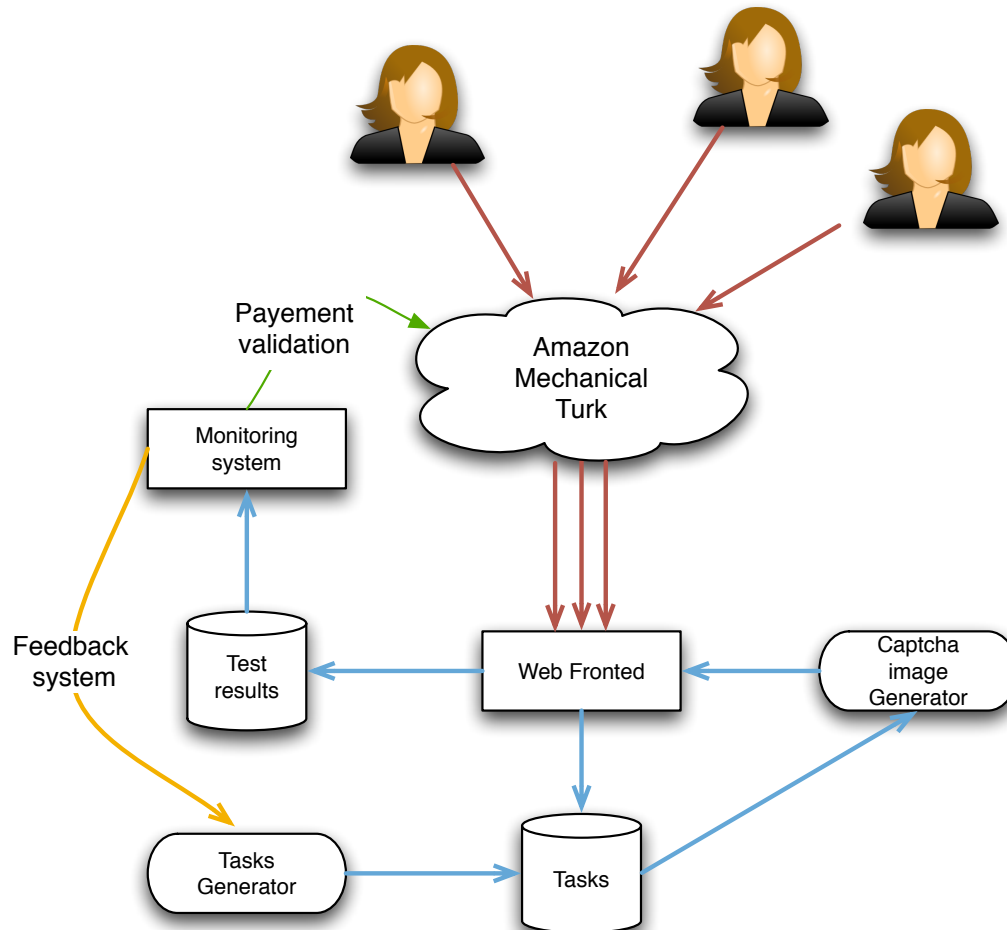
# Designing Better Captchas

# Think Lego again

- Decompose in features
- Analyze
  - feature in isolation
  - features interaction



# Evaluation system



# Experiment details

Round	Task	N possible	N sampled	N tests per sample	Total tests
1	Baseline (“Control”)	1	1	1000	<i>1000</i>
2	Real world captchas	8	8	1000	<i>8000</i>
3	Features in isolation	496	496	200	<i>99200</i>
4	<b>2</b> feature interactions	60950	60950	5	<i>304750</i>
5	<b>3</b> feature interactions	1 303 224	25000	10	<i>250000</i>
6	<b>4</b> feature interactions	113 951 684	25000	10	<i>250000</i>
	Total				<i>912150</i>



# Some of the features tested

3tr2bb

Blurring

tx3soh

Text color

0AGP22

Font

frbxth

Background color

02qeh

Collapsing

o a v y n

Tilting

azx<sup>t</sup>1g

Waving

0zt99n

Distortion

~~0izjw6~~

line

~~ul8mgx~~

line angle

~~jsf8be~~

line shape

~~4umt35~~

nb line

5h3lxk

line coverage

~~pe3prq~~

line position

~~0abqn6~~

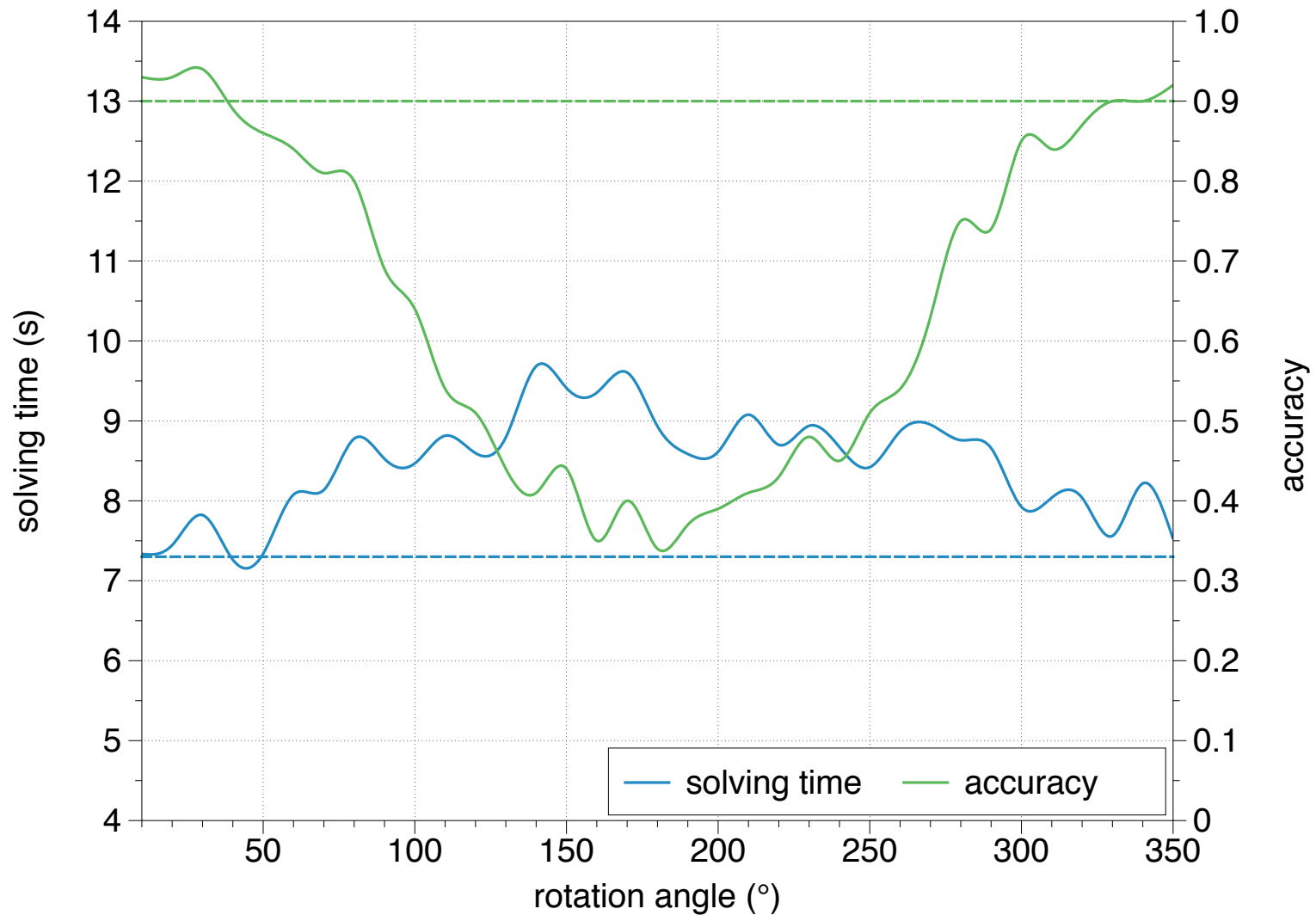
line size

~~k2r576~~

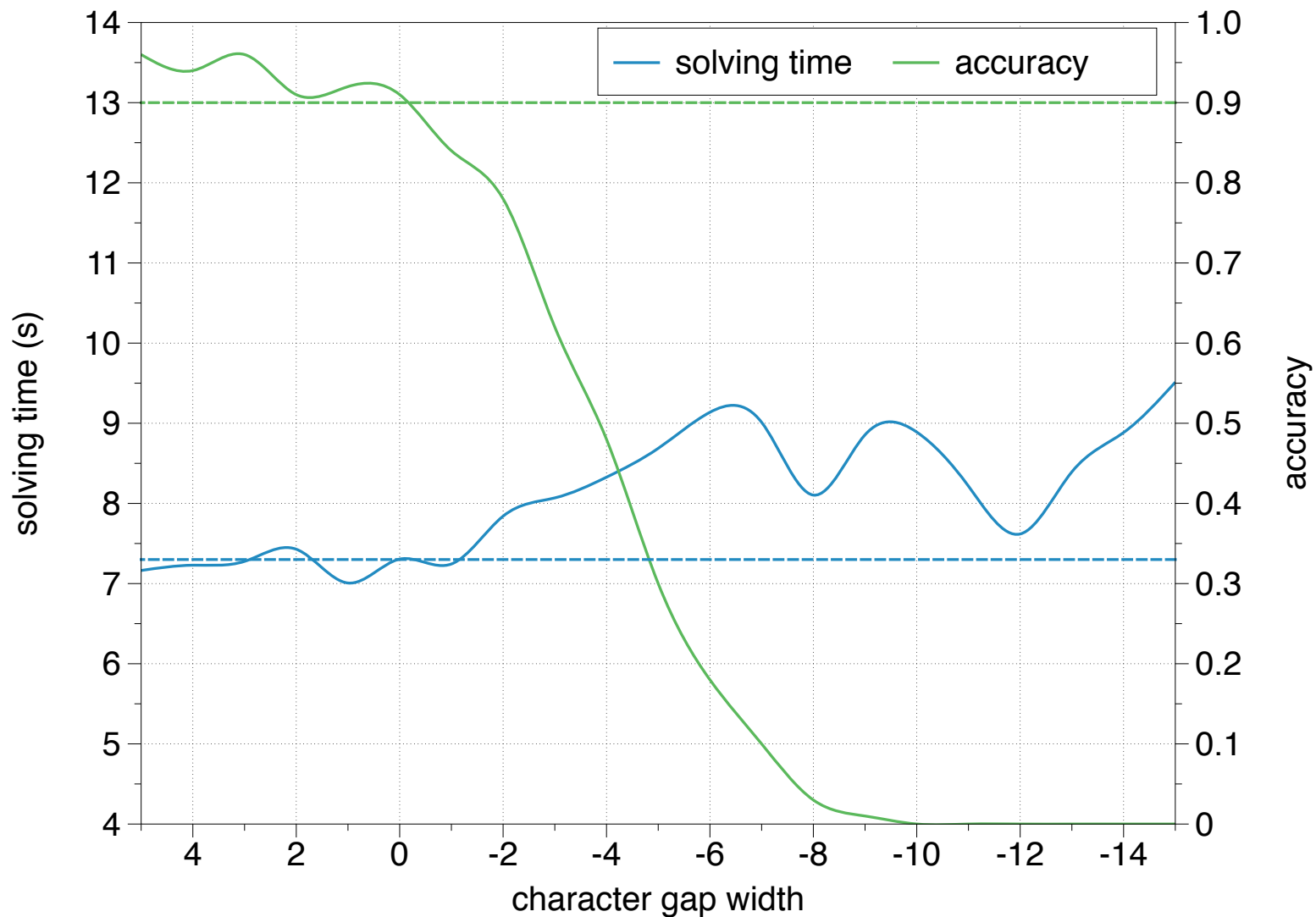
Noise



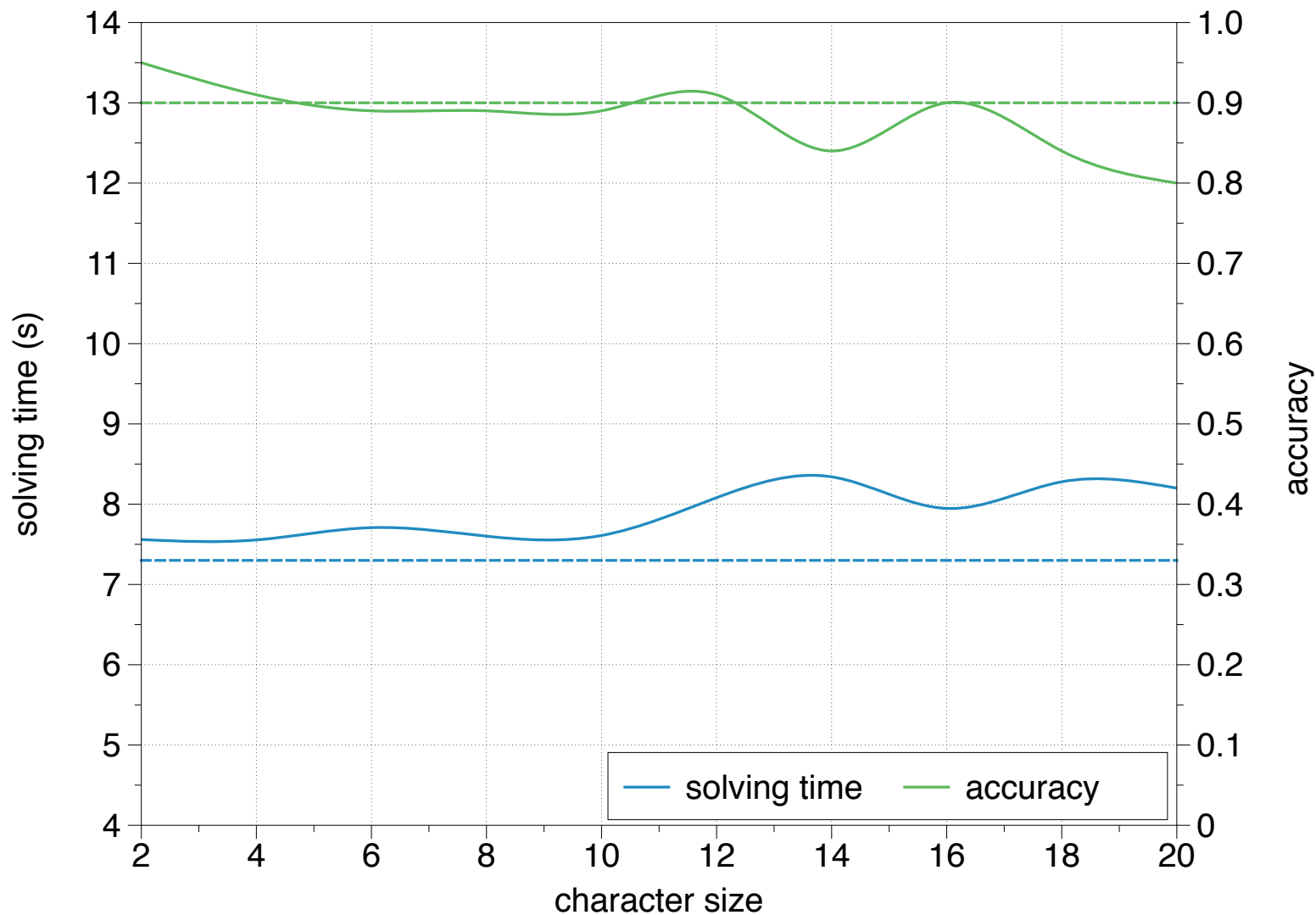
# Angle of rotation



# Collapsing

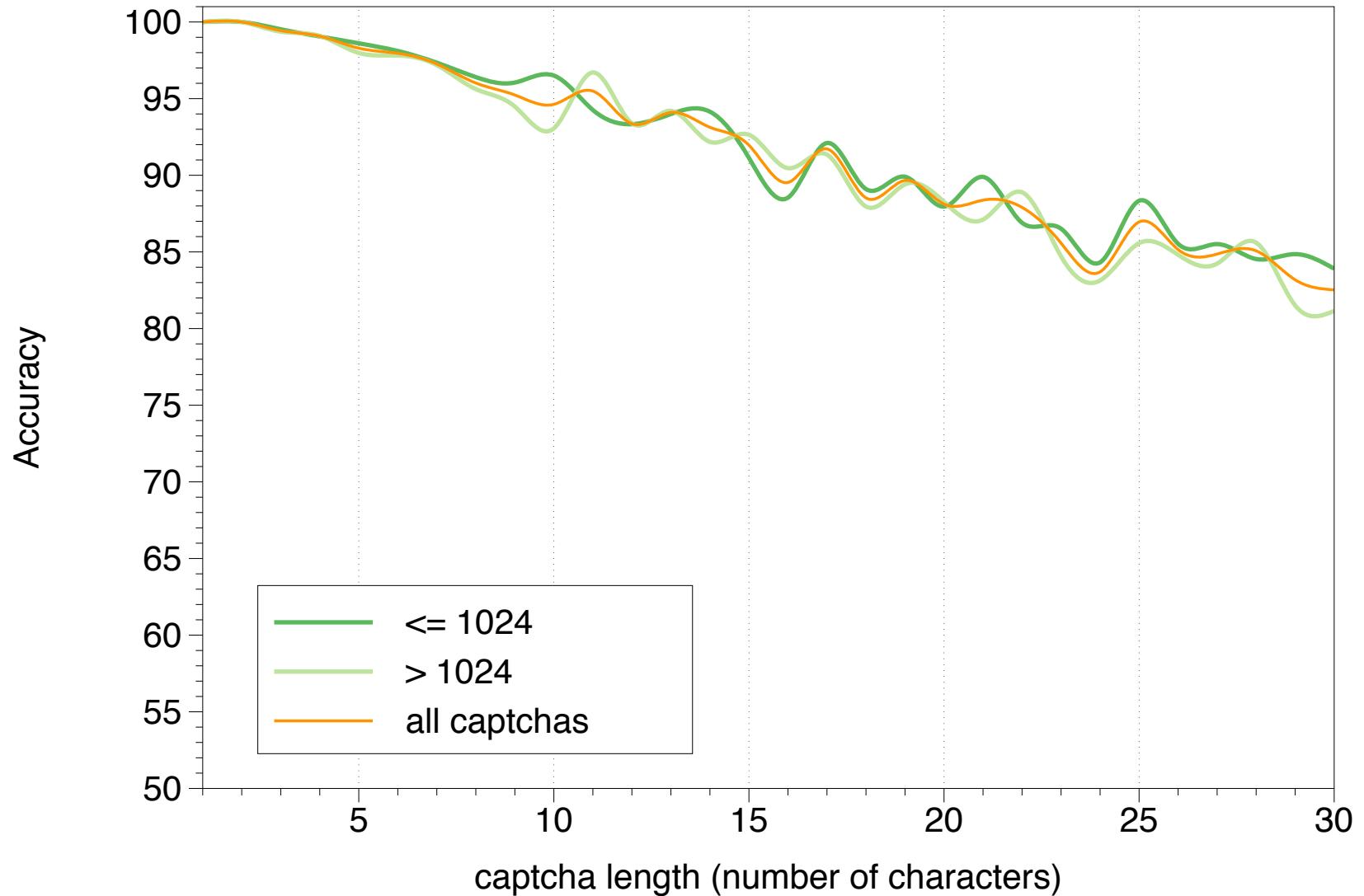


# Character size

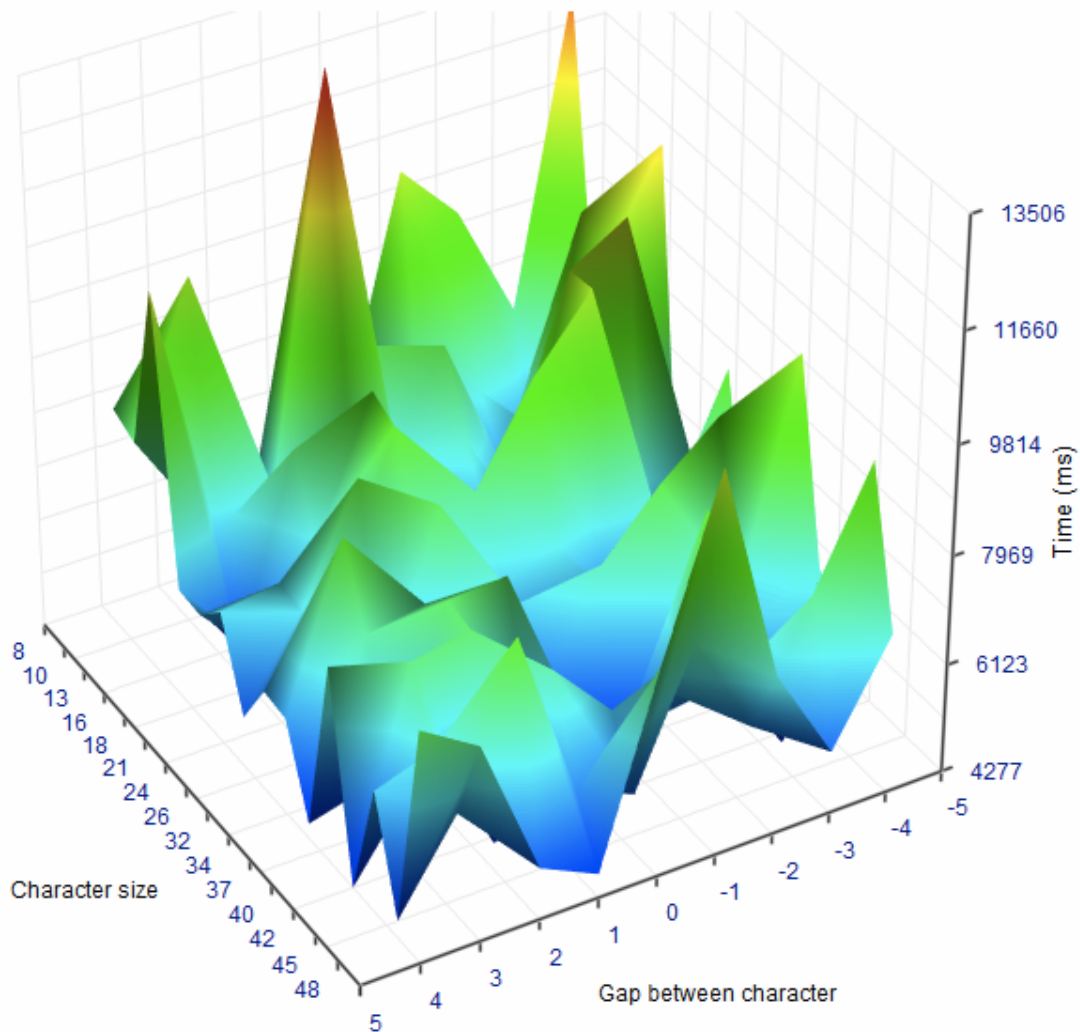




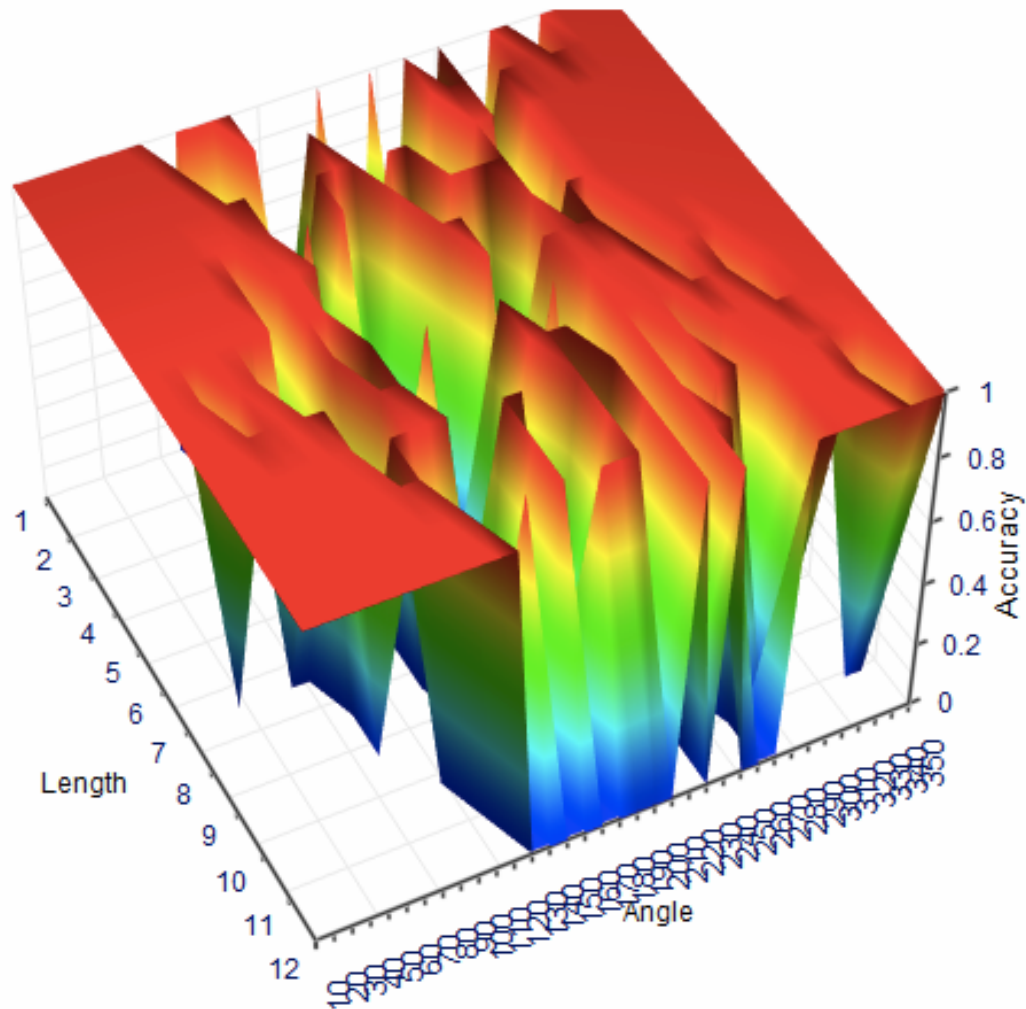
# Resolution invariant



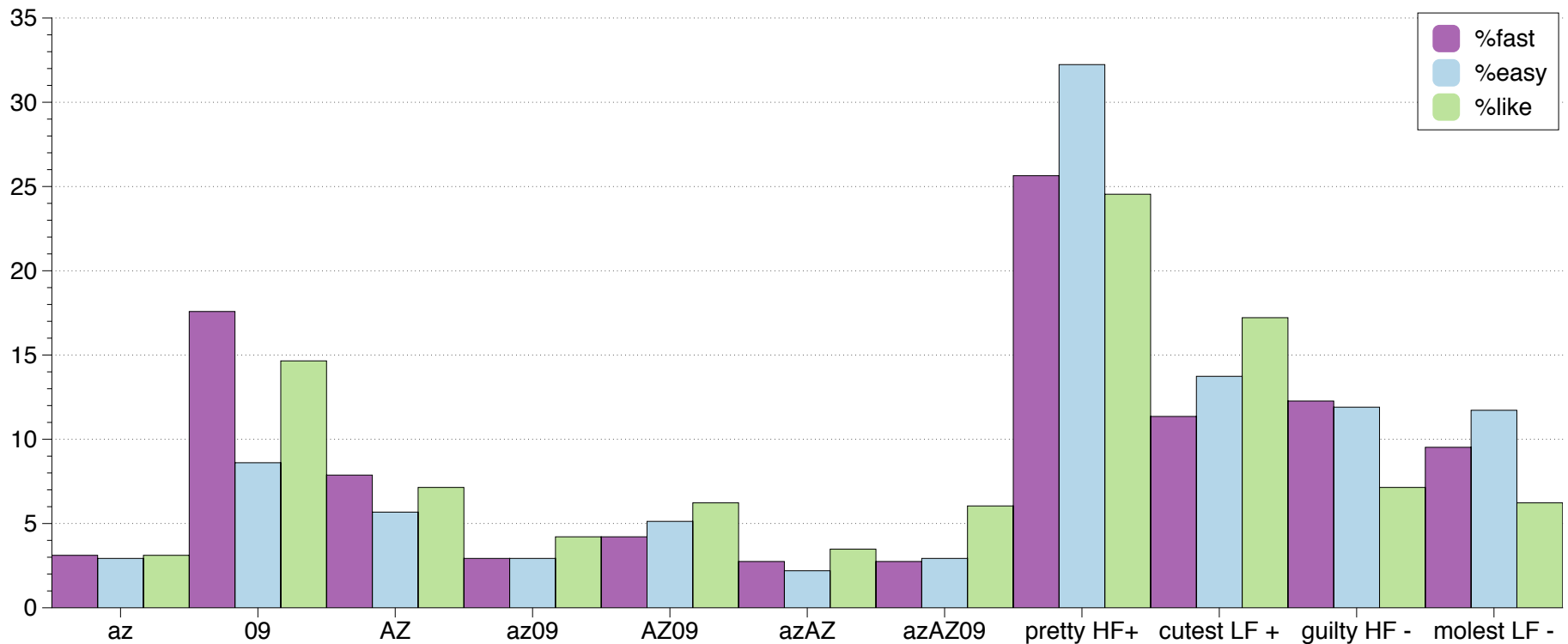
# 2D interactions



# Length vs Angle interaction



# Perception Does Not Match Number



# The New Wikipedia

- Use digit
- Wave the captcha
- Use random length (5-7)
- Use random size (34-50)
- Rotate letter (-25/ 25)
- Add a line for a super secure version



# End result

trustothor

Accuracy

Solving time



# End result

trustothor

Accuracy 84.8%

Solving time



# End result

trustothor

Accuracy 84.8%

Solving time 7.8s





# End result

trustothor

0120

0120

Accuracy 84.8%

Solving time 7.8s



# End result

trustothor

0120

0120

Accuracy

84.8%

89.2%

82.6%

Solving time

7.8s



# End result

trustothor

0120

0120

Accuracy

84.8%

89.2%

82.6%

Solving time

7.8s

4.9s

5.3s



# End result

trustother

0120

0120

↑  
confusing

Accuracy

84.8%

89.2%

82.6%

Solving time

7.8s

4.9s

5.3s



# End result

trustothor

01-20

00-08

01-20

00-08

↑  
confusing

Accuracy 84.8%

89.2%

82.6%

Solving time 7.8s

4.9s

5.3s



# End result

trustothor

01-20

00-08

01-20

00-08

↑  
confusing

Accuracy

84.8%

89.2%

97%

82.6%

92.2%

Solving time

7.8s

4.9s

5.3s



# End result

trustothor

01-20

00-08

01-20

00-08

↑  
confusing

Accuracy

84.8%

89.2%

97%

82.6%

92.2%

Solving time

7.8s

4.9s

4.9s

5.3s

5.2s





# How to Break Audio-Captcha



# Audio Captchas

Visual code | [Audio code](#) [Help](#)



Type the code shown  [Try a new code](#)

## The not-so-fine print

For [added security](#), please enter the verification code hidden in the image.



[Refresh the image](#) | [Listen to the verification code](#)



stop spam.  
read books.



# Audio Captchas

Visual code | [Audio code](#) [Help](#)



Type the code shown  [Try a new code](#)

## The not-so-fine print

For [added security](#), please enter the verification code hidden in the image.



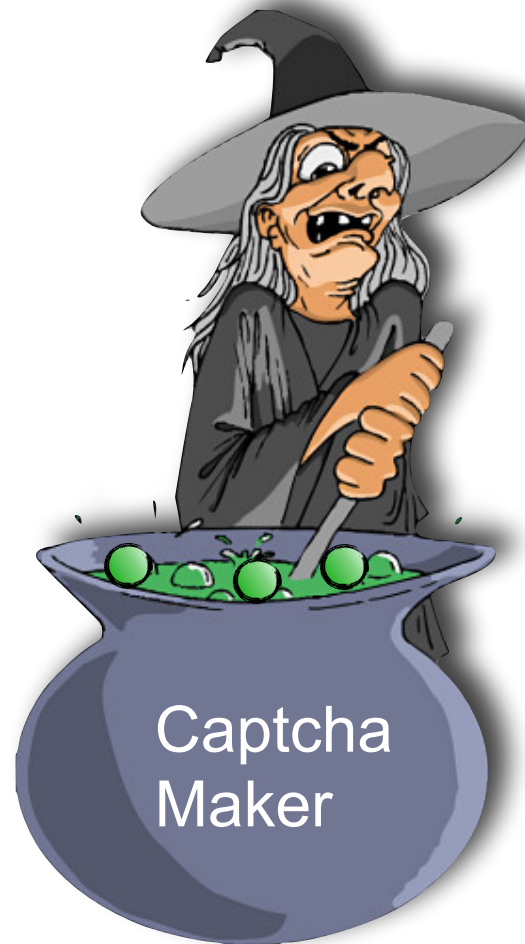
[Refresh the image](#) | [Listen to the verification code](#)



stop spam.  
read books.



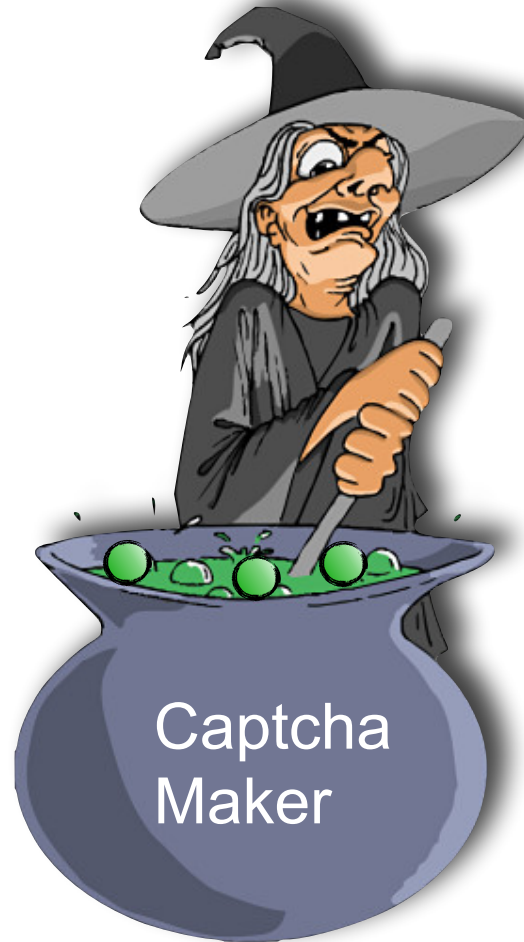
# Creating Audio Captcha



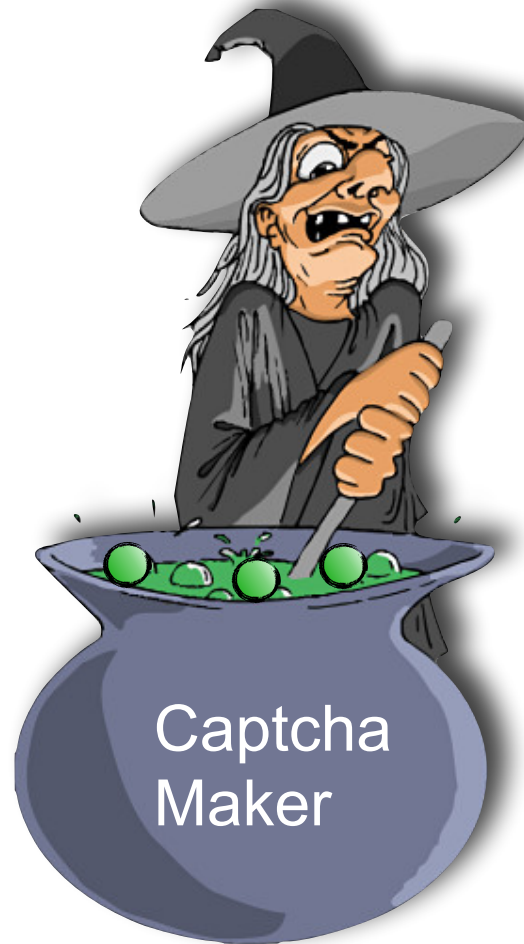
# Creating Audio Captcha



Voices



# Creating Audio Captcha



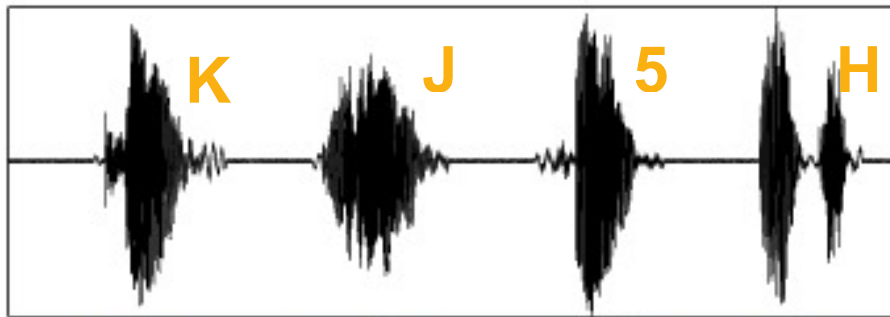
# Creating Audio Captcha



Super  
secure captcha



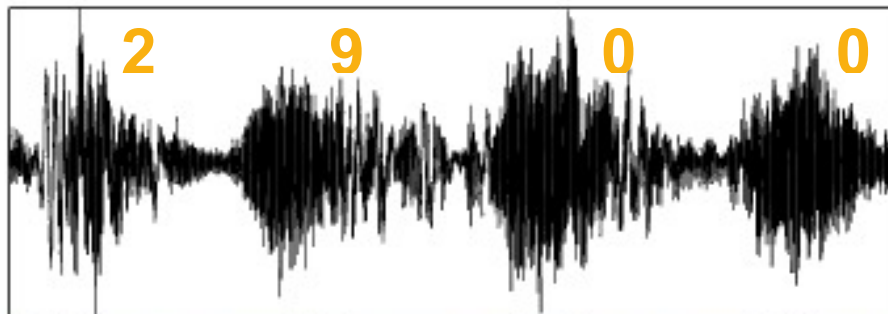
# Noise intensity (RMS/SNR)



Authori



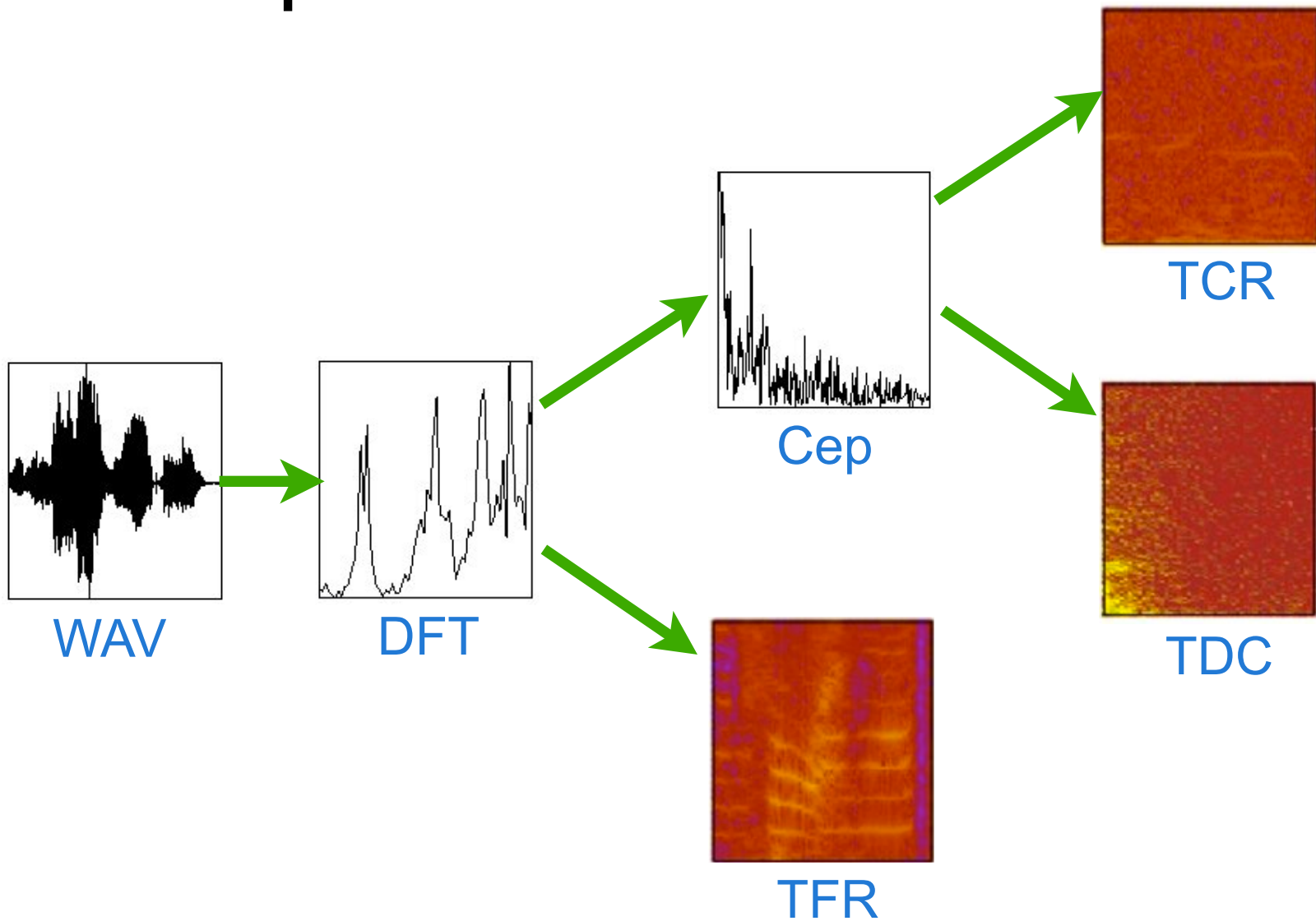
Dia



Micros

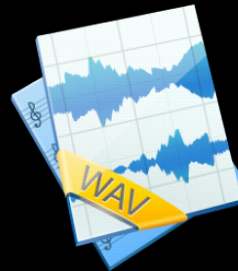


# Sound representation

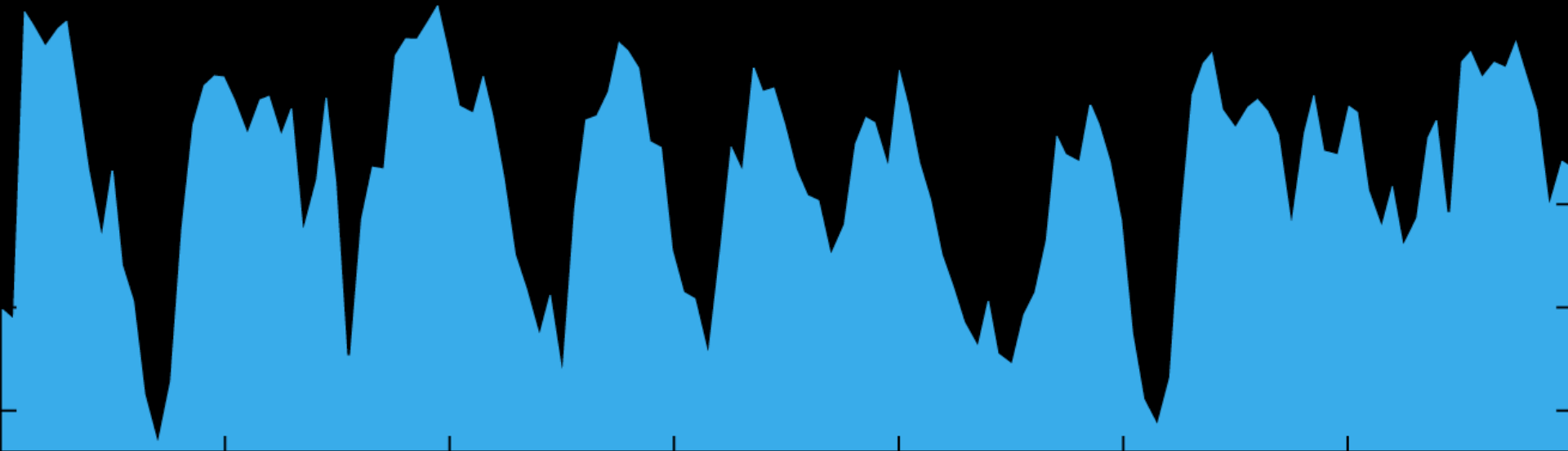




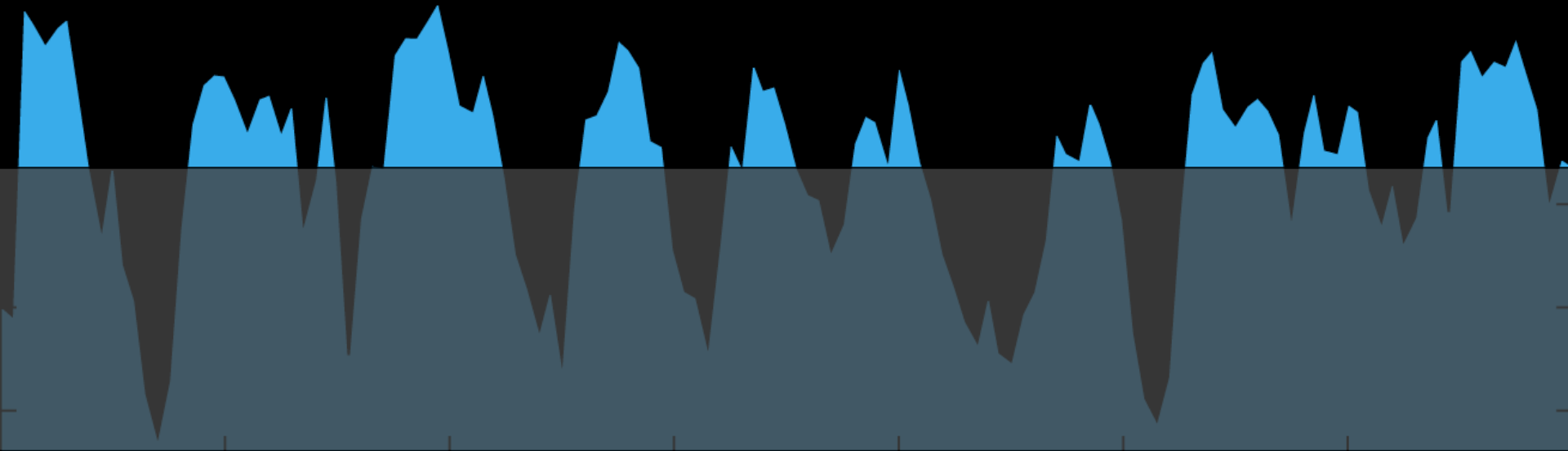
# Solving an audio captcha



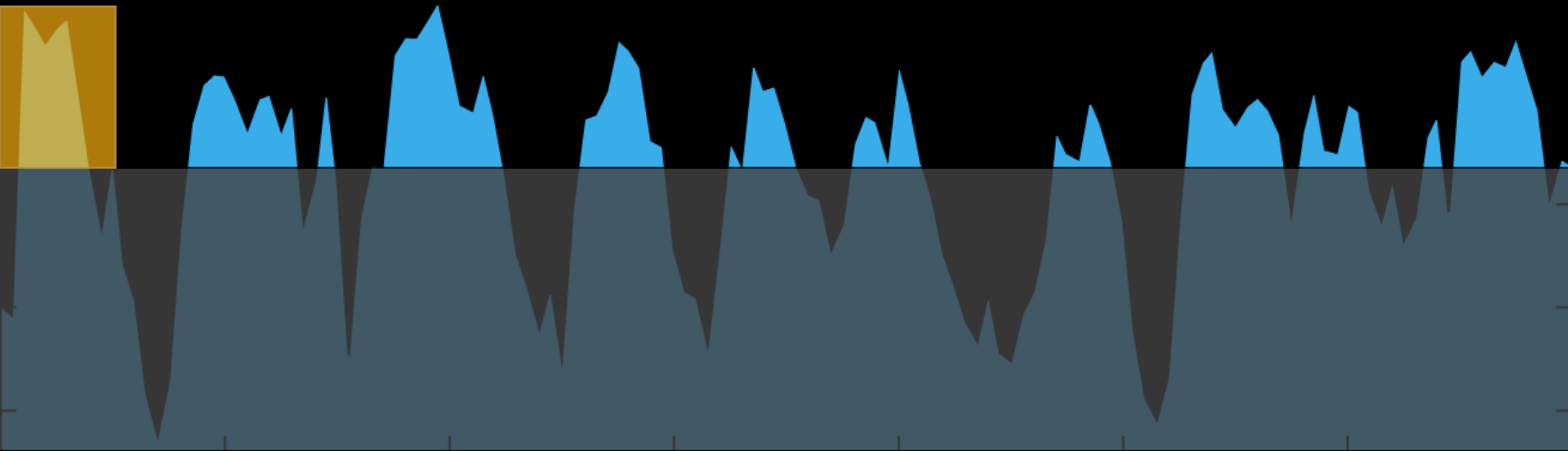
# Solving an audio captcha



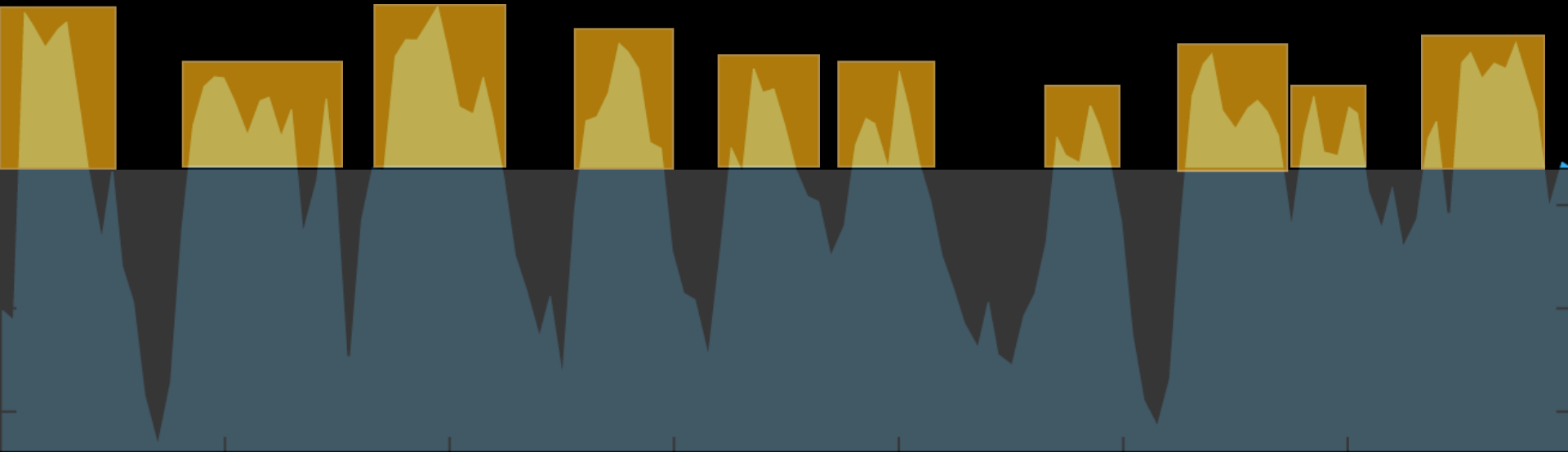
# Solving an audio captcha



# Solving an audio captcha

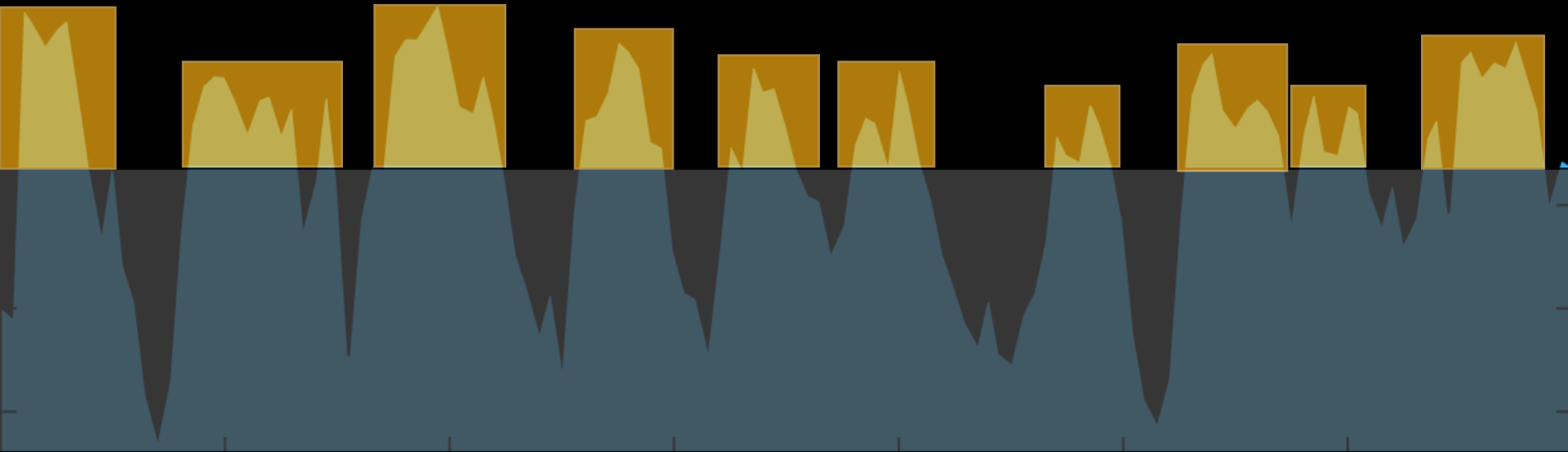


# Solving an audio captcha

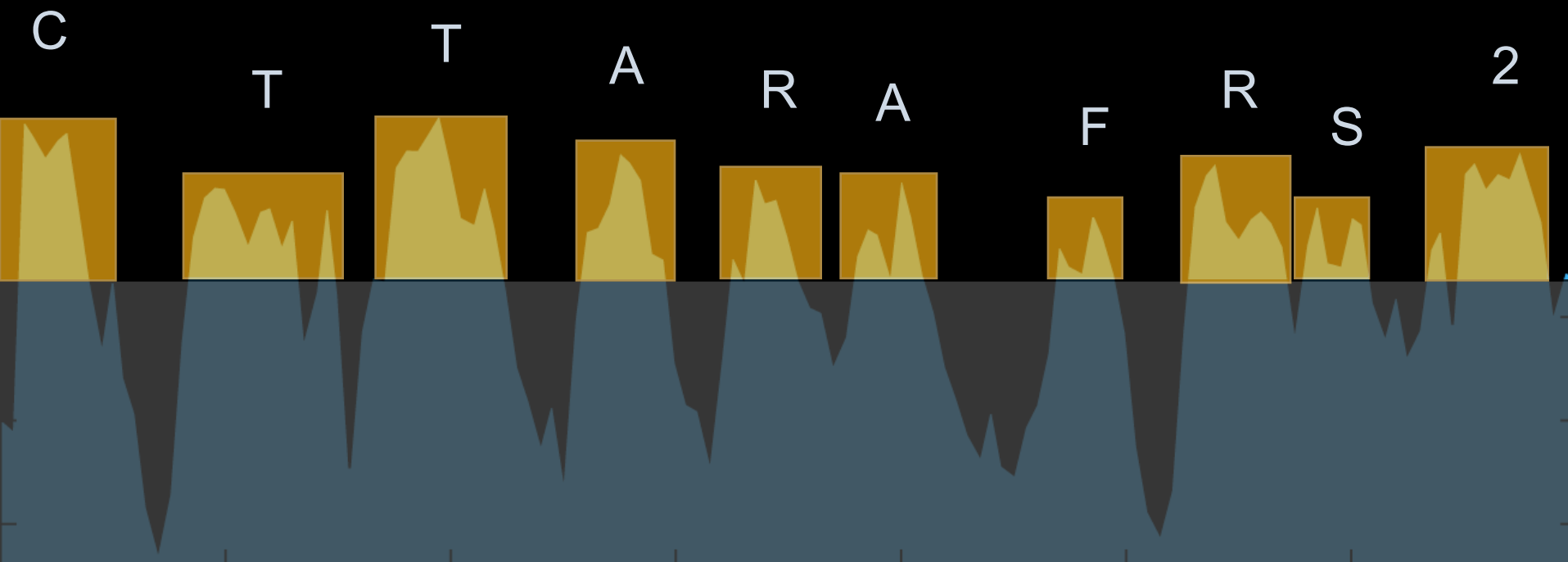


# Solving an audio captcha

C



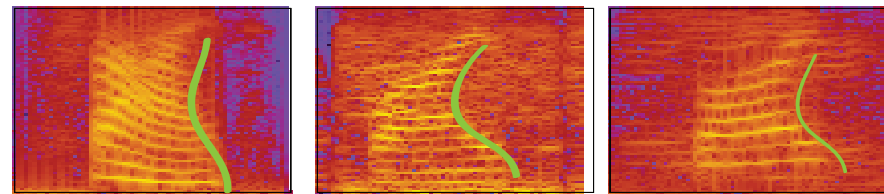
# Solving an audio captcha



# Dealing with random noise

- Statistical learning
- Supervised learning
- RLS (Regularized least square) classifier

5:

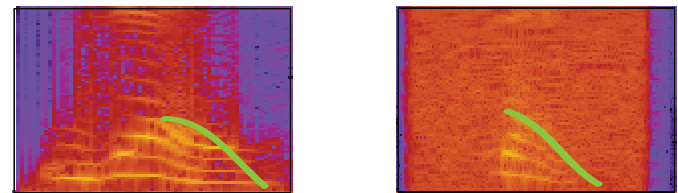


Authorize

eBay

Recaptcha

J:



Authorize

Digg





# Authorize.Net<sup>®</sup>

a CyberSource solution



# YAHOO!<sup>®</sup>

# Microsoft<sup>®</sup>

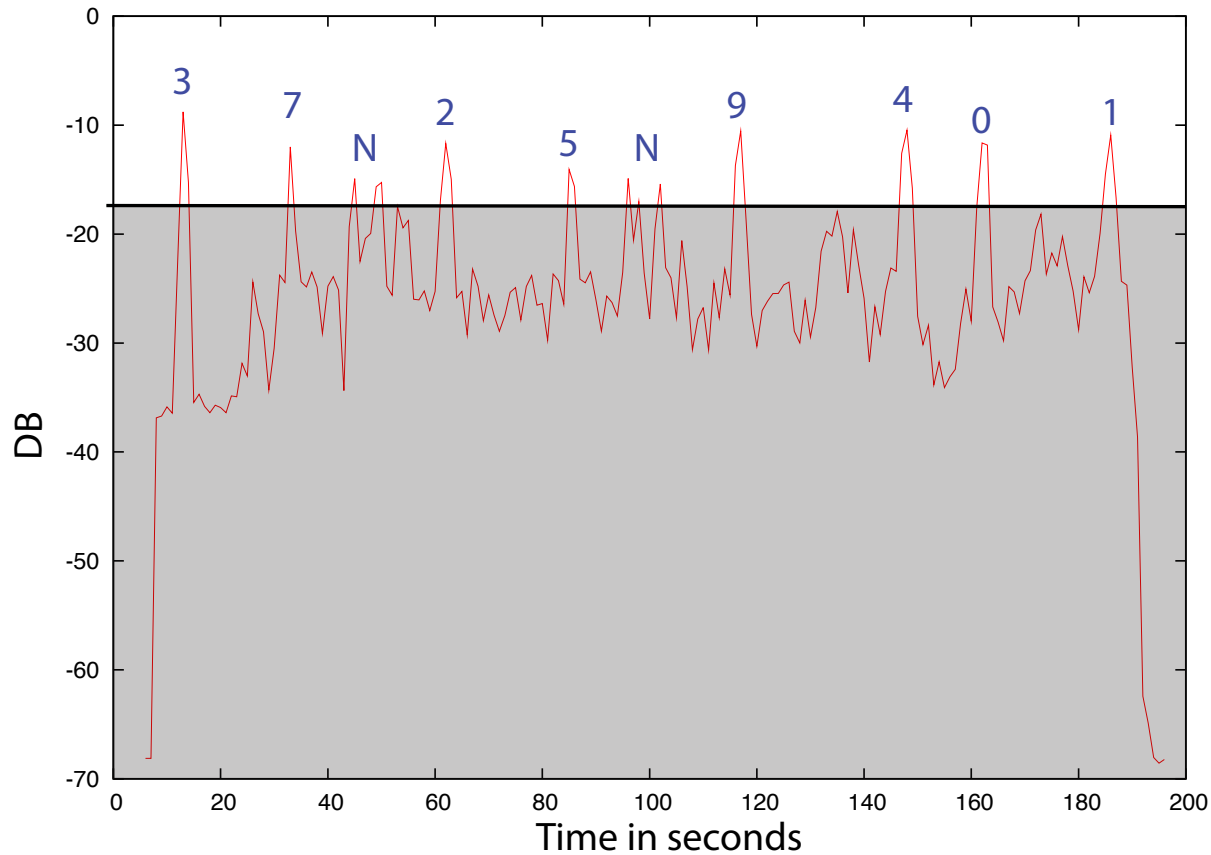


# Results

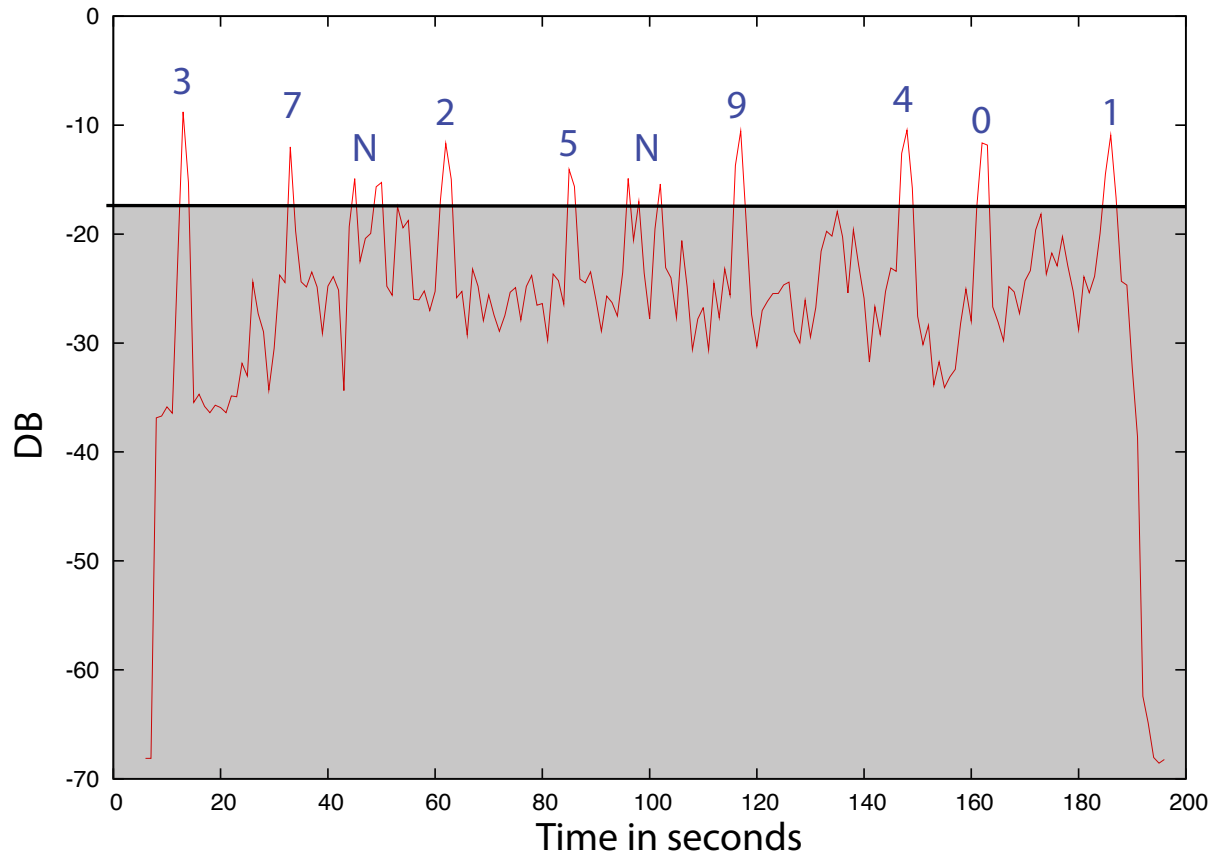
	Length	Coverage	Digit	Captcha
Authorize	5	100	97	<b>89.2%</b>
Digg	5	100	76	<b>41.4%</b>
eBay	6	85.6	92.5	<b>82.9%</b>
Microsoft	10	80.6	89.6	<b>48.9%</b>
Recaptcha	8	99.9	40.5	<b>1.5%</b>
Yahoo	7	99.1	74.7	<b>45.4%</b>



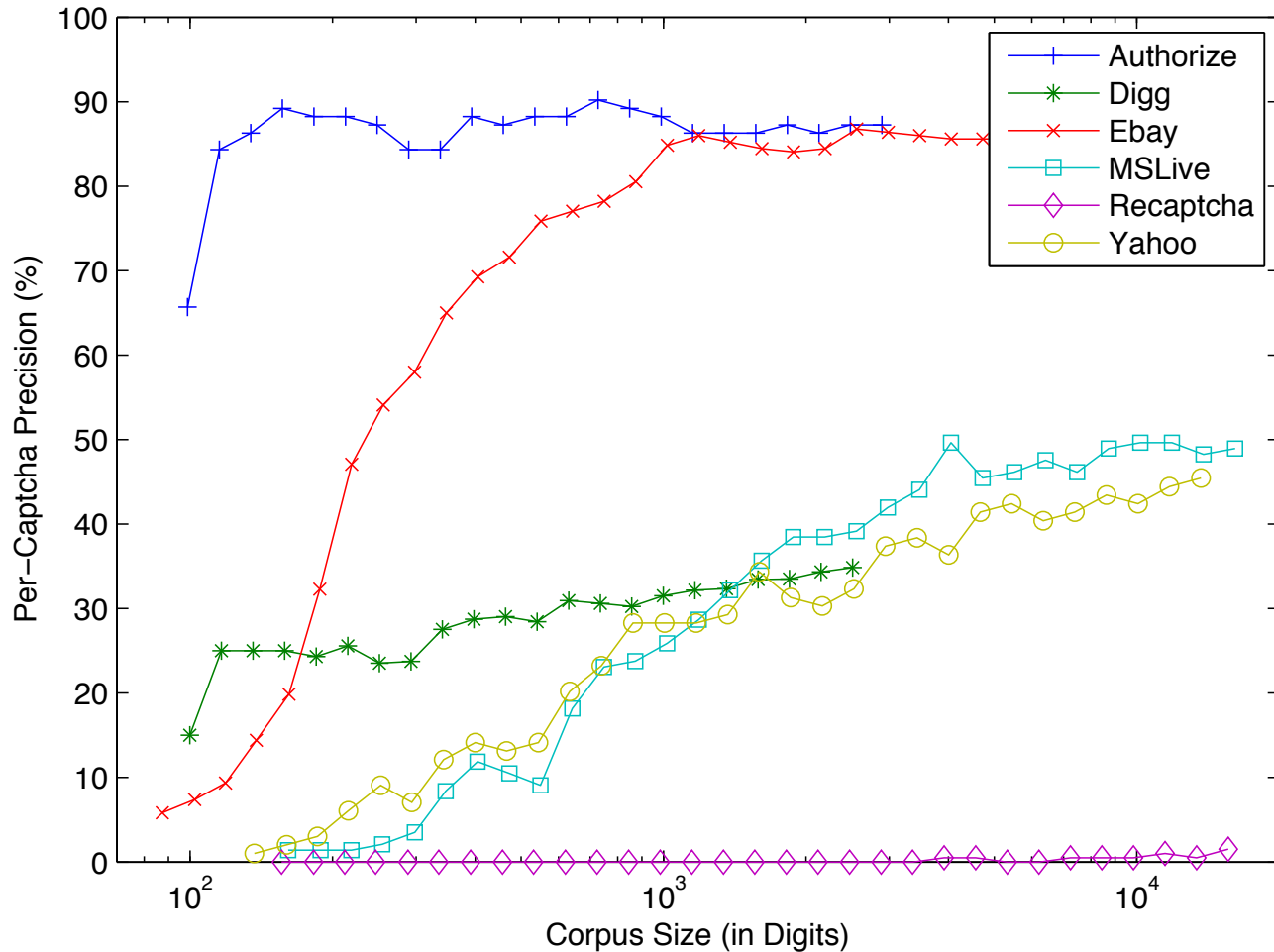
# Recaptcha semantic noise



# Recaptcha semantic noise

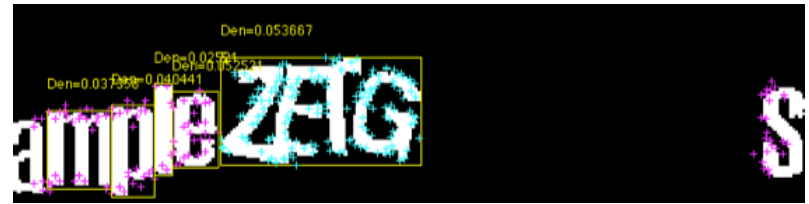


# How many captchas do you need ?



# Video captcha

- Interesting direction -> more design space
- Good for human
- Good for computer :(
  - Working on it



See blog post for more information: <http://elie.im/blog>

# Apply

- Within 3 months
  - Make sure you have a strong captcha scheme (use mine if you want)
  - Ensure that your site is accessible
- Within 6 months
  - Log your captchas failure rate and monitor them
  - Have a backup captcha scheme in case your scheme is broken



Thank you  
Questions ?



Follow-me !  
Twitter: @elie

Captcha research: <http://elie.im/tag/captcha>

