

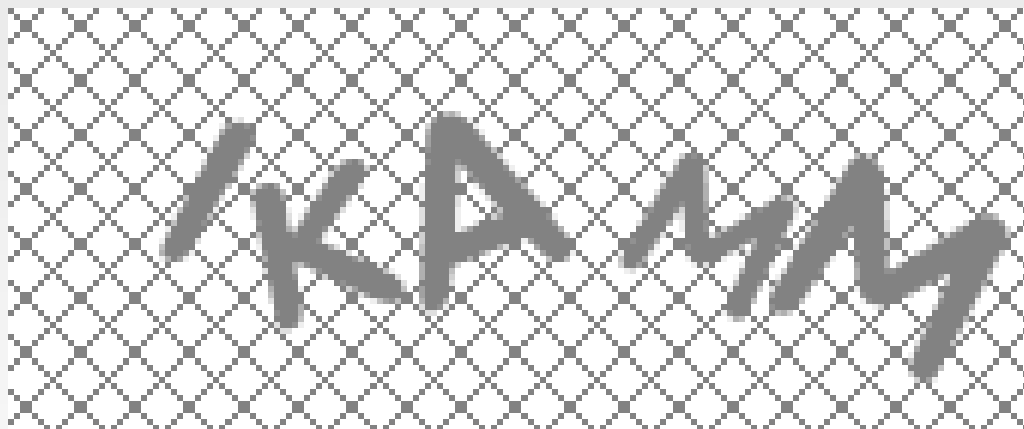
SOOM.cz

Seznam.cz
CAPTCHA API
BREAKER

by Emkei

Seznam Captcha Breaker

- CAPTCHA API
 - <http://captcha-api.seznam.cz/>
 - <http://registrace.seznam.cz/>



Seznam Captcha Breaker

- - Statické pozadí
- - Malá množina písmen [A-N]*
- - Žádná deformace
- - Velké odsazení jednotlivých písmen

- + Rotace a změna velikosti písmen
- + Barva písma splývá s pozadím
- + Vlastní font

Seznam Captcha Breaker

- Vlastní font od Suitcase Type Foundry
 - Seznam Script
 - Seznam Text
- WhatTheFont!
 - <http://new.myfonts.com/WhatTheFont/>

Seznam Captcha Breaker

- ImageMagick
- <http://www.imagemagick.org/>
 - compare
 - convert

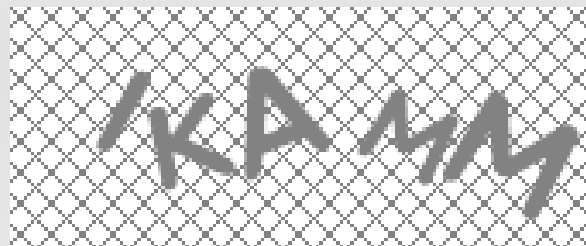
- Ukázka odstranění statického pozadí:

```
compare -compose src background.gif captchaImage.gif  
difference.gif
```

```
convert difference.gif -fill black +opaque white  
-paint 2 result.gif
```

Seznam Captcha Breaker

- Originál



- Odstranění pozadí



Seznam Captcha Breaker

- Monochromatické barvy



- Segmentace (Flood Fill)



Seznam Captcha Breaker

- Segmentace



- Rotace (+/- 45°)



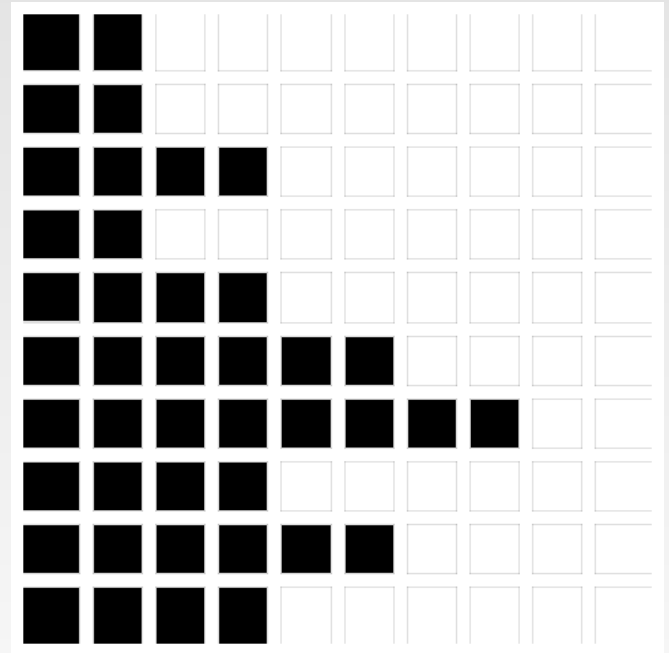
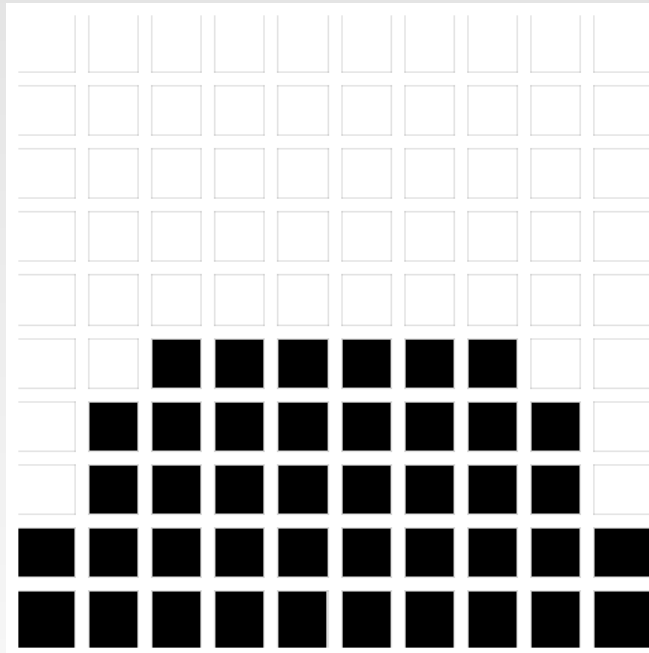
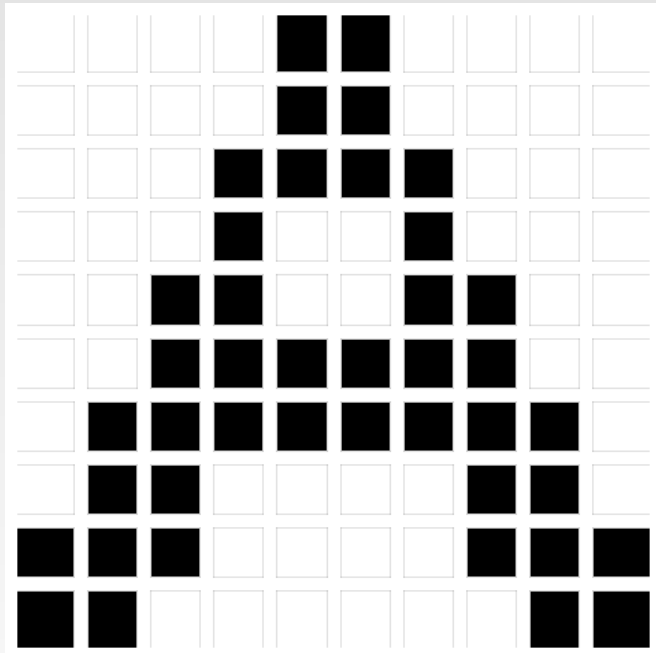
Seznam Captcha Breaker

- OCR
 - GOCR/JOOCR
 - <http://jocr.sourceforge.net/>
 - Múd učení (neuronová síť)

```
gocr -m 16 -m 130 -m 256 -p ./seznam/ -i image.pnm
```

Seznam Captcha Breaker

Princip fungování OCR
(horizontální a vertikální histogram)



Seznam Captcha Breaker

- Funkce „Aktualizuj kód“
 - <http://registrace.seznam.cz/newCaptcha>
 - <http://registrace.seznam.cz/captchalimage?hash=X>

5+ variant každé captcha!

- Proof of Concept

Seznam Captcha Breaker



Seznam Captcha Breaker

Děkuji za pozornost!

Emkei, SOOM.cz, 2011