

Text Detection in Natural Scenes

Matej Straka
mstrakam@gmail.com

Obsah

Predstavenie

ciele a výzvy

problémy

aplikácie

Metódy detekcie textu

ako to celé funguje

detektory, deskriptory

klasifikácia

ICDAR 2011 Robust Reading Competition

Canny Edge

Shape context

SIFT (Scale Invariant Feature Transform)

HoG (Histogram of Gradients)

Úvod



LE RALLY
SALLE
BAR TABAC LE RALLY
un ç*
s):Kit-
num;

Ciele a výzvy

Detekcia a lokalizácia textu v rôznom prostredí

Nezávislosť od:

Veľkosti písma

Fontu

Farby

Jazyka

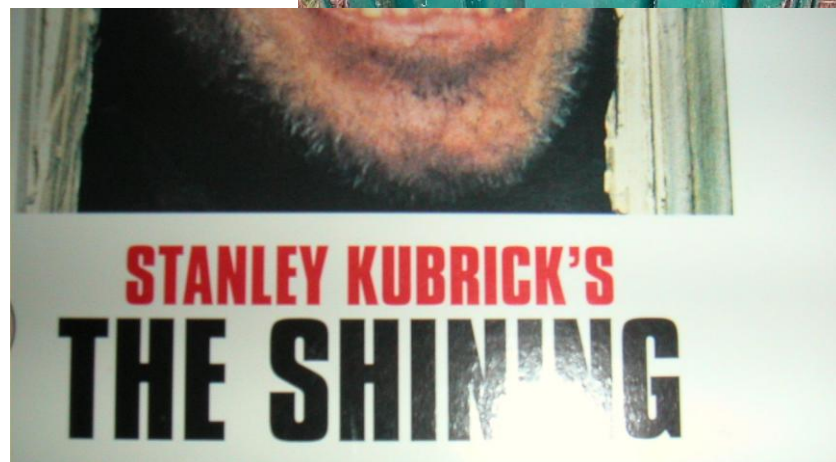
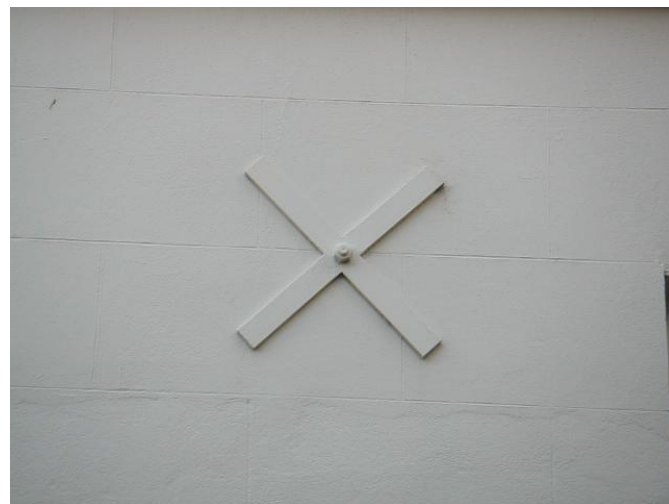
Pozadia textu

Polohy



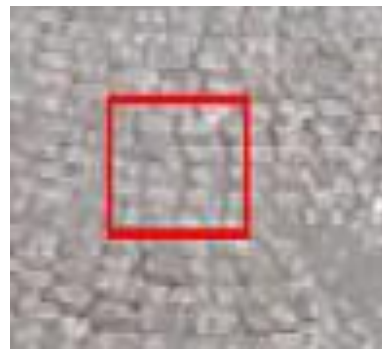
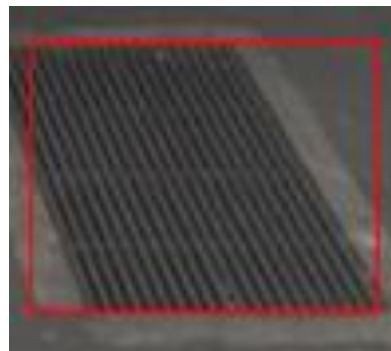
Problémy

- Osvetlenie
- Zaostreňenie
- Zašumený obraz
- Komplexné prostredie
- Nie na každom obrázku je text



Problémy

- Označovanie:
stromov
zábradlí
okien
tehlových múrov



Aplikácie

- Načo je vôbec dobré rozpoznávanie textu?
- Detekcia SPZ
- Rozpoznávanie dopravných značiek
- Lokalizácia robotov
- Pomôcka pre slabozrakých



Aplikácie

- Aplikácie pre klasifikáciu obrázkov
- Marketing, cielená reklama



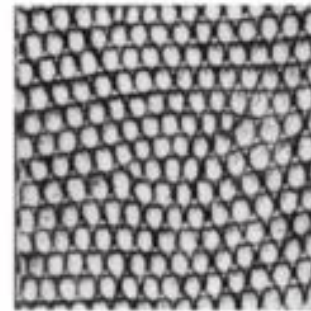
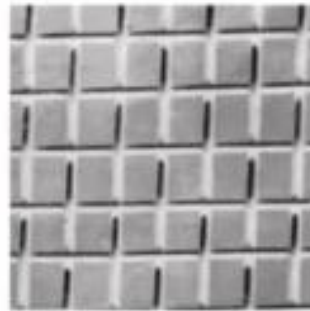
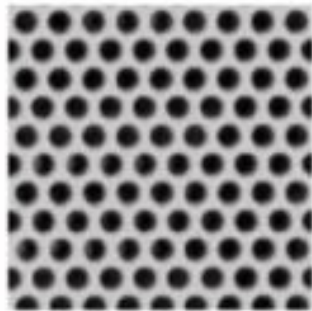
Ako to funguje ?



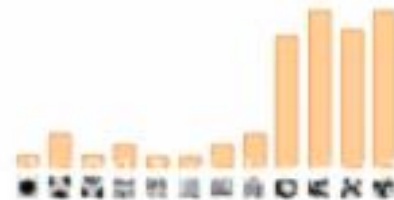
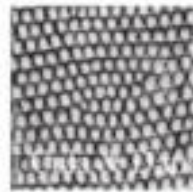
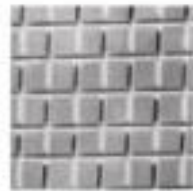
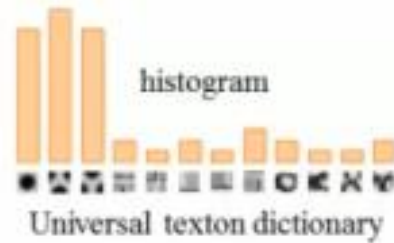
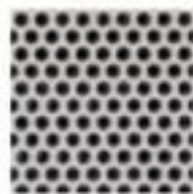
LE RALLY
SALLE
BAR TABAC LE RALLY
un q*
s):Kit-
num;

Ako to funguje ?

- Elementy opisujúce prostredie
- Zaujímavé časti obrázku



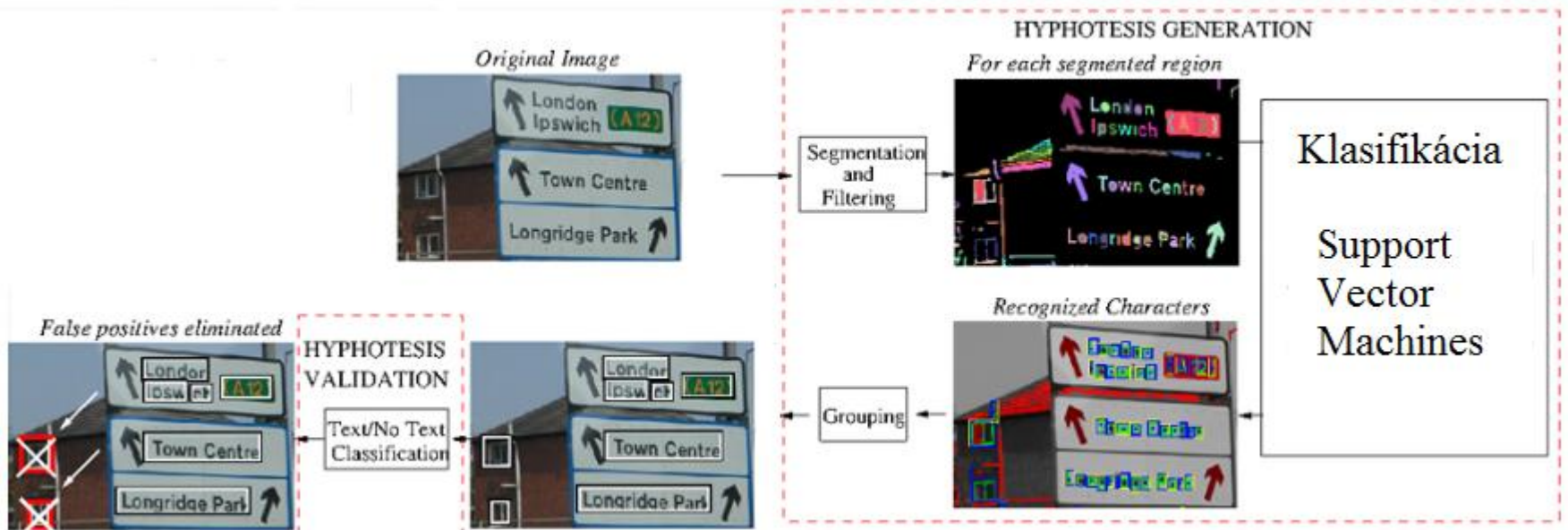
Ako to funguje ?



Ako funguje detekcia textu?

- Detekcia
- Deskripcia
- Klasifikácia
- Verifikácia
- Binarizácia
- Rozpoznávanie písmen/slov pomocou OCR

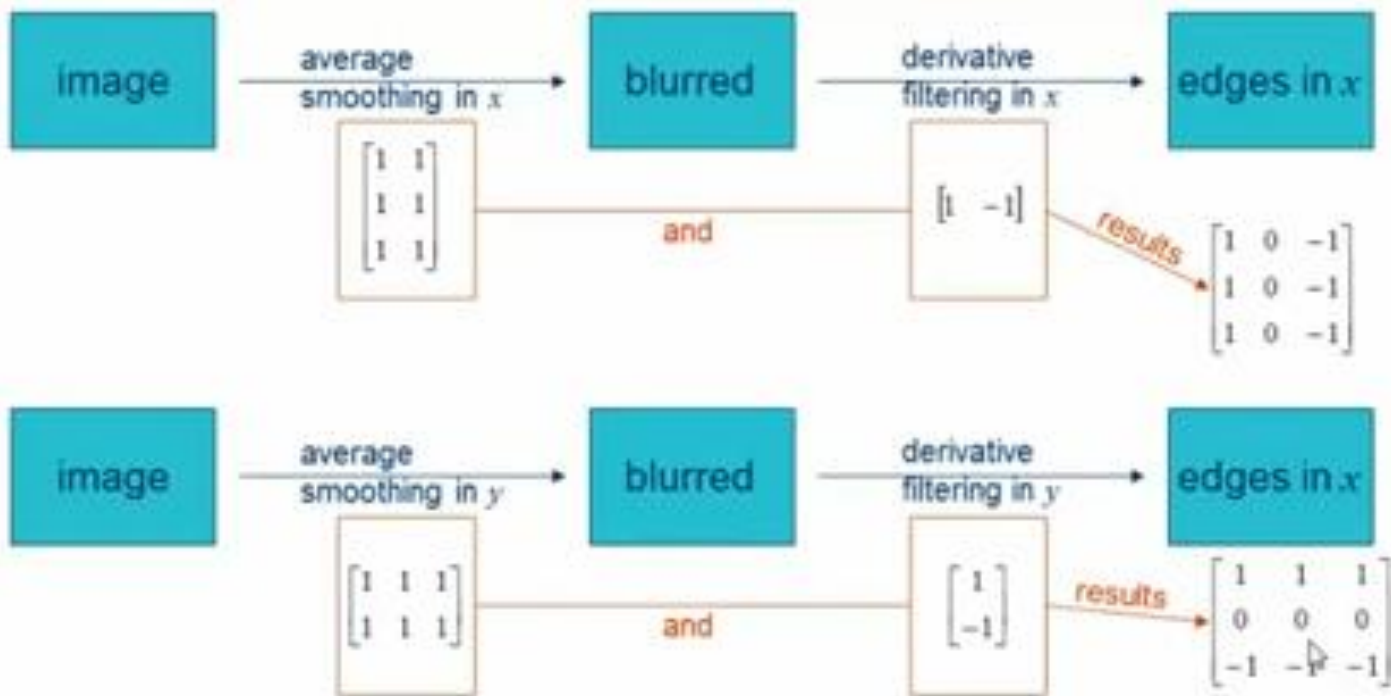
Model



Detekcia

- Nájdienie častí kde by mohol byť text
- Detektory:
 - Sobel
 - Prewitt
 - Canny Edge
 - Harris Detector
 - SIFT (Scale Invariant Feature Transform)

Detekcia hrán



Detekcia hrán

- Canny Edge detektor:
- Vyhladenie s Gausovským filtrom

$$S = I * g(x, y) = g(x, y) * I$$

- Výpočet derivácií

$$\nabla S = \nabla (g * I) = (\nabla g) * I$$

$$\nabla S = \begin{bmatrix} g_x \\ g_y \end{bmatrix} * I = \begin{bmatrix} g_x * I \\ g_y * I \end{bmatrix}$$

- Výpočet veľkosti a orientácie gradientu

$$\text{magnitude} = \sqrt{(S_x^2 + S_y^2)}$$

$$\text{direction} = \theta = \tan^{-1} \frac{S_y}{S_x}$$

Detekcia hrán



Deskripcia

- Opis detekovaných častí obrázka - vektory
- Deskriptory:
 - SIFT
 - HoG (Histogram of Gradients)
 - Shape Context

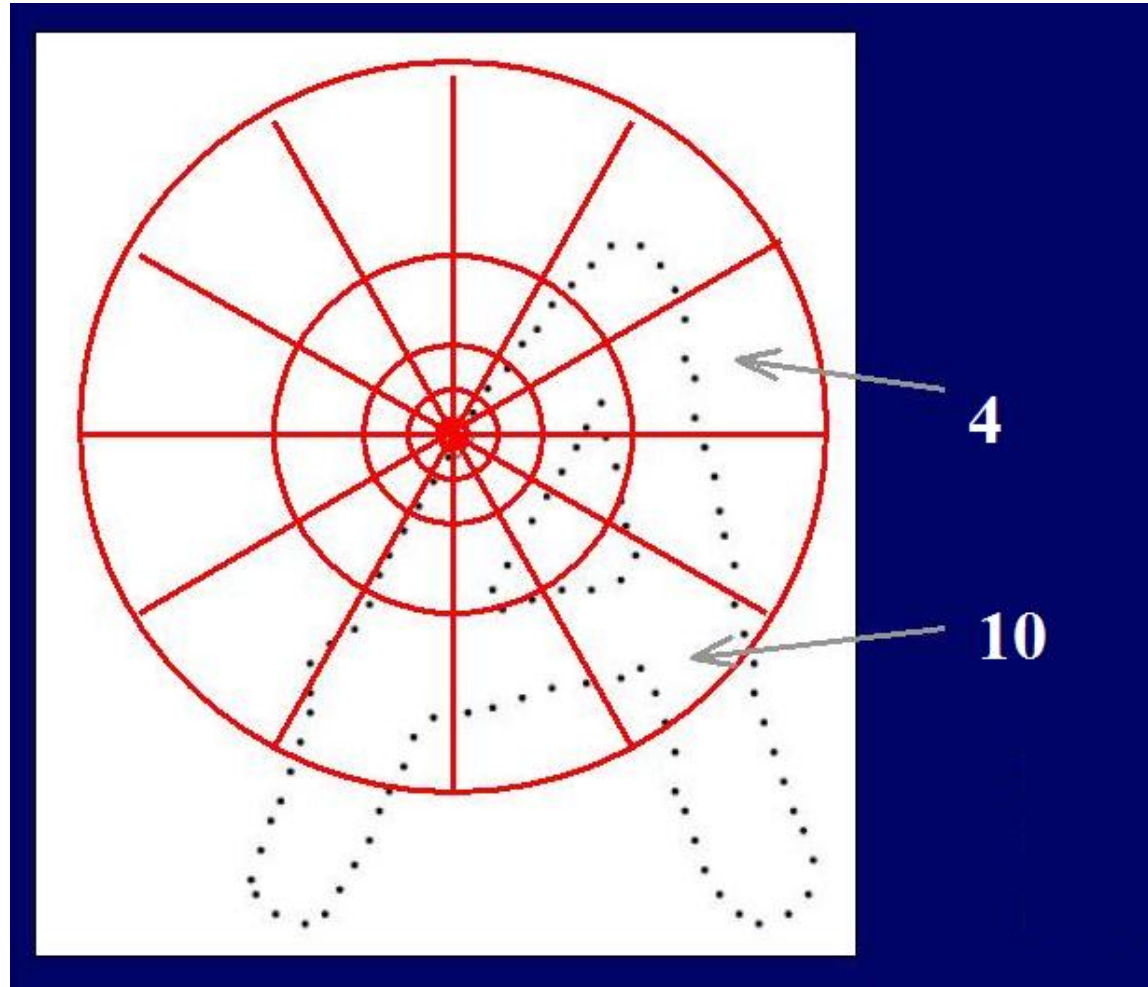
Deskripcia

- Shape context

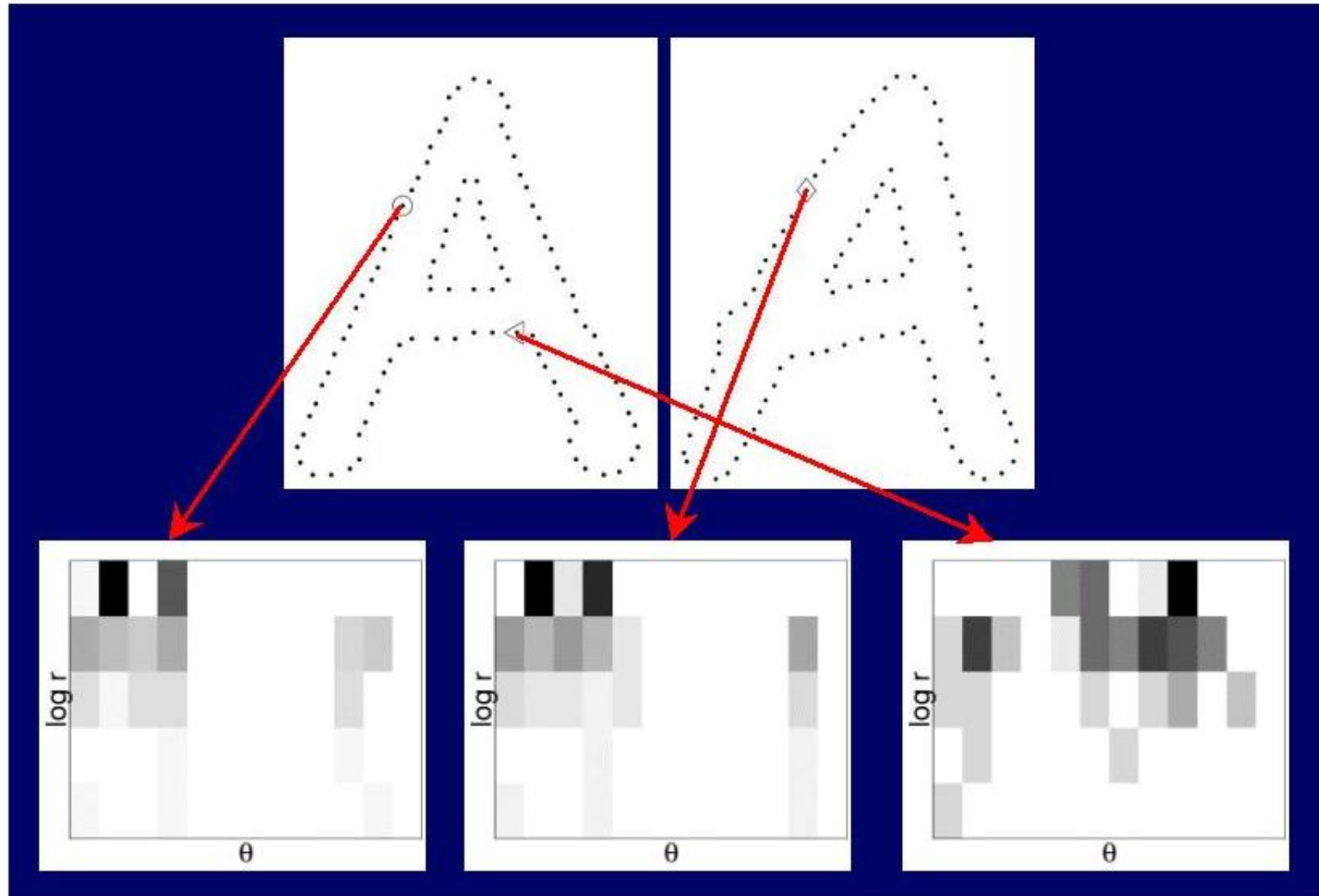
$$h_i(k) = \#\{q \neq p_i : (q - p_i) \in \text{bin}(k)\}$$

- Minimalizácia:

$$C_S = \frac{1}{2} \sum_{k=1}^K \frac{[g(k) - h(k)]^2}{g(k) + h(k)}$$

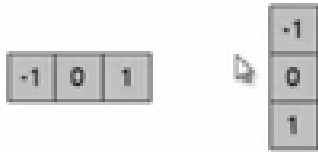


Deskripcia



Deskripcia

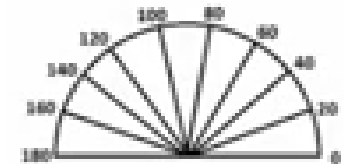
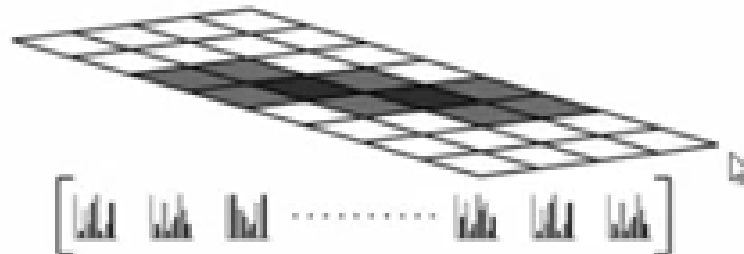
- HoG:
- Výpočet centrováných horizontálnych a vertikálnych gradientov, veľkosti, orientácie



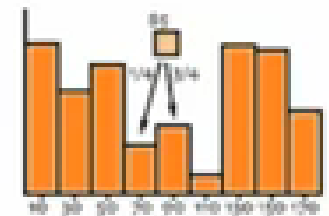
- Rozdelenie obrázka na bloky – čiastočné prekrývanie
- Bloky sa rozdelia na bunky

Deskripcia

- Vytvorenie histogramu
- Z histogramu sa vytvorí vektor - deskriptor



9 Bins



Bin centers

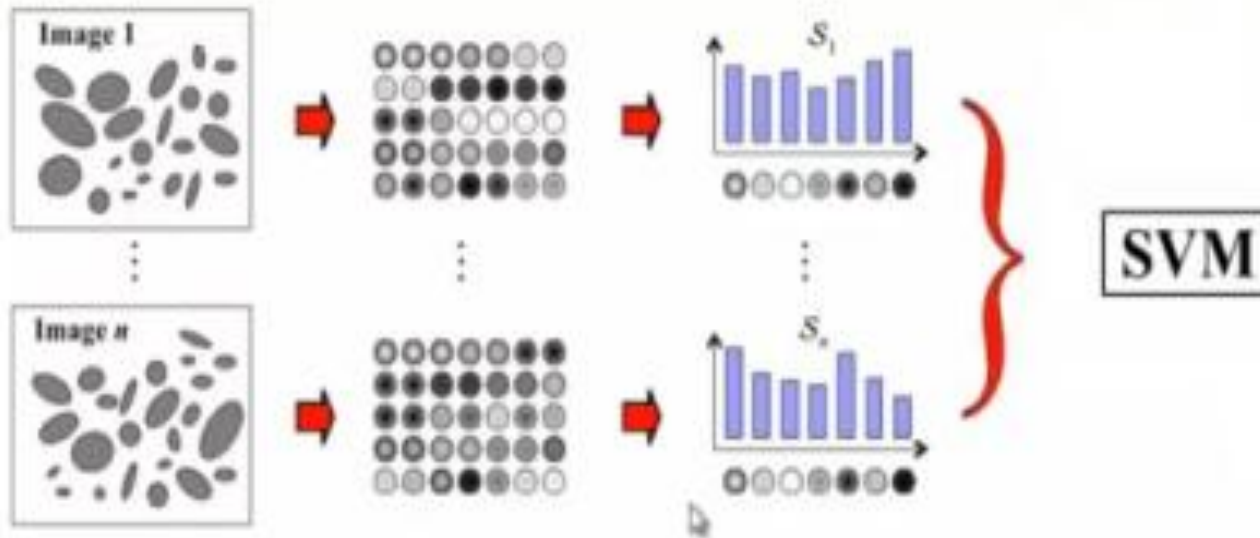
Klasifikácia

- Pozitívne trérovacie vzory - s textom



- Negatívne trérovacie vzory - bez textu
- Klasifikácia: SVM, neurónové siete

Klasifikácia



Klasifikácia

- Klasifikácia



- Lokalizácia spojenie písmen do slov a ozačenie



Postup detekcie textu



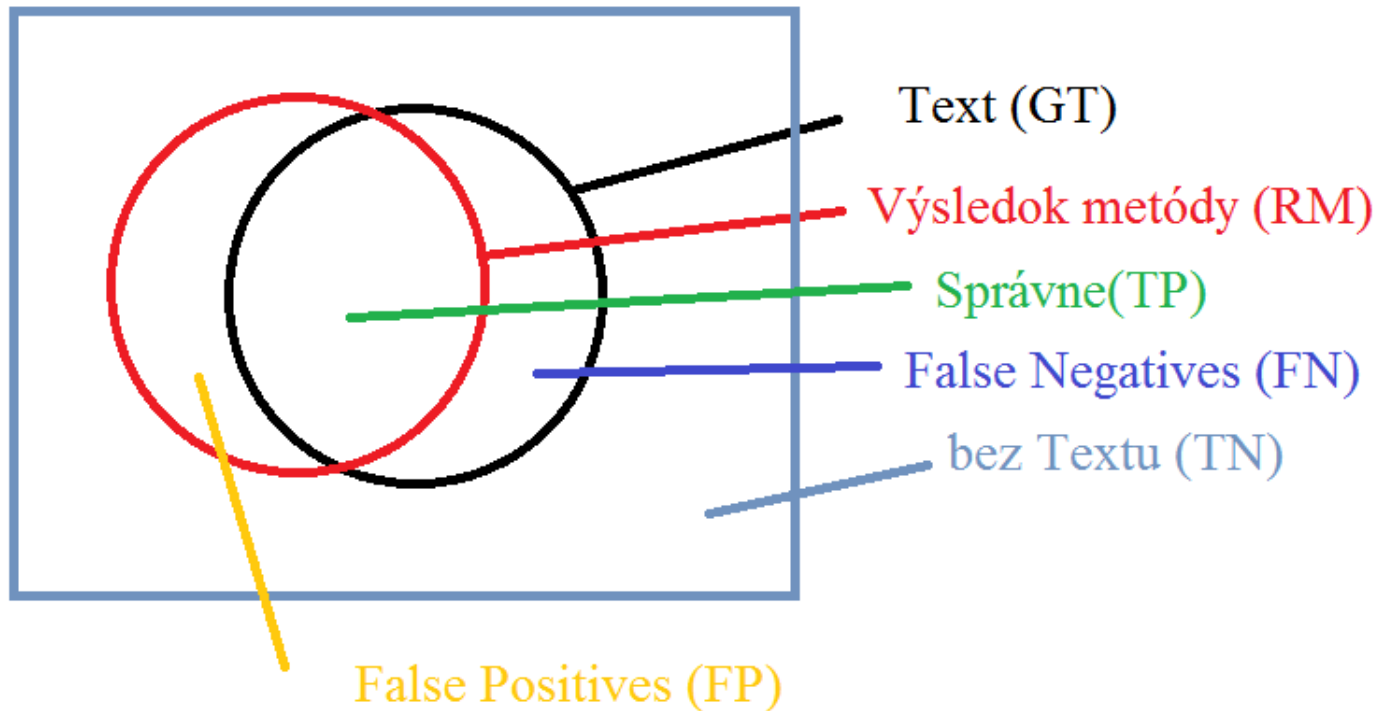
2011 Robust Reading Competition

- 9 metód
- Identifikovať text a označiť ho
- Dataset:
- Voľne dostupný
- Trénovacie (20. apríl)
- Testovacie (1. Jún)
- Výsledky (3. Jún)

2011 Robust Reading Competition

$$\text{precision} = \frac{GT \cap RM}{RM} = \frac{TP}{RM}$$

$$\text{recall} = \frac{GT \cap RM}{GT} = \frac{TP}{GT}$$



2011 Robust Reading Competition

Method	Recall	Precision
Kim's Method	62.47	82.98
Yi's Method	58.09	67.22
TH-TextLoc System	57.68	66.97
Neumann's Method	52.54	68.93
TDM_IACS	53.52	63.52
LIP6-Retin	50.07	62.97
KAIST AIPR System	44.57	59.67
ECNU-CCG Method	38.32	35.01
Text Hunter	25.96	50.05

Ďakujem za pozornosť

Zdroje

- Digital image processing - Rafael C. Gonzales, 3th edition
- Efficient Shape Matching Using Shape Contexts - G. Mori, J. Malik, S. Belongie
- Pattern classification – R.O Duda, P.E. Hart, D.G. Stork, 2th edition
- ICDAR 2011 Robust Reading Competition Challenge 2: Reading Text in Scene Images - A. Shahab, F. Shafait, A. Dengel
- A Computational Approach to Edge Detection - J. Canny
- Detecting Texts of Arbitrary Orientations in Natural Images - Cong Yao, Xiang Bai, Wenyu Liu, Yi Ma, Zhuowen Tu
- Distinctive Image Features from Scale-Invariant Keypoints - David G. Lowe
- Ďalšie internetové zdroje, [youtube.com](https://www.youtube.com), [wikipedia.org](https://www.wikipedia.org)