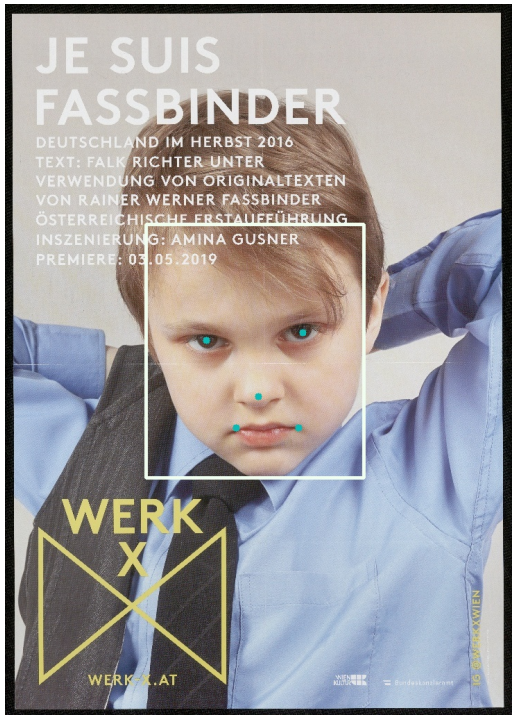
Porovnání metod pro detekci obličejů bez omezujících podmínek

Autorka: Bc. Soňa Drocárová
Vedúci práce: Ing. Adam Novozámský, Ph.D.
Konzultantka: Ing. Jitka Kostková, Ph.D.

FJFI ČVUT

12.9.2022

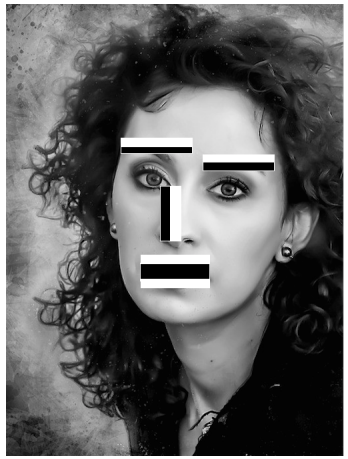
Detekcia tváre

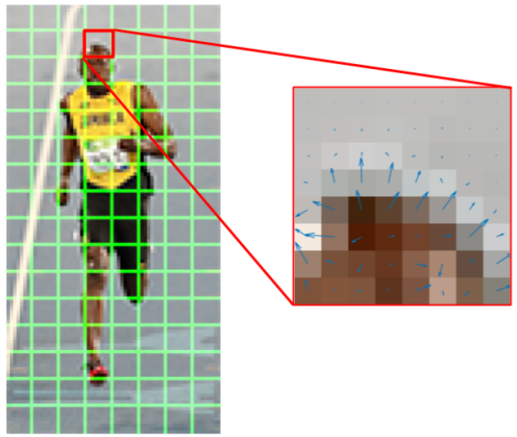


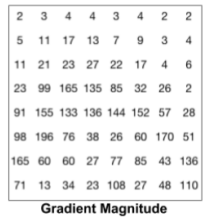
Sťažujúce podmienky

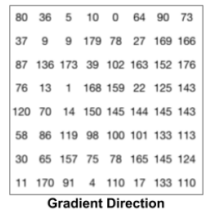
Algoritmy dekecie

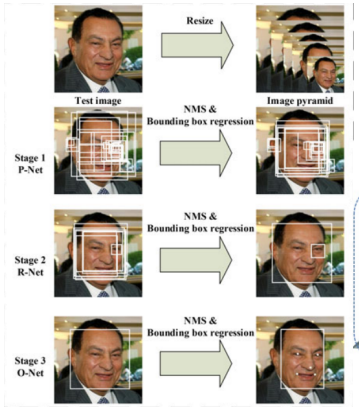
- Viola - Jones
- HOG+SVM
- MTCNN
- RetinaFace

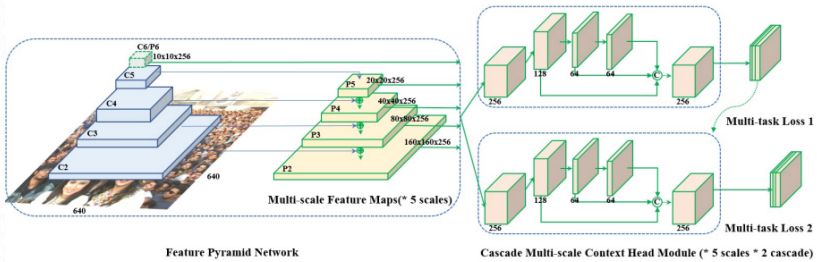












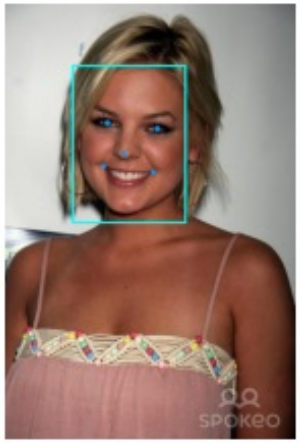
Datasety

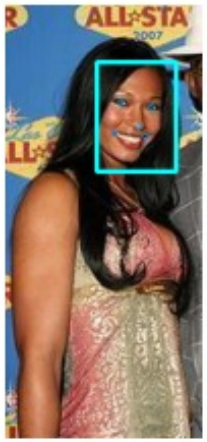
- WFLW dataset[1] - 2 500 obrázkov

[1] Wayne Wu et al. "Look at Boundary: A Boundary-Aware Face Alignment Algorithm". In: *CVPR*. Jún 2018.

Datasety

- CelebA dataset[2] - skoro 20 000 obrázkov





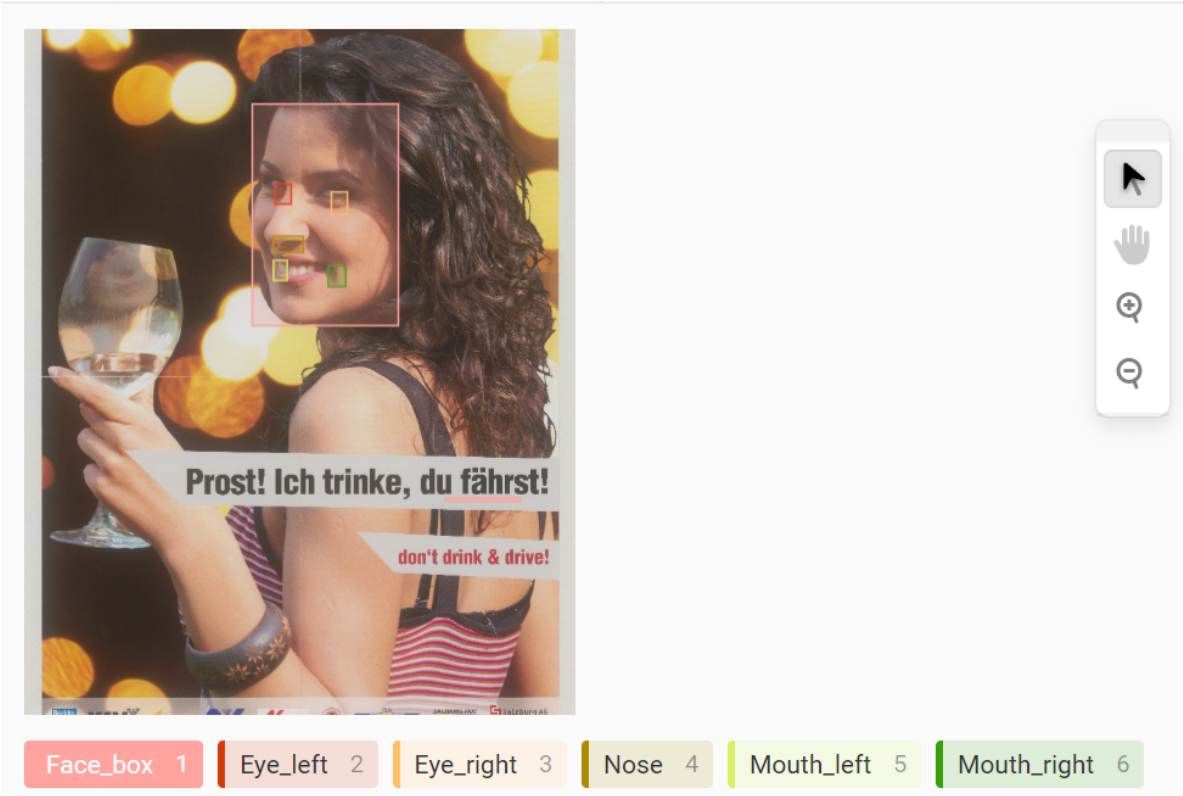
[2] Ziwei Liu et al. "Deep Learning Face Attributes in the Wild". In: *Proceedings of International Conference on Computer Vision (ICCV)*. Dec. 2015.

Datasety

- vlastný dataset z pridelených dát Viedenskej knižnice (označených skoro 130 obrázkov)

Datasety

- vlastný dataset z pridelených dát Viedenskej knižnice (označených skoro 130 obrázkov)



Metriky pre detekciu orientačných bodov tváre

Normalized Mean Error

$$\mathrm{NME} = \frac{1}{K}\sum_{k=1}^{K}\mathrm{NME}_k,$$

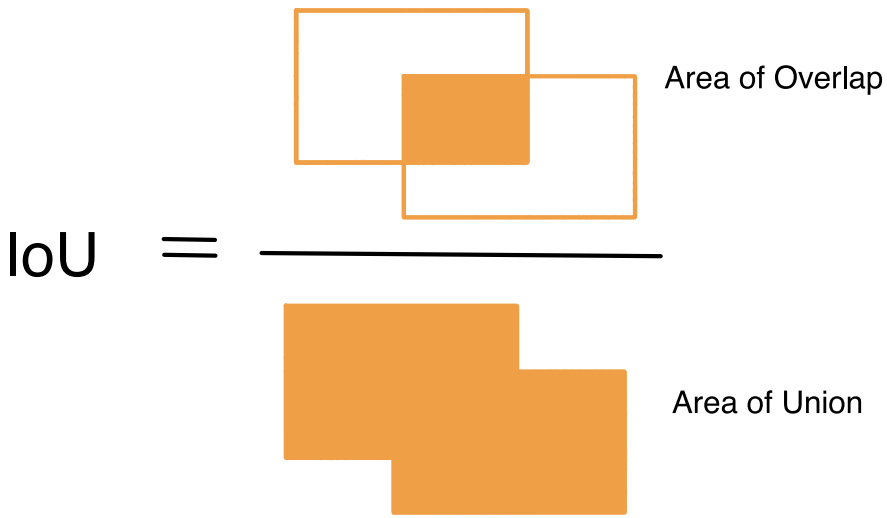
$$\mathrm{NME}_k = \frac{1}{N_L}\sum_{i=1}^{N_L}\frac{\left\|Y_i - \hat{Y}_i\right\|}{d} \times 100$$

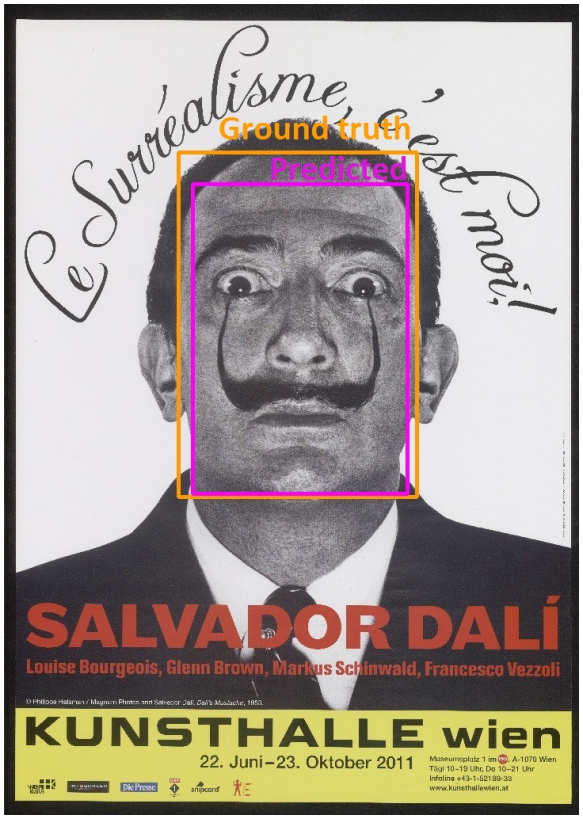
Metriky pre detekciu orientačných bodov tváre

Failure Rate

$$FR = \frac{1}{K} \sum_{k=1}^{K} [NME_k \geq 10\ \%]$$

Metriky pre detekciu tváre





Metriky pre detekciu oblasti tváre

$$\text{Precision} = \frac{tp}{tp + fp}$$

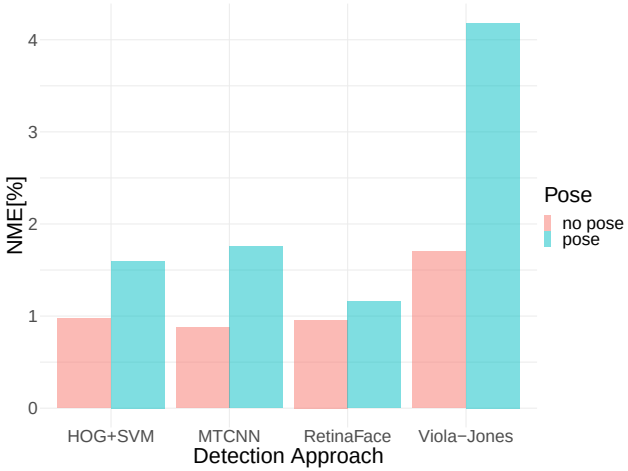
$$\text{Recall} = \frac{tp}{tp + fn}$$

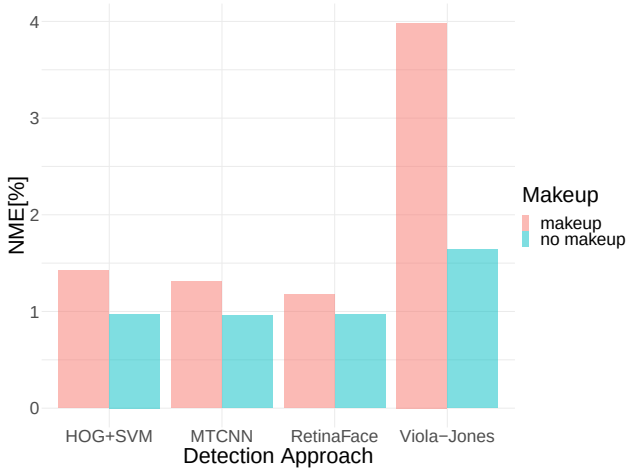
$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

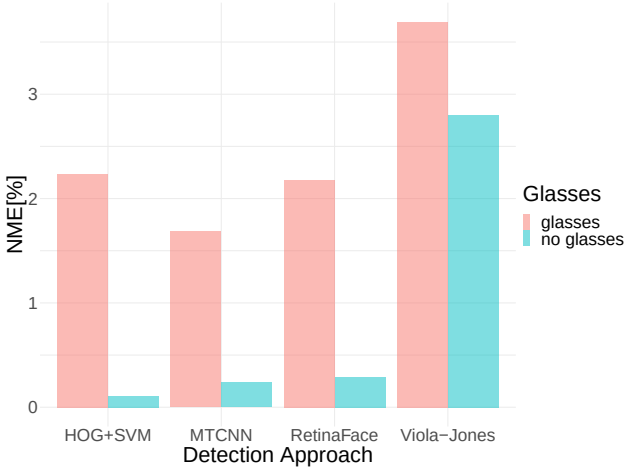
Normalized Mean Error [%]

Face Detection Approach	**Dataset**		
	WFLW	*CelebA*	*Vienna City Library*
RetinaFace	0.9843	1.7279	1.657
MTCNN	0.985	1.1971	1.4929
Viola-Jones	1.8003	2.7953	2.5119
HOG+SVM	1.0089	1.6744	1.9811

Normalized Mean Error



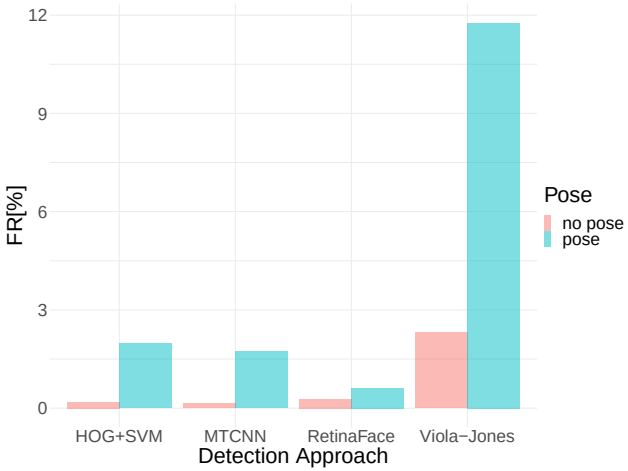


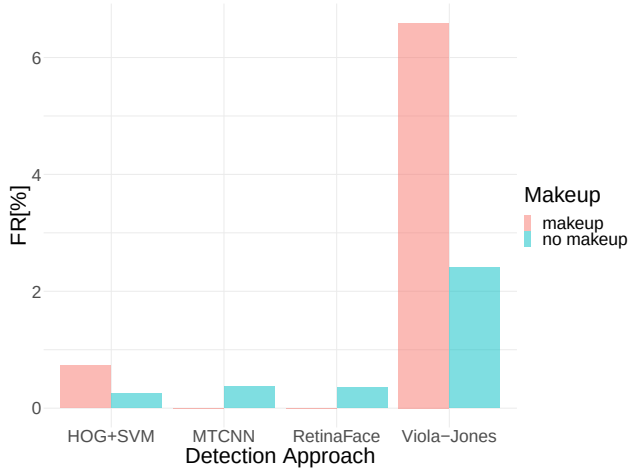


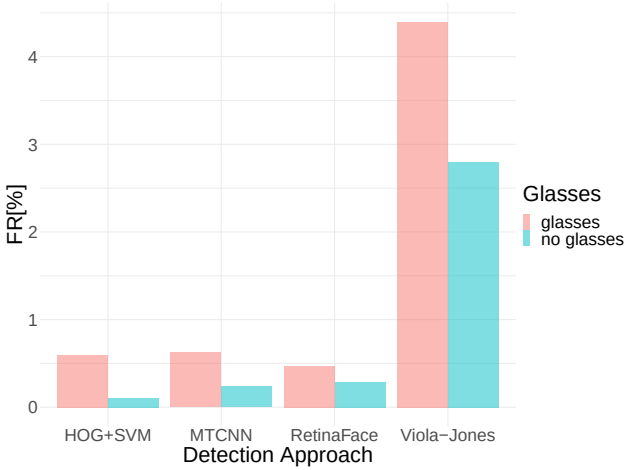
Failure Rate [%]

Face Detection Approach	**Dataset**		
	WFLW	*CelebA*	*Vienna City Library*
RetinaFace	0.3238	0.30122	0.7143
MTCNN	0.3454	0.2621	0
Viola-Jones	2.6946	2.9215	2.4793
HOG+SVM	0.2894	0.1374	0.7752

Failure Rate







Detekcia oblasti tváre

Approach	Dataset	Precision	Recall	F1 score
MTCNN	CelebA	96.45%	99.40%	97.90%
RetinaFace	CelebA	95.04%	99.79%	97.36%
HOG+SVM	CelebA	99.27%	94.80%	96.98%
Viola-Jones	CelebA	91.24%	87.06%	89.10%
MTCNN	VCL	97.74%	94.89%	96.30%
RetinaFace	VCL	88.05%	97.90%	92.72%
HOG+SVM	VCL	100.00%	95.56%	97.73%
Viola-Jones	VCL	88.32%	91.67%	89.96%

Detekcia oblasti tváre

Detekcia oblasti tváre pre rozmazané snímky

Approach	Precision	Recall	F1 score
MTCNN	94.28%	98.02%	96.11%
RetinaFace	91.01%	99.11%	94.89%
HOG+SVM	99.39%	64.68%	78.37%
Viola-Jones	93.95%	75.80%	83.90%

Detekcia oblasti tváre pre tváre s okuliarmi

Approach	Precision	Recall	F1 score
MTCNN	96.89%	98.92%	97.90%
RetinaFace	95.61%	99.23%	97.39%
HOG+SVM	99.40%	90.83%	94.92%
Viola-Jones	91.88%	75.78%	83.06%

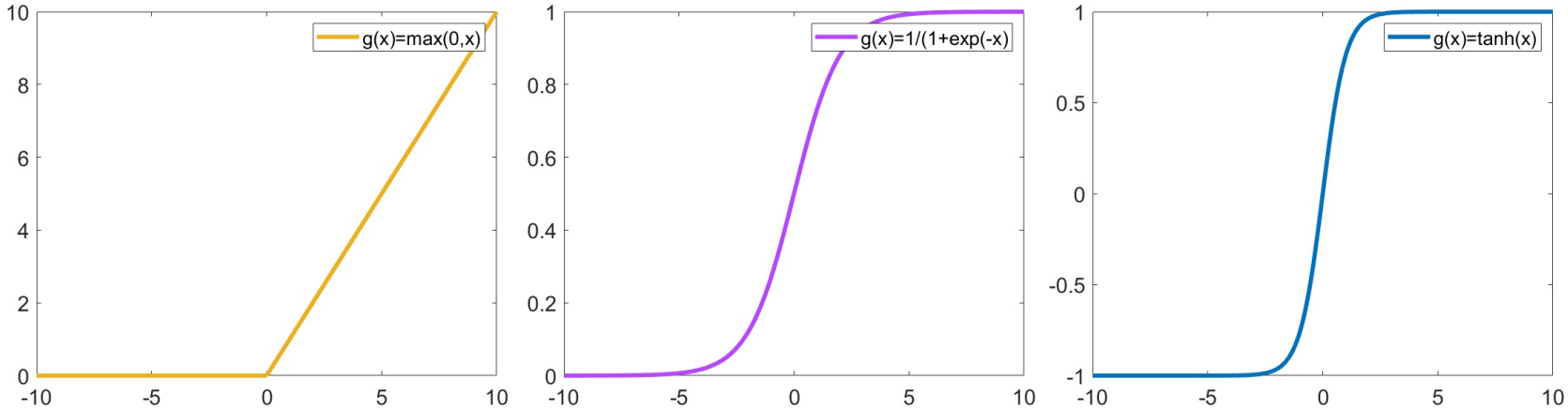
Zhrnutie

- Cieľ práce
- Metódy detekcie tváre
- Výsledky v rôznych sťažujúcich podmienkach

Ďakujem za pozornosť.

Dotazy oponenta

1. Obrázek 1.7: aktivační funkce, sigmoida a hyperbolická tangenta, neodpovídají obrázku. Rovněž ani popisek k sigmoidě „1/(1+exp(x))“ není správně. Mohla byste načrtnout, jak vypadají aktivační funkce sigmoida, hyperbolická tangenta a ReLU a popsat v čem se liší?



Dotazy oponenta

2. Rovnice 4.2: Podle čeho jste se rozhodla stanovit hodnotu normalizačního koeficientu d jako vzdálenost center očí? Jaký vliv na tu konstantu má (horizontální) rotace hlavy?

$$\mathrm{NME}_k = \frac{1}{N_L} \sum_{i=1}^{N_L} \frac{\left\| Y_i - \hat{Y}_i \right\|}{d} \times 100$$