

Rozpoznávání tváří pomocí konvolučních

neuronových sítí

Autor: Ing. Jan Kněžík; Vedoucí: Ing. Radovan Fusek, Ph.D.

VŠB - Technická univerzita Ostrava

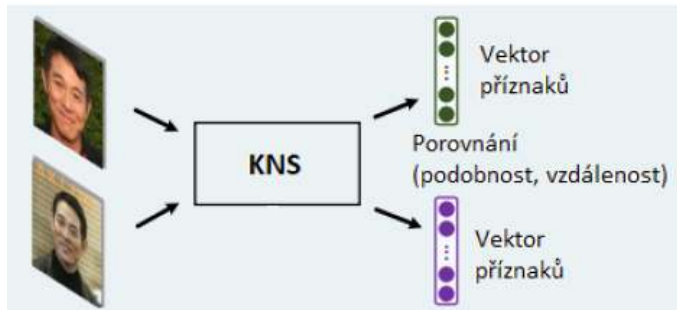
Motivace

Rozpoznávání tváří je jedna z metod pro identifikaci nebo verifikaci jedince, kterou lze využít zejména v sektoru bezpečnosti. K tomu se v posledních letech skoro výhradně používá hluboké učení, přesněji konvoluční neuronové sítě. Klasifikace obrazů pomocí konvolučních neuronových sítí dosahuje daleko lepší přesnosti než tradiční přístupy ve strojovém učení, ale také je pro natrénování sítě potřeba velkého množství trénovacích obrazů.

Pro natrénování kvalitního modelu je mimo jiné zapotřebí správně zvolit architekturu sítě, optimalizátor, regularizační techniky a chybovou funkci.

Postup

V praktickém použití systému pro rozpoznávání obličejů se často přidávají nové osoby a nebo je pro osoby k dispozici pouze pár obrázků. Z toho důvodu je zde sít' natrénována na osobách, které se newyskytují v testovacím datasetu. To znamená, že natrénována sít' slouží při testování pro extrakci příznaků. Tato práce je zaměřena na verifikaci obličejů, takže extrahované vektory příznaků jsou poté porovnány podle kosinovi podobnosti.



Předzpracování obrazů

Trénovací obrazy (903 876 obličejů, 10 559 osob) i testovací obrazy (6000 párů obličejů) jsou předzpracovány v následujících krocích:

- Detekce obličejů pomocí *Multi-task Cascaded CNN*
- Zarovnání tváří s transformační maticí
- Ořezání tváře
- Horizontální zrcadlení (pro zvětšení trénovacího datasetu a vytvoření robustnějších vektorů příznaku z testovacích obrazů)
- Normalizace pixelů do rozmezí od -1 do 1



Trénování

Při trénování modelů bylo experimentováno s různými variantami architektur *AlexNet*, *Inception*, *VGG* a *ResNet* a jejich dalšími modifikacemi. Pro vytvoření více diskriminačních příznaků jsou navíc otestovány chybové funkce *Center loss* a *ArcFace*, které byly vyvinuty přímo pro rozpoznávání obličejů.

Výsledky

Vybraný testovací dataset pro verifikaci obličejů se nazývá *Labeled Faces in the Wild* a obsahuje 3000 pozitivních a 3000 negativních párů. Lidská přesnost na tomto datasetu je 97,5 %.

Nejvyšší dosažená přesnost v této práci je **98,87 %** s modifikovanou architekturou **Preactivation ResNet** s 34 vrstvami a chybovou funkcí **ArcFace**. Také bylo dokázáno, že všechny kroky při předzpracování jsou prospěšné pro výslednou přesnost.