

Adaptácia neurónových sietí pre problém počítačového videnia v reálnom čase

Vedúci práce: Mgr. Alexander Szabari PhD.

Bc. Tomáš Kekeňák

V posledných rokoch sme mohli byť svedkami veľkého rozmachu strojového učenia a umelej inteligencie. Najväčšie pokroky sa sa podarilo dosiahnuť v oblasti hĺbkového učenia (deep learning) pomocou hlbokých neurónových sietí. Tieto prediktívne algoritmy sa podarilo aplikovať na množstvo takých problémov, kde v minulosti algoritmy strojového učenia mali veľmi zlú úspešnosť.

Veľké úspechy sa dosiahli v oblastiach počítačového videnia, rozpoznávania hlasu a textových analýz. Tieto úspechy sú najmä následkom toho, že teraz máme k dispozícii oveľa väčšiu a lacnejšiu výpočtovú silu a je dostupné obrovské množstvo tréningových dát pre tieto algoritmy.

My sa v tejto práci budeme venovať počítačovému videniu, konkrétnejšie detekcii tváří. Detekcia tváří má obrovské využitie v reálnom živote, či už ide o hľadanie zločincov, bezpečnostné systémy budov, odblokovania nášho smartphonu až po vyhľadávania mien osôb na fotke na sociálnych sieťach.

Detekcia tváří, ako aj iné podobné algoritmy je ale výpočtovo veľmi zložitá, lebo neurónové siete, ktoré riešia tento problém majú obrovské množstvo skrytých neurónov a konvolučných filtrov, aby dosiahli dostatočnú presnosť pri predikcii.

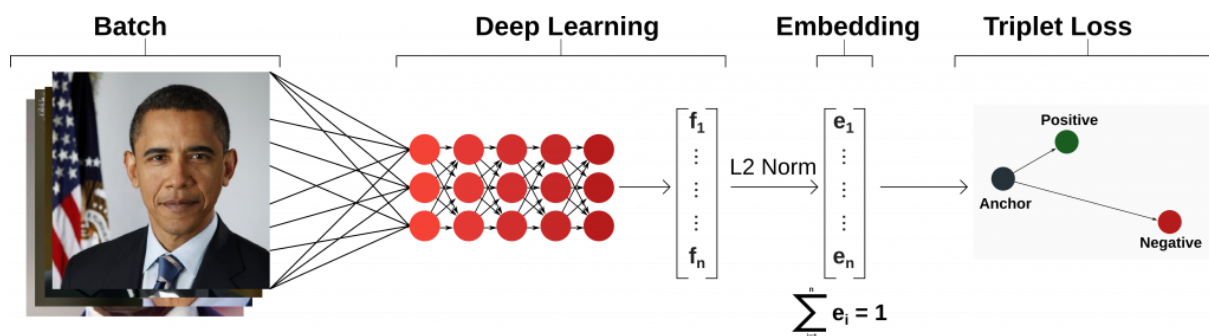
V aplikáciách je ale žiadúce, aby takéto detekčné systémy mohli fungovať aj na mobilných telefónoch alebo na rôznych embedded zariadeniach. Tieto zariadenia nie sú výpočtovo také výkonné, čo má za následok, že mnoho výpočtov trvá veľmi dlho a aplikácie nestíhajú fungovať v reálnom čase. Ďalším dôležitým aspektom je, že veľakrát tieto zariadenia fungujú na baterku, ktorá sa rýchlo vybije pri takýchto zložitých výpočtoch.

Preto by bolo dobré, ak by sme vedeli zjednodušiť modely na detekciu tváří, teda zmenšiť počet neurónov, konvolučných filtrov alebo aj počet skrytých vrstiev v danom modeli. Práve touto tematikou sa budeme zaoberať v tejto práci.

Rozpoznávanie tváří

Najprv sa pozrime ako vlastne funguje rozpoznávanie tváří pomocou neurónových sietí. Existuje viacero modelov, prístupov, ale najlepšie výsledky sa dosahujú s nasledovným:

1. Najprv sa určí alebo určia tváre na danom obrázku, alebo obrázku z videa.
2. Tieto obrázky sa normalizujú a dajú sa na vstup hlbokoj neurónovej siete, ktorá na výstupe vytvorí vektor prislúchajúci danej tvári. Pre tento výstupný vektor by malo platiť, že jeho vzdialenosť od vektorov tváří tej istej osoby (len napr. z inej fotky) by mala byť čím najmenšia. Naopak, vzdialenosť vektora od vektorov tváří iných osôb by mala byť čím väčšia.
3. Tento výstupný vektor sa porovná s tvármi (presnejšie vektormi tváří) osôb v databáze a takto sa určí o ktorú osobu ide. Tu je viacero možností. Najjednoduchšia je, že sa použije binárny klasifikátor (napr. SVM), ktorý rozhodne pre každú osobu v databáze, či ide o daného človeka alebo nie.



V práci sa budeme zaoberať rôznymi možnými prístupmi a porovnáme ich, ktoré by mohli byť najvhodnejšie na použite v real-time aplikáciách.

Minimalizácia neurónových sietí

Neurónové siete sa skladajú s vrstiev neurónov, kde jednotlivé neuróny sú podľa rôznych kritérií poprepájané s neurónmi nasledujúcej vrstvy pomocou váh. Týchto neurónov a váh sú v jednom modeli tisíce a nie každý neurón prispieje k finálnemu výsledku rovnakým dielom. Práve túto vlastnosť využíva väčšina minimalizačných metód. Málo dôležité časti siete sa vynechajú, keďže majú minimálny alebo niekedy aj žiadny efekt na výsledok a len spomaľujú celkový výpočet. Toto sa nazýva orezávanie (pruning).

Pri orezávaní siete máme viacero možností:

1. Vynecháme niektoré váhy.
2. Vynecháme niektoré neuróny.
3. Vynecháme celé konvolučné filtre.

Ďalším problémom je, na základe akých kritérií určíme, ktorý neurón je najmenej dôležitý. Sú metódy, ktoré určia len dôležitosť neurónov lokálne, teda len v rámci danej vrstvy, iné sa pozerajú na celú sieť. V našej práci sa budeme zaoberať týmito rôznymi prístupmi, ktorý je najvhodnejší pre náš problém a aj ich experimentálne vyhodnotíme.

Použitá literatúra:

- Parkhi, Omkar & Vedaldi, Andrea & Zisserman, Andrew. (2015). *Deep Face Recognition*. 1. 41.1-41.12. 10.5244/C.29.41.
- Florian Schroff and Dmitry Kalenichenko and James Philbin.(2015). *FaceNet: A Unified Embedding for Face Recognition and Clustering*. 1503.03832
- J Deng, J Guo, and S Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. *preprint arXiv:1801.07698*, 2018.
- Y. He, X. Zhang, and J. Sun. Channel pruning for accelerating very deep neural networks. In *ICCV*, 2017
- Ruichi Yu, Ang Li, Chun-Fu Chen, Jui-Hsin Lai, Vlad I. Morariu, Xintong Han, Mingfei Gao, Ching-Yung Lin, Larry S. Davis. (2017) *NISP: Pruning Networks using Neuron Importance Score Propagation*