

Aplikace neuronových sítí v OCR

Aleš Togner

*Katedra informatiky a počítačů, Ostravská univerzita v Ostravě 30. dubna 22, 70103, Ostrava,
Email: R12345@student.osu.cz*

Abstrakt

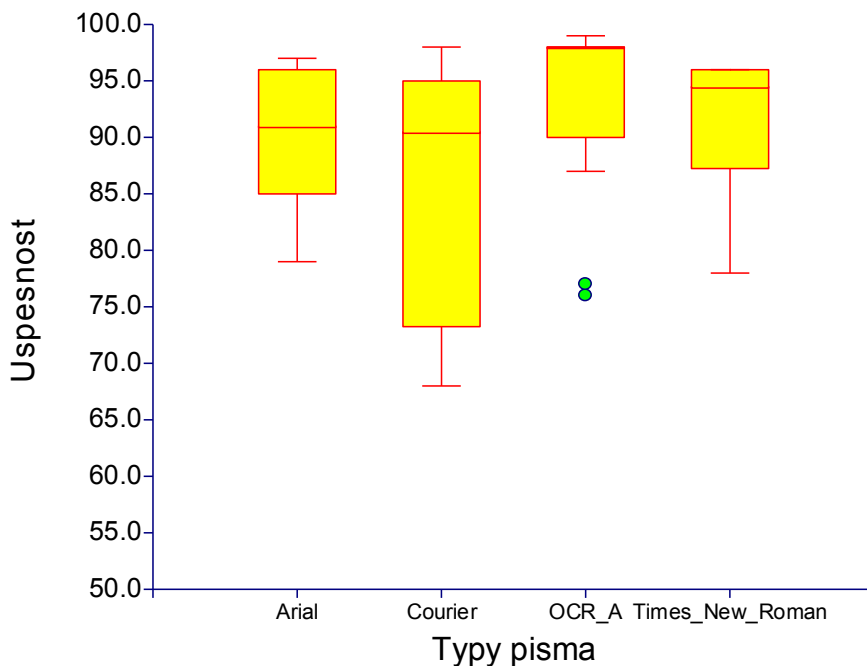
Tento příspěvek se zabývá problematikou optického rozpoznávání znaků (OCR) a její realizací s použitím neuronových sítí, přičemž jeho hlavním cílem je vytvoření programu na rozpoznávání znaků v obraze. První část pokrývá jednotlivé kroky OCR, mezi které patří předzpracování obrazu, segmentace obrazu a detekce rysů jednotlivých znaků. Druhá část se věnuje neuronovým sítím a jejich použití v OCR. Třetí, praktická část, popisuje vytvořený program na rozpoznávání znaků z uživatelského a programátorského hlediska. V rámci praktické části byla otestována vytvořená aplikace OCR na různých typech písma a s různými hodnotami parametrů neuronové sítě. V závěru bylo provedeno zhodnocení získaných výsledků experimentální části.

OCR (Optical Character Recognition, česky přeloženo jako optické rozpoznávání znaků) je technologie, umožňující konverzi netextových dokumentů, jako např. digitálních fotografií, některých souborů PDF nebo oskenovaných papírových dokumentů do textového formátu, ve kterém je možno provádět úpravy nebo vyhledávání. Ta se dělí na předzpracování, při kterém se zdrojový obraz upravuje tak, aby znaky šlo snáze identifikovat, segmentaci obrazu na jednotlivé podoblasti, ve kterých se znaky vyskytují a detekci rysů v nalezených podoblastech. [Cheriet et al., 2007] Podle rysů je poté třeba nějakým způsobem zjistit, o který znak se jedná. Jednou z metod mohou být neuronové sítě.

Neuronová síť je výpočetní model, který je na rozdíl od jiných, jako např. von Neumannovská architektura počítače, schopný se učit. Funkce sítě se neprogramuje zadáním přesného postupu, ale síť sama abstrahuje charakter funkce ze vzorových příkladů. Svoji funkcí připomíná inteligenci člověka, na jehož biologických neuronech je model postaven. Existuje mnoho různých typů neuronových sítí, mohou být cyklické, kde jsou neurony propojené v kruhu nebo acyklické, kde neurony jsou uspořádány ve vrstvách a signál se šíří do vyšších vrstev. Učení může probíhat s učitelem (je známý správný výsledek příkladu prezentovaného síti) nebo bez něj (samoorganizace). [Volná, 2008] Na základě získaných výsledků z analýzy současného stavu byla pro implementaci vybrána vícevrstvá neuronová síť s adaptačním algoritmem backpropagation, především díky její možné aplikaci na mnoho různých úloh OCR, což umožní snadno navázat další funkčnosti v případě rozšíření programu. [Vasudeva, 2012]

Vytvořená aplikace byla realizována v programovacím jazyce C++ s použitím frameworku Qt pro implementaci metod obstarávající práce s obrazem a uživatelského rozhraní. Grafické funkce knihovny jsou použity pouze pro manipulaci s jednotlivými pixely, veškeré grafické algoritmy byly implementovány jako součást práce. Grafy realizuje přídatná knihovna QCustomPlot.

Testování aplikace proběhlo na čtyřech různých typech písma: Arial, Times New Roman, OCR A Extended a Courier. Souhrnné výsledky jsou uvedeny na obrázku 1. Testy neuronové sítě proběhly na následujících parametrech: parametr učení, reprezentace výstupní vrstvy, velikost vstupní množiny, velikost skryté vrstvy, a zda na znaku proběhlo ztenčení, nebo ne.



Obrázek 1 – výsledky testování podle typů písma

Nejlépe podle očekávání dopadlo písmo OCR A, jehož medián úspěšnosti se přinejhorším rovnal nejvyšší úspěšnosti ostatních písem. Nejhůře naopak dopadlo písmo Courier, což bylo s největší pravděpodobností způsobeno velkým poklesem úspěšnosti u menších velikostí písma. V jistých případech sice dosáhlo úspěšností srovnatelných s OCR A, ale při testování s méně vhodnými parametry došlo k větším poklesům než u ostatních písem.

Klíčová slova: OCR, neuronové sítě, rozpoznávání znaků, počítačové vidění

Poděkování

Zde bych chtěl poděkovat doc. Evě Volné za vedení práce, pozitivní přístup, rychlou komunikaci a časté konzultace.

Literatura

1. CHERIET, Mohamed, Cheng-Lin LIU, Ching SUEN a Nawwaf KHARMA. Character Recognition Systems: A Guide for Students and Practitioners. U.S.A. : John Wiley & Sons, 2007. ISBN 978-0-471-41570-1.
2. VOLNÁ, Eva. OSTRAVSKÁ UNIVERZITA V OSTRAVĚ. *Neuronové sítě 1*. 2. vydání. 2008, 86 s. Dostupné z: http://www1.osu.cz/~volna/Neuronove_site_skripta.pdf
3. VASUDEVA, Nisha, Hem Jyotsana PARASHAR a Singh VIJENDRA. Offline Character Recognition System Using Artificial Neural Network. *International Journal of Machine Learning and Computing*. 2012, roč. 2, č. 4. Dostupné z: <http://www.ijmlc.org/papers/165-C01330-002.pdf>