

Matej Hrinčár - Konvoluční neuronové sítě a jejich využití při detekci objektů

Autor se v diplomové práci zabývá automatickou detekcí objektů na konkrétním případě hokejistů na hrací ploše, a to jak na statických obrázcích, tak ve videu. Tento úkol řeší pomocí poměrně nového komplexního modelu inspirovaného zpracováním zrakových vjemů u lidí – konvolučních neuronových sítí.

Diplomová práce sestává z vlastního textu práce, jehož součástí je uživatelská dokumentace, z programů *Detect*, *Train* a *OznacHracovCS* a množiny připravených trénovacích dat (na příloženém DVD).

V prvních třech kapitolách textu práce autor popisuje použité modely neuronových sítí a algoritmy používané k jejich učení – začíná modelem formálního neuronu, pokračuje jednoduchou vrstevnatou neuronovou sítí, dále popisuje architekturu Neocognitron, a posléze model konvolučních sítí, který z principů architektury Neocognitron vychází.

Ve čtvrté kapitole se autor zabývá problémem předzpracování obrázků a videa - odstraněním šumu, způsoby řešení prokládání obrazu, deformacemi a transformacemi používanými k jejich odstranění a také detekcí hran. Jak autor uvádí, model konvolučních neuronových sítí je možné použít na obrázky, které neprošly předzpracováním – různé urovně některých popsanych procesů předzpracování však můžeme pozorovat přímo v interních stavech sítí.

Klíčové jsou kapitoly 5 a 6, kde autor popisuje přípravu dat (zvolení vhodného poměru šířky ku výšce vzorů), a výběr konkrétního modelu (navrhuje celkem pět různých architektur konvolučních sítí a předkládá výsledky testování jejich vlastností – přesnost, rychlost detekce, odolnost vůči šumu – na připravených datech; poté ručně optimalizuje jejich parametry). Vybraný model potom autor podrobně analyzuje.

V sedmé kapitole autor představuje jedno z možných rozšíření detektoru – rozdělení hráčů do týmů podle barvy dresů pomocí shlukovacího algoritmu *k-means*.

Zbývající dvě kapitoly se zabývají výběrem vhodné existující implementace modelu konvolučních neuronových sítí (knihovny *EbLearn* a *Torch 7*) a vlastní implementací učení a detekce hráčů v *Torch 7*.

V závěru práce autor nastiňuje další možná rozšíření – např. sledování již jednou detekovaných hráčů.

Příložené DVD obsahuje implementaci programů *Train* a *Detect* napsaných ve skriptovacím jazyce Lua knihovny *Torch 7*, několik tisíc ručně označených trénovacích vzorů a pomocný software *OznacHracovCS* pro usnadnění získávání těchto vzorů napsaný v jazyce C#. Programy je možné spustit po nainstalování knihovny *Torch 7*.

Autor při zpracování tohoto tématu důkladně analyzoval množství postupů a modelů, jak v oblasti zpracování obrazu, tak v oblasti samotných konvolučních sítí. Vyzdvihla bych např. interpretaci vnitřních stavů konvoluční sítě a jejich porovnání s jevy zmíněnými v kapitole o předzpracování obrazu. V práci autor téma vždy dále rozvíjí směrem, který je podložený výsledky předchozí analýzy. Řešení, které nakonec zvolil, provádí automatickou detekci objektů vcelku uspokojivě, v reálném čase je dokonce možné provádět detekci hráčů na videu (byť pouze s malým rozlišením).

Práce je napsána ve slovenském jazyce, je pečlivě vypracovaná, jednotlivé algoritmy a postupy jsou přehledně popsány, autor popisované jevy vhodně ilustruje obrázky. Co se obsahové stránky týče, neshledala jsem žádné závažné nedostatky, mám pouze pár drobných připomínek:

- str. 45 – z průběhu křivek na obrázcích 6.1. a 6.2 je patrné, že těsně před ukončením učení se chyba na trénovacích vzorech začala ještě dále snižovat; bylo by zajímavé vidět, zda by se výsledky zlepšovaly dál i po 400 iteracích
- str. 16 – obrázek 3.4. se na stránku zřejmě nevešel celý
- str. 59 – špatný nadpis u Algoritmu 7.1 („Algoritmus učenia vrstevnatej siete“ místo „k-means“)
- autor možná přikládá příliš důležitosti době běhu algoritmu učení, které se provádí pouze jednorázově

Celkově hodnocenou práci považuji za zdařilou a doporučuji, aby byla uznána jako práce diplomová.

V Rostokách u Prahy, 11. ledna 2013

Mgr. Klára Pešková