



Systém pro indexaci a vyhledávání v obrazových datech

Martin Majer¹

1 Úvod

Cílem této práce bylo vytvořit systém pro indexaci a vyhledávání v obrazových datech, který je schopný ke vstupnímu obrázku vyhledat v databázi obsahově (vizuálně) podobné obrázky. Místo klasických metod rozpoznání obrazu na základě nízkourovňových příznaků (poměr jednotlivých RGB složek, hrany, tvary,...) byla využita konvoluční neuronová síť, která dokáže zachytit sémantiku obrázku.

2 Indexace a vyhledávání

Obrázek je do systému nahrán pomocí webového rozhraní² (z pevného disku nebo pomocí URL adresy). Následně dochází k předzpracování obrázku pro neuronovou síť: přepočítání rozměrů pro zachování poměru stran, zmenšení na nové rozměry, vyříznutí středu obrázku o velikosti 227×227 px a převod do barevného modelu BGR. Neuronová síť pak pro takto upravený obrázek spočítá příznakový vektor (dimenze 1000). Každý prvek tohoto vektoru vyjadřuje pravděpodobnost příslušnosti obrázku do dané klasifikační třídy.

Obrázek je opět převeden zpět do barevného modelu RGB a přidán do databáze. Pro nově přidávaný příznakový vektor a pro vektory již přidávaných obrázků je pak vytvořen k-dimenzionální strom na základě jejich euklidovské vzdálenosti - podobné obrázky budou tedy ve stromu blízko u sebe. Díky této vlastnosti postačí pro jejich vyhledání nalezení nejbližších sousedů nově přidávaného obrázku. Podobné obrázky jsou pak odeslány zpět do webového rozhraní.

Pro indexaci a vyhledávání byly otestovány tři metody: distanční matice, k-dimenzionální strom z knihovny scikit-learn a k-dimenzionální strom z knihovny FLANN, který byl nakonec využit ve finální verzi systému. Porovnání rychlosti indexace a vyhledávání jednotlivých metod je uvedeno v následující tabulce.

Metoda	Indexace	Vyhledávání
distanční matice	> 7min	208μs
KDTree sklearn	0.91s	15.3ms
KDTree FLANN	35ms	178μs

Tabulka 1: Rychlost indexace 10000 obrázků a vyhledávání pomocí jednotlivých metod

¹ student 2. ročníku bakalářského studijního programu Aplikované vědy a informatika, obor Kybernetika a řídicí technika, e-mail: mmajer@students.zcu.cz

²náhled webového rozhraní je dostupný na <http://home.zcu.cz/~mmajer/svk2015>

3 Vyhodnocení

K vyhodnocení přesnosti vyhledávání obrázků pomocí neuronové sítě byly náhodně vybrány 3 obrázky z 5 tříd z celkových 908 tříd v databázi³. Procentuální úspěšnost vyhledávání pak byla spočtena na základě toho, kolik je vyhledáno obrázků ze stejné třídy, jako je hledaný obrázek. Ačkoliv se na první pohled může zdát, že přesnost vyhledávání je nízká, je třeba brát v úvahu, že mnohé objekty a scény se objevují ve více třídách (např. ložnice, hotelový pokoj, pokoj na koleji), a tak se ve skutečnosti ve výsledcích objeví více podobných obrázků, než udává následující tabulka. Druhým důvodem pro nízkou procentuální úspěšnost je, že obrázky v jedné třídě se mohou velmi lišit (např. čelní pohled na kokpit, pilot v popředí a kokpit v pozadí). Nicméně je třeba podotknout, že neuronová síť pro určité třídy funguje lépe, než pro ostatní.

Class	Image Count	Image	Matching	Different	X-% Match	Mean
cockpit	695	1	20	675	2.88%	5.76%
		2	60	635	8.63%	
		3	40	655	5.76%	
beach	1194	1	66	1128	5.53%	4.52%
		2	18	1176	1.51%	
		3	78	1116	6.53%	
volcano	126	1	65	61	51.59%	55.03%
		2	68	58	53.97%	
		3	75	51	59.52%	
mountain	736	1	148	588	20.11%	16.58%
		2	73	663	9.92%	
		3	145	591	19.70%	
podium	143	1	9	134	6.29%	3.03%
		2	3	140	2.10%	
		3	1	142	0.70%	

Tabulka 2: Procentuální úspěšnost vyhledávání obrázku v rámci jedné třídy

4 Využité technologie

K implementaci systému byl využit programovací jazyk Python, knihovny FLANN a OpenCV a frameworky Caffé, Flask a Bootstrap. Databáze momentálně obsahuje přes sto osm tisíc obrázků převážně indoorových a outdoorových scén rozdělených do 908 tříd.

Literatura

Muja, M., and Lowe, D., 2013. *Fast Library for Approximate Nearest Neighbors - User Manual*.

Dostupné z: http://www.cs.ubc.ca/research/flann/uploads/FLANN/flann_manual-1.8.4.pdf

Krizhevsky, A., Sutskever, I., and Hinton, G. E., 2011. *ImageNet Classification with Deep Convolutional Neural Networks*.

Dostupné z: <http://www.cs.toronto.edu/~fritz/absps/imagenet.pdf>

³ systém využívá databázi SUN, dostupná na <http://groups.csail.mit.edu/vision/SUN>