

Google Landmark Retrieval Challenge

Petra Martinjak
Matija Šelendić
Stjepan Teskera
Stjepan Zbiljski

Uvod i motivacija

- Problem dohvaćanja sličnih slika u bazi - slike znamenitosti
- Prije - dohvaćanje slike pomoću metapodataka
- CBIR - dohvaćanje slike na temelju sadržaja
- Široko područje primjene - data-mining, medicina, prevencija zločina, edukacija
- Preko milijun slika
- Više od 15 000 znamenitosti

Cilj projekta

- Svaka slika - jedinstven id, nema metapodataka
 - CBIR
-
1. Za upit u obliku id-a slike znamenitosti, koje slike predstavljaju istu znamenitost?
 2. S kolikom preciznošću znamo da te slike doista predstavljaju istu znamenitost?

Pristup rješavanju problema

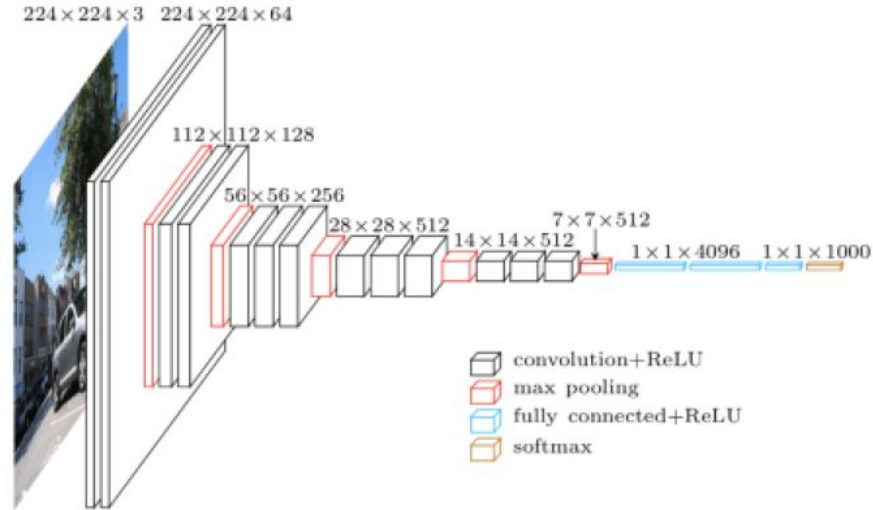
Podaci i obrada podataka

- Skup index - baza - 1 091 756 slika
- Skup query - upiti - 114 943 slike
- Oba skupa - id i url
- Kaggle skripta za preuzimanje slika - .jpg, 256x256, 128x128

Pristup rješavanju problema

Feature extraction

- Smanjuje opterećenje
- ImageNet - pretrenirane neuronske mreže - veća točnost, izbjegavanje overfittinga
- ResNet - točniji rezultati, veća dubina, manje parametara, veće slike
- Kompromis - VGG16 - može se koristiti na manjim slikama
 - ➔ model - pohrana slika samo kao njihovih najvažnijih značajki



Pristup rješavanju problema

Usporedba značajki

- Za sliku upita pronaći njoj najbližije slike iz baze
- Udaljenost vektora značajki - slika upita, slike iz baze
- Euklidska udaljenost, kosinusova sličnost
- Faiss biblioteka - bruteforce
- Rezultat - svaki redak - id slike upita, id-ovi njoj najbližijih slika iz baze

Korištene metode i algoritmi

- Pomoćna skripta - urllib i PIL biblioteke - preuzimanje slika i rukovanje istima
- Prva glavna skripta - Keras biblioteka - izrada modela i pomoću njega pretvorba slika u vektore značajki
- Druga glavna skripta - faiss biblioteka - međusobne udaljenosti vektora značajki
- NumPy, pandas

Rezultati

$$mAP@100 = \frac{1}{Q} \sum_{q=1}^Q \frac{1}{\min(m_q, 100)} \sum_{k=1}^{\min(n_q, 100)} P_q(k) rel_q(k)$$

- Public ocjena greške - nad 34% podataka testnog skupa
- Private ocjena greške - nad 66% podataka testnog skupa
- Kosinusova sličnost - public rezultat 0.4%, private rezultat 0.9%
- Euklidska udaljenost - public rezultat 4.3%, private rezultat 6.6%
- Pobjednici natjecanja - public rezultat 62.5%, private rezultat 62.7%

Diskusija

- VGG nije adekvatna mreža, ResNet je bolji
- Potreba za većim slikama
- Euklidska udaljenost daje bolje rezultate od kosinusove sličnosti
- Potreba za radom na GPU umjesto CPU

Zaključak

- Cilj: za query sliku pronaći njoj najbližnije slike iz baze
- 1. Korak - feature extraction:izvlačenje najbitnijih značajki svih slika
 - VGG16 - manje slike, manje slojeva, manja točnost
 - ResNet - veće slike, manje parametara, veća točnost
- 2. Korak - usporedba značajki
 - Faiss biblioteka
 - Euklidska udaljenost, kosinusova sličnost
- Rezultat: ne baš zadovoljavajući
 - Euklidska udaljenost: 6,6%
 - Kosinusova sličnost: 0.9%

Daljnja istraživanja

- Ograničenja u tehnologiji → jednostavniji model, manje slike, lošiji rezultati
- ResNet → kompleksniji model, veće slike, bolji rezultati (gotovo sigurno)
- Najbolji rezultati - ResNet + k-najbližih susjeda s euklidskom udaljenosti + query ekspanzija difuzijom

Hvala na pažnji!

Pitanja?