

Universitatea Transilvania din Braşov
Facultatea de Matematică şi Informatică
Specializarea Informatică

Proiect de disertaţie

Tehnici de Machine Learning în procesarea şi recunoaşterea imaginilor

Autor:

Draghia Alin-Madalin

Profesor coordonator:

Lect. dr. Sasu Lucian Mircea

Braşov

Iulie 2014

Cuprins

1	Introducere	1
1.1	Motivație	1
1.2	Enunțul problemei	3
1.3	Tehnologii folosite	3
1.4	Structura Lucrării	4
2	Recunoasterea obiectelor	5
3	Învățarea Automata	6
3.0.1	Învățarea Supervizată	6
4	Concluzii și posibilitati de dezvoltare	8
	Bibliografie	9

Capitolul 1

Introducere

Aceasta lucrare își propune să prezinte, din punct de vedere atât teoretic cât și practic în ce constă dezvoltarea unui algoritm de recunoaștere a obiectelor în imagini.

1.1 Motivație

Oamenii pot recunoaște o mulțime de obiecte într-o imagine fără să depună prea mult efort, chiar dacă în aceste imagini obiectele prezintă variații de perspectivă, de dimensiune, sunt translatate, rotite sau chiar obstrucționate.

Recunoașterea obiectelor este una dintre principalele aplicații ale viziunii artificiale¹ și procesarea de imagini.

Viziunea artificială reprezintă procesul invers al celui de formare a imaginii și se ocupă cu recuperarea de informații din imagini cu ajutorul metodelor matematice, geometrice, statistice și a teoriei învățării automate².

Viziunea artificială și învățarea automată sunt două domenii aflate în plină dezvoltare și sunt de mare interes atât în cadrul academic cât și în industria software.

Recunoașterea obiectelor este una dintre aplicațiile fundamentale ale viziunii artificiale. Deși de-a lungul timpului au fost dezvoltați mulți algoritmi, recunoașterea obiectelor este încă departe de a putea fi considerată o pro-

¹Computer vision

²Machine learning

1.1. MOTIVAȚIE

blema rezolvata.

1.2. ENUNȚUL PROBLEMEI

Dezvoltarea rapida a sistemelor de calcul a permis utilizarea acestor algoritmi în tot mai multe aplicații. Câteva dintre aplicațiile recunoașterii de obiecte sunt:

- Industriale: recunoașterea și verificarea cip-urilor pe o placa electronica, numărarea de obiecte pe o banda rulanta
- Securitate: recunoașterea unui intrus folosind o camera de supraveghere
- Medicale: recunoașterea diferitelor tumori într-o imagine de tomografie
- Fotografie: focalizare automata pe fete
- Internet: căutare google după imagini, marcarea automata a fetelor într-o poza de pe facebook

1.2 Enunțul problemei

Se scrie o librărie software cu ajutorul carea sa se antreneze și sa se folosească algoritmi de recunoaștere a obiectelor în imagini.

Algoritmul va învăța sa recunoască obiecte folosindu-se de un set de exemple pozitive cat și negative.

Se scrie o aplicatie care antrenează un algoritm de recunoaștere și îl salvează modelul învățat pe disc și una care încarcă modelul și îl aplica pe o imagine data.

1.3 Tehnologii folosite

Limbajul C++

Limbajul C++ este un limbaj de programare general care și este compilat în cod-mașina. Este un limbaj multi-paradigma, cu verificare statica a tipurilor. Suporta programarea procedurala, orientata pe obiecte și generica. Limbajul oferă facilitati de manipulare a memoriei la nivel scăzut. Fiind proiectat inițial ca un limbaj pentru programarea de sisteme(sisteme integrate, kernel sisteme de operare), performanța și eficienta sunt trăsături principale.

Dat fiind faptul ca este și compatibil cu limbajul C, utilizatorii C++ au la dispoziție o gama larga de librarii software din cele mai diverse ramuri de aplicații de care se pot folosii.

Limbaajul Python

Limbajele C++ și Python

Pentru realizarea lucrării am ales sa folosesc C++ și Python din mai multe motive:

C++ și Python sunt doua limbaje de programare atât de diferite încât putem spune ca se afla în capete diferite ale axei limbajelor de programare.

- C++ este compilat în cod-mașina, Python este interpretat
- Python are sistemul de tipuri dinamic și este recunoscut pentru flexibilitate
- C++ are sistemul de tipuri static și este recunoscut pentru eficienta
- Python eliberează automat memoria

Pentru multi programatori, aceste diferențe înseamna ca cele doua limbaje se completează perfect.

Librăria OpenCV

Librăria OpenCV este cea mai populara librărie de procesare de imagini

Librăria Boost

Librăria scikit-learn

Librăria Qt

1.4 Structura Lucrării

Capitolul 2

Recunoasterea obiectelor

Capitolul 3

Învățarea Automata

Învățarea automata, o ramura a inteligenței artificiale, este preocupată de construirea și studierea unor sisteme care pot învăța din date.

În 1959, Arthur Samuel a definit învățarea automata ca: "Domeniul de studiu care da calculatoarelor abilitatea de a învăța fără să fie explicit programate"[Sim13].

Tom M. Mitchell o oferă o definiție mai formală: "Se spune despre un program ca a învățat din experiența E cu privire la o clasă de acțiuni T și o măsura de performanță P , în cazul în care performanțele sale la sarcina T , măsurate prin P se îmbunătățesc cu experiența E ." [Mit97]

Algoritmi de învățare automata se pot caracteriza în funcție de tipul de date cu care este antrenat și tipul de răspunsului dorit:

Supervizată: Atunci când algoritmi sunt antrenați cu un set de date cu răspuns cunoscut și se dorește răspunsul în cazul unor date noi.

Nesupervizată: Atunci când algoritmi sunt antrenați cu un set de date cu răspuns necunoscut și se dorește găsirea unei structuri în setul de date.

În continuare vom discuta despre învățarea supervizată.

3.0.1 Învățarea Supervizată

Învățarea supervizată este sarcina învățării automate de a învăța o funcție de evaluare din datele marcate de antrenament. Datele de antrenament sunt

constituite dintr-un set de exemple de antrenament, fiecare exemplu consta într-o pereche de valori de intrare și o valoare de ieșire dorită. Un astfel de algoritm analizează datele de antrenament și produce o funcție, care poate fi folosită pentru rezolvarea problemei în cazul unor date care nu au mai fost văzute.

Învățarea supervizată constă în învățarea unei funcții $f : X \rightarrow Y$, unde X reprezintă domeniul datelor de intrare, iar Y domeniul datelor de ieșire, pentru un set de date de antrenare $\mathcal{D} = \{(x_i, y_i) | x_i \in X, y_i \in Y\}_{i=1}^n$ se dorește găsirea unei funcții $f(x_i) = y_i$ și pentru $x_i \notin \mathcal{D}$, adică date cu care nu a fost antrenat algoritmul.

În funcție de tipul de lui y , algoritmi de învățare supervizată pot fi clasificați astfel:

- Regresie: atunci când răspunsul este o valoare numerică

Exemplu:

$$y \in \mathbb{R}$$

- Clasificare: atunci când răspunsul este o valoare categorică.

Exemplu:

$$y \in \{0, 1\}$$

sau

$$y \in \{alb, negru, gri\}$$

Capitolul 4

Concluzii și posibilitati de dezvoltare

Bibliografie

- [Mit97] Tom Michael Mitchell. *Machine Learning*. McGraw-Hill Education, 1997.
- [Sim13] Phil Simon. *Too Big to Ignore: The Business Case for Big Data*, volume Volume 72 of Wiley and SAS Business Series. Wiley, 2013.