

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΕΡΓΑΣΤΗΡΙΟ ΕΠΕΞΕΡΓΑΣΙΑΣ ΣΗΜΑΤΩΝ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

ΥΠΟΛΟΓΙΣΤΙΚΗ ΟΡΑΣΗ

ΣΤΑΤΙΣΤΙΚΗ ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΕΙΚΟΝΩΝ- ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΕΙΚΟΝΩΝ

Διδάσκων Αναπληρωτής Καθηγητής Εμμανουήλ Ζ. Ψαράκης
Επικουρικό έργο :
Υποψήφιος Διδάκτωρ Αριστείδης Μπίφης
Υποψήφια Διδάκτωρ Αγγελική Φέκα

Πάτρα Δεκέμβριος 2019

ΣΤΟΙΧΕΙΩΔΗΣ ΘΕΩΡΙΑ

Ένα σημαντικό ερευνητικό αντικείμενο στο πεδίο της Όρασης Υπολογιστών (Computer Vision) και συγκεκριμένα αυτό της Οπτικής Εκμάθησης (Visual Learning), είναι η κατασκευή στατιστικών μοντέλων της εμφάνισης (appearance) των εικόνων, βασιζόμενοι στα εικονοστοιχεία τους [1]. Μία ευρέως διαδεδομένη εφαρμογή στο πεδίο αυτό αποτελεί η διαδικασία εύρεσης των εικονοστοιχείων και περιοχών σε μια εικόνα ή ακολουθία εικόνων, τα οποία χρωματικά αντιστοιχούν σε δέρμα (skin). Η διαδικασία αυτή ονομάζεται ανίχνευση δέρματος (skin detection) και συνήθως αποτελεί μέρος της προ επεξεργασίας, με σκοπό την εύρεση περιοχών που αντιστοιχούν στο πρόσωπο και τα άκρα ανθρώπων σε εικόνες [2].

Ένας ανιχνευτής δέρματος (skin detector) μετασχηματίζει ένα εικονοστοιχείο στο επιθυμητό χρωματικό χώρο και στη συνέχεια χρησιμοποιείται ένας κατηγοριοποιητής δέρματος (skin classifier) για την κατηγοριοποίηση του εικονοστοιχείου σε μία από τις δύο κατηγορίες, skin και non-skin. Ένας ταξινομητής δέρματος, ορίζει ένα όριο απόφασης της κλάσης χρώματος του δέρματος στο χρωματικό χώρο, που βασίζεται και επομένως εξαρτάται στο σύνολο των δεδομένων εκπαίδευσης.

ΚΑΤΗΓΟΡΙΟΠΟΙΗΤΗΣ NAIVE BAYES

Υποθέτουμε $f(w_i|\mathbf{x})$ την πιθανότητα ενός δοσμένου εικονοστοιχείου μιας εικόνας να ανήκει στην κλάση w_i , δοθέντος ότι το χρώμα του είναι \mathbf{x} . Έτσι, $P(w_1)$ και $P(w_2)$ είναι οι αντίστοιχες a posteriori πιθανότητες για τις κλάσεις skin και non-skin.

Αν ορίσουμε ως C_{ij} το κόστος απόφασης ενός διανύσματος χαρακτηριστικών $\mathbf{x} \in w_i$ όταν στην πραγματικότητα $\mathbf{x} \in w_j$. Τα C_{11} και C_{22} αντιπροσωπεύουν τα κόστη των ορθών κατηγοριοποιήσεων, ενώ οι λανθασμένες αντιπροσωπεύονται από τα C_{12} και C_{21} , αντίστοιχα.

Ο βέλτιστος κατηγοριοποιητής θα αναθέσει το \mathbf{x} στην κλάση που ελαχιστοποιεί το κόστος.

Ο κανόνας απόφασης είναι:

$$R_1(\mathbf{x}) = C_{11} \cdot f(w_1|\mathbf{x}) + C_{12} \cdot f(w_2|\mathbf{x}) \quad (1)$$

και

$$R_2(\mathbf{x}) = C_{21} \cdot f(w_1|\mathbf{x}) + C_{22} \cdot f(w_2|\mathbf{x}) \quad (2)$$

Έτσι,

$$\begin{array}{ccc} \mathbf{x} \in w_2 & & \\ R_1(\mathbf{x}) & > & R_2(\mathbf{x}) \\ & < & \\ \mathbf{x} \in w_1 & & \end{array} \quad (3)$$

Συνδυάζοντας τις σχέσεις (1) και (2):

$$\begin{array}{c} \mathbf{x} \in w_2 \\ \frac{f(w_1|\mathbf{x})}{f(w_2|\mathbf{x})} > \frac{C_{12} - C_{22}}{C_{21} - C_{11}} \\ \mathbf{x} \in w_1 \end{array} \quad (4)$$

Εφαρμόζοντας τον τύπο του Bayes:

$$f(w_i|\mathbf{x}) = \frac{f(\mathbf{x}|w_i) \cdot P(w_i)}{p(\mathbf{x})} \quad (5)$$

Με χρήση της (5), η (4) μπορεί να γραφεί ως:

$$\begin{array}{c} \mathbf{x} \in w_2 \\ \frac{f(\mathbf{x}|w_1)}{f(\mathbf{x}|w_2)} > r \\ \mathbf{x} \in w_1 \end{array} \quad (6)$$

όπου:

$$r = \frac{(C_{12} - C_{22}) \cdot P(w_2)}{(C_{21} - C_{11}) \cdot P(w_1)}.$$

Στις παραπάνω σχέσεις, οι $f(\mathbf{x}|w_i)$ είναι οι συναρτήσεις πυκνότητας πιθανότητας των κλάσεων skin ($i = 1$) και non-skin ($i = 2$), οι $P(w_i)$ είναι οι a priori πιθανότητες των κλάσεων w_i και το r αντιπροσωπεύει το κατώφλι απόφασης.

Τα κόστη λάθους ταξινόμησης, C_{12} και C_{21} ρυθμίζονται ανάλογα με την εφαρμογή (πώς θα ορίζατε αυτά τα κόστη στην περίπτωση του προβλήματος της ανίχνευσης αεροπλάνου από ραντάρ), ενώ τα κόστη ορθής ταξινόμησης, C_{11} και C_{22} τυπικά θεωρούνται μηδέν (0).

ΔΙΑΔΙΚΑΣΙΑ

1. Βασιζόμενοι στην παραπάνω διαδικασία και τη βιβλιογραφία που σας δίνεται, εκπαιδεύστε έναν κατηγοριοποιητή Naive Bayes 2 κλάσεων (skin , non-skin) για την κατηγοριοποίηση του συνόλου εικόνων που βρίσκονται στο αρχείο `train.zip` στον RGB χώρο.
2. Χρησιμοποιήστε το σύνολο των εικόνων που βρίσκονται στο αρχείο `test.zip` για την αξιολόγηση του συστήματος. Σχεδιάστε την καμπύλη ROC του κατηγοριοποιητή.
3. Επαναλάβετε την διαδικασία που περιγράφεται στο 1, αλλάζοντας τον χρωματικό χώρο από RGB σε HSV.

4. Επαναλάβετε τη διαδικασία που περιγράφεται στα 1,2 και 3, χρησιμοποιώντας τον κατηγοριοποιητή SVM. Για το σκοπό αυτό χρησιμοποιείστε την έτοιμη συνάρτηση της Matlab [5].
5. Σχολιάστε συνολικά την απόδοση των δύο κατηγοριοποιητών που υλοποιήσατε.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] Michael J. Jones, James M. Rehg, (2002), Statistical Color Models with Application to Skin Detection, International Journal of Computer Vision, 46(1), 81-96, doi:10.1023/a:1013200319198
- [2] Ahmed Elgammal, Crystal Muang, Dunxu Hu, (2009), Skin Detection { a Short Tutorial, Department of Computer Science, Rutgers University, Piscataway, NJ, 08902, USA.
- [3] W.R. Tan, C.S. Chan, Y. Pratheepan and J. Condell., (2012), A Fusion Approach for Efficient Human Skin Detection., IEEE Transactions on Industrial Informatics, vol. 8(1): 138-147.
- [4] Chai D., Phung S.L., Bouzerdoum A., (2003), A Bayesian skin/non-skin color classifier using non-parametric density estimation. Proceedings of the 2003 International Symposium on Circuits and Systems-ISCAS '03, doi: 10.1109/iscas.2003.1206010
- [5] <https://www.mathworks.com/help/stats/support-vector-machine-classification.html>