

Predefinisani projekat 2

Klasifikacija cveća

Opis problema

Potrebno je prepoznati na slikama tip cveća koji se na njima nalazi. Postoji 5 različitih vrsta cveća: *Snowdrop*, *LilyValley*, *Bluebell*, *Tigerlily* i *Fritillary*. Trening skup sadrži 320 slika, a test skup 80. Potrebno je postići tačnost na test skupu veću od 55%.



Primer slika cveća

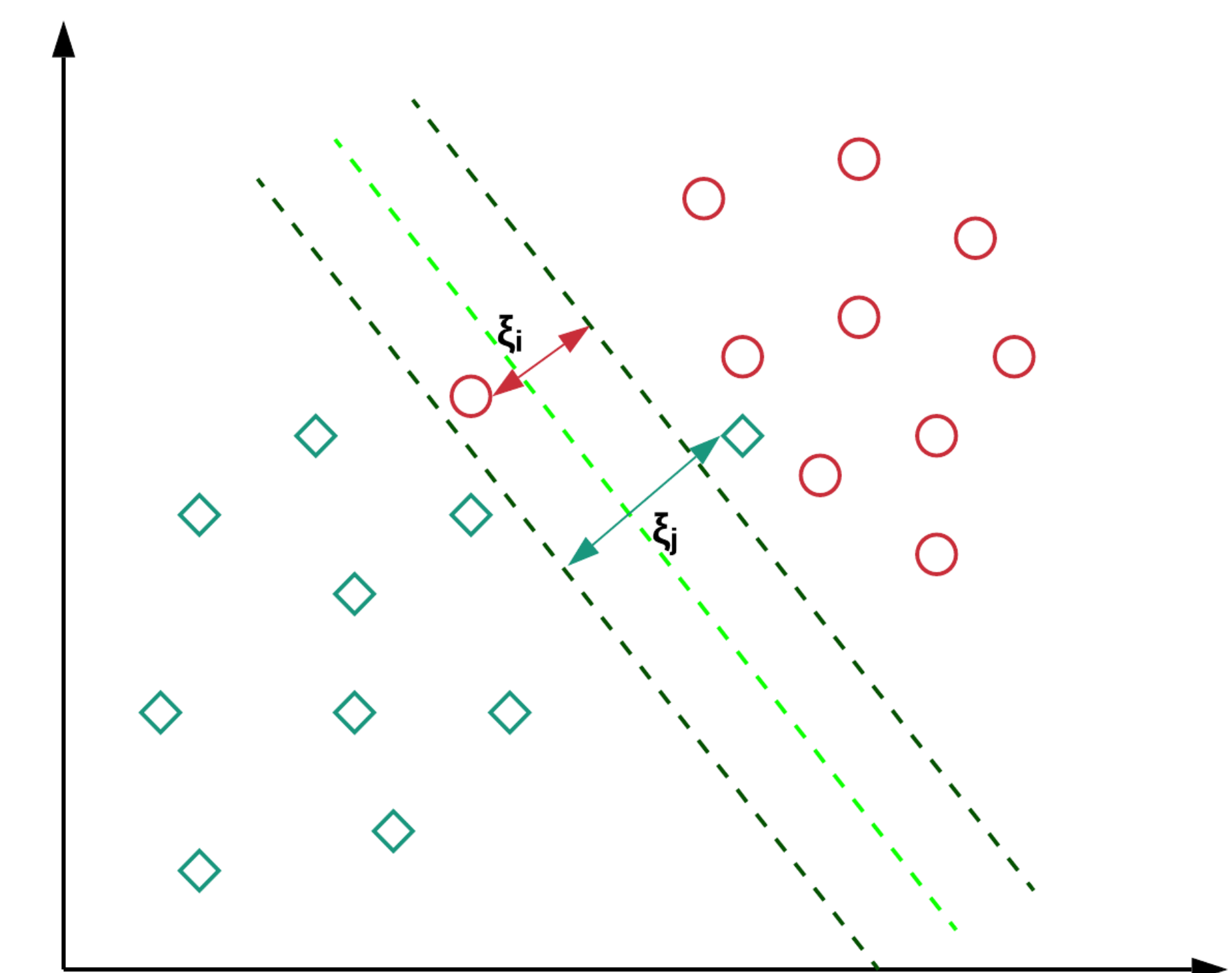
Rešenje

Sve slike, i trening i test skupa, su prilikom učitavanja svedene na dimenzije 500x500, jer je to najmanja dimenzija koju neka slika ima i po x i y osi. Takođe slike su prevedene iz RGB modela u grayscale(nijanse sivih). Za izdvajanje osobina sa slike korišten je HOG deskriptor. Veličina jedne ćelije je 32x32 piksela za koju se računa histogram gradijenata koji sadrži 9 intervala od po 20 stepeni svaki(0...20...40... ... 180). Kako bi rešio problem varijacije u osvetljenju slike izvršena je normalizacije histrograma gradijenta. Normalizacija je izvršena nad blokom dimenzije 3x3 ćelije. To se vrši tako što se svaka vrednost iz vektora izračunatih histrograma gradijenata podeli sa magnitudom vektora. Pomeraj bloka se vrši za veličinu jedne ćelije.

Treniranje modela vršeno je pomoću SVM klasifikatora koji koristi linearni kernel. Pošto je ovo višeklasni problem, potrebno je podesiti parametar decision_function_shape na "ovo"(one vs one) i to znači da će napraviti binarni klasifikator za svaki par klasa. U našem slučaju biće napravljeno 10 klasifikatora, jer imamo 5 klasa. ($n = \text{broj klasa}$, $n \cdot (n-1)/2$ klasifikatora)

Zaključak

Korištenjem HOG deskriptora i SVM klasifikatora postignuta je tačnost od 60%. Za decision_function_shape isproban je i "ovr"(one vs rest), ali "ovo" radi za nijansu brže.



SVM klasifikacija

Student: Slaven Garić SW47-2015

Predmet: Soft kompjuting