

# Projektbeszámoló

## Képfeldolgozás Haladóknak MSc gyakorlat

Kiss Marcell Tamás

2015. október 18.

### 1. Feladatspecifikáció

A projekt témája hivatalosan a „Logófelismerés” nevet viseli. A feladatot úgy értelmeztük, hogy tetszőleges, kapott input képre képesek legyünk meghatározni a képen lévő legnagyobb sikerrel detektált alkoholos ital logóját (részletekért lásd a github repository-ban lévő eredeti beszámolót:

<https://github.com/abrhm/unbeerables/blob/master/docs/report/report.pdf>).

### 2. Vállalt feladat

A sikeres logófelismerés céljából elengedhetetlen, hogy valamilyen gépi tanulási módszerrel, előző ismeretek alapján képesek legyünk új képek klasszifikálására. Ennek érdekében SVM (Support Vector Machine) alkalmazása mellett döntöttünk. Magának a tanulási adatbázisnak felállítása, az SVM tanítás beállítása és a képek detektálása során szerzett feature vektorok megfelelő kezelése lett az én feladatom megegyezés alapján.

Az SVM egy lineáris gép, amely egy lineáris döntési felületet definiálva képes osztályozni kapott feature vektorokat úgy, hogy az SVM kritériumfüggvényeként a döntési felülethez megadott szegélyt (felület körüli üres terület, amit a felülethez lévő legközelebbi pontok határoznak meg) szeretnénk maximalizálni [1].

Mivel az egyes képeken detektált pontok eltérő számúak lehetnek, így nem minden vektor ugyanakkora, tehát közvetlenül nem adható át az SVM-nek. Ezért egy olyan megvalósítást tervezek, ami a Bag of Words alapján működik és használatos ilyen esetekben (Bag-of-Features) [2].

Mivel a fejlesztést Matlab-ban tesszük meg, ezért a beépített 2-osztályos SVM-et fogom használni tanításra és klasszifikálásra [3][4].

Eddig a tanító adatbázis felállításával foglalkoztam, 5 különböző fajta logót keresve. Az adatbázis felépítésénél hangsúlyt helyeztem arra, hogy legyenek könnyen felismerhető

képek az effektív tanulás érdekében (csak logó, stock fotók, nincs forgatva, transzformálva). Az ilyen képeken kívül olyanokat is kerestem, amelyek nehezebben felismerhetőek, pl. a logó el van forgatva, vagy homályosabb a kép bizonyos pontokon.

További tesztelés során, ahogy haladunk előre, ezeket a képeket feldaraboljuk, hogy lássuk, mennyire képes a program detektálni azokat a logókat, amik az eredeti képek háttérében vannak, vagy nem a fókuszban. Ezeken kívül saját képeket is tervezünk készíteni tesztelés céljából.

Képek megtekinthetőek itt: <https://github.com/abrhm/unbeerables/tree/master/imgs>

### 3. Források

1. Farkas Richárd, Gépi tanulási módszerek – Lineáris gépek előadás: <http://www.inf.u-szeged.hu/~rfarkas/ML15/gepitanulas.html>
2. Cordelia Schmid, Bag-of-features for category classification
3. <https://www.mathworks.com/help/stats/classificationsvm-class.html>
4. <https://www.mathworks.com/help/stats/fitcsvm.html>