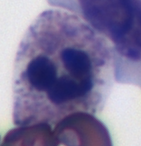
# I. mérföldkő

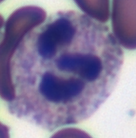
# Megoldandó feladat ismertetése

Hematológiai vérkenet felvételeken a vörösvértestek (Red Blood Cell – RBC), és a fehérvérsejtek (white blood cell – WBC) szegmentálása (háttértől való elkülönítés), az egyes sejtek középpontjának detektálása, megszámlálása, az RBC-k és WBC-k elkülönítése. A WBC-k detektálásának összevetése manuális annotációkkal a következő sejttípusokra:

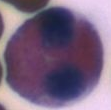
1. Band neutrophiles bne



1. Segmented neutrophiles ne



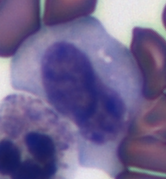
1. Eosinophiles eo



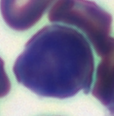
1. Basophiles ba



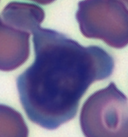
1. Monocytes mo



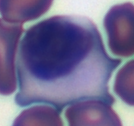
1. Lymphocytes ly



1. Large granular lymphocytes lgly



1. Reactive lymphocytes rly



További feladat a fehérvérsejtek osztályba sorolása – klasszifikáció.

Az 5+1 kijelölt osztály a következő:

* Basophiles (jelölés: ba)
* Eosinophiles (jelölés: eo)
* Lymphocytes (összevont osztály, jelölések: ly, rly, lgly
* Monocytes (jelölés: mo)
* Neutrophiles (összevont osztály, jelölések: ne, bne)
* Unknown (osztályokba nem sorolható)

Feltételezések

A vérkenet alapú automatikus sejtdetektálás esetén a következő feltételezésekkel élünk:

1. Háttér egységes színű, alacsony szaturációjú, magas intenzítású (lehetőleg fehér).
2. RBC és WBC-k színe a festési eljárás után különböző (a kapott mintákon az RBC inkább pirosas, a WBC inkább kékes tartományban van).
3. Az RBC-k számossága jóval nagyobb a WBC-k számosságánál (100:1 nagyságrend).
4. Normál esetben az RBC-k egyenletesen töltik ki a képet, nincsen sűrűsödés, nincs átfedés, nem képeznek konglomerált egységeket, valamint kellő számosságúak (>150).
5. Szintén normál esetben az RBC-k alakja kör, minimális bikonkáv, vagy bikonvex deformáció megengedett.
6. A WBC-k sejtmagjának színe jól definiált, stabil érték (legalább is egy adott festési, megvilágítási eljárásnál). Ezt az értéket a detektálás előtt is meg tudjuk határozni.

### Kenetfelvétel elfogadási kritériumai

A kenetfelvételeket a következő elfogadási szempontok szerint vizsgáljuk:

1. felbontás – a digitális felvétel mérete (>1000x1000pix)
2. élesség (később definiálva)
3. expozíció – túlexponált háttér területe (<10%)
4. kontraszt (később definiálva)
5. inhomogenitás - a háttérmegvilágításban lévő inhomogenitás (később dfiniálva)
6. RBC sejtkoncentráció – se túl kicsi, se túl nagy TOPOLÓGIAI feltétel
7. zajtartalom (később definiálva)
8. CHROMATIKUS és SZATURÁCIÓS feltételek (később definiálva)

# A feladatot megoldó szoftverek ismertetése

## Front-End

## Back-End

A szoftver képfeldolgozást végző moduljának fejlesztői verziója kerül átadásra. A fejlesztői verzió Python környezetben került implementálásra

### Forrásfile-ok

Az átadott **WBC** mappában találhatóak a szükséges forrásfile-ok és mappastruktúra:

* cell\_detector.py – a sejtdetektálást végző modul
* cell\_classifier.py – a sejt osztályozást végző modul
* evaluate.py – kiértékelést végző modul
* model mappa – cnn\_model.dnn : konvolúciós neurális hálózat tanított model
* input – input file-ok mappája
* dist – standalone verzió

A mappa továbbá tartalmaz egy \diag és \output mappát, melyben a modulok futtatásakor keletkező releváns és diagnosztikai kimenetek kerülnek mentésre.

### Installálás

A fejlesztői modulok futtatásához fejlesztői módban szükség van a Python környezet telepítésére. Python 3.0-nál nagyobb verzió szükséges. A későbbiekben a modulok fejlesztői környezet nélkül is futtathatóak lesznek.

1. Anaconda Python csomagkezelő telepítése

Anaconda telepítése:

URL: https://www.continuum.io/downloads

Python 3.5 version – 64-BIT INSTALLER – ezt kell választani

Minden felajánlott opciót el kell fogadni.

1. Python fejlesztői környezet és szükséges csomagok telepítése

Python telepítése

URL: <https://www.python.org/downloads/>

DOWNLOAD Python 3.5.0\* – ezt kell választani

Command prompt segítségével a forrásfile-ok mappájába (WBC mappa) kell navigálni.

A forrásfile-ok mappájában (WBC mappa) a következő paranccsal lehet installálni a szükséges Python csomagokat:

*pip install -r requirements.txt*

#### CNTK installálás

A CNTK keretrendszer a neurális hálózatok tanítását és modellként való futtatását támogatja. A sejtosztályozáshoz (cell\_classfier) szükséges installálni.

A következő linkről ki kell választani a használt architektúrának megfelelő telepítőt:

URL: <https://github.com/Microsoft/CNTK/wiki/Setup-Windows-Python>

CPU-s verziót válasszunk!

Ez után command promptból a következő parancsot kell indítanunk:

*pip install*[*https://cntk.ai/PythonWheel/CPU-Only/cntk-2.0.beta15.0-cp35-cp35m-win\_amd64.whl*](https://cntk.ai/PythonWheel/CPU-Only/cntk-2.0.beta15.0-cp35-cp35m-win_amd64.whl)

### Python forrásfile-ok futtatása

##### cell\_detector.py

A cell\_detector program egy fejlesztői verzió – ennek egy standalone változata a végleges alkalmazás képfeldolgozási back-end-jeként fog funkcionálni. A cell\_detector a vörös- és fehérvérsejtek detektálását valósítja meg.

Command prompt/Anaconda prompt segítségével a forrásfile-ok mappájába (WBC mappa) kell navigálni. Ott a következő paranccsal lehet egy kenetmintát kiértékelni:

*python cell\_detector.py -i <image\_file> -s True*

Az <image\_file> a feldolgozandó minta (.bmp, vagy .jpg) file teljes elérési útvonalát jelenti. Az eredmény az \output mappába kerül elmentésre. A \diag mappában diagnosztikai eredmények kerülnek elmentésre.

Egy teljes kenetmintákat tartalmazó mappa kiértékelése batch módban történik a következő paranccsal:

*python cell\_detector.py -i <image\_dir> -s True -o <diag\_out\_dir\_name> -b batch*

Az <image\_dir> a kenetmintákat tartalmazó mappa teljes elérési útvonala, a <diag\_out> egy mappaazonosító, mely nevű mappát az \output és a \diag mappákban elkészítünk és oda mentjük a megfelelő kimeneteket.

##### cell\_classifier.py

A cell\_classifier program egy fejlesztői verzió – ennek egy standalone változata a végleges alkalmazás képfeldolgozási back-end-jeként fog funkcionálni. A cell\_classifier a vörös- és fehérvérsejtek detektálását valósítja meg és a fehérvértesteket osztályozza 5 kategóriába.

Command prompt/Anaconda prompt segítségével a forrásfile-ok mappájába (WBC mappa) kell navigálni. Ott a következő paranccsal lehet egy kenetmintát kiértékelni:

*python cell\_classifier.py -i <image\_file> -s True*

Az <image\_file> a feldolgozandó minta (.bmp, vagy .jpg) file teljes elérési útvonalát jelenti. Az eredmény az \output mappába kerül elmentésre. A \diag mappában diagnosztikai eredmények kerülnek elmentésre.

Egy teljes kenetmintákat tartalmazó mappa kiértékelése batch módban történik a következő paranccsal:

*python cell\_ classifier.py -i <image\_dir> -s True -o <diag\_out\_dir\_name> -b batch*

Az <image\_dir> a kenetmintákat tartalmazó mappa teljes elérési útvonala, a <diag\_out> egy mappaazonosító, mely nevű mappát az \output és a \diag mappákban elkészítünk és oda mentjük a megfelelő kimeneteket xml és képi formátumokba.

##### evaluate.py

Miután elkészültek az automatikus detekciók a következő paranccsal tudjuk összehasonlítani ezeket a manuális annotálásokkal:

*python evaluate.py -m <image\_dir> -a <diag\_out\_dir> -s True*

Itt a <diag\_out\_dir> teljes elérési utat jelent.

### Standalone verziók futtatása

A cell\_classifier.exe és az evaluate.exe a fenti forrásfile-ok standalone futtatható verziói. Paraméterezésük megegyezik a fenti példákkal:

*cell\_ classifier.exe -i <image\_dir> -s True -o <diag\_out\_dir\_name> -b batch*

*evaluate.exe -m <image\_dir> -a <diag\_out\_dir> -s True*

### Algoritmusok leírása

A képfeldolgozó modul adaptív módon közelíti meg a sejtdetekciós feladatot, a lehető legkevesebb előfeltételezést használva. A feltételezések:

* Háttér alacsony szaturációjú, magas intenzítású
* Vörösvértestek magasabb szaturációjú, sötétebb kör alakú, vagy minimálisan konvex módon deformálódott alakzatok
* Vörösvértestek mérete szűk eloszlást mutat
* A vörösvértestek kitöltése a képen sem nem extrém kicsi, sem nem extrém nagy
* Fehérvérsejtek sejtmagja a legnagyobb szaturáció, egybefüggő objektumok a képeken.

Nem szükséges megadni a következőket:

* Vörösvértestek átlagos mérete (számítjuk)
* Fehérvérsejt magok színe (számítjuk)

A képi objektumok – sejtek szegmentációját adaptív módon végezzük. A pusztán két móduszú pixel-állapotot (előtér-háttér) a következő zajok terhelik:

* Nem egyenletes megvilágítás
* Optikai és digitális zajok
* Egyéb objektumok: thrombociták, …
* RBC sejtfal sötétebb, szaturáltabb
* RBC sejtek közepe világosabb
* WBC sejtplazma
* WBC sejtmag

A szegmentációs algoritmusok egy kétlépcsős klaszterezési eljáráson alapulnak. Az első lépcsőben a hátteret és az előteret választjuk szét a lehető legpontosabban, lehetőleg a detektálandó sejteket szétválasztva még akkor is, ha összeérnek. A második lépcsőben a fehérvérsejt sejtmagjait szegmentáljuk.

Az RBC detekciónál topológiai műveleteket (distance transform, watershed, local minima, morfológia) végzünk az előtér maszkon és ezek segítségével detektáljuk a sejtek középpontjait.

A WBC detekciónál a fehérvérsejt maszkon komponens analízis (morfológiai műveletek) segítségével szűrjük a potenciális WBC középpontokat.

A WBC-k sejtmagjának befoglaló téglalapját meghatározzuk. Fragmentálódott sejtmagok esetén a részeket egyesítjük egy algoritmussal, mely megkeresi a közeli komponenseket és vizsgálja a lehetséges kapcsolódásokat. A detektált sejtek befoglaló téglalapját (bounding boksz) a sejtmag maszk inerciatenzora alapján határozzuk meg - ezeket a befoglaló téglalapokat használjuk a WBC-k osztályozásához mintának. A tanításhoz a befoglaló téglalapokat 90,180,270 fokos szögben is elforgatjuk. A sejtmag RGB képének kék csatornája alapján gamma korrekciót végzünk a tanító képeken.

#### WBC osztályozás neurális hálózatokkal

A detektált fehérvérsejtek osztályozása konvolúciós neurális hálózatokkal történik. A neurális hálózat minta alapján tanul és állítja be a modell paramétereket. A hálózat a következő rétegekből áll:

convolutional\_layer\_1 = Convolution((5,5), 8, init=glorot\_uniform(), activation=relu, pad=True, strides=(1,1))(input)

pooling\_layer\_1 = MaxPooling((2,2), strides=(1,1))(convolutional\_layer\_1 )

convolutional\_layer\_2 = Convolution((9,9), 16, init=glorot\_uniform(), activation=relu, pad=True, strides=(1,1))(pooling\_layer\_1)

pooling\_layer\_2 = MaxPooling((3,3), strides=(2,2))(convolutional\_layer\_2)

convolutional\_layer\_3 = Convolution((13,13), 32, init=glorot\_uniform(), activation=relu, pad=True, strides=(1,1))(pooling\_layer\_2)

pooling\_layer\_3 = MaxPooling((4,4), strides=(2,2))(convolutional\_layer\_3)

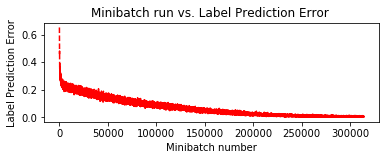
fully\_connected\_layer = Dense(128, init=glorot\_uniform())(pooling\_layer\_3)

dropout\_layer = Dropout(0.5)(fully\_connected\_layer)

output\_layer = Dense(out\_dims, init=glorot\_uniform(), activation=None)(dropout\_layer)

A kép szélét metsző automatikus annotációk "None" címkét kapnak mutatva, hogy detektáljuk itt a sejteket, de nem osztályozzuk. Azokat a wbc-ket, melyeknek a középpontja a kép szélétől 50 pixelnél nagyobb távolságban van nem detektáljuk. Az „un” kategóriába az 5 alap wbc osztályba nem eső detekciókat soroljuk.

A tanítás folymatát a következő ábra mutatja:



A tanítómintán 4.5%-os, a tesztmintán 4.9%-os klasszifikációs hibát mértünk.

### Detekciós eredmények

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Azonosító | n\_WBC | n\_Detected | n\_MatchWBC | n\_Image\_with\_annotations | n\_Image |  | Ratio\_TP | Ratio\_FP |  |  |  |  |  |  |
| 1 | 144 | 142 | 125 | 62 | 62 |  | 0,868056 | 0,119718 |  |  |  |  |  |  |
| 2 | 135 | 145 | 128 | 90 | 90 |  | 0,948148 | 0,117241 |  |  |  |  |  |  |
| 3 | 105 | 105 | 92 | 79 | 79 |  | 0,87619 | 0,12381 |  |  |  |  |  |  |
| 4 | 85 | 88 | 81 | 68 | 68 |  | 0,952941 | 0,079545 |  |  |  |  |  |  |
| 5 | 50 | 51 | 50 | 41 | 41 |  | 1 | 0,019608 |  |  |  |  |  |  |
| 6 | 59 | 62 | 55 | 45 | 45 |  | 0,932203 | 0,112903 |  |  |  |  |  |  |
| 7 | 63 | 70 | 61 | 47 | 47 |  | 0,968254 | 0,128571 |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  | Nincs annotáció | |  |  |  |
| 9 | 133 | 145 | 125 | 97 | 97 |  | 0,93985 | 0,137931 |  | Más nagyítás | |  |  |  |
| 11 |  |  |  |  |  |  |  |  |  | Rossz annotációk, rossz minőségű felvételek | | | | |
| 17 | 134 | 155 | 123 | 99 | 99 |  | 0,91791 | 0,206452 |  |  |  |  |  |  |
| 18 | 111 | 115 | 103 | 101 | 99 |  | 0,927928 | 0,104348 |  | Kézi |  |  |  |  |
| 20 | 29 | 28 | 27 | 24 | 24 |  | 0,931034 | 0,035714 |  |  |  |  |  |  |
| 21 | 149 | 154 | 142 | 108 | 108 |  | 0,95302 | 0,077922 |  |  |  |  |  |  |
| 25 | 45 | 48 | 44 | 47 | 47 |  | 0,977778 | 0,083333 |  | Más nagyítás | |  |  |  |
| 32 | 574 | 613 | 557 | 493 | 493 |  | 0,970383 | 0,091354 |  |  |  |  |  |  |
| 34 | 154 | 176 | 141 | 94 | 94 |  | 0,915584 | 0,198864 |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SUM | 1970 | 2097 | 1854 | 1495 | 1493 |  | 0,941117 | 0,11588 |  |  |  |  |  |  |

### Klasszifikációs eredmények

A hálózat tanítása 11588 darab, WBC-t tartalmazó tanítóképen történt, mely képhalmazt felbontottunk 75-25 arányban tanító és teszt halmazra.

A 11588 kép 2897 darab annotált, wbc detekciót tartalmazó képből és azok 90,180, 270 fokos elforgatott változataiból állt elő.

Sztenderd eljárás szerint a klasszifikáációs képhalmazt 75-25%-ban felosztottuk tanító és tesz képhalmazra. A teszt halmazon a klsszifikációs eredmények a következőképp alakultak:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | un | NE | EO | BA | MO | LY |  |
| un | 170 | 7 | 7 | 0 | 19 | 10 | 213 |
| NE | 11 | 1667 | 1 | 1 | 6 | 3 | 1689 |
| EO | 7 | 3 | 93 | 0 | 0 | 1 | 104 |
| BA | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| MO | 6 | 4 | 1 | 2 | 186 | 17 | 216 |
| LY | 7 | 1 | 0 | 1 | 26 | 639 | 674 |
|  | 201 | 1682 | 102 | 5 | 237 | 670 | 2897 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | Rows: annotations |  | | Cols: automatic classifications | |   Trained on | 8691 | Train success rate | 0.96 |
|  |  |  |  |
| Correct classification | 2756 | Test success rate | 0.95 |
| All classification | 2897 |  |  |
|  |  |  |  |
| Total # of images: | 11588 |  |  |