

DIPLOMAMUNKA

Lakatos Károly

Debrecen
2010

Debreceni Egyetem
Informatikai Kar

Érhálózat detektáló algoritmusok összehasonlítása retinaképekben

Témavezető:
Dr. Hajdu András
egyetemi docens

Készítette:
Lakatos Károly
programtervező matematikus

Debrecen
2010

Tartalomjegyzék

Tartalomjegyzék.....	1
Bevezetés.....	3
Szemfenéki elváltozások.....	3
Szemészeti szűrővizsgálatok.....	3
A diplomamunka célja.....	4
RGB képek, az érhálózat szegmentálás általános leírása.....	5
Illesztett szűrők.....	8
Az algoritmus megvalósítása.....	10
A kernelek meghatározása.....	10
Az algoritmus értékelése.....	13
Érhálózat detektálás a másodrendű entrópia felhasználásával.....	14
Az algoritmus folyamatábrája.....	15
Zöld színcsatorna kiválasztás.....	16
Maszkgenerálás.....	16
Illesztett szűrők alkalmazása.....	18
Mediánszűrés.....	18
Átmenetmátrix meghatározás.....	19
Küszöbölés a másodrendű entrópiával.....	21
Utófeldolgozás.....	23
Az algoritmus értékelése.....	24
Megvalósítás.....	26
A program felépítése.....	27

Az érhálózat detektorok összehasonlítása.....	28
Szenzitivitás.....	29
Specificitás.....	29
Pontosság.....	29
További mérőszámok.....	30
A detektor pontosságának vizsgálata különböző paraméterek esetén.....	31
A detektor összehasonlítása más algoritmusokkal.....	32
Összefoglalás.....	35
Függelék.....	36
Köszönetnyilvánítás.....	39
Irodalomjegyzék.....	40

Bevezetés

Magyarországon néhány évvel ezelőtt a lakosság 5 százaléka, körülbelül félmillió ember volt cukorbeteg, ez a szám napjainkra 760 ezerre nőtt. Világszerte a cukorbetegek száma eléri a 220 millió főt, a legfrissebb felmérések szerint számuk 2025-re 333 millió lesz.

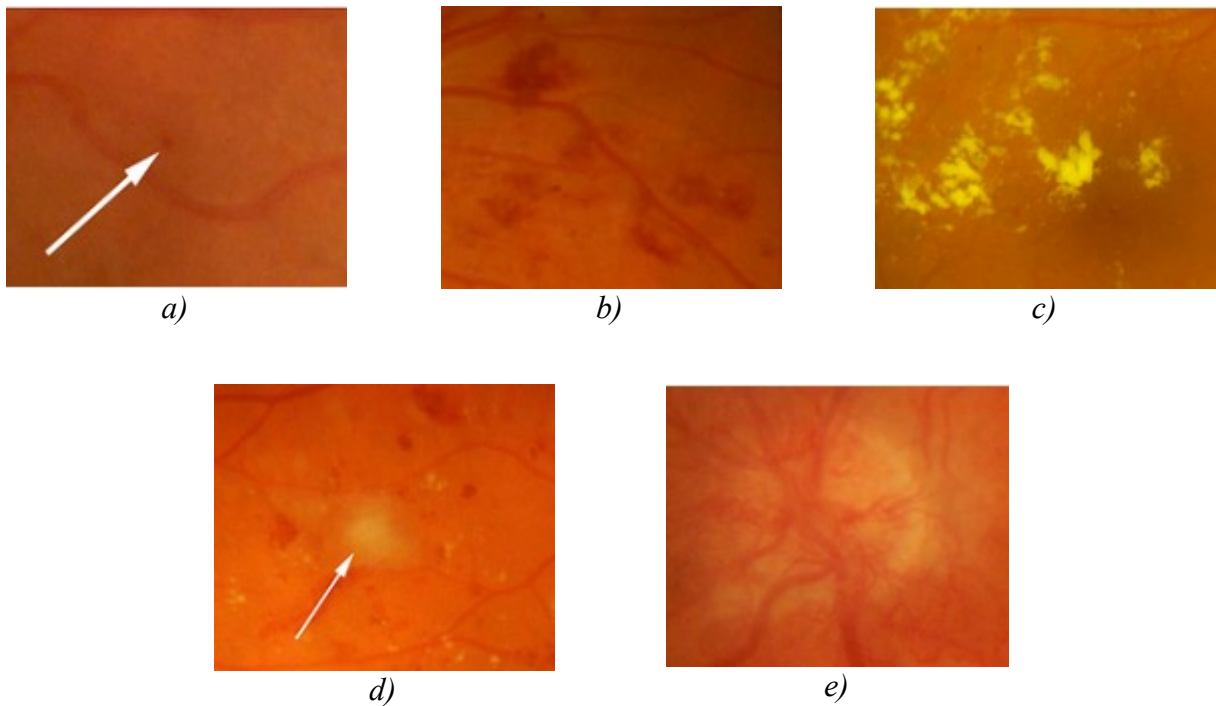
A diabéteszes retinopátia a szem, ezen belül a retina megbetegedése. 1-es típusú cukorbetegknél 15 év után kb. 90%-ban, 2-es típusú diabéteszben 25%-ban jelentkezik. Európában a szerzett vakságok 30%-a a cukorbetegség miatt következik be.

Szemfenéki elváltozások

A legkorábban detektálható szövődmények a mikroaneurizmák, amelyek apró duzzanatok a retina hajszálerein. Ezek gyengítik az érfalat és a meggyengült érfalon át savó szivároghat ki, mely retinaödémát és az ún. kemény exudátumok létrejöttét eredményezi. Ezt nevezik háttérretinopátiának, amely 20 éves diabétesztartam után a cukorbetegek túlnyomó többségénél megtalálható. A retinopátia előrehaladtával, a mikroaneurizmákban a vér megalvadhat és lágy exudátumok keletkezhetnek és elzárhatják a hajszálereket. Ez tovább fokozza az érburjánzásra való hajlamot, újabb és újabb erek jöhetnek létre, melyek tökéletlenek, elpattanhatnak és kisebb-nagyobb bevérzéseket okozhatnak. Ha az újonnan képződött erek benőnek a szem üvegtestébe és ott okoznak bevérzést, akkor a látás elhomályosul. Ezt a stádiumot proliferatív retinopátiának nevezik [1].

Szemészeti szűrővizsgálatok

A szemfenékről készített képek elemzését végző szakemberek gyakran elemző központba, ún. Reading Center-be tömörülnek, ahol a szemészeti vizsgálóhelyekről kapott képeket elemzik, majd az eredményt visszaküldik a vizsgálatot kérő orvosnak. Az elemzések hatékonyságának növelése érdekében igény volt egy olyan előszűrő rendszer kifejlesztésére, amely elkülöníti egymástól az egészséges és a diabéteszes retinopátiát ábrázoló szemfenéki képeket. Egy ilyen előszűrő rendszerben fontos szerepet játszik a helyesen detektált érhálózat.



1. ábra: A szemfenék különböző elváltozásai:

a) mikroaneurizma b) retinaödéma c) kemény exudátum

d) lágy exudátum e) proliferatív retinopátia

A diplomamunka célja

Diplomamunkámban először be szeretnék mutatni részletesen két érhálózat detektort, az illesztett szűrőket és az arra épülő másodrendű entrópia alapján detektáló módszert. Ez az algoritmus elég pontos ahhoz, hogy a későbbiekben fel lehet használni különböző szemfenéki elváltozások automatikus felismeréséhez.

A diplomamunka második részében pedig megadok néhány, a szegmentált érhálózat jóságát, pontosságát mérő értéket, amelyekkel össze lehet hasonlítani az algoritmust akár saját magával különböző paraméterek esetén, akár más, a szakirodalomban megtalálható algoritmus eredményével.

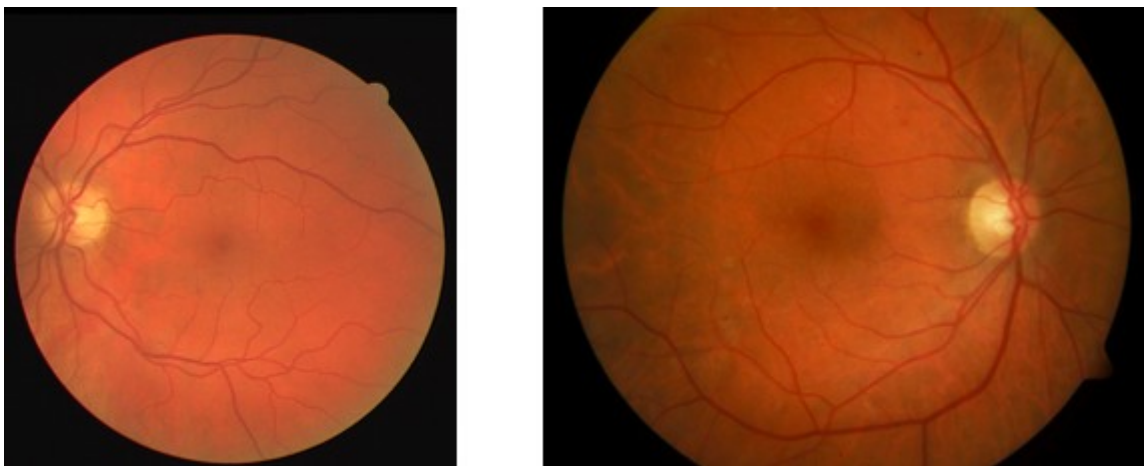
RGB képek, az érhálózat detektálás általános leírása

A bemutatandó érhálózat detektáló algoritmusok input képei az alábbi két, kutatási célra szabadon felhasználható adatbázisból származnak:

1. DRIVE (Digital Retinal Images for Vessel Extraction) [2]
2. DIARETDB1 (Standard Diabetic Retinopathy Database) [3]

A retináról készített felvételek általában előtétlencsével, speciális fundus kamerákkal készülnek és multispektrális (RGB) képet eredményeznek. Az RGB esetében minden egyes pixelhez három színcsatornán (vörös, zöld, kék) rendelnek intenzitásokat, mely intenzitások 0-255-ig vehetnek fel egész értékeket, így színcsatornánként 8 bit/pixel színinformációt tárolnak.

A DRIVE adatbázis 40 db képet tartalmaz, melyek 565×584 pixel felbontásúak, míg a DIARETDB1 89 db, a DRIVE-nál egy kicsit rosszabb minőségű, azonban sokkal nagyobb felbontású (1500×1152 pixel) képet. A DRIVE esetében minden képhez tartozik egy szemész szakorvos által manuálisan szegmentált érhálózat is, amelyek a diplomamunka későbbi részében fontos szerepet játszanak a detektorok pontosságának mérésében, mivel ezekhez hasonlítjuk különböző módszerekkel az algoritmusok által szolgáltatott output képeket.



2. ábra: Egy-egy kép a felhasznált adatbázisokból.

Az első kép a DRIVE-ból, a második a DIARETDB1-ből.

Az RGB képek több szempontból is hátrányosak a detektálást illetően:

1. A fundus képek színeloszlása nagy változatosságot mutat a betegek között, többek között az etnikai hovatartozásuk miatt, ugyanis a pigmentek jelen vannak a retinában is, az ún. pigmentált rétegben (retinal pigment epithelium, RPE), ezért a sötét pigmenttel rendelkező embereknél sokkal sötétebb a retinájukról készített felvétel, mint a világosabb pigmenttel rendelkező társaiknál.
2. Az erek színe nagyban függ attól, hogy hol helyezkedik el a retinán, ugyanis a vastagabb erekre sokkal nagyobb hatással van a fényvisszaverődés, amit a kamera vakuja okoz a látóidegfőről történő tükröződés által. A kisebb, vékonyabb erekre ez nem vonatkozik annyira és sokkal kevesebb fényt is kapnak.
3. A színes képeken történő detektálás gyakran eredményez hibás érhálózatot.

E hátrányok miatt a színes, RGB képet először szürkeskálásra kell alakítanunk, ami egy olyan kép, melyben minden pixelt egy 0-255-ig terjedő egész szám határoz meg. Általánosságban elmondható, hogy a szemfenék képeken a zöld színcsatorna hordozza a legtöbb információt az érhálózattal kapcsolatban, mivel itt különülnek el a legjobban az erek a háttértől (az erek sötétebbek, míg a háttér világosabb), ezért a bemutatandó algoritmusok is mindig ezt a színcsatornát használják az erek detektálására. Mivel a zöld színcsatorna is csak 0-255-ig vehet fel egész értékeket, ezért ez már tekinthető szürkeskálásnak is, így az algoritmusoknak elegendő az, ha a nekik inputként megadott RGB képből kiválasztják a zöld színcsatornát.



3. ábra: A 2. ábra első képének vörös, zöld és kék színcsatornája.

Az algoritmusok kimenete mindig egy olyan logikai értékeket tartalmazó kép (bináris kép), melyben az 1-es jelöli azt, ha az illető pixel eleme az érhálózatnak és a 0-s pedig azt mondja meg, hogy az adott pixel nem érpont, azaz:

$$bv(x, y) = \begin{cases} 1, & \text{ha } I(x, y) \in \text{érhálózat} \\ 0, & \text{egyébként} \end{cases}$$

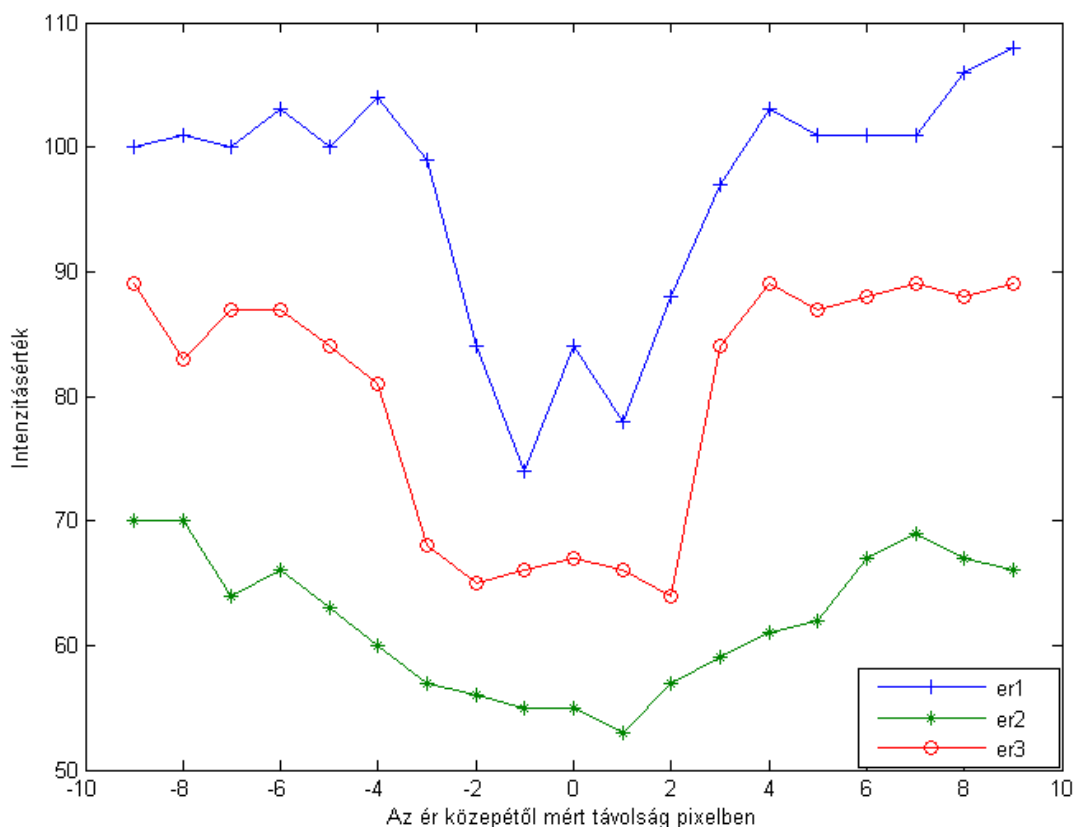
ahol:

- $bv(x, y)$: az érhálózatot reprezentáló bináris kép
- $I(x, y)$: az eredeti szemfenék kép

Illesztett szűrők

Chaudhuri 1989-ben megfigyelte [4], hogy a retinaképeken az ereknek három fontos tulajdonsága van:

1. Az ereknek általában kicsi a görbületük, lineáris szakaszokkal jól meg lehet őket közelíteni.
2. Az erek tükröződése sokkal kisebb, mint más felületek tükröződése a retinán, ezért viszonylag sötétebbek a képeken. A 4. ábrán látható, hogy függőleges irány mentén a pixelek szürkeskálás intenzitásértékeit alapul véve, az erek, mint völgyek modellezhetők. A görbe lokális minimuma az adott ér közepét határozza meg.



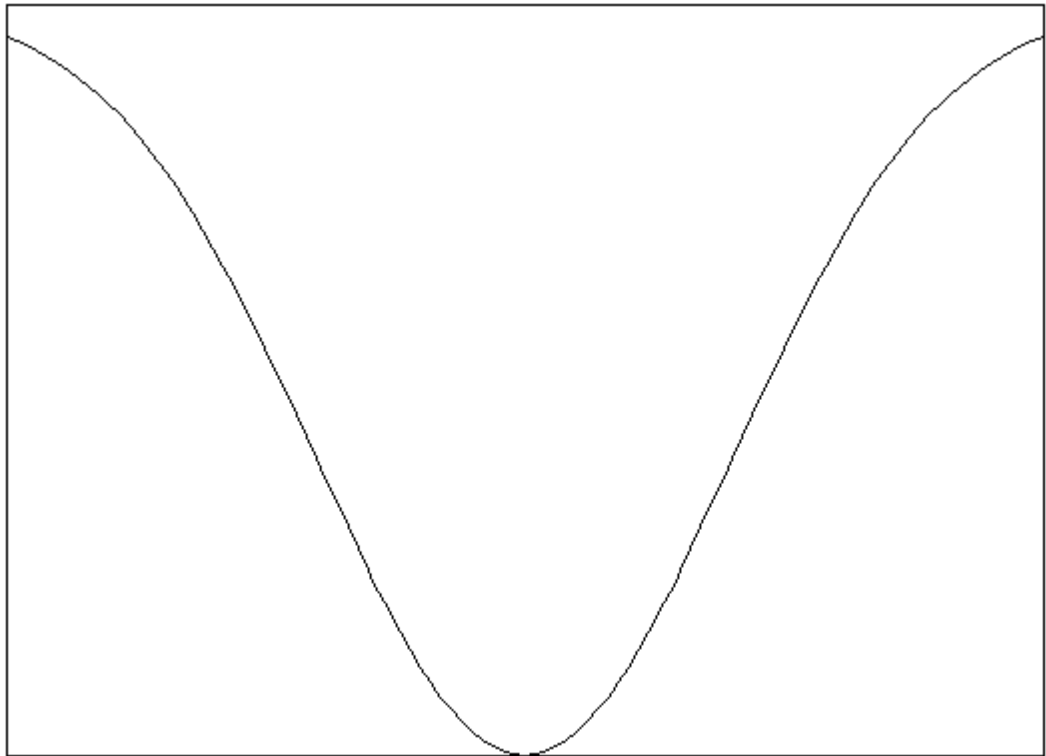
4. ábra: Három különböző ér függőleges irányban vett intenzitásértékei.

Ebből arra következtettek, hogy az erek jól közelíthetők egy megfelelő Gauss görbével:

$$G(x, y) = \frac{-1}{2\pi\sigma^2} e^{-\frac{d^2}{2\sigma^2}}$$

ahol:

- d : a függőleges irányú távolság az (x, y) pont és az ér elméleti közepe között.
- σ : a Gauss görbe paramétere, jelen algoritmus esetében $\sigma=2$ választással.



5. ábra: A megfelelő Gauss görbe $\sigma=2$ paraméterrel.

3. A látóidegfőből kilépve, az erek szélessége fokozatosan csökken, általánosságban elmondható róluk, hogy 2-10 pixel (36-180 μm) szélesek.

Az algoritmus megvalósítása

A kétdimenziós képeken az erek tetszőleges szöget bezárhatnak az y tengellyel, ezért a konvolúciós kernelt (Gauss görbét) minden lehetséges irányban el kell forgatni és a kapott eredményeket össze kell hasonlítani: minden egyes pixelre a maximális értéket kell kiválasztani.

A kernel matematikailag a következőképpen határozható meg:

$$K(x, y) = -e^{-\frac{x^2}{2\sigma^2}} \quad |y| \leq \frac{L}{2}$$

ahol az L az eret approximáló szakasz hossza, amit tapasztalati úton, normál és rendellenes szemfenék képek összehasonlításával határoztak meg, jelen algoritmus esetében $L=9$. Ez a kernel azonban csak azokat az ereket képes detektálni, amelyek párhuzamosak (vagy legalábbis nagyon kis szöget zár be) az y tengellyel, ezért a kernelt el kell forgatni.

Figyelembe kell még vennünk azt is, hogy a retinaképek általában zajjal terheltek, ezért a kernel értékeinek a meghatározásakor ügyelnünk kell arra, hogy a zajra a konvolúció eredménye nulla legyen. Ezt legegyszerűbben úgy tehetjük meg, hogy a konvolúciós kernel valamennyi értékéből kivonjuk a konvolúciós kernel értékeinek az átlagát.

A kernelek meghatározása

Legyen $p(x, y)$ a kernel egy pontja és θ_i legyen a detektálni kívánt ér és az y tengely által bezárt szög. A forgatás mátrixával összeszorozva az eredeti kernelt megkaphatjuk az elforgatott kernelt.

Az i fokkal történő forgatás mátrixa:

$$r_i = \begin{pmatrix} \cos \Theta_i & -\sin \Theta_i \\ \sin \Theta_i & \cos \Theta_i \end{pmatrix}$$

Így az i fokkal elforgatott kernel a

$$p_i = p * r_i \quad (i=0, 15, \dots, 165)$$

összefüggéssel határozható meg.

Ahhoz, hogy az összes lehetséges irányú eret detektálni tudjuk, 15° -onként kell forgatnunk a kernelt, így összesen 12 különböző kernellel kell a szemfenék képet konvolválnunk és közülük pixelenként a maximális értéket kiválasztani.

0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0
0	0	0	0	0	4	4	3	0	0	0	0	0	0	0	0
0	0	0	0	4	4	3	2	0	0	0	0	0	0	0	0
0	0	0	4	4	3	2	0	-2	-4	-5	-6	0	0	0	0
0	4	4	3	2	0	-2	-4	-5	-6	-5	-4	0	0	0	0
4	4	3	2	0	-2	-4	-5	-6	-5	-4	-2	0	0	0	0
0	3	2	0	-2	-4	-5	-6	-5	-4	-2	0	2	3	4	0
0	0	0	-2	-4	-5	-6	-5	-4	-2	0	2	3	4	4	0
0	0	0	-4	-5	-6	-5	-4	-2	0	2	3	4	4	0	0
0	0	0	0	-6	-5	-4	-2	0	2	3	4	4	0	0	0
0	0	0	0	0	-4	-2	0	2	3	4	4	0	0	0	0
0	0	0	0	0	0	0	2	3	4	4	0	0	0	0	0
0	0	0	0	0	0	0	3	4	4	0	0	0	0	0	0
0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0

6. a)

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	4	3	2	1	-2	-5	-6	-5	-2	1	2	3	4	0	0
0	4	3	2	1	-2	-5	-6	-5	-2	1	2	3	4	0	0
0	4	3	2	1	-2	-5	-6	-5	-2	1	2	3	4	0	0
0	4	3	2	1	-2	-5	-6	-5	-2	1	2	3	4	0	0
0	4	3	2	1	-2	-5	-6	-5	-2	1	2	3	4	0	0
0	4	3	2	1	-2	-5	-6	-5	-2	1	2	3	4	0	0
0	4	3	2	1	-2	-5	-6	-5	-2	1	2	3	4	0	0
0	4	3	2	1	-2	-5	-6	-5	-2	1	2	3	4	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

6. b)

6. ábra: Két kernel, az a) a 45° -os, a b) a függőleges iránynak megfelelő.

$$(\sigma=2, L=9)$$

Mivel a Gauss görbe x irányban mindkét oldalán végtelenül hosszú, ezért le kellett vágni ezeket az $u=\pm 6$ határokon, azaz a görbét csak a $[-6, 6]$ intervallumon értelmezzük.

Ekkor egy pontnak a környezetét az alábbi módon határozhatjuk meg:

$$N = \left\{ (u, v) \mid |u| \leq 6, |v| \leq \frac{L}{2} \right\}$$

Tehát ekkor az i -edik kernel egy értéke:

$$K_i(x, y) = -e^{\frac{-u^2}{2\sigma^2}} \quad \forall p_i \in N$$

Ha A -val jelöljük az N elemeinek a számát, akkor a kernel átlagát az alábbi összefüggéssel határozhatjuk meg:

$$m_i = \frac{\sum_{p_i \in N} K_i(x, y)}{A}$$

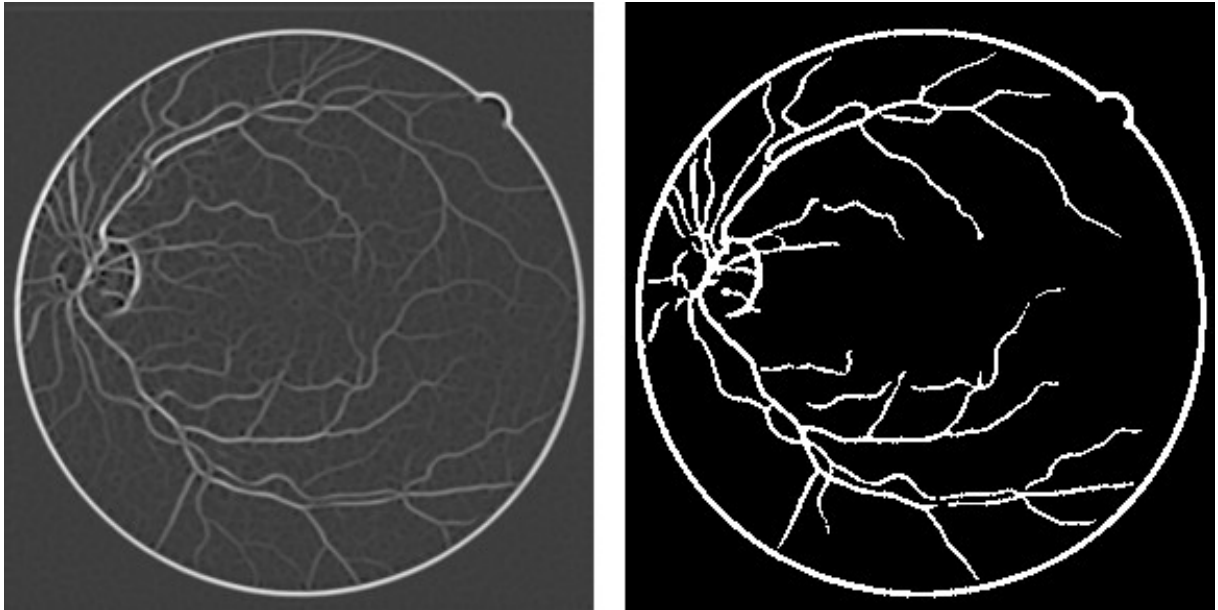
Így már kiszámíthatóak az algoritmus által is használt konvolúciós kernelek:

$$K'_i(x, y) = K_i(x, y) - m_i \quad \forall p_i \in N$$

Az m_i -kre a már korábban említett, zajokra vonatkozó megállapítás miatt volt szükség, ebben az esetben a zajos pixelekre a konvolúció 0-val tér vissza.

Az algoritmus értékelése

Az illesztett szűrők algoritmusja sikeresen detektálta az ereket, de meg kell jegyezni, hogy az eredmény még szürkeárnyalatos kép, ezért szükség van egy küszöbölésre, amivel különválasztjuk egymástól az érponthoz és a háttérponthoz. Az eljárás eredménye a már korábban említett érhálózatot reprezentáló bináris kép.



7. ábra: Az illesztett szűrő eredménye és annak küszöböltje $K=80$ -al.

A küszöbölés lényege, hogy az algoritmus kimeneti képének pixeleit egy rögzített K küszöbszám alapján két osztályba soroljuk: ha az adott pixel intenzitásértéke nagyobb vagy egyenlő, mint a K küszöbszám, akkor a bináris képen a pixelhez 1-et rendelünk, egyébként pedig 0-t.

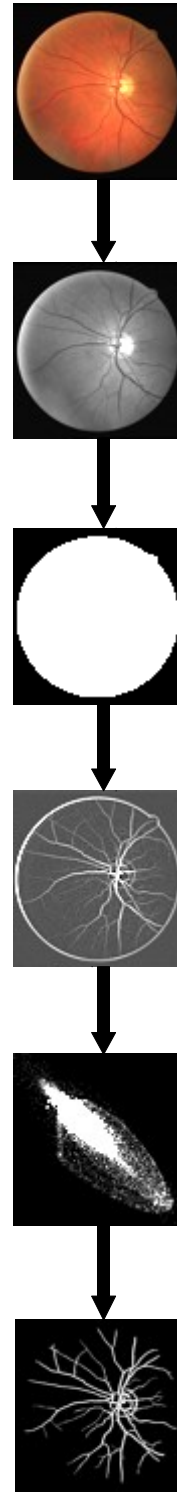
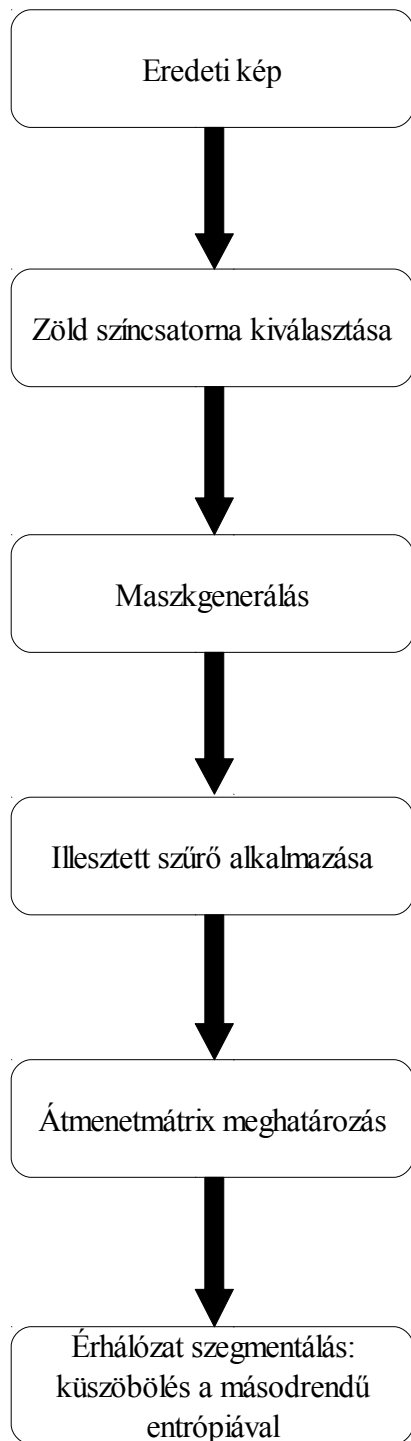
Azonban az optimális küszöb megválasztása nem olyan egyszerű dolog, mint elsőre tűnik. Más-más retinaképen más-más küszöbre van szükség és mindezt természetesen automatikusan kell megválasztani emberi beavatkozás nélkül. Erre az optimális küszöbválasztásra mutatok be most egy módszert.

Érhálózat detektálás a másodrendű entrópia felhasználásával

Ahogy már korábban láthattuk, az illesztett szűrőkkel sikeresen lehet detektálni a szemfenék képekről az érhálózatot, viszont probléma van vele a küszöbölésnél, amikor az optimális küszöböt kell meghatároznunk. A másik hátránya az, hogy átlagosan csak kb. 88%-os pontosságot ér el, ami azt jelenti, hogy a manuálisan szegmentált érhálózattal ekkora százalékban egyezik meg. Valószínűleg azért ilyen alacsony a találati aránya, mert az algoritmus nem képes detektálni a vékonyabb ereket, ami azzal jár, hogy ha a detektált érhálózatot arra szeretnénk felhasználni, hogy azon mikroaneurizmákat ismerjünk fel, akkor ezt egyszerűen nem tudjuk megtenni, mivel ezek az elváltozások éppen a hajszálereken jelentkeznek. Chaudhuri 1989-ben publikálta ezt a módszert, azóta számtalan olyan algoritmust mutattak be, amelyek sokkal jobb (93-94% körüli) pontosságot értek el. Ezek közül az egyik legjobbat M. Niemeijer tette közzé [5]: minden pixelhez meghatározott egy sajátosság vektort, ami a Gauss szűrőből, annak első- és másodrendű deriváltjából állt különböző értékekkel skálázva (1,2,4,8,16). Aztán ezekből a sajátosságvektorokból kNN osztályozót felhasználva állítják elő az érhálózatot. Ezt az eljárást már nevezhetjük nagyon pontosnak, mivel a DRIVE adatbázison 94,16%-os pontosságot ért el. Azt azonban figyelembe kell vennünk, hogy ez nem egy automatikus módszer, emberi beavatkozást igényel és az eredmény nagyban függ a tanító minta minőségétől is.

Most egy olyan technikát mutatok be, amely kihasználja azt, hogy az illesztett szűrők algoritmusával előállított MFR (Matched Filter Response) képen egy adott pixel intenzitásértéke nem független a környezetében lévő pixelek intenzitásértékeitől [6]. Az intenzitás-skála alapján bizonyos szabályok alapján meghatározzuk egy 256×256 méretű mátrixot, mely elemei közül a maximális másodrendű entrópiával rendelkezőt fogjuk kiválasztani optimális küszöbnek, amivel már elő tudjuk állítani az érhálózatot reprezentáló bináris képet. Az alkalmazott eljárás megfelelő paraméterválasztással több mint 95%-os pontosságot ér el.

Az algoritmus folyamatábrája



8. ábra: Az algoritmus folyamatábrája.

Zöld színcsatorna kiválasztása

Mint már korábban említettük a retinaképeken a zöld színcsatorna hordozza a legtöbb információt az érhálózattal kapcsolatban, mivel itt különülnek el a legjobban a háttértől, sokkal sötétebbek annál. Ennél a csatornánál megfigyelhető az is, hogy sokkal kisebb mértékben terhelt zajjal a többinél. További előny az is, hogy az RGB képnek csak egyik részét kell feldolgoznunk, ami jelentősen lecsökkenti az algoritmus futási idejét (kb. a harmadára).

Maszkgenerálás

A maszkgenerálás arra szolgál, hogy a szemfenék lényeges információt hordozó területét különválasszuk a kép háttérétől és külön vizsgáljuk tőle. Ezt a területet szokás ROI-nak (Region Of Interest) nevezni. A fundus felvételek esetében ROI-nak hívjuk azt a képrészt, ahol az érhálózat megtalálható.



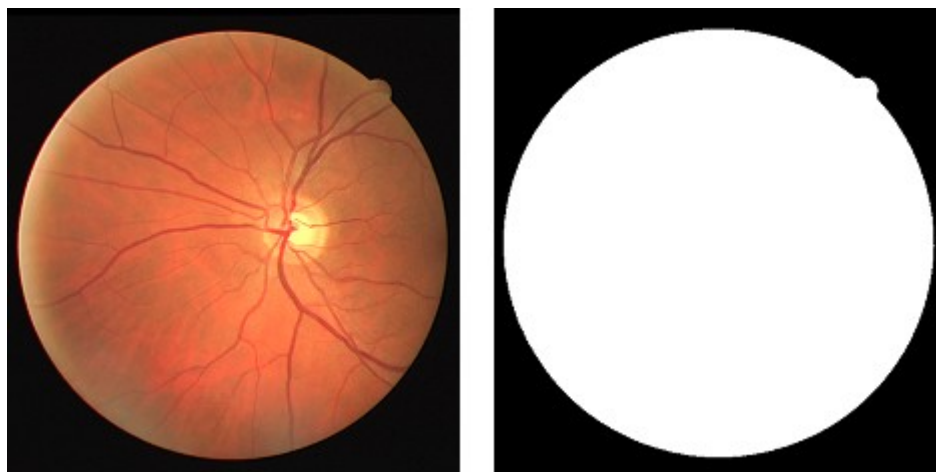
9. ábra: ROI, a bevonalkázott rész nem tartozik a ROI-hoz.

A felhasznált maszk egy logikai mátrix (bináris kép), amelynek a mérete megegyezik az eredeti képpel és a benne található érték 1-es, ha az adott pixel a színes képen a ROI-hoz tartozik, egyébként pedig 0-s.

Létezik egy olyan módszer, mellyel sikeresen lehet maszkot generálni: matematikai morfológiai operátorokat használunk fel a ROI meghatározására [7].

Az eljárás lépései:

1. Az eredeti színes kép vörös színcsatornájára egy küszöbölést hajtunk végre $k=35$ küszöbvel, mely meghatározása tapasztalati úton történt.
2. A küszöbölt képen egy nyitást hajtunk végre egy 3×3 -as négyzet alakú struktúraelemmel.
3. Az eredményen egy zárást hajtunk végre ugyanazzal a struktúraelemmel.
4. Végül még egy további erodálást alkalmazunk a kapott képre, szintén a 3×3 -as struktúraelemmel.



10. ábra: Eredeti kép és az annak generált maszk.

Az eredeti képet pixelenként összeszorozva a generált maszkkal elértük azt, hogy már csak a ROI-n belül végezzünk műveleteket.

Illesztett szűrők alkalmazása

Mint már korábban láthattuk az illesztett szűrők alkalmazásával sikeresen kiemelhetjük az ereket a képről. Továbbá azért is jó ennek a módszernek az alkalmazása, mivel az érhálózatnak az egészét határozza meg, ellentétben néhány éldetektáló algoritmussal, amelyek esetében a kapott erek megszakadhatnak, nem folytonosak. Fontos megjegyezni, hogy az illesztett szűrőket nem csak az irodalomban megtalálható $\sigma=2$ paraméterrel futtattuk le, hanem más σ értékekre is (pl.: $\sigma=1$, $\sigma=0.5$, $\sigma=0.1$), amely nagyban javította a találati arányokat, sokkal több vékony eret detektált, mind a DRIVE, mind a DIARETDB1 adatbázis esetén.

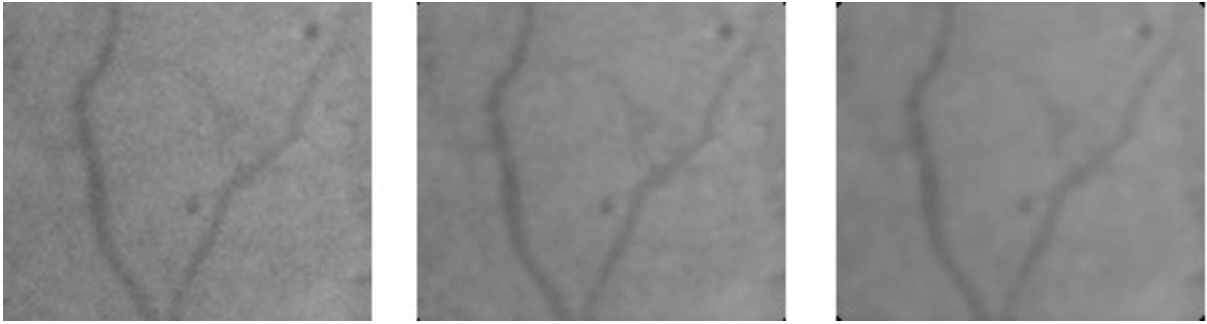
Ehhez a lépéshez hozzátartozhat még egy előfeldolgozási eljárás is, melynek a célja a képeken lévő zaj eltávolítása. Ez az előfeldolgozási eljárás a mediánszűrés.

Mediánszűrés

Az (x, y) koordinátájú pont $N \times N$ -es környezetét (N páratlan) vizsgálva az (x, y) koordinátájú pont intenzitását azon intenzitás értékkel helyettesítjük, amely – a környezetében lévő N^2 darab intenzitás értéket nagyság szerint rendezve – a rendezett sorozatban a középső helyen áll.

A mediánszűrés hatására a képben található kiugró intenzitásértékű képpontok fényessége közeledni fog a környezetük fényességéhez, szélsőséges esetben, ha egy homogén képrészletben található egy az átlagfényességnél sötétebb vagy világosabb képpont, az egyszerűen eltűnik a képből. A mediánszűrő ezen tulajdonsága miatt nagyon hatékonyan távolítja el az impulzusszerű zajokat.

Általában 5×5 -ös medián szűrőt használtunk a DIARETDB1 adatbázisból származó képeken, mely nagyon jól kiküszöbölte a zajokat.



11. ábra: Felnagyított képrészlet a DIARETDB1-ből.

Az első az eredeti, majd a 3×3 -as és az 5×5 -ös mediánszűrés eredménye.

Átmenetmátrix meghatározás

Az átmenetmátrixa egy szűrkeskálás képnek egy $L \times L$ -es (L a szűrkeskála elemeinek a száma) négyzetes mátrix a $W=[t_{ij}]$, amelynek az elemei megadják minden egyes intenzitáspárra az átmenetek számát, további információkat szolgáltatva ezzel a képen lévő struktúrákról. Az átmenetmátrix meghatározásához 4-szomszédságot használtunk, ami azt jelenti, hogy az I kép egy $I(x, y)$ pixelének környezetén az alábbi 4 pixelt értjük:

$$I(x-1, y), I(x+1, y), I(x, y-1), I(x, y+1)$$

Tegyük fel például, hogy $t_{32,35}=124$ egy I képre, akkor ez azt adja meg, hogy az I képen a 32 intenzitásból a 35 intenzitásba 4-szomszédságot használva 124 átmenet található.

Az átmenetmátrix egy t_{ij} eleme az alábbi módon határozhatjuk meg:

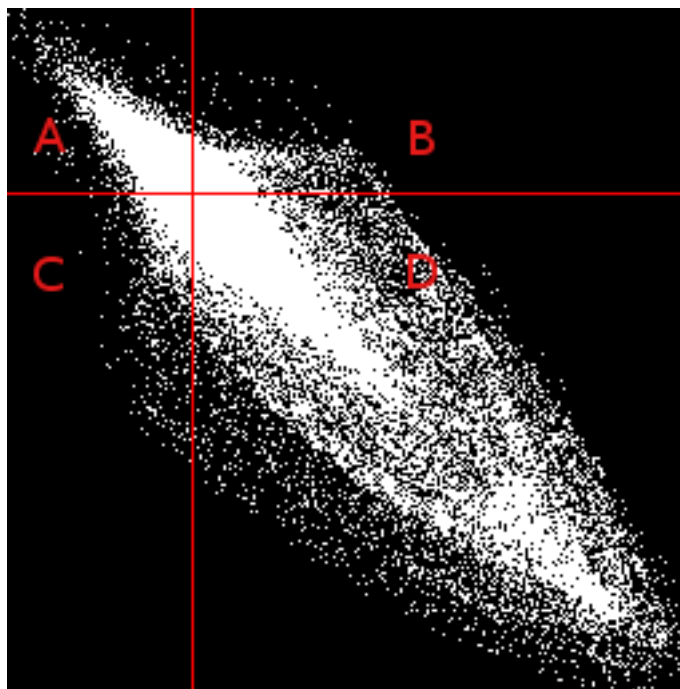
$$t_{ij} = \sum_{l=0}^{L-1} \sum_{k=0}^{L-1} \delta$$

ahol:

$$\delta = \begin{cases} 1, & \text{ha } \begin{cases} I(l, k) = i \text{ és } I(l, k+1) = j \\ \text{vagy} \\ I(l, k) = i \text{ és } I(l+1, k) = j \end{cases} \\ 0, & \text{egyébként} \end{cases}$$

Legyen K egy a küszöböléshez használt érték. Ekkor a K az átmenetmátrixot 4 részre osztja: A-ra, B-re, C-re és D-re. Ezek az alábbiaknak felelnek meg:

- A: érhálózat
- B: átmenet az érhálózat és a háttér között
- C: háttér
- D: átmenet a háttér és az érhálózat között



12. ábra: A K által 4 részre osztott átmenetmátrix.
($K=68$)

A következő részben meghatározzuk a küszöböléshez felhasználható optimális K küszöbszámot.

Küszöbölés a másodrendű entrópiával

A feladatunk tehát az, hogy meghatározzuk azt az optimális K küszöbszámot, amely megfelelően bontja szét az átmenetmátrixot 4 részre. Ehhez vezessük be az alábbi mennyiségeket [8]:

1. Egy adott t_{ij} átmenet valószínűsége:

$$p_{ij} = \frac{t_{ij}}{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} t_{ij}}$$

2. Az A részbe tartozás valószínűsége:

$$P_A = \sum_{i=0}^K \sum_{j=0}^K p_{ij}$$

3. A C részbe tartozás valószínűsége:

$$P_C = \sum_{i=K+1}^{L-1} \sum_{j=K+1}^{L-1} p_{ij}$$

4. Normalizált valószínűség A esetén:

$$P_{ij}^A = \frac{p_{ij}}{P_A} \quad i=(0,1, \dots, K) \quad j=(0,1, \dots, K)$$

5. Normalizált valószínűség C esetén:

$$P_{ij}^C = \frac{p_{ij}}{P_C} \quad i=(0,1, \dots, K) \quad j=(0,1, \dots, K)$$

6. Az objektum (érhálózat) másodrendű entrópiája:

$$H_A^{(2)}(K) = -\frac{1}{2} \sum_{i=0}^K \sum_{j=0}^K P_{ij}^A \log_2 P_{ij}^A$$

7. Az előzőhöz hasonlóan a háttér másodrendű entrópiája:

$$H_C^{(2)}(K) = -\frac{1}{2} \sum_{i=K+1}^{L-1} \sum_{j=K+1}^{L-1} P_{ij}^C \log_2 P_{ij}^C$$

8. A teljes másodrendű entrópia:

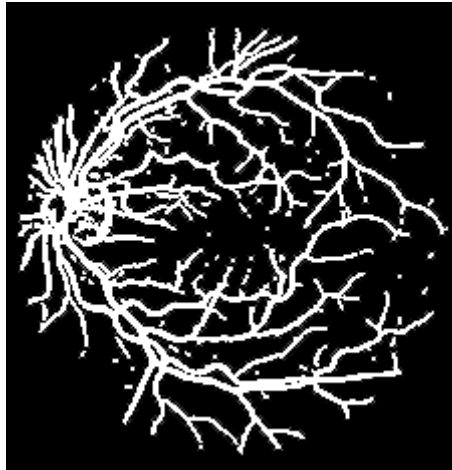
$$H_T^{(2)}(K) = H_A^{(2)}(K) + H_C^{(2)}(K)$$

Az optimális küszöbértéket az a K paraméter adja meg, amely esetén a $H_T^{(2)}(K)$ teljes másodrendű entrópia maximális, azaz:

$$K_{opt} = \arg \left\{ \max_{K \in \{0, 1, \dots, L-1\}} H_T^{(2)}(K) \right\}$$

Utófeldolgozás

Ahogy a 13. ábrán láthatjuk, az eredményül kapott képen találhatóak helytelenül azonosított pixelek. Ahhoz, hogy tiszta, teljes érhálózatot kapjunk, ezeket el kell távolítani, szükség van egy utófeldolgozási lépésre.

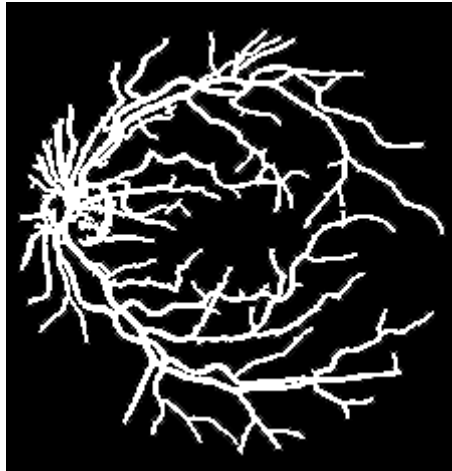


13. ábra: Az eljárás eredménye.

Az utófeldolgozást az alábbi módon tehetjük meg:

1. A 8-szomszédságot alapul véve megkeressük az így kapcsolódó objektumokat.
2. Minden objektumnak meghatározzuk a méretét, azaz azt, hogy hány pixelből tevődnek össze.
3. Azon objektumokat, amelyek egy meghatározott S értéknél kevesebb képpontból állnak eltávolítjuk a képről.

Az algoritmusunk esetében az $S=1000$ értéket választottuk, amely elegendő volt ahhoz, hogy kiküszöböljük a tévesen detektált objektumokat.



14. ábra: Az utófeldolgozás eredménye.

Az algoritmus értékelése

Az algoritmust mindkét korábban említett adatbázisra, a DRIVE-ra és a DIARETDB1-re is lefuttattuk.

A DRIVE esetén nem volt szükség előfeldolgozási módszerre, mivel nagyon jó minőségű képeket tartalmaz. Több σ paraméter kipróbálása után a legjobb eredményt a $\sigma=0.5$ paraméter adta, ellentétben a hasonló érhálózat detektorokat bemutató cikkek által használt $\sigma=2$ értékkel. Az algoritmus ez esetben nagyon gyors: képenként átlagosan 5 másodperc alatt detektálta az érhálózatot.

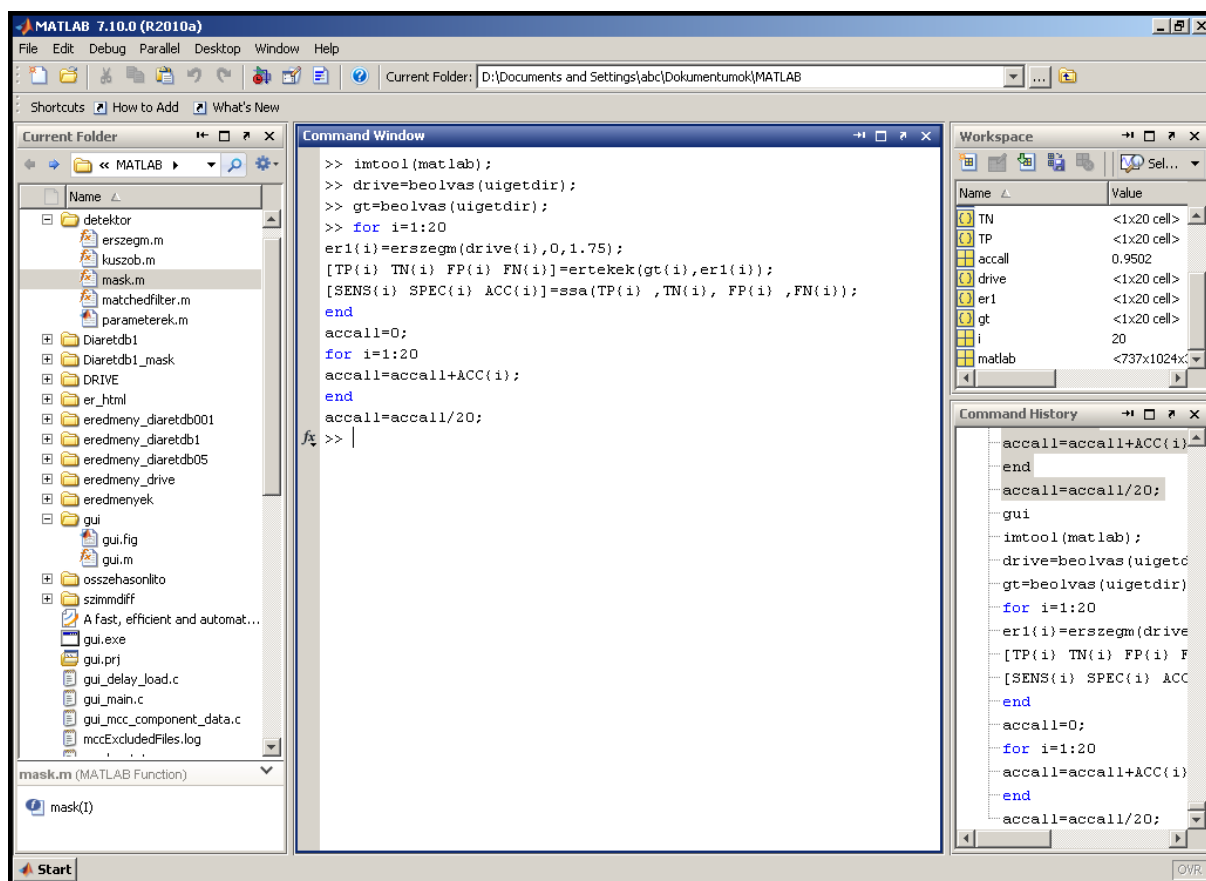
A DIARETDB1 adatbázison már használtuk a korábban bemutatott előszűrési módszert, a medián filterezést 5×5 -ös ablakkal, amellyel sikeresen el tudtuk távolítani a képeken lévő zajt. Ezeknél a képeknél manuálisan szegmentált érhálózat hiányában ugyan nem tudtuk mérni a pontosságot, azonban $\sigma=0.1$, $\sigma=0.5$ és $\sigma=1$ értékeket alkalmazva is kielégítő eredményeket kaptunk. Fontos továbbá megjegyezni, hogy ezen felvételek nagyobb felbontása, valamint rosszabb minősége miatt az optimális küszöbszámot el kellett tolnunk egy tapasztalati értékkel, $K_{\text{elt}}=4$ -gyel. A futási idő itt is nagyon gyors: kb. 26-27 másodpercre volt szükség képenként.

Utófeldolgozásra mindkét esetben szükség volt, mivel nélküle a hibásan detektált érponatok nagyban rontották a pontosságot.

Összességében elmondható az algoritmusról, hogy nagyon gyors, pontos és ami a legfőbb erénye: teljesen automatikus, nincs szükség emberi beavatkozásra.

Megvalósítás

Az algoritmusok megvalósítására egy olyan környezetet kerestünk, amellyel könnyen és gyorsan lehet implementálni a detektorokat, valamint rendelkezik olyan keretrendszerrel, ami segítséget nyújt a különböző képfeldolgozási eszközök, módszerek használatában. Ezen okok miatt a választás a MathWorks cég MATLAB programcsomagjára esett [9], melyben az eljárások nagy nehézség nélkül programozhatóak, a hozzá kapcsolódó Image Processing Toolbox pedig nagyon sok képfeldolgozáshoz felhasználható eljárást eleve tartalmaz, így azokat nem kell nekünk külön megvalósítanunk.



15. ábra: Képernyőkép a MATLAB-ról.

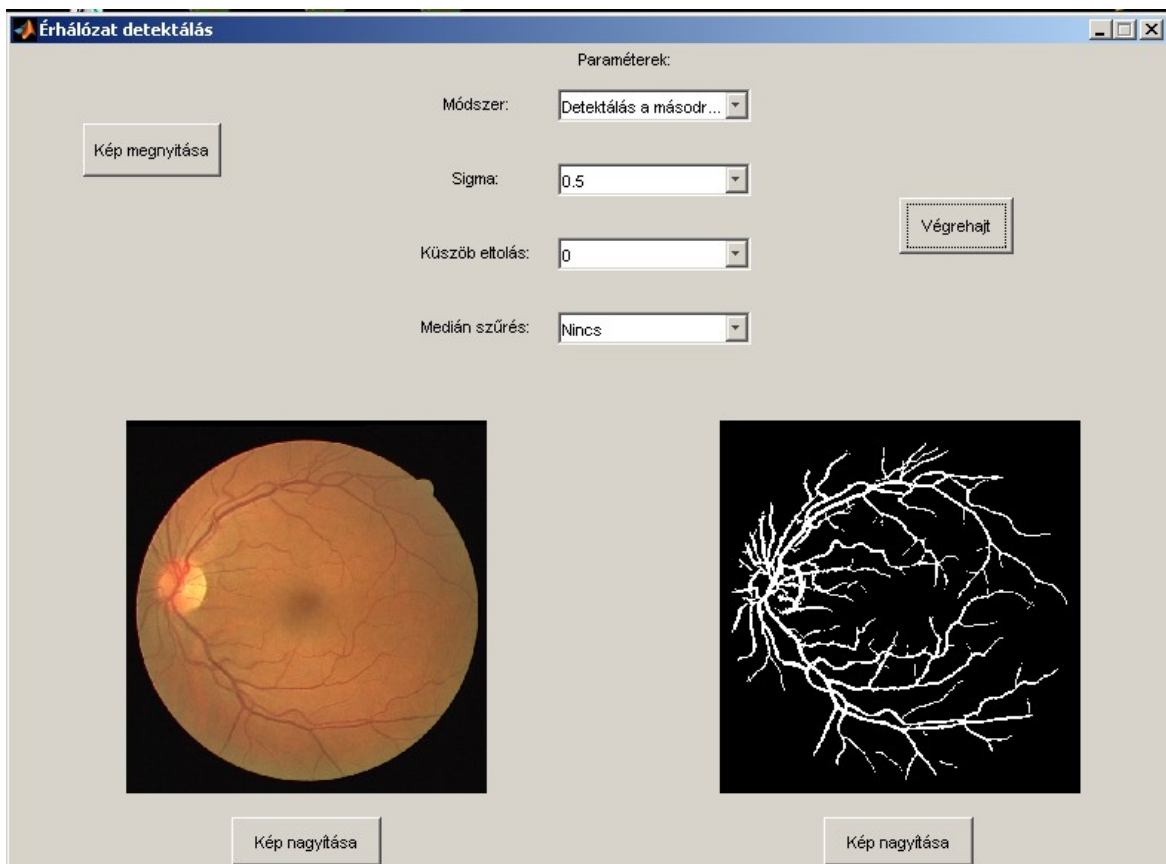
A MATLAB a forrásfájlokat, vagy más néven a szkripteket ún. .m fájlokban tárolja, külön fordítót nem használ, az utasításokat sorról-sorra hajtja végre.

A program felépítése

- *matchedfilter.m*: az illesztett szűrők algoritmusát megvalósító szkript. Bemenete egy szürkeárnyaltos kép és a σ paraméter, kimenete pedig az MFR kép.
- *mask.m*: egy adott szemfenék kép bináris maszkját generálja le.
- *kuszob.m*: első lépésében meghatározza az input kép átmenetmátrixát, majd abból kiszámolja a másodrendű entrópiákat minden intenzitásértékre és közülük visszaadja a maximálisat, azaz az optimális küszöböt.
- *erszegm.m*: ez a szkript végzi el az érhálózat detektálást, végrehajtja az elő- és az utófeldolgozást. A kimenete egy, az érhálózatot reprezentáló bináris kép.

Kétféleképpen lehet meghívni az előfeldolgozás szempontjából:

1. $bv=erszegm(I, keltolas, sigma)$: ebben az esetben nincs előfeldolgozás.
 2. $bv=erszegm(I, keltolas, sigma, 'median', N)$: ekkor az I kép zöld színcsatornáján végrehajt egy $N \times N$ -es mediánszűrést.
- *gui.m*: a felhasználói interfész megvalósítása.



16. ábra: Képernyőkép a programból.

Az érhálózat detektorok összehasonlítása

A DRIVE adatbázis esetén minden egyes képhez a rendelkezésünkre áll két szakértő (X és Y) által manuálisan szegmentált érhálózat, így lehetőségünk van a korábban bemutatott detektor jóságát mérni más-más paraméterek esetén, vagy azt különböző érhálózat szegmentáló algoritmusokkal összehasonlítani.

X által 577649 pixel lett megjelölve érként és 3960494 pixel háttérként, azaz az összes képpont 12.7%-a ér. Y esetén 556532 pixel az erek, míg 3981611 a háttérpixelék száma (12.3% ér). Az összehasonlítások alapjául az X eredményeit használtuk fel, tehát ezt tekintjük 100%-os pontosságúnak, amit a szakirodalom ground-truth-nak, vagy gold standard-nak nevez és a továbbiakban G-vel jelöljük, az érhálózat detektorok kimenetét pedig C-vel.

A fentiek alapján négy különböző érték határozható meg:

1. Valódi pozitív (True Positive, TP) nevezzük azt, ha mindkét képen, azaz G-ben és C-ben is egy adott pixel érként van megjelölve.
2. Valódi negatív (True Negative, TN), ha G-ben és C-ben is a pixel nem érpon, tehát ha háttérpixel mindkét esetben
3. Téves pozitív (False Positive, FP), ha egy pixel C-ben érpon, de G-ben nem. Ezt az értéket szokás elsőfajú hibának is nevezni.
4. Téves negatív (False Negative, FN), ha egy pixel C-ben nem érpon, pedig a G-ben az. Ezt pedig szokták másodfajú hibának nevezni.

Ebből a négy értékből származtathatóak különböző mennyiségek, melyek az összehasonlítás alapjául szolgálhatnak. Ezek közül a legszélesebb körben elterjedt mérőszámok: szenzitivitás, specificitás és a pontosság.

Szenzitivitás (Sens)

Más néven valódi pozitív arány (True Positive Rate, TPR), megadja, hogy csak az érhálózat pixeleit figyelembe véve mennyire pontos az algoritmus:

$$Sens = \frac{TP}{P} = \frac{TP}{TP + FN}$$

ahol P az összes pozitívaknak a száma.

Specificitás (Spec)

Más néven valódi negatív arány (True Negative Rate, TNR). Ez a mérőszám a háttérpixelek eltalálásának az arányát mutatja meg:

$$Spec = \frac{TN}{N} = \frac{TN}{TN + FP}$$

ahol N az összes negatívaknak a száma.

Pontosság (Accuracy, Acc)

Ez az érték pedig megadja a teljes pontosságát az algoritmusnak:

$$Acc = \frac{TP + TN}{P + N} = \frac{TP + TN}{TP + TN + FP + FN}$$

ahol:

P: az összes pozitívaknak a száma.

N: az összes negatívaknak a száma.

További mérőszámok

Az alábbiakban megadunk néhány, ritkában használt mérőszámot:

1. Téves pozitív arány (False Positive Rate, FPR):

$$FPR = \frac{FP}{TN + FP} = 1 - Spec$$

2. Téves negatív arány (False Negative Rate, FNR):

$$FNR = \frac{FN}{TP + FN} = 1 - Sens$$

3. Szimmetrikus differencia:

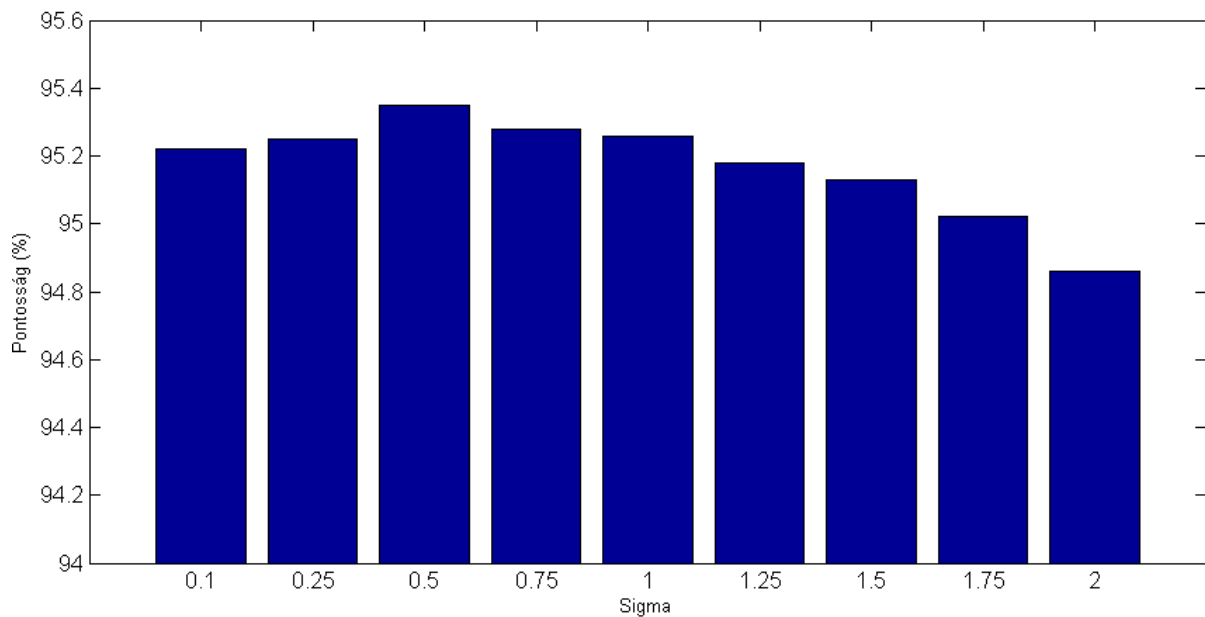
A szimmetrikus differencia két halmaz, A és B között van értelmezve az alábbi szerint:

$$S(A, B) = \frac{|A \setminus B| + |B \setminus A|}{|A \cup B|}$$

Ezzel a mennyiséggel is meg lehet határozni a detektálás jóságát. 0 értéket ad, ha mindkét halmaz elemei megegyeznek és 1-et, ha teljesen különböznek.

A detektor pontosságának vizsgálata különböző paraméterek esetén

A korábban bemutatásra került érhálózat detektornak van egy paramétere, a σ , amelynek a megváltoztatásával az algoritmus pontossága is megváltozik. A 17. ábrán látható a különböző σ értékek esetén mért pontosságok.



17. ábra: Az algoritmus pontossága különböző σ értékek esetén.

Megfigyelhető, hogy az algoritmus a $\sigma=0.5$ érték esetén a legjobb, ekkor a pontossága 95.35%-os. Ezen érték után a σ -t növelve a pontosság egyre csökken. A következő részben a detektort a $\sigma=0.5$ paraméterválasztás mellett összehasonlítjuk más algoritmusokkal.

A detektor összehasonlítása más algoritmusokkal

A továbbiakban a detektorunkat más, az irodalomban megtalálható algoritmussal hasonlítjuk össze, melyek közül kettőt, az illesztett szűrők módszerét, valamint M. Niemeijer kNN osztályozót használó algoritmusát már ismertettük. Lássunk most néhány további eljárást röviden:

1. Zana et al. [10]:

Ez a módszer matematikai morfológiai operátorokat használ az erek detektálására, amely alapvetően 3 lépésre bontható:

1. A lineáris részek meghatározása a nyitások supremuma által, ahol a nyitás struktúraeleme lineáris különböző irányokban.
2. Zajeltávolítás szintén a nyitások supremuma felhasználásával.
3. Ezután a zaj nélküli képről a nemkívánatos objektumok eltávolítása Laplace-szűrővel, majd egy speciális filter alkalmazása.

2. Jiang et al. [11]:

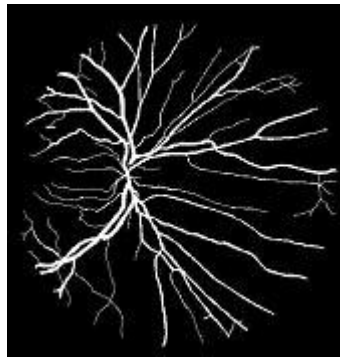
A retinaképeken többszintű lokális küszöbölést hajtanak végre, majd az így kapott bináris képek osztályozása történik meg aszerint, hogy rendelkeznek -e érszerű tulajdonságokkal.

3. Martinez-Perez et al. [12]:

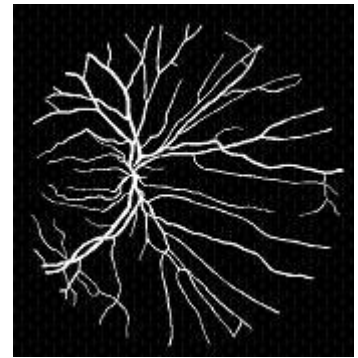
Az érhálózatot két tulajdonsággal jellemzik: az intenzitások gradiens értékével, valamint az erek erősségét mérő számmal. Ezen értékeket a kép különböző területein meghatározva, azok maximuma alapján a képpontokat két osztályba lehet sorolni: érhálózat pontok és háttérpontok.



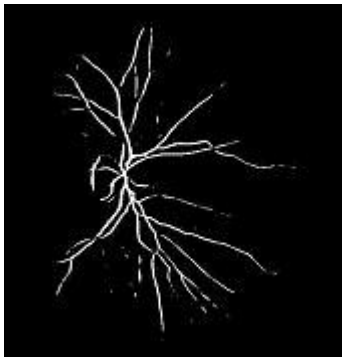
a) eredeti



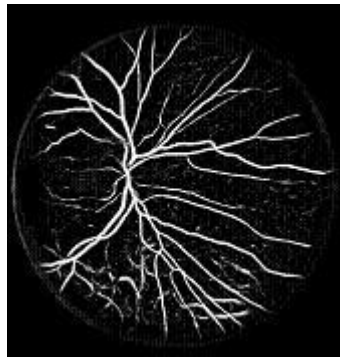
b) 1. manuális (gold standard)



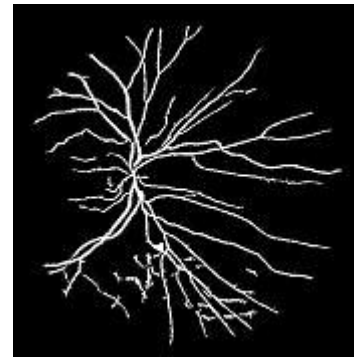
c) 2. manuális



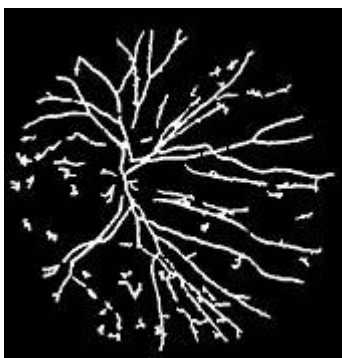
d) Chaudhuri et al.



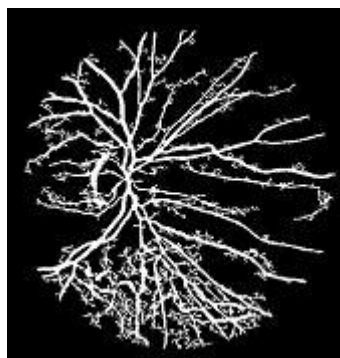
e) Niemeijer et al.



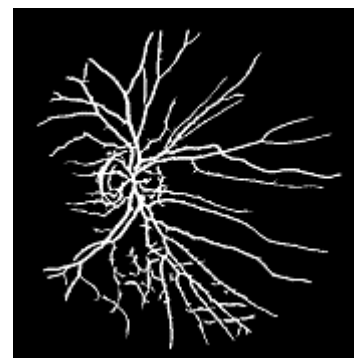
f) Zana et al.



g) Jiang et al.



h) Martinez-Perez et al.



i) másodrendű entrópiaalapú (Villalobos et al.)

18. ábra: A különböző algoritmusok eredménye egy DRIVE képre.

Módszer	Átlagos pontosság (%)
2. manuális	94,73
Chaudhuri et al.	87,73
Niemeijer et al.	94,16
Zana et al.	93,77
Jiang et al.	92,12
Martinez-Perez et al.	91,81
másodrendű entrópialapú (Villalobos et al.)	95,35

1. táblázat: A különböző módszerek átlagos pontossága a DRIVE adatbázis esetén.

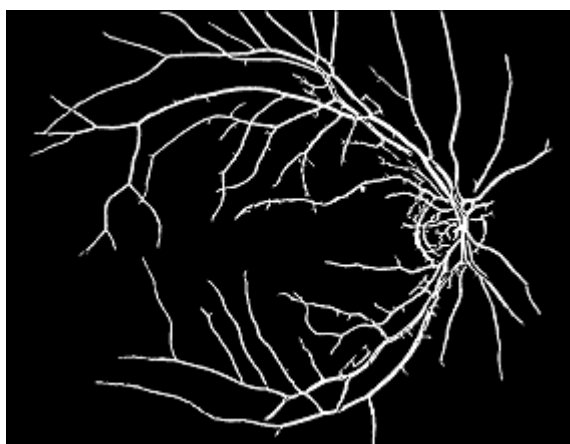
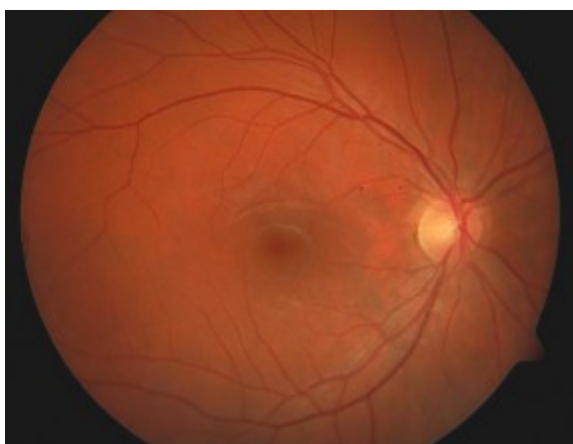
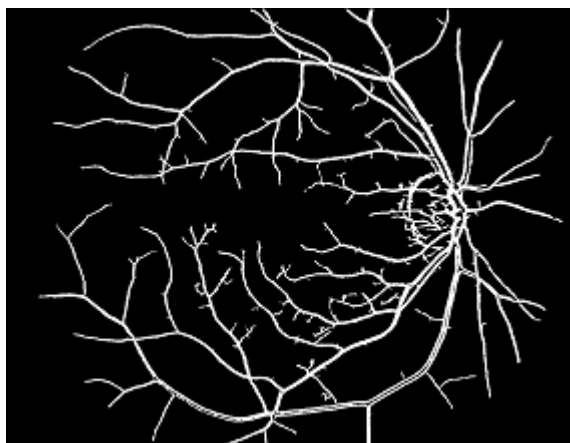
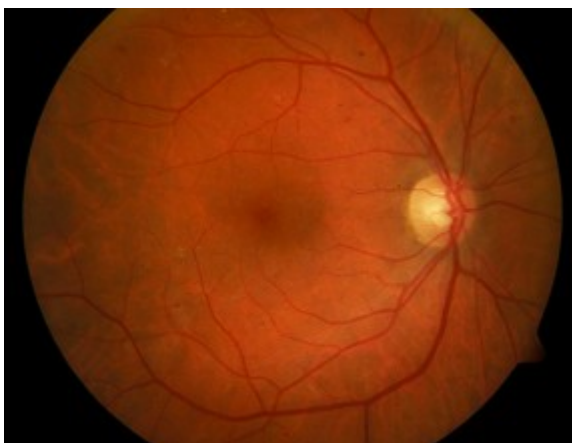
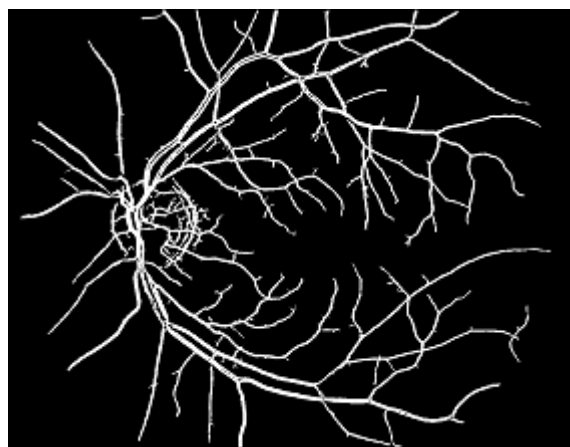
Összefoglalás

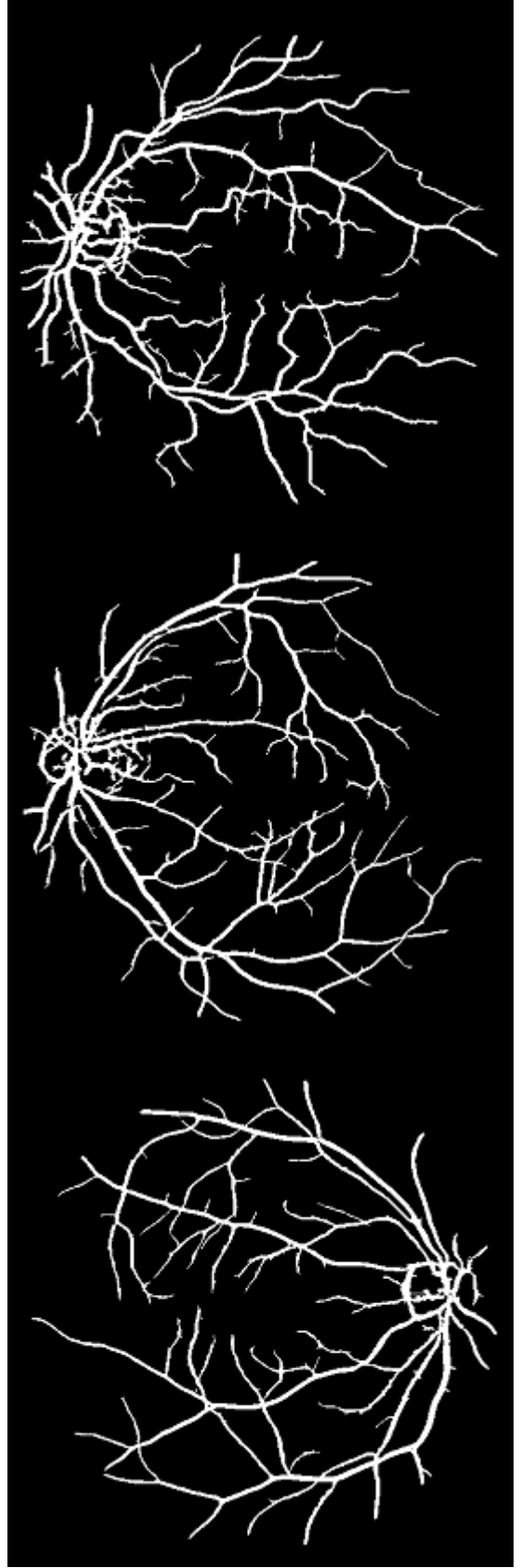
Diplomamunkámban bemutattam egy gyors és hatásos érhálózat detektáló algoritmust, amely teljesen automatikus, nem igényel semmiféle emberi közreműködést. A módszert megvizsgáltam különböző paraméterek mellett is, ami tovább növelte annak pontosságát, valamint összehasonlítottam további algoritmusokkal is. A pontosságot tovább is lehetne fokozni, ugyanis néhány output képen megfigyelhető, hogy a vakfolt környékét is tévesen az érhálózat részeként azonosítja az algoritmus. Véleményem szerint ezt úgy lehetne a legegyszerűbben kiküszöbölni, hogy először is meghatározzuk a vakfolt helyét a retinaképen, majd annak a környezetében külön detektálnánk az ereket valamilyen módszerrel.

Függelék A: a különböző mérőszámok értéke a DRIVE adatbázis esetén

Kép / Mérőszám	Szenzitivitás	Specificitás	Pontosság	Téves pozitív arány	Téves negatív arány	Szimmetrikus differencia
1	0,7961	0,9714	0,9558	0,0286	0,2039	0,3838
2	0,7427	0,9785	0,9544	0,0215	0,2573	0,3749
3	0,6655	0,9777	0,9466	0,0223	0,3345	0,4461
4	0,6583	0,9866	0,9564	0,0134	0,3417	0,4188
5	0,7335	0,9787	0,9557	0,0213	0,2665	0,392
6	0,576	0,9909	0,9505	0,0091	0,424	0,469
7	0,6943	0,9776	0,9517	0,0224	0,3057	0,4322
8	0,4434	0,9869	0,9401	0,0131	0,5566	0,6107
9	0,6633	0,9832	0,9573	0,0168	0,3367	0,4428
10	0,6621	0,9836	0,9571	0,0164	0,3379	0,4404
11	0,5583	0,9886	0,9523	0,0114	0,4417	0,4776
12	0,6982	0,9779	0,9537	0,0221	0,3018	0,4342
13	0,5898	0,9893	0,9502	0,0107	0,4102	0,4633
14	0,6814	0,9833	0,9589	0,0167	0,3186	0,427
15	0,7804	0,9639	0,9508	0,0361	0,2196	0,4685
16	0,655	0,9833	0,9537	0,0167	0,345	0,4392
17	0,4813	0,9899	0,947	0,0101	0,5187	0,5661
18	0,5556	0,9886	0,9543	0,0114	0,4444	0,5091
19	0,7505	0,9846	0,9652	0,0154	0,2495	0,3585
20	0,6643	0,9812	0,9579	0,0188	0,3357	0,463
Átlagos	0,6525	0,9823	0,9535	0,0177	0,3475	0,4509

Függelék B: retinaképek és a detektált érhálózat





Köszönetnyilvánítás

A dolgozat végén szeretnék köszönetet mondani témavezetőmnek, Dr. Hajdu Andrásnak, aki szakmai hozzáértésével, tanácsaival, segítőkészségével elősegítette a diplomamunka elkészültét.

Irodalomjegyzék

- [1] dr. Fövényi József: A cukorbetegségről mindenkinek, 1996,
http://www.informed.hu/_ebook/diab.pdf

- [2] DRIVE
<http://www.isi.uu.nl/Research/Databases/DRIVE/>

- [3] DIARETDB1
<http://www2.it.lut.fi/project/imageret/diaretdb1/>

- [4] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, Detection of blood vessels in retinal images using two-dimensional matched filters, *IEEE Transactions on Medical Imaging* 8(3), pp. 263-269, 1989.

- [5] M. Niemeijer, J.J. Staal, B. van Ginneken, M. Loog, M.D. Abramoff, "Comparative study of retinal vessel segmentation methods on a new publicly available database", in: SPIE Medical Imaging, Editor(s): J. Michael Fitzpatrick, M. Sonka, SPIE, 2004, vol. 5370, pp. 648-656.

- [6] Fabiola M. Villalobos-Castaldi, Edgardo Manuel Felipe Riverón, Luis Pastor Sánchez Fernández: A fast, efficient and automated method to extract vessels from fundus images. *J. Visualization* 13(3): 263-270, 2010

- [7] Frank ter Haar: Automatic localization of the optic disc in digital colour images of the human retina. M.S. Thesis, Utrecht University, Dec. 16, 2005

- [8] Pal NR, Pal SK (1989) Entropic thresholding. *Signal Process* 16:97–10

- [9] MATLAB,
<http://www.mathworks.com/products/matlab/>

- [10] F. Zana and J. Klein, "A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform," *IEEE Transactions on Medical Imaging* 18(5), pp. 419–428, 1999.

- [11] X. Jiang and D. Mojon, "Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25(1), pp. 131–137, 2003.

- [12] M. Martínez-Pérez, A. Hughes, A. Stanton, S. Thom, A. Bharath, and K. Parker, "Scale-space analysis for the characterisation of retinal blood vessels," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI'99*, C. Taylor and A. Colchester, eds., pp. 90–97, 1999.