

DIPLOMAMUNKA

Talajvizsgáló rendszer tervezése

Tüű-Szabó Boldizsár

**Villamosmérnöki MSc szak
Automatizálási rendszerek szakirány**

2014

Feladat-kiíró lap diplomamunkához

Hallgató adatai

név: Tüű-Szabó Boldizsár

Neptun-kód: TQ6ERT

szak: villamosmérnöki Msc szak

specializáció: automatizálási rendszerek szakirány

Cím: Talajvizsgáló rendszer tervezése

Feladatok leírása:

A dolgozat célja olyan rendszer kidolgozása, amely képi információk alapján meghatározza

talaj főbb tulajdonságait és típusát.

részfeladatok:

- 1) szakirodalom-kutatás
- 2) hasonló rendszerek áttekintése
- 3) hasonló módszerek értékelése
- 4) saját rendszer tervezése, prototípus-készítés

Győr, 2014.09.11.

Tormási Alex, egyetemi
tanársegéd

Dr. Kuczmann Miklós,
egyetemi tanár

Konzultációs lap szakdolgozathoz és diplomamunkához

Hallgató adatai

név: Tűő-Szabó Boldizsár

Neptun-kód: TQ6ERT

szak: villamosmérnöki MSc szak

specializáció: automatizálási rendszerek szakirány

Dátum	Tevékenység	Aláírás

A dolgozat a formai követelményeknek megfelel:

Dátum

Tormási Alex, egyetemi
tanársegéd

Diplomamunka értékelő lap

Hallgató adatai

név: Tüű-Szabó Boldizsár
szak: villamosmérnöki MSc
specializáció: automatizálási rendszerek szakirány
tagozat: nappali

Neptun-kód:TQ6ERT

A diplomamunka adatai

cím: Talajvizsgáló rendszer tervezése
nyelv: magyar
típus: nyilvános

A diplomamunka bírálatra bocsátható:

A bíráló:

név:
munkahely:
beosztás:

Dátum

Dr. Kuczmann Miklós, egyetemi tanár

A bíráló javaslata:

dátum

érdemjegy

A belső konzulens javaslata:

dátum

érdemjegy

Tormási Alex, egyetemi tanársegéd

A ZVB döntése:

dátum

érdemjegy

aláírás (ZVB elnök)

Nyilatkozat

Alulírott, Tüű-Szabó Boldizsár villamosmérnöki Msc szakos hallgató kijelentem, hogy a Talajvizsgáló rendszer tervezése című diplomamunka feladat kidolgozása a saját munkám, abban csak a megjelölt forrásokat, és a megjelölt mértékben használtam fel, az idézés szabályainak megfelelően, a hivatkozások pontos megjelölésével.

Eredményeim saját munkán, számításokon, kutatáson, valós méréseken alapulnak, és a legjobb tudásom szerint hitelesek.

Győr, 2014. 11.20.

hallgató

Államvizsga-tárgyak

Hallgató neve: Tüü-Szabó Boldizsár

Neptun kódja: TQ6ERT

Belső konzulens neve: Tormási Alex

Államvizsga-tantárgyak:

- Számítási intelligencia
- Digitális rendszerek
- Programozható vezérlőrendszerek

Képzés szintje: MSC

Tagozat: Nappali

Államvizsga időpontja: 2014/2015 tanév I. félév

Dátum: 2014.11.20.

.....
Hallgató aláírása

.....
Belső konzulens aláírása

Kivonat

Talajvizsgáló rendszer tervezése

A képfeldolgozás az utóbbi évtizedekben rohamosan fejlődik, és a tudomány fontos területévé vált. A képfeldolgozást számos területén alkalmazzák, például biometrikus azonosításra (ujjlenyomat felismerés, írisz felismerés stb.), de az orvostudományban is fontos szerepet kap. Diplomamunkám célja egy olyan rendszer kidolgozása volt, amely képi információk alapján meghatározza a talaj típusát (fű, homok, aszfalt). Ez az autonóm robotoknál kiemelt jelentőséggel bír. A robot képes lesz a képi információk alapján dönteni arról, hogy milyen útvonalon közlekedjen elkerülve ezzel a nehezebb terepeket, járhatatlan területeket.

A dolgozatom első részében a szakirodalom-kutatás eredményét ismertetem, a leggyakrabban használt textúra-leíró módszerek (ko-okkurencia mátrix, gray level difference módszer, LBP-operátor és változatai) valamint klasszifikációs módszerek (Naiv-Bayes-módszer, szupport vektor gépek, neurális hálózatok, k-legközelebbi szomszéd módszere) elméleti hátterét foglalom össze.

A dolgozatom második része a feladat gyakorlati megvalósításával foglalkozik. A textúrák leírására a circular LBP-operátort valamint az uniform LBP-operátort használtam. Az egyes talajtípusokról készült képek osztályozására több módszert is kipróbáltam (többosztályú szupport vektor gépek, előrecsatolt neurális hálózatok).

Ismertetem a neurális hálózat struktúrájának (neuronok, rejtett rétegek száma) illetve a neuronok aktivációs függvényének (lineáris, szigmoid stb.) kiválasztása során kapott eredményeket. A neurális hálózatok tanítására több tanulási algoritmust (backpropagation algoritmus és továbbfejlesztett változatai, Levenberg-Marquardt algoritmus, konjugált gradiens módszerek) alkalmaztam, és az egyes algoritmussal kapott eredményeket összehasonlítottam tanulási idő illetve epochszám alapján.

Abstract

Designing a terrain classification system

The image processing is developing rapidly in recent decades, and has become an important field for science. The image processing is used in many fields, such as biometric authentication (fingerprint recognition, iris recognition, etc.), But also plays an important role in medicine.

Aim of my thesis was to design a system that determines the type of the terrain (grass, sand, asphalt) based on information of images. Terrain classification is an important part of image processing. For example the autonomous robots can use it, they will able to determine the optimal route avoiding the difficult terrains.

In the first part of my thesis I describe the results of the literature search, I summarize the theoretical background of most used texture descriptor methods (co-occurrence matrix, gray level difference method, LBP operator and its variants) and classification methods (Naive-Bayes method, support vector machines, neural networks, k-nearest neighbor method).

The second part of the thesis deals with the my terrain classification system. I used the circular LBP-operator and uniform LBP-operator to describe the texture of the terrain pictures. I tried more classification methods (multiclass support vector machines, feedforward neural networks) for the classification of terrain images. I describe the process of selection of the neural network structure and activation functions (linear, sigmoid etc). For the learning of the neural network I tried more learning algorithms (backpropagation algorithm and improved versions, Levenberg-Marquardt algorithm, conjugate gradient methods), and I compared the results in terms of learning time and epoch number.

Tartalomjegyzék

1. Bevezetés.....	2
2. Textúrák leírása [1], [2], [3]	3
2.1. Statisztikai módszerek.....	3
2.1.1. Intenzitás hisztogram [1].....	3
2.1.2. Gray Level Difference módszer [2]	4
2.1.3. Ko-okkurencia mátrix (gray level co-occurrence matrix GLCM) [3].....	5
2.2. Local Binary Pattern [4].....	6
2.2.1. Circular LBP	7
2.2.2. Uniform LBP	9
2.2.3. Forgatás-invariáns LBP (rotation invariant LBP)	10
2.2.4. Local Ternary Pattern (LTP)	11
3. Klasszifikációs módszerek	13
3.1. Random Forest [5].....	13
3.2. Naiv-Bayes-módszer [6]	13
3.3. k-legközelebbi szomszéd módszere [7].....	15
3.4. Döntési fák	17
3.5. Szupport Vektor Gép [8], [9], [10].....	17
3.5.1. Lagrange duál feladat megoldása (általános eset) [10]	20
3.5.2. A lineáris SVM Lagrange duál feladatának megoldása	20
3.5.3. Puha margójú SVM (Soft margin SVM) [8].....	22
3.5.4. Nemlineáris SVM [8]	23
3.5.5. Több osztályú SVM	24
3.6. Neurális hálózatok [11], [12]	25
3.6.1. Perceptron [11], [12]	26
3.6.2. A perceptron tanulása [11], [12]	27
3.6.3. Többrétegű perceptron [11], [12]	28
3.6.4. Backpropagation algoritmus [11], [12]	29
3.6.5. Backpropagation algoritmus momentummal [11], [12].....	30
3.6.6. Backpropagation algoritmus adaptív tanulási paraméterrel [12]	31
3.6.7. Newton-módszer [11], [12]	32
3.6.8. BFGS (Broyden-Fletcher-Goldfarb-Shanno) eljárás [11], [12]	33
3.6.9. Levenberg-Marquardt algoritmus [11], [12]	33
3.6.10. Konjugált gradiens módszer [11], [12]	35
4. A feladat gyakorlati megvalósítása [13], [14], [15]	37

4.1. Talajfelismerő rendszerek áttekintése [13], [14].....	37
4.1.1. 3D LIDAR alapú talajfelismerés [13]	38
4.1.2. Talajfelismerés textúra leíró módszerekkel [14].....	38
4.2. Talaj Textúrájának leírása	39
4.2.1. Circular LBP(8,1).....	39
4.2.2. Uniform LBP	42
4.3. Klasszifikációs eljárások	43
4.3.1. Tanító pontok előállítása	43
4.3.2. Osztályozás többosztályú SVM-mel	43
4.3.3. Osztályozás előrecsatolt neurális hálózattal	46
5. Összefoglalás, továbbfejlesztési lehetőségek.....	53

1. Bevezetés

A képfeldolgozás az utóbbi évtizedekben rohamosan fejlődik, és a tudomány fontos területévé vált. A képfeldolgozást számos területén alkalmazzák, például biometrikus azonosításra (ujjlenyomat felismerés, írisz felismerés stb.), de az orvostudományban is fontos szerepet kap. Diplomamunkám célja egy olyan rendszer kidolgozása volt, amely képi információk alapján meghatározza a talaj típusát (fű, homok, aszfalt). Ez az autonóm robotoknál kiemelt jelentőséggel bír. A robot képes lesz a képi információk alapján dönteni arról, hogy milyen útvonalon közlekedjen elkerülve ezzel a nehezebb terepeket, járhatatlan területeket.

A dolgozatom első részében a szakirodalom-kutatás eredményét ismertetem, a leggyakrabban használt textúra-leíró módszerek (ko-okkurencia mátrix, gray level difference módszer, LBP-operátor és változatai) valamint klasszifikációs módszerek (Naiv-Bayes-módszer, szupport vektor gépek, neurális hálózatok, k-legközelebbi szomszéd módszere) elméleti hátterét foglalom össze.

A dolgozatom második része a feladat gyakorlati megvalósításával foglalkozik. A textúrák leírására a circular LBP-operátort valamint az uniform LBP-operátort használtam. Az egyes talajtípusokról készült képek osztályozására több módszert is kipróbáltam (többosztályú szupport vektor gépek, előrecsatolt neurális hálózatok).

Ismertetem a neurális hálózat struktúrájának (neuronok, rejtett rétegek száma) illetve a neuronok aktivációs függvényének (lineáris, szigmoid stb.) kiválasztása során kapott eredményeket. A neurális hálózatok tanítására több tanulási algoritmust (backpropagation algoritmus és továbbfejlesztett változatai, Levenberg-Marquardt algoritmus, konjugált gradiens módszerek) alkalmaztam, és az egyes algoritmussal kapott eredményeket összehasonlítottam tanulási idő illetve epochszám alapján.

2. Textúrák leírása [1], [2], [3]

Számos leíró kidolgoztak a kép tartalmának azonosítására. A leírókat a kép tulajdonságainak jellemzésére használják (pl: textúrák, színek, alakzatok).

Textúra a szín vagy intenzitások térbeli leírása egy képen vagy annak egy részletén.

Jack Sklansky így definiálta a textúra fogalmát: Egy képen egy területnek állandó textúrája van, ha a lokális statisztikák vagy egyéb lokális tulajdonságok állandóak, lassan változnak, vagy megközelítőleg periódikusak.

A textúrákat 2 csoportba lehet osztani:

- Sztochasztikus
- Determinisztikus

A determinisztikus textúrák jól meghatározott geometriai alakzatok ismétlődéséből állnak.

A textúra leírására számos technikát dolgoztak ki. MihranTuceryan és Anil K. Jain (1993) a textúra leírására használt módszereket 4 csoportba osztották:

- statisztikai
- geometriai
- modell alapú
- jelfeldolgozó módszerek

2.1. Statisztikai módszerek

Statisztikai módszereket gyakran alkalmaznak a textúra elemzés során.

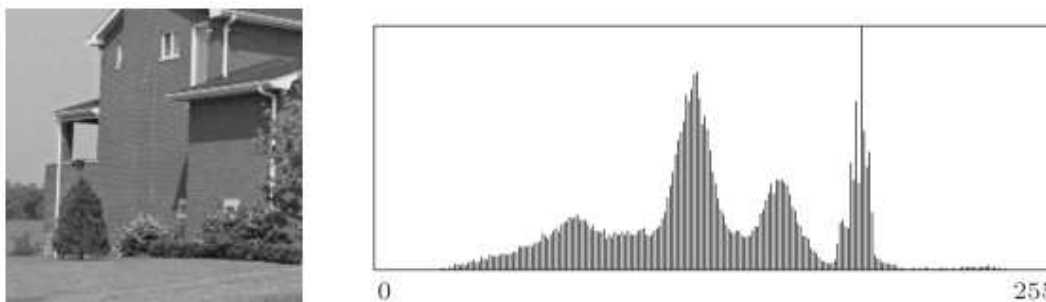
Statisztikai módszert először Herbert Kaizer alkalmazott 1955-ben, amikor légi felvételeket elemzett autokorreláció segítségével.

Az elsőrendű módszerek (pl: szórásnégyzet, négyzetes közép, entrópia, intenzitás hisztogram stb.) statisztikai számításokat végeznek az eredeti kép pixelértékeinek felhasználásával.

2.1.1. Intenzitás hisztogram [1]

A képfeldolgozási eljárásokban a képek intenzitáshisztogramja alapvető jelentőséggel bír. A hisztogram metrikusan skálázott tulajdonságok grafikus ábrázolása. Matematikai értelemben a hisztogram egy függvény, ami az egyes diszjunkt osztályokba tartozó megfigyeléseket számolja. A hisztogram, mint grafikon ennek egy ábrázolási módja. Az intenzitáshisztogram azért nagyon népszerű a képfeldolgozási eljárásokban, mert igen

szemléletesen mutatja az adott kép intenzitás-eloszlását. Tekintsünk tehát egy 256-szintű digitális képet, ahol a fekete és a fehér közti tartományt 256 árnyalatra bontottuk fel, úgy, hogy az $i+1$ -edik árnyalat az világosabb, mint az i -edik. Most a szürkeségi skálát 256 intervallumra osztjuk, ahol minden intervallumba egyetlen szürkeségi érték esik, azaz ($y_0 < y_1 < \dots < y_{256}$). A valószínűségi mintát maga a kép, illetve annak intenzitásértékei adják. A hisztogram i -edik oszlopának magasságát az $[y_i, y_{i+1})$ intervallumba eső elemek száma, azaz a képen lévő i -edik szürkeárnyalatú képpontok száma adja.



1. ábra kép intenzitás hisztogramja

Az elsőrendű módszerek előnye, hogy gyorsan és egyszerűen számíthatóak. Hátrányuk, hogy nem nyújtanak információt a pixelek egymás közötti kapcsolatáról. 2 teljesen különböző kép, amely egyaránt 50-50%-ban tartalmaz fekete és fehér pixeleket ugyanazt az intenzitás hisztogramot eredményezi, ezért a 2 kép elsőrendű statisztikai módszerekkel nem megkülönböztethető.

A másodrendű módszerek pixelek 2 (általában szomszédos) csoportja közötti kapcsolatot állapítják meg.

Harmad- vagy magasabb rendű módszerek is léteznek, de ritkán használják őket a megnövekedett számítási idő és az értelmezési nehézségek miatt.

2.1.2. Gray Level Difference módszer [2]

Ennél a módszernél az összes pixelre kiszámolják a pixel és egy eltolási vektorral meghatározott pixel értékkülönbségének abszolút értékét.

$$f'(x, y) = |f(x, y) - f(x + \delta x, y + \delta y)| \quad (2.1)$$

Majd ezen értékek felhasználásával egy m dimenziós P' valószínűségi vektort készítünk, amelynél az i -edik elem annak a valószínűsége, hogy $f'(x, y)$ értéke i . A vektor felhasználásával a tulajdonságok gyorsan és egyszerűen számíthatóak.

Kontraszt	$CON = \sum i^2 \cdot P'(i)$
Angular second Moment	$ASM = \sum P'(i)^2$
Entrópia	$ENT = -\sum P'(i) \cdot \log P'(i)$
Átlag	$MEAN = \frac{1}{m} \sum i \cdot P'(i)$

1. táblázat tulajdonságok számítása

2.1.3. Ko-okkurencia mátrix (gray level co-occurrence matrix GLCM) [3]

A ko-okkurencia mátrix a másodrendű statisztikai módszerek csoportjába tartozik. Az egyes szűrkeskálás szintek térbeli eloszlását fejezi ki.

$$C_{\Delta x, \Delta y}(i, j) = \sum_{p=1}^n \sum_{q=1}^m \begin{cases} 1, & \text{ha } I(p, q) = i \text{ és } I(p + \Delta x, q + \Delta y) = j \\ 0, & \text{különben} \end{cases} \quad (2.2)$$

A ko-okkurencia mátrixnál úgy próbáljuk jellemezni a textúrát, hogy valamilyen eltolási vektort választunk, és azt vizsgáljuk, hogy milyen világosság szintről milyenre hányszor megy a kép az adott vektor mentén. Így a létrejött mátrix n*n-es, ha n világossági szint van a képen.

Kontraszt: A helyi változásokat méri a GLCM mátrixban. A kontraszt értéke alacsony, ha a szomszédos pixelek közel azonos szürke skála értékkel rendelkeznek. A kontraszt értékkészlete $[0, (\text{size}(\text{GLCM}) - 1)^2]$. Ha az összes pixel értéke megegyezik, akkor a kontraszt 0.

$$\sum_{i,j=0} P_{i,j} (i - j)^2$$

Homogenitás: A homogenitás értéke akkor nagy, ha a legtöbb pixel szürke skála értéke azonos vagy nagyon közeli. A homogenitás értéke 0 és 1 között van. A homogenitás 1, ha a GLCM mátrix diagonális (olyan négyzetes mátrix, amelynek minden főátlón kívüli eleme 0).

$$\sum_{i,j=0}^{N-1, M-1} \frac{P_{i,j}}{1 + (i - j)} \quad (2.3)$$

Energia: a GLCM elemeinek négyzetösszege, értéke 0 és 1 közötti.

Akkor legkisebb az értéke, ha az összes elem azonos értékű. Az energia értéke 1, ha minden pixel azonos értékű.

$$\sum_{i,j=0}^{N-1, M-1} P_{i,j}^2 \quad (2.4)$$

Entrópia: a szürke skála eloszlásának véletlenszerűségét fejezi ki

$$\sum_{i,j=0}^{N-1,M-1} P_{i,j} (-\ln P_{i,j}) \quad (2.5)$$

Az entrópia akkor a legnagyobb, amikor a mátrix minden eleme egyenlő nagyságú.

Mivel az logaritmus függvény a 0-ra nincs definiálva, ezért az alábbi kiegészítést kell tenni:

$$0 \cdot (-\ln(0)) = 0$$

2.2. Local Binary Pattern [4]

Nagyon hatékony textúra leíró módszer, ami népszerűségét főként a számítási egyszerűségének köszönheti. Gyakran alkalmazzák objektum felismerésre és követésre. A biometrikus alkalmazások területén szintén népszerű (írisz felismerés, ujjlenyomat felismerés, tenyérlenyomat felismerés, járás felismerés stb.).

A másik pontos tulajdonsága a robosztusság a szürke skála szintek megváltozására pl: a megvilágítás megváltozása miatt.

Az eredeti LBP operátor megalkotása Timo Ojala nevéhez fűződik 1996-ban, aki szürke skálás képek leírására készítette el az operátort. Később azonban készítettek LBP-operátorokat színes (több csatornás) képek, videók és volumetrikus adatok jellemzésére is. A volumetrikus adatokat széles körben alkalmazzák orvosi alkalmazásoknál (3D CT és MRI).

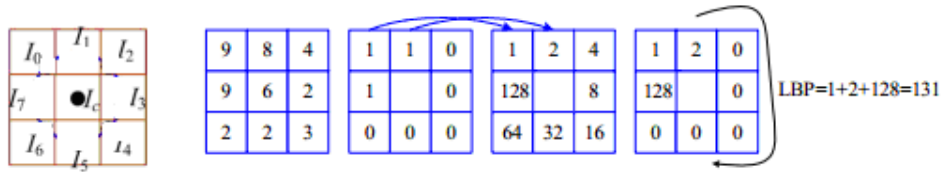
A szürke skálás kép minden pixele köré egy 3*3-as méretű ablakot helyezett. Minden pixelhez kiszámított egy 8 bites bináris számsorozatot az alábbiak szerint:

$$LBP(I_c) = \sum_{i=0}^{P-1} 2^i s(I_i - I_c) \quad (2.6)$$

i : a szomszédos pixelek sorszáma (0-7)

A szomszédos pixel értékéből kivonta a center pixel értékét ($A = I_i - I_c$). Ha a szomszédos pixel a nagyobb értékű ($A > 0$), akkor $s(A) = 1$, különben $s(A) = 0$.

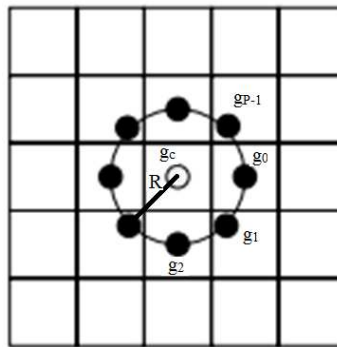
A 8-bites bináris számsorozatot kapunk, amelyek 256 különböző értéket vehetnek fel. A kapott eredményeket hisztogramon szokás ábrázolni, amit LBP-hisztogramnak neveznek.



2. ábra Local binary pattern számítása

2.2.1. Circular LBP

Olaja 2002-ben kiterjesztette az LBP operátort különböző távolságú szomszédos pixelek vizsgálatára. Extended LBP-nek vagy Circular LBP-nek nevezik, mivel az eredeti LBP-operátor általánosításának tekinthető. Egy tetszőleges, egész pixel sugarú kör kerülete mentén vesszünk fel P db pontot.



3. ábra Extended LBP

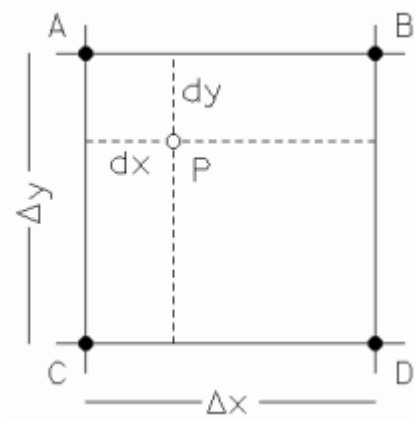
A középső pixel intenzitása: $g_c = I(x, y)$ (2.7)

A pontok intenzitása a kör kerülete mentén: $g_p = I(x_p, y_p)$ $p = 0, 1, \dots, P-1$ (2.8)

A mintavételi pontok koordinátái:

$$\begin{aligned} x_p &= x + R \cos(2\pi p / P) \\ y_p &= y - R \sin(2\pi p / P) \end{aligned} \quad (2.9)$$

Bilineáris interpolációt használunk, ha a mintavételi pont nem a pixel középpontja. A bilineáris interpoláció a síkban mindkét koordináta tengely mentén lineáris interpolációt alkalmaz.



4. ábra bilineáris transzformáció

A P pont intenzitás értéke az alábbi matematikai összefüggéssel számítható:

$$I_p = I_A + dx \cdot (I_B - I_A) + dy \cdot (I_C - I_A) + dx \cdot dy \cdot (I_A - I_B - I_C + I_D) \quad (2.10)$$

Az extended LBP számítása az együttes eloszlásból vezethető le. A helyi textúra a P+1 db ($P > 0$) pixel (középső pixel + a P db mintavételi pont a kör mentén) együttes eloszlásával jellemezhető.

$$T = t(g_c, g_0, g_1, \dots, g_{P-1}) \quad (2.11)$$

Információvesztés nélkül kivonhatjuk a középső pixel intenzitásértékét a szomszédos mintavételi pontok intenzitásértékeiből:

$$T = t(g_c, g_0 - g_c, g_1 - g_c, \dots, g_{P-1} - g_c) \quad (2.12)$$

Az együttes eloszlás az eloszlás szorzattá alakításával közelíthető feltéve, hogy a középső pixel intenzitása és az intenzitáskülönbségek statisztikailag függetlenek egymástól:

$$T \approx t(g_c) t(g_0 - g_c, g_1 - g_c, \dots, g_{P-1} - g_c) \quad (2.13)$$

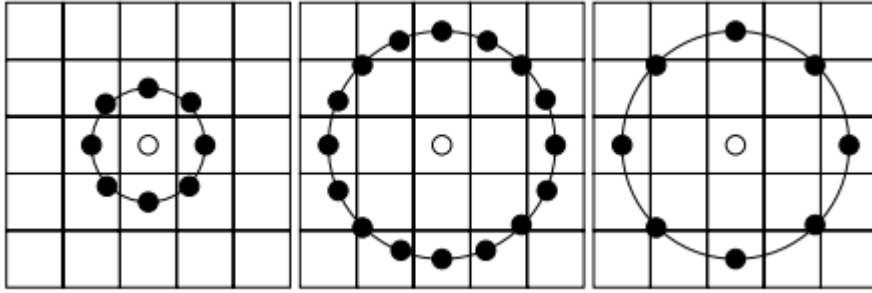
Mivel az első faktor (g_c) nem tartalmaz hasznos információt a helyi textúra leírását illetően, ezért a intenzitáskülönbségek együttes eloszlásával jellemezhetjük a helyi textúrát ($t(g_0 - g_c, g_1 - g_c, \dots, g_{P-1} - g_c)$).

Az eloszlás közelítésére Ojala az intenzitáskülönbségek előjelét vette figyelembe.

$$t(s(g_0 - g_c), s(g_1 - g_c), \dots, s(g_{p-1} - g_c)) \quad (2.14)$$

$$\text{ahol } s(z) \text{ lépcsős függvény } s(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases} \quad (2.15)$$

$$\text{Az LBP értéke az alábbi képlettel számolható ki: } LBP_{p,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (2.16)$$



5. ábra a, LBP(8,1)

b, LBP(16,1)

c, LBP(8,2)

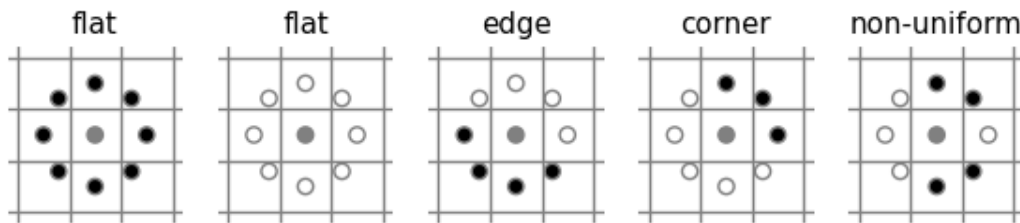
Az ábrán a leggyakrabban alkalmazott LBP-operátorok láthatóak. 8 szomszédos pont esetén $2^8=256$, 16 pont esetén $2^{16}=65536$ különböző LBP értéket kaphatunk egy adott pixelhez. Az LBP értékének a kép összes pixelére való kiszámítása után az eredményeket hisztogramon ábrázoljuk, tehát 256 illetve 65536 oszlopos hisztogramot fogunk kapni.

A LBP(8,1)-operátor hasonlít Ojala eredeti LBP-operátorára, hiszen ott is 8 környező pont intenzitásértékét hasonlította össze a középső pixel értékével, viszont az LBP(8,1) esetén nem minden pont esik a pixel középpontjába (bilineáris interpolációt alkalmazunk). Egyes alkalmazások során szükséges van arra, hogy az operátor forgatás-invariáns legyen. Ilyen operátor megalkotására az LBP(8,1) jobb kiindulási alapként szolgál.

2.2.2. Uniform LBP

Bevezették az uniform pattern fogalmat a zaj csökkentésére az LBP- hisztogramban. Az uniform LBP-operátor bevezetését az indokolta, hogy néhány bitsorozat sokkal többször fordul elő, mint a többi. Ojala 2002-ben megállapította, hogy LBP(8,1) esetén (8 a szomszédos pontok száma, a kör sugara 1 pixel hosszúságú) az előálló bitminták közel 90 %-a uniform. LBP(16,2) esetén ez az arány kb. 70 %-os.

Egy LBP akkor uniform, ha a bitmintán körkörösén végigmenve legfeljebb 2 db 0-1 vagy 1-0 átmenetet kapunk. Pl: 0000000 (0 átmenet) 01110000 (2 átmenet) and 11001111 (2 átmenet) uniformok, de pl. 11001001 (4 átmenet) and 01010010 (6 átmenet) nem azok.



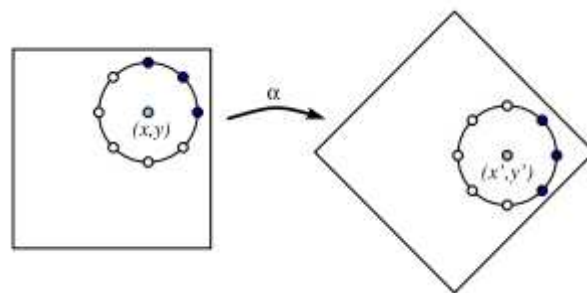
6. ábra uniform és nem uniform minták

Az ábrán 8 bites LBP-bitsorozatokra láthatunk példákat. A fekete pontok jelölik a 0 értékeket, a fehérek az 1 bitértékeket. Ha minden szomszédos pixel értéke kisebb mint a középső pixelé, akkor csak 0 bitekből álló LBP-bitsorozatot kapunk (fekete pontok). Ha viszont minden szomszédos pixel értéke nagyobb mint a középső pixelé, akkor csak 1 bitekből álló bitsorozatunk lesz (fehér pontok). Ezeket, a csak azonos értékű bitekből álló bitsorozatokat a szakirodalom flat jelzővel illeti.

8 szomszédos pont esetén 256 különböző bitmintát kaphatunk, amelyből 58 lesz uniform. A nem uniform mintákat nem különböztetjük meg egymástól, így 59 különböző címkénk lesz, szemben az eredeti LBP-vel, amikor 256 volt. 16 mintavételi pont esetén 65536 helyett csak 243 címkénk lesz, mivel 242 bitsorozat uniform.

2.2.3. Forgatás-invariáns LBP (rotation invariant LBP)

Sok esetben szükség van arra, hogy a textúrák jellemzésére forgatás-invariáns LBP-t használjunk.



7. ábra forgatás-invariáns LBP

Az ábrán jól megfigyelhető a kép forgatásának hatása az LBP-operátorra. A kör mentén elhelyezkedő szomszédos pontok is elfordulnak, így megváltozik a középső pixelhez tartozó LBP-érték.

Forgatás-invariáns leképezéssel az LBP-operátor (részben) forgatás-invariánssá tehető. A leképezés az összes LBP bináris kód legkisebb értékükbe való körkörös bitenkénti jobbra forgatását jelenti:

$$LBP_{P,R}^{ri} = \min_i ROR(LBP_{P,R}, i) \quad (2.17)$$

Így például a 10000110, 00110100, and 00001101 bináris kódokhoz a leképezés után ugyanaz, a 00001101 kód fog tartozni.

Ez az operátor csak a $\frac{a \cdot 360^\circ}{P}$ szöggel ($a=1,2,\dots,P-1$) való elfordulásokra lesz invariáns, de osztályozási feladatok eredményei azt mutatják, hogy az operátor más szögelfordulások esetén is nagy robosztusságot mutat.

2.2.4. Local Ternary Pattern (LTP)

Az LBP kibővítésének tekinthető. Az LBP-nél a küszöb a középső pixel értékénél volt, ezért érzékeny a zajokra, ennek csökkentésére vezették be az LTP-t.

Itt szintén egy pixel értékét hasonlítják össze a szomszédos pixelekkal, de bevezetnek egy küszöbértéket (k).

$$T = \begin{cases} 1 & T \geq (c + k) \\ 0 & T < (c + k) \text{ és } T > (c - k) \\ -1 & T \leq (c - k) \end{cases} \quad (2.18)$$

Ahol c a középső pixel értéke, T a szomszédos pixel értéke

Az így előálló ternáris kódot, 2 bináris kódra osztják, az első tartalmazza a pozitív értékeket, a második a negatív értékeket. A pozitív és a negatív értékeket tartalmazó kódokból külön-külön hisztogramot számolunk, így az LTP mérete az LBP méretének 2-szerese lesz.

Az ábrán láthatunk egy példát az LTP számítására 8 szomszédos pixel esetén.

94	38	54
23	50	78
47	66	12

1	-1	0
-1		1
0	1	-1

Ternary Pattern (k=5): 1(-1)01(-1)10(-1)

Part1=10010100, Part2=01001001

8. ábra LTP számítása

3. Klasszifikációs módszerek

A klasszifikációs módszerek osztályozási problémák megoldására szolgálnak. Ebben a fejezetben ismertetem a leggyakrabban használt klasszifikációs módszereket.

Legfontosabb klasszifikátorcsaládok:

- legközelebbi szomszéd módszer
- neurális hálók
- döntési szabályok, döntési fák
- Naiv-Bayes módszerek
- Szupport vektor gépek
- Random forest

3.1. Random Forest [5]

A Random Forest klasszifikátor egyszerre több kisebb méretű döntési fát is épít. Minden fa minden csúcsába k (konstans) kisorsolt változó közül valamilyen célfüggvény szerint legjobbat választja (az eredeti modell a Gini-index szerint dönt). A fák építését egy előre rögzített mélységig folytatja. Ha n db attribútumunk van célszerű $\log n$ -t választani a fák mélységének, így a modellünk $O(n)$ változót használ majd összesen. Egy ismeretlen rekord osztályáról a fák döntési szavazással döntenek.

3.2. Naiv-Bayes-módszer [6]

A naiv-Bayes osztályozás a felügyelt tanulás egy fajtája. Az egyik leggyakrabban alkalmazott és leghatékonyabb tanulási algoritmus, mert számos előnyös tulajdonsággal rendelkezik. Egyrészt nagyméretű problémákra is sikeresen alkalmazható, továbbá nem jelentenek problémát számára a zajos adatok sem.

A naiv-Bayes osztályozás először egy valószínűségi modellt épít fel a tanító példák alapján, majd egy ismeretlen minta osztálycímeként e modell alapján határozza meg.

Legyenek adottak:

$$A = \{X_1, X_2, \dots, X_n\} \quad \text{attribútumok halmaza} \quad (3.1)$$

$$V = \{C_1, C_2, \dots, C_n\} \quad \text{osztály-címkék halmaza} \quad (3.2)$$

Ezen felül adott a mintáknak egy N elemű halmaza úgy, hogy minden mintánál ismert az attribútumok értéke és a minta osztálya. A feladat egy még ismeretlen, attribútumokkal leírt objektum besorolása a megfelelő osztályba az tanító példák alapján felépített valószínűségi modell segítségével.

A naiv-Bayes osztályozás során abba a $c^* \in C$ osztályba soroljuk az $x_1, x_2, \dots, x_n$ jellemzőkkel leírt objektumot, melyre $P(c|x_1, x_2, \dots, x_n)$ valószínűség a legnagyobb, azaz

$$c^* = \arg \max_{c \in C} P(c|x_1, x_2, \dots, x_n) \quad (3.3)$$

Mivel a megfigyeléseinkből ezt közvetlen nem tudjuk kiszámolni, ezért olyan alakra kell hozni a képletet, hogy a benne szereplő valószínűségek könnyen megbecsülhetők legyenek. Ehhez alkalmazzuk először a Bayes-tételt:

$$c^* = \arg \max_{c \in C} P(c|x_1, x_2, \dots, x_n) = \arg \max_{c \in C} \frac{P(c|x_1, x_2, \dots, x_n) \cdot P(c)}{P(x_1, x_2, \dots, x_n)} \quad (3.4)$$

Ahhoz, hogy ennél egyszerűbb alakra hozzuk a képletet, fel kell tenni, hogy az osztály ismeretében az attribútumok egymástól függetlenek. Innen ered a naiv elnevezés.

$$c^* = \arg \max_{c \in C} \frac{P(c|x_1, x_2, \dots, x_n) \cdot P(c)}{P(x_1, x_2, \dots, x_n)} = \arg \max_{c \in C} \frac{P(c) \cdot \prod_{i=1}^n P(x_i|c)}{P(x_1, x_2, \dots, x_n)} \quad (3.5)$$

Mivel számunkra csak az a fontos, hogy melyik c osztály esetén adja a képlet a legnagyobb eredményt, ezért a nevezővel egyszerűsíthetünk, így az alábbi képlet adódik:

$$c^* = \arg \max_{c \in C} P(c) \cdot \prod_{i=1}^n P(x_i|c) \quad (3.6)$$

A $P(c)$ és a $P(x_i|c)$ valószínűségek a megfigyelési adatokból általában könnyen számíthatóak:

$$\begin{aligned} P(c) &\approx \frac{n_c}{N} \\ P(x_i|c) &= \frac{n_{c,x_i}}{n_c} \end{aligned} \quad (3.7)$$

ahol n_c azoknak a tanító mintáknak a száma, amelyek a c osztályba tartoznak, N a tanító minták száma, n_{c,x_i} azoknak a tanító mintáknak a száma, ahol az osztály c és az X_i attribútum-változó értéke x_i .

A $P(x_i|c)$ valószínűségeket célszerű m-estimate módszerrel közelíteni, mivel kevés tanító adat esetén előfordulhat, hogy egyes osztály(ok)hoz 0 becsült valószínűség tartozik. A gyakorlatban azonban általában egyetlen eseménynek sem 0 a valószínűsége. Az m-estimate módszer a 0 valószínűségeket 0-hoz közeli, de pozitív értékre korrigálja, így pontosabb becslést kapunk:

$$P(x_i|c) = \frac{n_{c,x_i} + m \cdot p}{n_c + m} \quad (3.8)$$

ahol p a priori (megfigyeléseket) megelőző valószínűsége az attribútumnak. Ha az attribútum-változó k -féle értéket vehet fel, akkor p értéke $1/k$,

m egy tetszőleges szám, ennyi véletlen mintát adunk hozzá a mintáink halmazához, minél kevesebb a tanító mintánk annál nagyobb m értéket célszerű választani

Ennek a módszernek az egyik, gyakran alkalmazott esete a Laplace-simítás, amikor $m = k$, vagyis

$$P(x_i|c) = \frac{n_{c,x_i} + 1}{n_c + k} \quad (3.9)$$

3.3. k-legközelebbi szomszéd módszere [7]

A k-legközelebbi szomszéd módszere egy „lusta” klasszifikáló eljárás, amely nem épít modellt. Alapelgondolása, hogy a hasonló attribútumú objektumok hasonló tulajdonságokkal bírnak. A hasonlóságot távolságfüggvénnyel mérjük. A tanuló adatbázist eltároljuk, és amikor egy ismeretlen objektumot kell klasszifikálnunk, akkor megkeressük a távolságfüggvény szerinti k darab legközelebbi pontot, és az objektumot abba a kategóriába soroljuk, amely a leggyakoribb a k szomszéd között (többségi szavazás). A módszer egyfajta lokális sűrűségfüggvény becslő eljárásnak is tekinthető.

A k érték megfelelő kiválasztása mindig az adatoktól függ. Minél nagyobb értéket választunk, annál kisebb a zaj az osztályozásnál, viszont az osztályok közötti határok kevésbé kivehetőek.

A távolság számítása

Tekintsünk egy elemre, mint egy pontra az m -dimenziós vektortérben. Az elemek közötti különbözőséget a vektoraik különbségének hosszával (normájával) definiáljuk. ($d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|$). Az Euklideszi- és a Manhattan-norma a Minkowski-norma speciális esetei.

$$\text{Euklideszi-norma: } L(\vec{z}) = \sqrt{z_1^2 + z_2^2 + \dots + z_m^2} \quad (3.10)$$

$$\text{Manhattan-norma: } L(\vec{z}) = |z_1| + |z_2| + \dots + |z_m| \quad (3.11)$$

$$\text{Minkowski-norma: } L(\vec{z}) = (|z_1|^p + |z_2|^p + \dots + |z_m|^p)^{1/p} \quad (3.12)$$

Hisztogramok távolsága: Számos módszert kidolgoztak a hisztogramok különbözőségének megállapítására.

Hisztogram metszés (Histogram intersection):

$$d_{\text{intersection}}(H_1, H_2) = \sum_i \min(H_1(i), H_2(i)) \quad (3.13)$$

Normalizált hisztogramok (minden i szürkeértékhez megadja az előfordulási valószínűségét) esetén az értéke 0 és 1 között van. Teljes egyezés esetén 1, teljes különbözőség esetén 0.

Chi-négyzet módszer:

$$d_{\text{chi-square}}(H_1, H_2) = \sum_i \frac{(H_1(i) - H_2(i))^2}{H_1(i) + H_2(i)} \quad (3.14)$$

ahol H_1, H_2 normalizált hisztogramok

A hisztogramok egyezése esetén az értéke 0, teljes különbözőség esetén végtelen értékű.

Bhattacharyya-távolság:

$$d_{\text{Bhattacharyya}}(H_1, H_2) = \sqrt{1 - \sum_i \frac{\sqrt{H_1(i) \cdot H_2(i)}}{\sqrt{\sum_i H_1(i) \cdot \sum_i H_2(i)}}} \quad (3.15)$$

ahol H_1, H_2 normalizált hisztogramok

A hisztogramok egyezése esetén az értéke 0, teljes különbözőség esetén 1.

- Kullback-Leibler divergencia:

$$d_{KL}(P \parallel Q) = \sum_i \ln\left(\frac{P(i)}{Q(i)}\right)P(i) \quad (3.16)$$

Q hisztogram távolsága P-től.

- Jenson-Shannon divergencia:

$$d_{JS}(P, Q) = \frac{1}{2}D(P \parallel M) + \frac{1}{2}D(Q \parallel M) \quad (3.17)$$

Ahol $M = \frac{1}{2}(P + Q)$ és $D(Q \parallel M)$ a Kullback-Leibler divergencia

3.4. Döntési fák

A döntési fák tanulása az egyik legegyszerűbb, mégis az egyik leggyakrabban alkalmazott felügyelt tanulási módszer. Az eddig tárgyalt algoritmusokhoz hasonlóan itt is bemeneti és kimeneti minták adottak, és ebből kell megtanulni a kimeneti értékeket képző függvényt. A bemeneteket attribútumokkal írjuk le, ezek és a kimenetek is lehetnek diszkrét vagy folytonosak, de a gyakorlaton csak diszkrét esetekkel foglalkozunk.

A döntési fa egy tesztsorozat elvégzése után jut el a döntéshez, ahol a fának:

- a belső csúcsai a bemeneti attribútumokhoz tartozó teszteknek,
- a csomópontokból kilépő ágak a tesztek lehetséges kimeneteleinek felelnek meg,
- a levélcsomópontok adják meg azt azokat az értékeket, amivel a döntési fa visszatér, ha az adott csomópontot elértük.

3.5. Szupport Vektor Gép [8], [9], [10]

Az SVM (Support Vector Machine) Vladimir N. Vapnik és Corinna Cortes többféle gépi tanulási feladatra alkalmazható modellje (1995).

Alapváltozata lineáris kétsztályos szeparálásra képes, de kiterjeszthető

- Egy- vagy többosztályos szeparálásra
- Nemlineáris szeparálásra
- Nemlineáris regressziós feladatokra

SVM lineárisan szeparálható adatokra:

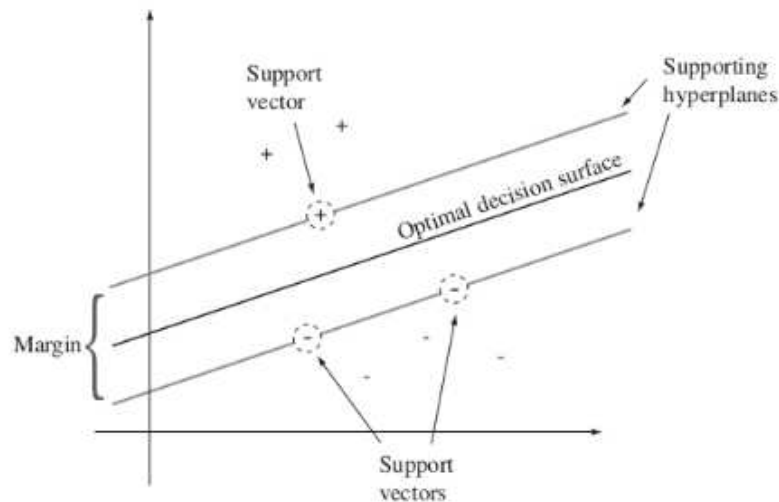
Legyen 2 különböző osztály: $y \in \{+1, -1\}$ (3.18)

A tanulóadatbázis álljon l darab tanítópontból (n dimenziós vektorok): $(\vec{x}, y)$ ahol $\vec{x} \in R^n$, $y \in \{+1, -1\}$ (3.19)

A tanítóhalmaz: $D: \{(\vec{x}_i, y) | \vec{x}_i \in R^n, y \in \{+1, -1\}\}$ (3.20)

A szeparáló hipersíkot úgy kell megválasztani, hogy minél messzebb legyen a 2 osztály pontjaitól, és pontosan középen legyen a 2 osztály pontjai között.

Margó-nak (margin) nevezzük a szeparáló hipersíkkal párhuzamos síkok által meghatározott térrészt, amely nem tartalmaz tanítópontokat. A hipersíkhöz legközelebb eső tanítópontok a szupport vektorok.

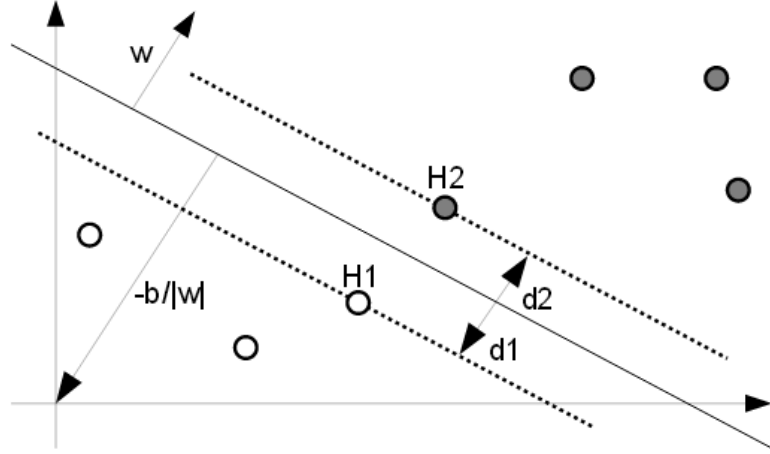


9. ábra szupport vektor gépek

Amennyiben a tanuló adatbázis lineárisan szeparálható a hipersík az alábbi egyenlettel fejezhető ki: $\vec{w} \cdot \vec{x} + b = 0$ (3.21)

ahol w a hipersíkra merőleges vektor (normálvektor)

$\frac{b}{|\vec{w}|}$ a hipersík távolsága az origó-tól



10. ábra elválasztó hipersík

A feladatunk az elválasztó hipersíkot úgy beállítani (w és b értékének megfelelő meghatározása), hogy a margó maximális legyen. A tanításra vonatkozó korlátok az alábbiak szerint alakulnak:

$$\bar{w} \cdot \bar{x}_i \geq 1 + b \quad \text{minden } (\bar{x}_i, y_i) \in D - re, \text{ ahol } y_i = +1 \quad (3.22)$$

$$\bar{w} \cdot \bar{x}_i \geq 1 - b \quad \text{minden } (\bar{x}_i, y_i) \in D - re, \text{ ahol } y_i = -1 \quad (3.23)$$

Ami $\bar{w} \cdot (y_i \bar{x}_i) \geq 1 - y_i b$ minden $(\bar{x}_i, y_i) \in D - re$ alakban írható fel a 2 korlátot egyesítve.

A margó maximalizálandó hossza $\frac{2}{|\bar{w}|}$, tehát a w vektor hosszától függ. Ez a feladat a négyzetgyökös kifejezés miatt bonyolult, de szerencsére a w vektor hossza ekvivalens $\frac{1}{2}|\bar{w}|^2$ kifejezéssel. Az optimalizálási feladat tehát a $\frac{1}{2}|\bar{w}|^2$ minimalizálását jelenti, amely egy kvadrátikus programozási feladat. Kvadrátikus programozási feladatnak nevezzük az olyan optimalizálási feladatot, amelyben a feltételek lineárisak, a célfüggvény pedig egy kvadrátikus függvény. A kvadrátikus programozás hátránya a nagy számításigény, mivel az összes tanulópont korlátot jelent. Egy optimalizálási feladathoz hozzárendelhető egy vele szoros kapcsolatban levő másik optimalizálási feladat. Az eredeti feladatot primál feladatnak nevezik, a primál feladathoz kapcsolódó optimalizálási feladatot duál feladatnak nevezik. Duál feladatot többféleképpen lehet definiálni, de szupport vektor gépek esetén a Lagrange duál feladatot szokták használni.

3.5.1. Lagrange duál feladat megoldása (általános eset) [10]

Legyen a primál feladat az alábbi:

$$\begin{aligned} f(x) &\rightarrow \min \\ g_i(x) &\leq 0 \quad i=1,2,\dots,m \\ h_i(x) &= 0 \quad i=1,2,\dots,k \\ x &\in X \end{aligned} \quad (3.24)$$

ahol $f(x)$ a primál feladat célfüggvénye

$g_i(x), h_i(x)$ feltételek

A primál feladathoz tartozó Lagrange duál feladat az alábbiak szerint írható fel:

$$\begin{aligned} \Theta(u_1, \dots, u_m, v_1, \dots, v_k) &\rightarrow \max \\ u_i &\geq 0 \quad i=1, \dots, m \end{aligned} \quad (3.25)$$

$$\text{ahol } \Theta(u_1, \dots, u_m, v_1, \dots, v_k) = \inf\{L(x, u, v) : x \in X\} = \inf\left\{f(x) + \sum_{i=1}^m u_i g_i(x) + \sum_{i=1}^k v_i h_i(x) : x \in X\right\}$$

(3.26)

$L(x, u, v)$ a Lagrange-függvény, az egyes feltételekhez rendelt $u_1, \dots, u_m, v_1, \dots, v_k$ együtthatókat Lagrange szorzóknak nevezzük.

A Lagrange duál feladat célfüggvénye tehát a Lagrange függvénynek minden Lagrange szorzóhoz tartozó infimuma az X halmazon. A Lagrange duál feladat feltételi halmaza egyszerű, hiszen csak az egyenlőtlenségekhez tartozó Lagrange szorzókra ír elő nemnegativitást.

3.5.2. A lineáris SVM Lagrange duál feladatának megoldása

Az előző részben felírtam hogyan kell meghatározni egy primál feladathoz tartozó Lagrange duál feladat megoldását általános esetben.

Alkalmazzuk az ott felírt összefüggéseket az SVM-hez tartozó kvadratikus programozási feladat Lagrange duál feladatának meghatározására és megoldására!

Az eredeti kvadratikus programozási feladat:

$$\begin{aligned} \min_{\vec{w}, b} [\Phi(\vec{w}, b)] &= \min_{\vec{w}, b} \frac{1}{2} |\vec{w}|^2 \\ y_i \cdot (\vec{w} \cdot \vec{x}_i - b) - 1 &\geq 0 \quad \text{a feltételek} \end{aligned} \quad (3.27)$$

Az ehhez tartozó Lagrange duál feladat így írható fel:

$$\max_{\alpha \geq 0} [\Theta(\vec{\alpha}, \vec{w}, b)] = \max_{\alpha \geq 0} \min_{\vec{w}, b} \left\{ \frac{1}{2} |\vec{w}|^2 - \sum_{i=1}^n \alpha_i [y_i (\vec{w} \cdot \vec{x}_i - b) - 1] \right\} \quad (3.28)$$

ahol $\alpha_i \geq 0$ az egyes feltételekhez rendelt Lagrange szorzók

Az ellentétes irányú szélsőértékek miatt ez szemléletesen azt jelenti, hogy $\Theta(\vec{\alpha}, \vec{w}, b)$ nyeregpontját keressük. Mivel $\Phi(\vec{w}, b)$ konvex pontosan 1 ilyen nyeregpont van.

Bizonyítható, hogy a $(\vec{w}^*, b^*, \vec{\alpha}^*)$ megoldás, akkor és csak akkor lesz az eredeti feladatnak is szélsőértéke, ha teljesülnek a Karush-Kuhn-Tucker feltételek:

$$\begin{aligned} \frac{\partial \Theta}{\partial \vec{w}}(\vec{w}^*, b^*, \vec{\alpha}^*) &= \vec{0} \\ \frac{\partial \Theta}{\partial b}(\vec{w}^*, b^*, \vec{\alpha}^*) &= 0 \\ \alpha_i^* (y_i (\vec{w}^* \cdot \vec{x}_i - b^*) - 1) &= 0 \\ y_i (\vec{w}^* \cdot \vec{x}_i - b^*) - 1 &\geq 0 \\ \alpha_i^* &\geq 0 \end{aligned} \quad (3.29)$$

Ha minden feltétel teljesül, akkor

$$\Theta(\vec{\alpha}^*, \vec{w}^*, b^*) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \vec{x}_i \cdot \vec{x}_j \quad (3.30)$$

A Karush-Kuhn-Tucker feltételek felhasználásával végül a maximális margó megtalálásának duális alakja:

$$\begin{aligned} \max_{\vec{\alpha}} [\Theta(\vec{\alpha}^*, \vec{w}^*, b^*)] &= \max_{\vec{\alpha}} \left(\sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \vec{x}_i \cdot \vec{x}_j \right) \quad (3.31) \\ \sum_{i=1}^l \alpha_i y_i &= 0 \\ \alpha_i &\geq 0 \end{aligned} \quad \text{feltételek}$$

Az egyik Karush-Kuhn-Tucker feltétel szerint:

$$\alpha_i^* (y_i (\vec{w}^* \cdot \vec{x}_i - b^*) - 1) = 0 \quad (3.32)$$

Ebből az következik, hogy:

- $\alpha_i^* = 0$ vagy
- $y_i (\vec{w}^* \cdot \vec{x}_i - b^*) - 1 = 0$.

Ha $\alpha_j > 0$ egy $(\vec{x}_j, y_j) \in D$ tanulópontra, akkor az a tanulópontra rajta van az egyik szupport-síkon.

Látható, hogy a duális megoldásban a hipersíkhöz legközelebb eső tanulópontokat határozzuk meg (szupport vektorok), amelyek a szupport-síkokat korlátozzák. Ezzel szemben a primál feladat megoldása a szupport-síkokat határozza meg.

Ha klasszifikátort arra szeretnénk használni, hogy egy pontot a 2 osztály valamelyikébe besoroljon, akkor az alábbi döntési függvénnyel határozhatjuk meg a pont osztályát:

$$f(\vec{x}) = \text{sgn}(\vec{w}^* \cdot \vec{x} - b^*) = \text{sgn}\left(\sum_{i=1}^l \alpha_i^* y_i \vec{x}_i \cdot \vec{x} - \sum_{i=1}^l \alpha_i^* y_i \vec{x}_i \cdot \vec{x}_{sv+} + 1\right) \quad (3.33)$$

ahol $\vec{x}_{sv+}$ egy tetszőleges szupport vektor a +1 osztályból

3.5.3. Puha margójú SVM (Soft margin SVM) [8]

Az előző részben ismertetett SVM megoldása optimális esetben érvényes, amikor az adatok lineárisan szeparálhatóak, a valóságban viszont az adataink zajosak lehetnek, és így az adatainkat nem lehet tökéletesen szétválasztani egy hipersíkkal. Általában vagy csak olyan lineáris szeparálás lehetséges, ahol a két osztály között csak nagyon kis biztonsági sáv alakítható ki, vagy lineáris szeparálás hibamentesen nem is lehetséges. A kis biztonsági sávot növelhetnénk, ha nem ragaszkodnánk ahhoz, hogy minden tanítópont a sávon kívül helyezkedjen el. Ha a pontok többsége a sávon kívül lesz, de néhány mintapontnál megengedjük, hogy a sávon belülré kerüljenek (sőt esetleg azt is megengedjük, hogy egyes mintapontok az elválasztó hipersík „rossz” oldalára kerüljenek), akkor ezen az áron a sáv mérete esetleg jelentősen növelhető.

A probléma megoldására puha margókat (soft margin) alkalmazunk. Az osztályozási feladatra vonatkozó feltételeket gyengítjük kiegészítő (slack) változók bevezetésével:

$$y_i(\vec{w} \cdot \vec{x}_i - b) > 1 - \varepsilon_i \quad \text{ahol } i = 1, \dots, l \text{ és } \varepsilon_i > 0 \quad (3.34)$$

Egy ε_i kiegészítő változó az elválasztó hipersík hibáját adja meg az x_i vektor esetén:

- Ha $0 < \varepsilon_i < 1$, akkor az x_i vektor a margón belülré esik, de a hipersík megfelelő oldalára
- Ha $\varepsilon_i > 1$, akkor az x_i vektor a hipersík nem megfelelő oldalára esik

A szupport vektorok azok az x_i vektor, amelyekre teljesül:

$$y_i(\vec{w} \cdot \vec{x}_i - b) = 1 - \varepsilon_i \quad \text{ahol } i = 1, \dots, l \text{ és } \varepsilon_i > 0 \quad (3.35)$$

Az optimalizálási feladat a következő lesz a kiegészítő változók bevezetésével:

$$\min_{\vec{w}, \varepsilon, b} \left\{ \frac{1}{2} \|\vec{w}\|^2 + C \sum_{i=1}^l \varepsilon_i \right\} \quad (3.36)$$

$$y_i(\vec{w} \cdot \vec{x}_i - b) > 1 - \varepsilon_i \quad \text{feltételek, ahol } i = 1, \dots, l \text{ és } \varepsilon_i > 0$$

A célfüggvényben szereplő C egy pozitív értékű paraméter, amit a felhasználó határoz meg a megoldás során, regularizációs paraméternek nevezzük.

A maximális margó megtalálásának duális alakja a Karush-Kuhn-Tucker feltételek felhasználásával határozható meg. Látható, hogy a képlet megegyezik a lineáris SVM-nél kapott összefüggéssel, nem tartalmazza a kiegészítő változókat, csak a feltételek változtak meg:

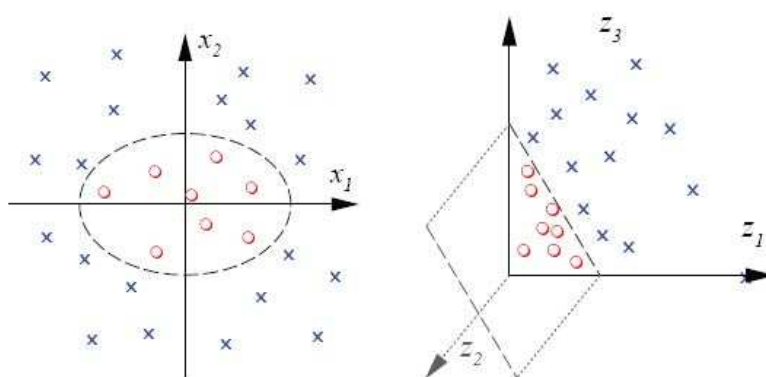
$$\max_{\vec{\alpha}} [\Theta(\vec{\alpha}^*, \vec{w}^*, b^*)] = \max_{\vec{\alpha}} \left(\sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \vec{x}_i \cdot \vec{x}_j \right) \quad (3.37)$$

$$\begin{aligned} \sum_{i=1}^l \alpha_i y_i &= 0 \\ 0 \leq \alpha_i &\leq C \end{aligned} \quad \text{feltételek}$$

3.5.4. Nemlineáris SVM [8]

Abban az esetben, ha a tanulóponatok nem szeparálhatóak lineárisan az eredeti tanítóponokat magasabb dimenzióba transzformáljuk remélve, hogy ott már lineárisan szeparálhatóak lesznek.

A kernel trükk lényege, hogy olyan problémáknál, melyeknél a változók skaláris szorzata szerepel csak, a skaláris szorzatot lecserélhetjük a nagyobb dimenziós térben vett skaláris szorzatra. Ennek a kiszámításához nincs szükség a nagyobb dimenziós térbe képező függvényre ($\varphi(\vec{x})$ -re).



11. ábra kernel trükk alkalmazása

Kernel függvénynek nevezünk egy olyan függvényt, amely vektorok egy transzformált skaláris szorzatát az eredeti adattérben számítja ki.

$$K(\vec{x}, \vec{y}) = \varphi(\vec{x}) \cdot \varphi(\vec{y}) \quad (3.38)$$

Egy Kernel függvény szimmetrikus a változóiban: $K(\vec{x}, \vec{y}) = K(\vec{y}, \vec{x})$ (3.39)

A lineáris SVM-nél felírt döntési függvény az alábbi alakra változik Kernel függvény $K(\vec{x}_i, \vec{x}) = \varphi(\vec{x}_i) \cdot \varphi(\vec{x})$ használatával:

$$f(\vec{x}) = \text{sgn}(\vec{w}^* \cdot \varphi(\vec{x}) - b^*) = \text{sgn}\left(\sum_{i=1}^l \alpha_i^* y_i \varphi(\vec{x}_i) \cdot \varphi(\vec{x}) - b^*\right) = \text{sgn}\left(\sum_{i=1}^l \alpha_i^* y_i K(\vec{x}_i, \vec{x}) - b^*\right) \quad (3.40)$$

A szükséges és elégséges feltételeket ahhoz, hogy egy függvény Kernel függvény legyen a Mercer-tétel fogalmazza meg:

Egy változóiban szimmetrikus folytonos $K(\vec{x}, \vec{y})$ függvény akkor és csak akkor fejezhető ki $K(\vec{x}, \vec{y}) = \varphi(\vec{x}) \cdot \varphi(\vec{y})$ alakban, ha minden olyan $g(x)$ függvény esetén, amelyre $\int g(x)^2 dx$ véges, teljesül $\int \int K(x, y) g(x) g(y) dx dy > 0$.

A leggyakrabban használt Kernel függvények:

- Polinomiális (homogén): $K(\vec{x}_i, \vec{x}_j) = (\vec{x}_i \cdot \vec{x}_j)^d \quad (3.41)$

- Polinomiális (inhomogén): $K(\vec{x}_i, \vec{x}_j) = (\vec{x}_i \cdot \vec{x}_j + 1)^d \quad (3.42)$

- Gauss (RBF): $K(\vec{x}_i, \vec{x}_j) = \exp(-\gamma |\vec{x}_i - \vec{x}_j|^2) \quad (3.43)$

ahol $\gamma > 0$ (3.44)

- Hiperbolikus tangens: $K(\vec{x}_i, \vec{x}_j) = \tanh(\kappa \vec{x}_i \cdot \vec{x}_j + c)^d \quad (3.45)$

ahol $\kappa > 0$ és $c < 0$

3.5.5. Több osztályú SVM

Az SVM-et eredetileg bináris osztályozási feladatokra dolgozták ki, de 2-nél több osztályos problémák megoldására is alkalmassá tehető.

2 alap stratégiát különböztetünk meg a több osztályból álló problémák SVM-mel történő megoldására:

- a többosztályos osztályozási feladat megfogalmazása feltételes szélsőérték-számítási feladatként
- a többosztályos osztályozási feladat felbontása több bináris osztályozási feladattá

A többosztályos feladat több bináris feladattá történő felbontása esetén egy vektor osztályát az egyes bináris SVM-ek előrejelzéseinek összegzése alapján határozzák meg. Ehhez többségi

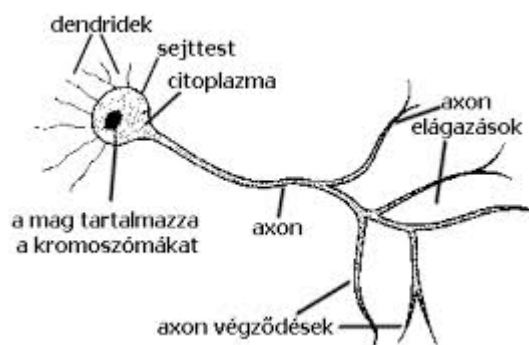
szavazást szokás alkalmazni, a vektor abba az osztályba fog tartozni, amelyik a legtöbb szavazatot kapta.

A gyakorlatban 2 féle megoldás terjedt el:

- „egy az egy ellen” (one against one) : Az összes különböző osztálypárhoz 1-1 bináris osztályozót készítünk, így $k > 2$ osztály esetén $\frac{k(k-1)}{2}$ darab osztályozónk lesz. Minden osztályozó csak az adott 2 osztályhoz tartozó vektorokat használja fel.
- „egy a többi ellen” (one against others): olyan bináris osztályozók alkotása, amelyek 1 osztályt tekintenek + osztályként, az összes többit – osztályként, így összesen k darab bináris osztályozónk lesz.

3.6. Neurális hálózatok [11], [12]

A mesterséges neurális hálózatok működésének megértéséhez szükséges megismerni az agyban található neurális hálózatokban végbemenő folyamatokat, hiszen a mesterséges neurális hálózatok ezek mintájára épülnek.



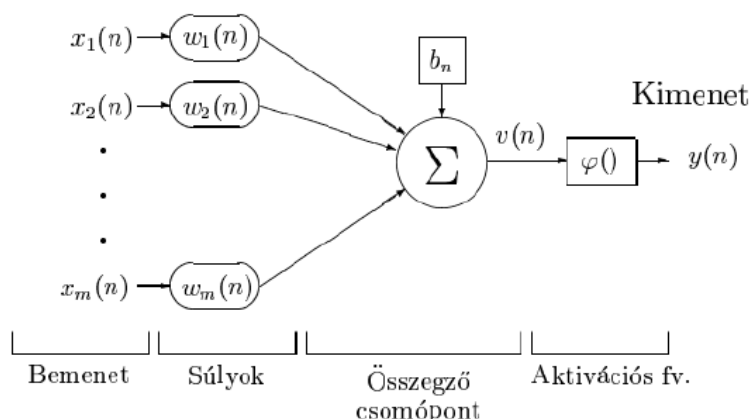
12. ábra Neuron felépítése

Neuronnak nevezzük az idegsejt és nyúlványainak együttesét. A neuronok az idegrendszer legkisebb elemét jelentik. A neuronok ingerérzékeny sejtek, amelyeknek az ingerfelvétel és az idegi ingerek vezetése a feladatuk.

A dendritek azok a nyúlványok, amelyek az információ felvételére és a sejttest irányába történő vezetéséért felelősek. A dendritek az agyi neuronok bemenetei, amelyek más neuronokhoz csatlakoznak. Az információ-feldolgozást a neuronok sejtmagja végzi. Az axon egy cső alakú nyúlvány, amely az ingerületeket a sejttől távolodó irányba vezeti. Az axon végződésai több másik neuron dendritjeihez csatlakozhatnak.

3.6.1. Perceptron [11], [12]

A neurális hálózat elemi egységekből, neuronokból épülnek fel, melyek valamilyen lokális feldolgozást hajtanak végre. A neuronok valamilyen struktúra szerint összekapcsolódva hálózatot alkotnak. A neurális hálózat tanulása minták alapján megy végbe, a megszerzett tudás a neuronok közötti összeköttetésekben, súlyokban tárolódik.



13. ábra Perceptron felépítése

A mesterséges neuron az agyi neuron másolata, egy több bemenetű, egy kimenetű elem, ahol a bemenetek súlyozva egy összegző bemenetére kerülnek.

A Rosenblatt vagy egyszerű perceptron egy olyan hálózat, amely a megfelelő beállítás illetve tanítás után egy lineáris szeparálási probléma megoldására alkalmas.

A neuron a biológiai neuronhoz hasonlóan reagál a feldolgozott inputokra, a súlyozott összeg egy aktivációs függvényen megy keresztül. A bemeneti súlyozott összeget ingernek (excitation) is szokták nevezni.

A bemeneti súlyozott összeg M db bemenet esetén:

$$v(n) = \sum_{i=1}^M w_i x_i + b_n \quad (3.46)$$

A súlyozott összeghez hozzáadunk egy b eltolási (bias) értéket. Az eltolási érték nélkül az osztályokat elválasztó hipersík az origón minden esetben áthaladna. Az eltolási érték lehetővé teszi, hogy a hipersík b -vel való eltolásával olyan adathalmazok osztályozását is meg tudjuk oldani, amit egy origón áthaladó hipersíkkal nem lehetett volna. Általában a $b(n)$ torzítást is súlyként szoktuk értelmezni és $w(n)$ -ként hivatkozunk rá. A torzításhoz tartozó bemeneti értéket pedig $x(n) = 1$ -nek vesszük.

Az aktivációs függvény kimenetén megjelenő érték továbbítódik a hozzá kapcsolódó neuronokhoz. A kimeneti jelet válasznak (activation) is szokták nevezni.

Az aktivációs függvénynek az esetek többségében nemlineáris függvényt kell választani. Lineáris aktivációs függvény használatával a neuron egy lineáris függvényt valósítana meg, amely sok esetben azt eredményezné, hogy a hálózat hibája nem csökkenne le kellőképpen.

A leggyakrabban alkalmazott aktivációs függvények az alábbiak:

- Egységugrás függvény:
$$y(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (3.47)$$

- Előjelfüggvény:
$$y(x) = \begin{cases} 1, & x > 0 \\ -1, & x \leq 0 \end{cases} \quad (3.48)$$

- Sigmoid függvény
$$y(x) = \frac{1}{1 + e^{-x}} \quad (3.49)$$

Az egyes rétegekhez tartozó neuronok a bemenettel vagy az előző réteg neuronjainak kimenetével illetve a következő réteg bemenetével vannak összekötve. A legtöbb hálózat esetében a rétegek teljesen össze vannak kötve, tehát a réteg minden kimenete össze van kapcsolva a következő réteg összes bemenetével.

Az egyes rétegeket feladatuk alapján különböztetjük meg:

- Bemeneti réteg: Ebben a rétegben azok a neuronok találhatóak, amelyek a bemeneti jelek továbbítását végzik a további rétegek felé.
- Rejtett réteg: A tulajdonképpeni információfeldolgozást ebben a rétegben található neuronok végzik. A rejtett rétegek száma több is lehet.
- Kimeneti réteg: Ebben a rétegben található neuronok továbbítják az információt a külvilág felé.

3.6.2. A perceptron tanulása [11], [12]

A tanulási folyamat tanulóadatokkal történik. A tanulóadatoknál a bemenetekhez tartozó kimeneti érték ismert. A bemeneti értékeket a perceptron bemeneteire adjuk, és kiszámoljuk a neuron kimenetét. A neuron kimeneti értékét összehasonlítjuk az elvárt kimeneti értékkel. A két érték közötti különbséget a hálózat hibájának nevezzük. A tanulás során az a célunk, hogy a hálózat hibája lehetőleg minél kisebb értékre csökkenjen. Az ideális az lenne, ha a hiba 0-ra csökkenne, de általában ez nem valósítható meg.

Jelölje az $x(n)$ bemenethez tartozó ideális kimenetet $d(n)$, illetve $y(n)$ jelentse a hálózat által az $x(n)$ bemenetre adott kimenetét, ekkor a neurális hálózat négyzetes hibája a következőképpen számítható ki:

$$\varepsilon = (d(n) - y(n))^2 \quad (3.50)$$

A tanulás szabályai:

- ha az adott bemenet mellett számított kimenet jó ($y_i = d_i$), akkor a súlytényezőket nem változtatjuk;
- ha $y_i = 0$, de $d_i = 1$, akkor növeljük az "aktív" inputok (melyek zérusnál nagyobbak) súlytényezőit;
- ha $y_i = 1$, de $d_i = 0$, akkor csökkentjük az "aktív" inputok súlytényezőit.

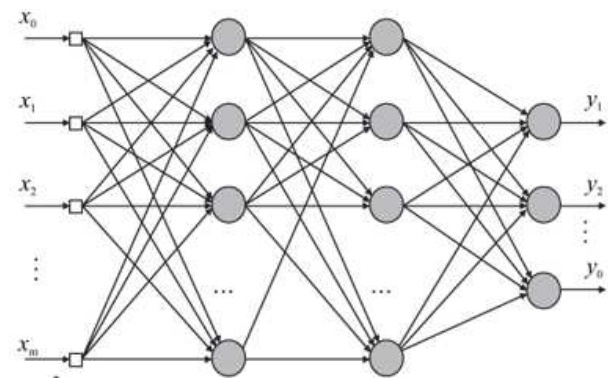
A fenti növelések és csökkentések mértéke sokféle lehet, tipikusan:

- rögzített növekmény
- a $(d_i - y_i)$ –vel arányos növekmény
- az előző kettő kombinációja

3.6.3. Többrétegű perceptron [11], [12]

Egy egyszerű, lineárisan szeparálható feladat egyetlen perceptronnal is megoldható, viszont bonyolultabb, lineárisan nem szeparálható feladatok megoldásához többrétegű perceptron hálózat szükséges. A többrétegű perceptron hálózat áll egy bemeneti rétegből, rejtett réteg(ek)ből és egy kimeneti rétegből.

A rejtett réteg és az abban elhelyezkedő perceptronok száma lényegesen befolyásolhatja, általában javíthatja, a hálózat approximációs képességét, hiszen egy bonyolultabb függvénykapcsolatot reprezentál a bemenet és a kimenet között. Viszont egy nagyobb hálózat tanítása hosszabb időt vesz igénybe.



14. ábra többrétegű előrecsatolt hálózat felépítése

3.6.4. Backpropagation algoritmus [11], [12]

Elsőrendű gradiens módszer, amely a hiba elsőrendű deriváltjait használja iteratív módon a hálózat súlyainak beállítására. A módszer hátránya, hogy könnyen beragadhat egy lokális minimumba.

A hálózat által meghatározott kimenet és az elvárt kimenet közötti különbség az alábbi összefüggéssel írható fel, amit a neurális hálózat hibájának nevezünk:

$$E = \frac{1}{2} \sum_{i=1}^m (d_i(n) - y_i(n))^2 \quad (3.51)$$

A backpropagation algoritmust használjuk arra, hogy a hiba minimumát megtaláljuk.

Ebben az összefüggésben d_i a i . mintához tartozó ideális kimenet, y_i az i . mintára számított kimenet, m pedig a minták darabszáma.

A súlyok meghatározására szolgáló iterációs eljárás két fázisból áll:

- 1) propagáció: az aktuális súlyokkal a bemenő jelekből a kimeneti jelek kiszámítása.
- 2) backpropagáció: a bemeneti jelekhez tartozó tanulóhalmazbeli kimeneti jelek elő írt értékei és a valóságos kimenetek alapján a súlyok módosítása, a kimeneti rétegtől a bemeneti réteg felé, azaz visszafelé haladva.

Leggyakrabban szigmoid aktivációs függvényt szoktak alkalmazni, mivel ez a függvény deriválható.

- a) minden kimeneti neuronra kiszámítjuk a módosult súlyokat a gradiensek meghatározása után szigmoid aktivációs függvény használata esetén:

$$\begin{aligned} \frac{\partial E}{\partial w_{j,k}} &= y_j \delta_k \\ \delta_k &= y_k (1 - y_k) (d_k - y_k) \\ w_{l,k}^* &= w_{l,k} + \eta \delta_k y_l \end{aligned} \quad (3.52)$$

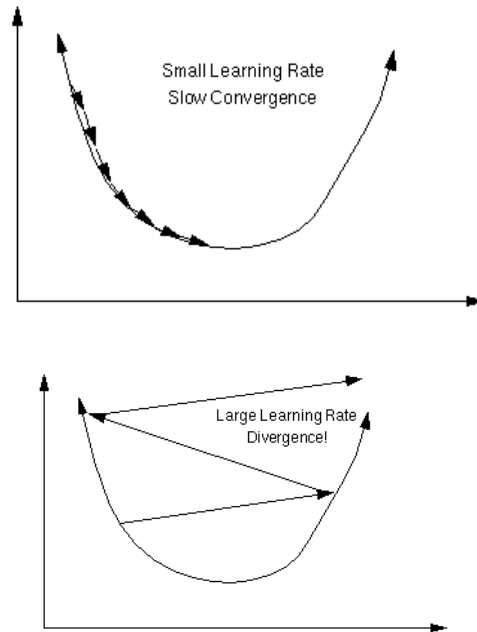
$\eta > 0$ tanulási paraméter

- b) a rejtett rétegekben a következő réteg értékeire támaszkodva számítjuk a módosult súlyokat:

$$\begin{aligned} \frac{\partial E}{\partial w_{i,j}} &= y_i \delta_j \\ \delta_j &= y_j (1 - y_j) \sum_k \delta_k w_{j,k} \\ w_{l,j}^* &= w_{l,j} + \eta \delta_j y_l \end{aligned} \quad (3.53)$$

$\eta > 0$ tanulási ráta

A tanulási ráta megfelelő beállítása nagyon fontos, értékét a programozónak kell meghatározni. Ha túl kicsinek választjuk, akkor az algoritmus lassan fog konvergálni a hiba minimuma felé. A túl nagy tanulási ráta hatására az algoritmus instabillá válhat, oszcillálhat.



15. ábra Tanulási ráta hatása

3.6.5. Backpropagation algoritmus momentummal [11], [12]

A backpropagation algoritmus momentummal alkalmas arra, hogy elkerüljük a hiba lokális minimumba történő beragadását. Ennek elkerülésére meg kell növelni a tanulási ráta értékét, viszont az instabil viselkedést okozhat. Az algoritmus előnye, hogy nagyobb tanulási rátát használhatunk a stabilitás megtartása mellett mint a backpropagation algoritmusnál. Momentum értékének 0 és 1 közötti értéket kell választani.

A módosult súlyok az alábbi összefüggésekkel számíthatóak:

a) kimeneti neuronok esetén:

$$\begin{aligned} w_{l,k}^*(t+1) &= w_{l,k}(t) + \mu(w_{l,k}(t) - w_{l,k}(t-1)) + \eta \delta_k y_l \\ \delta_k &= y_k(1 - y_k)(d_k - y_k) \end{aligned} \quad (3.54)$$

μ a momentum értéke

b) rejtett réteg neuronjai esetén:

$$\begin{aligned}
w_{l,k}^*(t+1) &= w_{l,k}(t) + \mu(w_{l,k}(t) - w_{l,k}(t-1)) + \eta \delta_k y_l \\
\delta_j &= y_j(1 - y_j) \sum_k \delta_k w_{j,k}
\end{aligned}
\tag{3.55}$$

μ a momentum értéke

3.6.6. Backpropagation algoritmus adaptív tanulási paraméterrel [12]

Az eredeti backpropagation algoritmus során a tanulási ráta értéke a számítás során a programozó által meghatározott érték. A backpropagation algoritmus teljesítményét nagyban befolyásolja a tanulási ráta értéke.

Az algoritmus teljesítményét javítani lehet adaptív tanulási ráta alkalmazásával, mivel az optimális tanulási ráta a tanulási folyamat során változik.

Az algoritmus célja, hogy a lehető legnagyobb legyen a tanulási ráta a tanulás során a stabilitás megtartása mellett.

A tanulási ráta módosítása az alábbi folyamat szerint történik:

- Ha az új hiba egy előre definiált aránnyal meghaladja az utolsó iterációban számolt hibát (általában 1,04):
 - A súlyokat nem módosítjuk
 - A tanulási rátát csökkentjük meghatározott részére csökkentjük (általában 0,7)
- Ha az új hiba kisebb mint az utolsó iterációban számolt hiba
 - A súlyokat módosítjuk
 - A tanulási rátát növelni fogjuk (általában 1,05 szörösére)
- Az új hiba nem kisebb mint a régi és nem haladja meg a meghatározott aránnyal
 - A súlyokat módosítjuk
 - A tanulási ráta változatlan marad

Túltanulás: A modell túl jól illeszkedik a tanuló példákra, ennek következtében nem tud jól általánosítani, azaz új, a tesztelés során nem tanult példák esetén nem működik jól a modell. Ha a neurális hálózat paramétereinek száma sokkal kisebb mint a tanuló adatok száma, akkor elég kicsi az esély a túltanulásra. A neurális hálózat alul illetve túltanulásának elkerülését tehát alapvetően a megfelelően megválasztott rejtett rétegbeli csomópontok számával biztosíthatjuk.

A túltanulás akkor jelentkezhet, ha korlátozott száma tanuló adat áll a rendelkezésünkre. Több módszer ismert ilyen esetekben is a túltanulás elkerülésére.

Korai leállás módszere (early stopping):

A módszer alkalmazása esetén a rendelkezésre álló adatokat 2 részre osztjuk, egy tanuló halmazra és egy validációs halmazra. A tanuló halmazon addig történik a tanítás, míg a validációs halmazon is csökken a hiba. Ha a tanuló halmazon a hálózat hibáját nagyon lecsökkentjük, a hálózat elkezd túltanulni a tanuló adatokat, és a validációs hiba növekedni fog.

Regularizáció:

A regulációs módszer lényege, hogy a hálózat hibájának csökkentése mellett minimalizálni próbáljuk a hálózat által megvalósított függvény gradiensét is. Ez az ötlet arra a tapasztalatra támaszkodik, hogy a túltanulás esetén a hálózat a mérési hibákat is megtanulva, nagy lokális gradienseket produkál az adatpontokban (hullámozás). Ezért azután a minimalizálandó hibafüggvényt kiegészítjük egy a gradiensek nagyságával arányos ún. regularizációs taggal. Ennek eredményeként a súlyok kisebbek lesznek, és csökken a túltanulás valószínűsége és a hálózat válasza simább lesz.

3.6.7. Newton-módszer [11], [12]

A Newton-módszer a másodrendű gradiens módszerek csoportjába tartozik, a hibafüggvény másodrendű deriváltjait használja fel a hálózat súlyainak meghatározására. A Newton-módszer négyzetes függvények minimumát 1 lépésben találja meg. Általában gyorsabb konvergenciát mutatnak mint a backpropagation algoritmus, de hátránya a nagyobb számítási igény (Hesse-mátrix számítása miatt), valamint előfordulhat oszcilláció vagy divergencia. Ezzel szemben a backpropagation algoritmus a tanulási ráta megfelelő választása esetén garantáltan konvergálni fog.

A Newton-módszer azt feltételezi, hogy a gradiens vektor komponensei a súlyok értékétől függenek, illetve a súlyok lineárisan függetlenek.

Tekintsük a hibafüggvény másodfokú Taylor-polinomját a w_k pontban!

$$q(w) = E(w_k) + \nabla E(w_k)(w - w_k) + \frac{1}{2}(w - w_k)H(w_k)(w - w_k) \quad (3.56)$$

$$H = \begin{vmatrix} \frac{\partial^2 E}{\partial w_1^2} & \frac{\partial^2 E}{\partial w_1 \partial w_2} & \cdots & \frac{\partial^2 E}{\partial w_1 \partial w_N} \\ \frac{\partial^2 E}{\partial w_2 \partial w_1} & \frac{\partial^2 E}{\partial w_2^2} & \cdots & \frac{\partial^2 E}{\partial w_2 \partial w_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 E}{\partial w_N \partial w_1} & \frac{\partial^2 E}{\partial w_N \partial w_2} & \cdots & \frac{\partial^2 E}{\partial w_N^2} \end{vmatrix} \quad (3.57)$$

A H a Hesse-mátrixot jelöli, amely az E hibafüggvény másodrendű deriváltjait tartalmazza.

Határozzuk meg $q(w)$ minimumát! A minimum szükséges feltétele:

$$\nabla q(w) = 0 \quad (3.58)$$

Elvégezve a gradiens-képzést az alábbi összefüggést kapjuk:

$$\begin{aligned} \nabla E(w_k) + H(w_k)(w - w_k) &= 0 \\ w_{k+1} &= w_k - H(w_k)^{-1} \nabla E(w_k) \end{aligned} \quad (3.59)$$

A Hesse-mátrix kiszámítása bonyolult, hosszadalmas, ezért olyan módszereket dolgoztak ki, amelyek a Hesse-mátrixok közelítik.

Ezeket a módszereket kvázi Newton-módszereknek nevezik. A kvázi Newton-módszereknek a másodrendű deriváltakat nem számítjuk ki, hanem a közelítésüket számoljuk ki egy egyszerű képlettel.

3.6.8. BFGS (Broyden-Fletcher-Goldfarb-Shanno) eljárás [11], [12]

Vezessük be az y_k jelölést a gradiensek különbségére, az s_k jelölést pedig a w súlyvektorok különbségére!

$$\begin{aligned} s_k &= w_{k+1} - w_k \\ y_k &= \nabla E(w_{k+1}) - \nabla E(w_k) \end{aligned} \quad (3.60)$$

Induló lépésként kiindulunk egy tetszőleges $w_k \in R^n$ vektorból és egy szimmetrikus pozitív definit B_k mátrixból.

A k . lépés az alábbiak szerint alakul:

1. Megoldjuk a $B_k s_k = -\nabla E(w_k)$ lineáris egyenletrendszert az s_k különbségvektorra.
2. Kiszámoljuk a w_{k+1} vektor értékét a $w_{k+1} = w_k + s_k$ összefüggés alapján.
3. Az alábbi formulákkal kiszámoljuk az y_k vektort és a Hesse mátrixot közelítő következő B_{k+1} mátrix értékét:

$$B_{k+1} = B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{(B_k s_k)(B_k s_k)^T}{(B_k s_k)^T s_k} \quad (3.61)$$

3.6.9. Levenberg-Marquardt algoritmus [11], [12]

A Levenberg-Marquardt algoritmus az egyik leggyakrabban használt, leghatékonyabb tanuló algoritmus, amely sokkal hatékonyabb mint a backpropagation algoritmus. A módszer kiválóan alkalmas neurális hálózatok súlyainak beállítására. Az eljárás ötvözi a backpropagation eljárás és a Gauss-Newton módszer előnyeit: a backpropagation algoritmus stabilitását illetve a Gauss-Newton módszer gyors konvergenciáját.

Ezt az iterációs eljárást a legkisebb négyzetes függvényközelítési probléma megoldására használják, így használható nemlineáris függvények négyzetösszegének minimalizálására is.

$$E(w) = \frac{1}{2} \sum_{i=1}^m (f_i(w))^2 \rightarrow \min \quad (3.62)$$

A Levenberg-Marquardt algoritmus a Hesse mátrix közelítésére az alábbi összefüggést adja, így biztosítva a Hesse mátrix invertálhatóságát:

$$H \approx J^T J + \mu I \quad (3.63)$$

ahol J a Jacobi-mátrix, I egységmátrix

A súlyok új értéke az alábbi összefüggéssel számítható:

$$\begin{aligned} \underline{w}_{k+1} &= \underline{w}_k - (J_k^T J_k + \mu I)^{-1} J_k \underline{e}_k \\ \underline{e}_k &= \underline{y}_k - \underline{t}_k \end{aligned} \quad (3.64)$$

ahol $\underline{e}_k$ a hibavektor értéke a k . lépésben

A μ ún. regularizációs paraméter a keresés irányát és a módosítás nagyságát egyaránt befolyásolja. Ha a regularizációs paraméter kicsi értékű (közel 0), akkor az algoritmus a Gauss-Newton módszerhez konvergál, ha az értéke a végtelenbe tart, akkor pedig a legmeredekebb lejtő típusú megközelítést adja.

Az algoritmus k . iterációs lépése az alábbiak szerint alakul:

1. Adott $\underline{w}_k$ súlyok valamint a μ_k regularizációs paraméter

Inicializálásként tetszőleges pozitív μ értéket választhatunk

2. J_k és $\underline{e}_k$ meghatározása

3. $\underline{s}_k = -(J_k^T J_k + \mu I)^{-1} J_k \underline{e}_k$ kiszámítása

4. r_k az ún megbízhatósági régió számítása
$$r_k = \frac{E(\underline{w}_k) - E(\underline{w}_k - \underline{s}_k)}{E(\underline{w}_k) - \frac{1}{2} \|J_k \underline{s}_k + \underline{e}_k\|^2} \quad (3.65)$$

5. a μ paraméter értékét az r_k értékétől függően állítjuk

o Ha $r_k < 0.25$ akkor $\mu_{k+1} = 4\mu_k$

o Ha $r_k > 0.75$ akkor $\mu_{k+1} = \frac{1}{2}\mu_k$

o Egyébként $\mu_{k+1} = \mu_k$

6. Ha $r_k \leq 0$ akkor $\underline{w}_{k+1} = \underline{w}_k$, különben $\underline{w}_{k+1} = \underline{w}_k + \underline{s}_k$

3.6.10. Konjugált gradiens módszer [11], [12]

A konjugált gradiens módszer egy iterációs módszer, amely alkalmazható többrétegű perceptron hálózatok tanító algoritmusaként.

Az átlagos négyzetes hiba $\underline{w}_k$ körüli Taylor-sora:

$$E(\underline{w}_k + \Delta \underline{w}_k) = E(\underline{w}_k) + g(\underline{w}_k)^T \Delta \underline{w}_k + \frac{1}{2} \Delta \underline{w}_k^T H(\underline{w}_k) \Delta \underline{w}_k + \dots \quad (3.66)$$

ahol $g(\underline{w})$ a gradiens vektor, $H(\underline{w})$ pedig a Hesse mátrix

A konjugált gradiens módszer alkalmazható, hiszen a fenti Taylor-sor alapján közelítőleg egy

$$f(x) = \frac{1}{2} x^T A x - b^T x + c \quad (3.67)$$

alakú kvadratikus képletet kell minimalizálni. Először kiszámítjuk a gradienst a backpropagation eljárással, majd a konjugált gradiens módszer segítségével úgy módosítjuk a súlyokat, mintha a közelítő kvadratikus forma minimumát keresnénk.

Az iterációs eljárás az alábbiak szerint írható le:

- 1) Inicializálás: Az első iterációban a minimum-keresés a legmeredekebb lejtő irányában (negatív gradiens) kezdődik.

$$\underline{p}_0 = -\underline{g}_0 \quad (3.68)$$

- 2) k. lépés:

- a. Keressük az új súlyokat

$$\underline{w}_k = \underline{w}_{k+1} + \eta_k \underline{p}_{k-1} \quad (3.69)$$

alakban. Az η_k úgy határozzuk meg, hogy $E(\underline{w}_k)$ minimális legyen a keresési irány mentén. Ez egy egyenes mentén való keresés, ami numerikusan megoldható.

- b. Ezután a $\underline{w}_k$ pontban megkeressük a hibafüggvény gradiensét.

Ha $|\underline{g}_k| \leq \delta |\underline{g}_0|$, ahol $\delta > 0$ előre megadott kis érték, akkor megállunk.

Ellenkező esetben folytatjuk a következő lépéssel, viszont a következő lépés nem a $-\underline{g}_k$ irányban történik, hanem a konjugált gradiens módszer szerint meghatározott irányban. A keresési irány akkor lesz konjugált, ha merőleges a korábbi gradiens-változásokra.

$$\underline{p}_k = -\underline{g}_k + \beta_k \underline{p}_{k-1} \quad (3.70)$$

A β_k számítása a Fletcher-Reeves-formula szerint:

$$\beta_k = \frac{\underline{g}_k^T \underline{g}_k}{\underline{g}_{k-1}^T \underline{g}_{k-1}} \quad (3.71)$$

A β_k számítása a Polak-Ribière-formula szerint:

$$\beta_k = \frac{\Delta \underline{g}_{k-1}^T + \underline{g}_k^T}{\underline{g}_{k-1}^T \underline{g}_{k-1}} \quad (3.72)$$

4. A feladat gyakorlati megvalósítása [13], [14], [15]

A talaj felismerő rendszer programjait Matlab-ban készítettem el. Ebben a fejezetben ismertetem a textúrák leírására használt módszereket, valamint az egyes klasszifikációs módszerekkel kapott eredményeket.

4.1. Talajfelismerő rendszerek áttekintése [13], [14]

A talajfelismerés a képfeldolgozás egyik fontos területét jelenti, például az autonóm robotoknál kiemelt jelentőséggel bír.

Az autonóm robotok az alkalmazások széles palettáját lefedik. Az autonómia azt jelenti, hogy a robotok tevékenységüket részben vagy teljesen maguk irányítják.

Például a hadászat és a katasztrófavédelem előszeretettel alkalmaz autonóm robotokat felderítési céllal, élen járnak az autonóm robot fejlesztésében. Az autonóm robotok katonai alkalmazása olyan környezetben különösen indokolt, amely ember számára valamilyen okból nem megközelíthető (harci gázzal vagy nukleáris sugárzással szennyezett), vagy nagyon kockázatosan megközelíthető őrzött terület.

A talajfelszín felismerése nélkülözhetetlen az autonóm robotok biztonságos közlekedéséhez. A robotnak képes kell lennie a környező terepviszonyok feltérképezésére, ezek ismeretében dönteni tud az optimális útvonaláról elkerülve az akadályokat és nehezen járható terepeket.

A robotnak tehát az általa készített képi információk alapján meg kell tudnia határozni a talaj típusát annak jellemzőit.

A talajfelismerési technikák 2 csoportba oszthatóak:

- Retrospektív talajfelismerés: a gyűjtött adatokból meghatározzuk a talaj típusát, amelyen éppen halad a robot
- Prospektív talajfelismerés: a robot környezetében lévő talajviszonyokat határozzuk meg

A prospektív talajfelismerés történhet a környezet geometriájának meghatározásával a LIDAR szenzorok (Light Detection And Ranging) vagy sztereo kamerák által szolgáltatott információk alapján. A sztereo kamerák az emberi binokuláris látást módszerével készítenek képet a környezetükről.

4.1.1. 3D LIDAR alapú talajfelismerés [13]

A LIDAR (lézer szkennelés) egy hatékony technológia geometriai adatok előállítására. A LIDAR lézerrel térképezi fel a kérdéses területet, a fényvisszaverődés méréseiből háromdimenziós profilt alkotva leképezi a felület jellemzőit és érzékeli a tárgyakat. Az objektumokat több millió elemből álló információhalmaz (pontfelhő) határozza meg.

A 3D LIDAR alapú talajfelismerés lépései:

1. szűrés: A pontfelhő sok hibás mérési eredményt tartalmaz, amiket el kell távolítanunk. Az extrém magas illetve alacsony intenzitású pontok hibás mérésre utalnak, ezeket könnyen meghatározhatjuk. A zaj csökkentésére eltávolítjuk azokat a pontokat, amelyek fix sugarú környezetében egy adott számnál kevesebb szomszédja van. A szűrés általában a pontok 20-25 %-át távolítja el.
2. Talaj síkjának meghatározása: A szűrés után a talaj síkjának a detektálása történik. Ha a talaj síkja tartalmazza a legtöbb pontot, akkor egyszerű dolgunk van. Ebben az esetben például a RANSAC (iteratív robosztus becslés) algoritmussal könnyen meghatározhatjuk a talaj síkját.
Több sík esetén már nehezebb meghatározni a talaj síkját, főleg, ha nem a talaj síkja tartalmazza a legtöbb pontot. Egyszerű és hatékony megoldás ebben az esetben, ha csak egy referencia síkhoz közeli pontokat vesszük számításba, és ennek a síknak a paramétereit frissítjük képkockánként.
3. A talajviszonyok meghatározása: A pontok talaj síkjától való távolságából (maximális távolság, távolságok szórása) meg tudjuk határozni, hogy sík vagy göröngyös a talaj.
4. A talaj típusának meghatározása: Az intenzitásból számított tulajdonságok alapján meg tudjuk határozni a talaj típusát. (intenzitás minimuma, intenzitás maximuma, maximum és minimum intenzitás különbsége, intenzitás átlaga, mediánja, szórása). Mivel az intenzitás értéke a beesési szögtől is függ, ezért a talaj típusának meghatározásához a beesési szöget, a lézer távolságát és az adatpontok számát is figyelembe vesszük. Ezen tulajdonságok értékei alapján az osztályozás valamely klasszifikációs módszerrel történik.

4.1.2. Talajfelismerés textúra leíró módszerekkel [14]

A prospektív talajfelismerés történhet textúra leíró módszerekkel is. Ebben az esetben a talajról készített képi információk alapján történik a talaj típusának meghatározása. Előnye, hogy ebben az esetben egy kamera is elegendő a terep vizsgálatához.

A talaj textúrájának jellemzésére számos módszer került már alkalmazásra:

- Ko-okkurencia mátrixok (2.1.3. fejezet)
- Markov-modellek
- LBP-operátorok (2.2. fejezet)

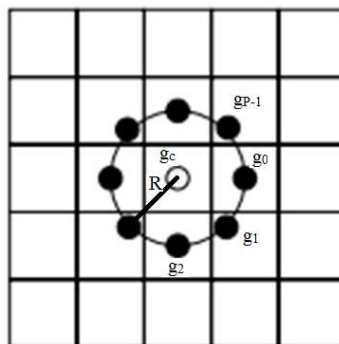
Az osztályba a kiszámolt képi információk felhasználásával valamilyen klasszifikációs módszerrel történik.

4.2. Talaj Textúrájának leírása

Az egyes talajtípusokról készített képek textúrájának leírására a circular LBP-operátort (LBP(8,1)) valamint ennek egy továbbfejlesztett változatát az uniform LBP-t használtam.

4.2.1. Circular LBP(8,1)

Az LBP(8,1) esetén 8 környező pont intenzitásértékét, amelyek egy kör mentén helyezkednek el (kör sugara egy pixeltávolság) hasonlítjuk össze a középső pixel értékével, viszont mivel nem minden pont esik a pixel középpontjába, ezért ezen pontok intenzitásértékének (g_1, g_3, g_5, g_7) meghatározására bilineáris interpolációt alkalmazunk.



16. ábra LBP(8,1) pixeleinek elhelyezkedése

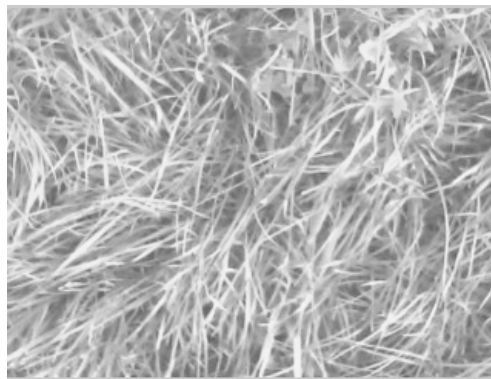
A LBP(8,1) kiszámításának a lépései az alábbiak:

- 1) A talajról készült kép beolvasása: A képek betöltésére a Matlab-ban az `imread(...)` függvény szolgál. A függvény visszatérési értéke egy képmátrix, amelynek mérete a kép típusától függ. Szürkeárnyaltos képek esetében egy $M \times N$ -es mátrixot kapunk., ahol M a vízszintes sorokban található pixelek számát, N a függőleges oszlopokban található képpontok számát jelöli. Színes képek esetében, ahol három színcsatorna van (piros, zöld, kék), ott egy $M \times N \times 3$ -as képmátrixot kapunk eredményül.



17. ábra Beolvasott színes kép

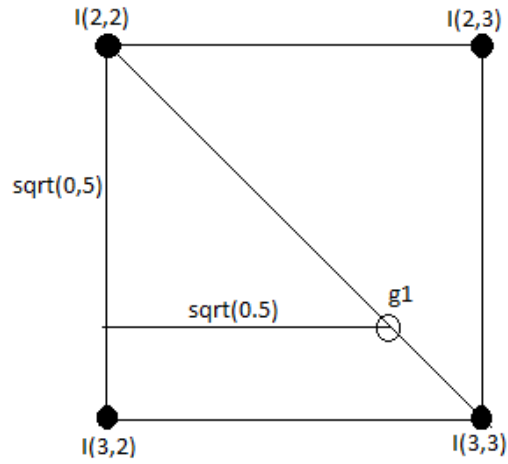
- 2) Szürkeárnyaltos kép előállítás: Ha a beolvasott kép színes volt, akkor átalakítjuk szürkeárnyaltossá (rgb2gray(...)) utasítás), mivel az LBP-operátort szürkeárnyaltos képeken fogjuk alkalmazni. A szürkeárnyaltos képeken 256 szürkeárnyalat szerepelhet, mivel 8-as bitmélységgel rendelkeznek.



18. ábra Szürkeárnyaltos kép

- 3) Az egyes pixelekhez tartozó LBP-értékek meghatározása: Az egyes pixelekhez tartozó eredményeket egy mátrixban tároljuk. Ha a szürkeárnyaltos képünk MxN-es méretű, akkor az eredmények tárolására szolgáló mátrixunk (M-2)x(N-2) méretű lesz, hiszen a 2-2 szélső sorban illetve oszlopban levő pixelekhez nem tudunk LBP-értéket számítani, mivel nem tudunk körük egy 1 pixeltávolságnyi sugarú kört rajzolni. A páros sorszámú (g_0, g_2, g_4, g_6) pixelek értékei könnyen meghatározhatóak, hiszen pixelközéppontra esnek. A páratlan sorszámú (g_1, g_3, g_5, g_7) pixelek értékének meghatározása bilineáris interpolációval történik. Például a 2. sor 2. pixeléhez tartozó g_1 értékének meghatározása az alábbiak szerint történik:

$$g_1 = I(2,2) + \sqrt{0,5}[I(2,3) - I(2,2)] + \sqrt{0,5}[I(3,2) - I(2,2)] + 0.5[I(2,2) - I(2,3) - I(3,2) + I(3,3)] \quad (4.1)$$

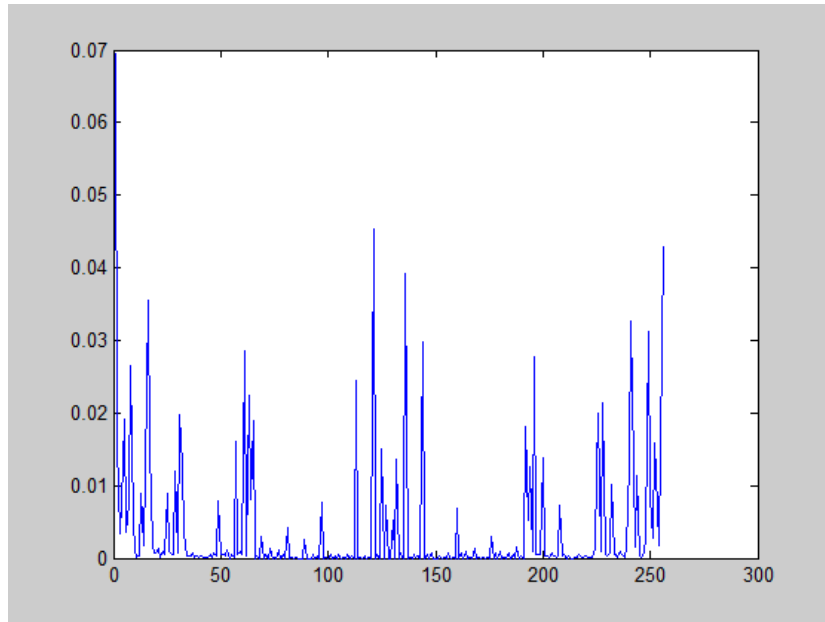


19. ábra Az $I(2,2)$ középső pixelhez tartozó g_1 értékének meghatározása

A 8 környező pont értékének kiszámítása után, összehasonlítjuk az értéküket a középső pont intenzitásával, és kiszámoljuk az adott ponthoz tartozó LBP-értéket:

$$LBP_{8,1}(x_c, y_c) = \sum_{p=0}^7 s(g_p - g_c) 2^p \quad (4.2)$$

Az LBP-mátrixot sorvektorra (256 elemű, mivel az LBP 8 bites, ezért 256 különböző értéket vehet fel) rendeztem a reshape(...) utasítással, majd a kapott eredményt normalizált LBP-hisztogramon ábrázoltam.



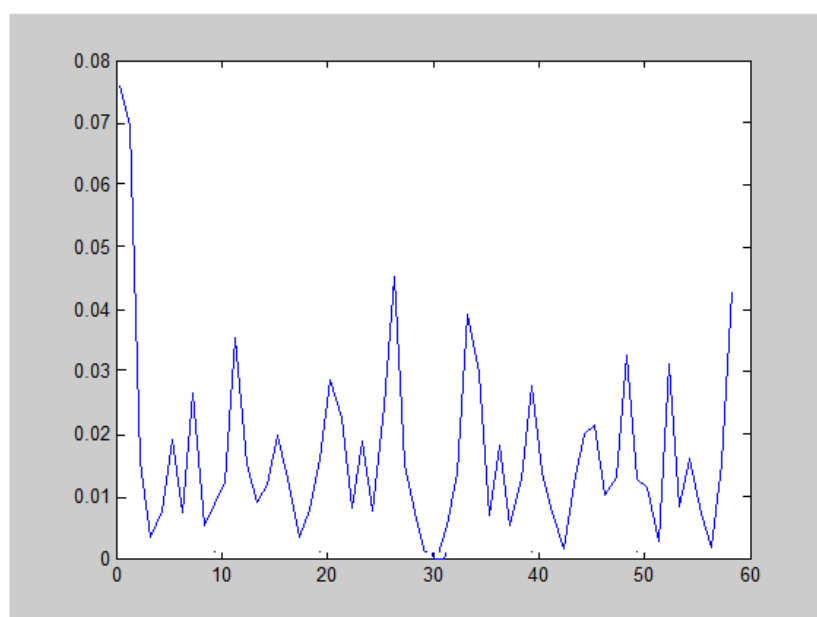
20. ábra Normalizált LBP(8,1)-hisztogram

4.2.2. Uniform LBP

Ojala 2002-ben megállapította, hogy LBP(8,1) esetén (8 a szomszédos pontok száma, a kör sugara 1 pixel hosszúságú) az előálló LBP-értékek közel 90 %-a uniform, így nem veszünk sok információt azzal, ha a nem uniform mintákat egyetlen címkével illetjük. Egy LBP-bitminta akkor uniform, ha a bitmintán körkörösén végigmenve legfeljebb 2 db 0-1 vagy 1-0 átmenetet kapunk. Pl: 00000000 (0 átmenet) 01110000 (2 átmenet) and 11001111 (2 átmenet) uniformok, de pl. 11001001 (4 átmenet) and 01010010 (6 átmenet) nem azok.

LBP(8,1) esetén a 256 lehetséges bitmintából 58 lesz uniform, így a textúra jellemzése uniform LBP esetén 59 elemű vektorral történik, szemben az LBP(8,1) esetén alkalmazott 256 elemű vektorral.

Az uniform LBP meghatározásához a már kiszámított LBP(8,1) vektorból indultam ki. Mind a 256 bitmintára meghatároztam az 0-1 és 1-0 átmenetek összegét. A 2-nél több átmenetű bitmintához tartozó minták számát összegeztem, és az 59 elemű vektor első elemében eltároltam, tehát az 59 elemű vektor első eleme tartalmazza a nem uniform mintákat. Az utána következő 58 elem tartalmazza az uniform mintákat. A vektor 2. eleme tartalmazza a 00000000 bitsorozathoz tartozó minták számát, 3. eleme 00000001 bitsorozathoz tartozó minták számát, egészen az utolsó elemig, ami 11111111 bitsorozathoz tartozó minták számát mutatja. A kapott eredményt normalizáltam, és hisztogramon ábrázoltam. Ojala megállapításához hasonlóan az eredményeimnél szintén csupán 10 % körüli az uniform minták száma.



21. ábra Normalizált uniform LBP-hisztogram

4.3. Klasszifikációs eljárások

A tesztadatok osztályozására 2 klasszifikációs eljárást próbáltam ki, az SVM-mel illetve előrecsatolt neurális hálózattal történő osztályozást. Ebben a fejezetben áttekintem ezen módszerek paramétereinek beállítását, valamint az osztályozás során kapott eredményeket.

4.3.1. Tanító pontok előállítása

Mindhárom osztály (fű, homok, aszfalt) 25-25 db tanító képet tartalmaz. Mindegyik képhez kiszámoltam a 256 elemű normalizált LBP(8,1) vektorokat, illetve 59 elemű normalizált uniform LBP vektorokat, és egy-egy mátrixban eltároltam, így egy 75x256 valamint egy 75x59 méretű mátrixot kaptam, amiket elmentettem egy-egy xls kiterjesztésű fájlba.



22. ábra Képek a tanító halmazból

4.3.2. Osztályozás többosztályú SVM-mel

A tesztképek osztályba sorolására többosztályú SVM módszert kell alkalmazni, mivel 3 osztályunk van. A fejezetben ismertetett módszerek közül az „egy a többi ellen” módszert alkalmaztam.

Mindhárom osztályhoz tartozik egy-egy bináris osztályozó, amely a tesztképet az adott osztályba sorolja be vagy az összes többi osztályba. (Például a fűhöz tartozó osztályozó esetén a fű alkotja az egyik osztályt, a másik osztályt az aszfalt és a homok együttesen alkotják.) A tesztadat abba az osztályba fog tartozni, amelyik a legtöbb szavazatot kapta. A módszer hátránya, hogy abban az esetben, ha több osztály is azonos számú szavazatot kapott, akkor nem tudjuk a tesztadatot osztályba sorolni.

SMO algoritmus [16]

A szupport vektor gépek gyakorlati felhasználásának gondot okoz az SVM tanításához szükséges kvadratikus programozás nagy algoritmikus komplexitása, ezért kidolgoztak olyan módszereket, amelyekkel a tanítási időt csökkenteni lehet. Ezen módszerek közül az SVM tanítására az SMO (Sequential Minimal Optimization) algoritmust használtam. Az SMO egy egyszerű, hatékony iterációs módszer, amely gyorsan meg tudja oldani a SVM kvadratikus programozási feladatát. Az SMO a korábbi módszerekkel ellentétben a lehető legkisebb méretű optimalizálási problémát oldja meg minden lépésben. Az SVM esetén ez 2 Lagrange-szorzó optimalizálását jelenti lépésenként, mivel a Lagrange szorzóknak teljesíteniük kell egy lineáris egyenlőség feltételt (2.30). Az SMO minden lépésben kiválaszt 2 Lagrange-szorzót, amelyeket optimalizál, majd frissíti az SVM-et az új optimalizált értékeknek megfelelően. A következő 2 optimalizálandó változó kiválasztása heurisztikusan történik. A módszer előnye, hogy teljesen kikerüli a kvadratikus programozás számítását, mivel a 2 Lagrange-szorzó analitikusan is optimalizálható.

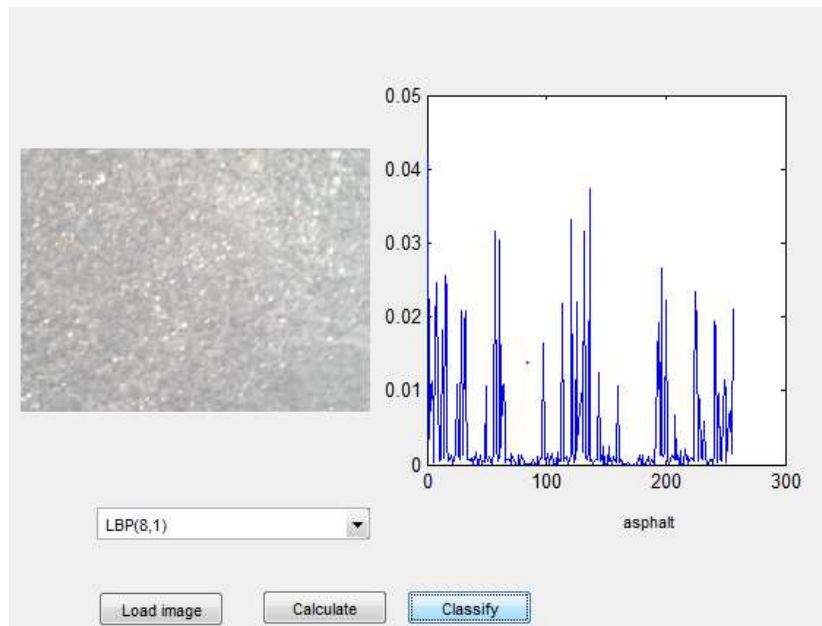
A tesztképek osztályozásához készítettem egy grafikus felhasználói felületet (GUI) a Matlab-ban. A 22. ábrán láthatjuk a program kinézetét.

A 'Load image' nyomógomb megnyomásával egy, a számítógépünkön mentett tesztképet tudunk beolvasni, amely a jobb oldali mezőben meg fog jelenni.

A legördülő menüből ki tudjuk választani, hogy melyik LBP-operátort használjuk majd az osztályozásra. A 2 választási lehetőségünk az LBP(8,1) valamint az uniform LBP.

A 'Calculate' nyomógomb megnyomásával a program kiszámítja a beolvasott képhez tartozó LBP-vektor értékét, és a bal oldali mezőben normalizált hisztogramon ábrázolja.

A 'Classify' nyomógommbal végezhetjük el a tesztképünk osztályozását. Az `xlsread(...)` utasítással beolvassuk az adott LBP-operátorhoz tartozó 75 elemű tanító adathalmazt, amit korábban egy xls fájlban tároltunk el. A tesztkép osztályozásához szükség van a tesztminták osztályainak beolvasására is, amit egy 75 elemű oszlopvektorban tároltam. Az osztályozásnál 4 eset lehetséges (fű, homok, aszfalt, szavazás döntetlen eredménye miatt nem eldönthető). Az osztályozással kapott eredményt a normalizált hisztogram alatt elhelyezkedő statikus szövegdoboz jeleníti meg.



23. ábra Az SVM osztályozó program kinézete

75 tesztmintán (25 fű, 25 homok, 25 aszfalt) (24. ábra) megvizsgáltam az osztályozó pontosságát uniform LBP használata esetén. A kapott eredményeket a 2. táblázatban rögzítettem. Az eredményeket összehasonlítva látható, hogy puha margójú SVM alkalmazásával nagyobb pontosságot lehet elérni. A puha margójú SVM abban különbözik az eredeti SVM-től, hogy az osztályozás feltételeit kiegészítő változókkal gyengítjük, így megengedjük, hogy néhány tanítópont a margón belülre kerüljön. A felhasználó által beállítandó C regularizációs paraméterrel a kiegészítő változók súlyát adjuk meg. Az osztályozás során több regularizációs paramétert is kipróbáltam, és a regularizációs paraméter 0,5 értékénél kaptam a legnagyobb pontosságot.



24. ábra Néhány kép a tesztképek közül

Osztályozás pontossága	Többosztályú SVM	Többosztályú puha margójú SVM C=0,05	Többosztályú puha margójú SVM C=0,5
fű	96%	96%	96%
homok	80%	84%	88%
aszfalt	88%	88%	92%

2. táblázat Osztályozás pontossága többosztályú SVM esetén

4.3.3. Osztályozás előrecsatolt neurális hálózattal

A tesztképek osztályba sorolása előrecsatolt neurális hálózattal is megvalósítható. Uniform LBP használata esetén a bemeneti vektor 59 elemű lesz, a kimeneti vektor pedig 3 elemű, hiszen 3 osztályunk van.

A tanítási halmazunk 75 mintából áll (4.2.1. fejezet).

Hálózat tanítása

A hálózat tanítása történhet egyenként (szekvenciális tanítás) vagy epochonként (köteget (batch) tanítás). Soros tanítás esetén a súlymódosításokat tanító pontonként, köteget tanítás esetén a teljes tanító készlet (epoch) felhasználása után végezzük. Az első esetben tehát a súlymódosítások száma megegyezik a tanító pontok számának és a tanítási ciklusok számának szorzatával, míg a batch eljárásnál csak ciklusonként korrigálunk, ahol egy ciklus a teljes tanító készlet egyszeri felhasználását jelenti. A szekvenciális tanítás előnye, hogy kevesebb tárolási helyet igényel, valamint könnyű implementálni.

A feladat során a szekvenciális tanítást használtam, a hálózat hibájának számítására az átlagos négyzetes hibát választottam.

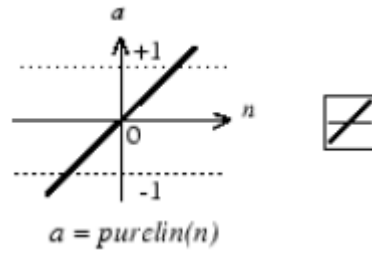
Aktivációs függvény kiválasztása

A neurális hálózat teljesítőképessége szempontjából fontos, hogy megfelelő aktivációs függvényt válasszunk ki. Általában gyorsabb a tanulás, ha az aktivációs függvény antiszimmetrikus. ($\varphi(-n) = -\varphi(n)$)

A Matlab programban az alábbi aktivációs függvények közül lehet választani:

- Lineáris aktivációs függvény

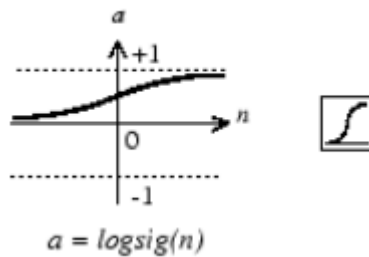
$$\varphi(n) = n \quad (4.3)$$



25. ábra Lineáris aktivációs függvény

- Logisztikus aktivációs függvény

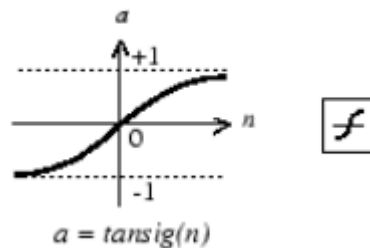
$$\varphi(n) = \frac{2}{1 + e^{-n}} \quad (4.4)$$



26. ábra Logisztikus aktivációs függvény

- Tangens hiperbolikus aktivációs függvény

$$\varphi(n) = \frac{2}{1 + e^{-2n}} - 1 \quad (4.5)$$



27. ábra Tangens hiperbolikus aktivációs függvény

A 3. táblázatban a különböző aktivációs függvényekkel kapott eredményeket ismertetem. Azt vizsgáltam meg, hogy 0,1 értékű tanulási paraméter mellett 1000 epoch tanulási ciklus alatt milyen értékre csökken le az átlagos négyzetes hiba backpropagation tanulási algoritmus alkalmazásával. Az átlagos négyzetes hiba meghatározásánál 10 futtatás eredményének átlagát vettem. Az eredményekből jól látszik, hogy a leggyorsabb tanítást a rejtett réteg(ek)ben tangens hiperbolikus aktivációs függvény alkalmazásával lehet elérni. A

kimeneti réteg(ek)ben lineáris, illetve tangens hiperbolikus aktivációs függvény is jó megoldásnak mutatkozik. A kapott eredményeket figyelembe véve a rejtett réteg(ek)ben tangens hiperbolikus, a kimeneti rétegben lineáris aktivációs függvényt választottam.

	Rejtett réteg	Kimeneti réteg	Rejtett réteg	Kimeneti réteg	Rejtett réteg	Kimeneti réteg	Rejtett réteg	Kimeneti réteg	Rejtett réteg	Kimeneti réteg	Rejtett réteg	Kimeneti réteg
rejtett réteg neuronjainak száma	Logisztikus	Lineáris	Logisztikus	Logisztikus	Logisztikus	tangens hip.	tangens hip.	Lineáris	Tangens hip.	Logisztikus	Tangens hip.	Tangens hip.
1	0,1919		0,2731		0,2492		0,1514		0,277		0,21	
3	0,1716		0,2597		0,1844		0,0823		0,2434		0,1001	
5	0,16		0,2567		0,1518		0,073		0,2442		0,0672	
7	0,1477		0,249		0,1225		0,0678		0,2409		0,063	
10	0,1278		0,248		0,1189		0,0635		0,2432		0,0604	
5+5	0,1538		0,2561		0,1453		0,0513		0,2425		0,0558	
10+10	0,1242		0,2495		0,1239		0,0452		0,2387		0,0353	

3. táblázat a hálózat átlagos négyzetes hibájának értékei különböző aktivációs függvényekkel 1000 epoch tanulási ciklus után (tanulási paraméter 0,1)

A neurális hálózat ideális méretének meghatározása

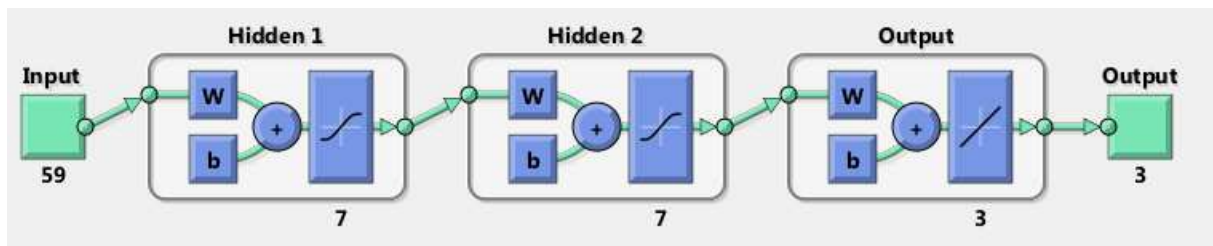
A neurális hálózat ideális méretének meghatározására még nem dolgoztak ki általános módszert, annak meghatározása próbálgatásos úton lehetséges. Általánosan elmondható, hogy több réteg alkalmazása könnyítheti a feladat megoldását, illetve rétegenként kevesebb neuron is elegendő lehet. Kétféle módszer terjedt el a hálózat méretének meghatározására. Az első szerint egy nagyobb hálózatból kiindulva próbáljuk meg csökkenteni annak méretét, a másik szerint egy kisebb hálózat bővítésével próbálunk eljutni a megfelelő megoldásig.

A neurális hálózat rejtett rétegeinek, illetve neuronjainak ideális számának meghatározásához a 75 elemű tanító adathalmazt 2 részre osztottam: 80 %-a a hálózat tanítására szolgál, 20 %-a validációs minta. A túltanulás elkerülésére a hálózat tanítása le fog állni, ha a validációs minta hibája 10 egymást követő epochban növekedni fog. Megvizsgáltam, hogy a különböző hálózati struktúrákkal (rejtett rétegek száma, neuronok száma) mekkora értékűre lehet csökkenteni a validációs hibát (4. táblázat). A legkisebb validációs hibát 7-7 neuronból álló 2 rejtett rétegű neurális hálózathoz kaptam.

Rejtett réteg neuronjainak száma	Validációs hiba
1	0,1602
3	0,1216
5	0,0891
7	0,0787
10	0,0601
5+5	0,0510
7+7	0,0480
10+10	0,0514
15+15	0,0642

4. táblázat a validációs hiba értéke az egyes hálózati struktúráknál

A választott struktúrájú neurális hálózat 59 bemenettel és 3 kimenettel rendelkezik. 2 rejtett rétegében egyaránt 7-7 neuron található, a kimeneti rétegben pedig 3 neuron. A rejtett réteg neuronjainak tangens hiperbolikus aktivációs függvényük, a kimeneti réteg neuronjainak lineáris aktivációs függvényük van.



28. ábra A választott neurális hálózat struktúrája

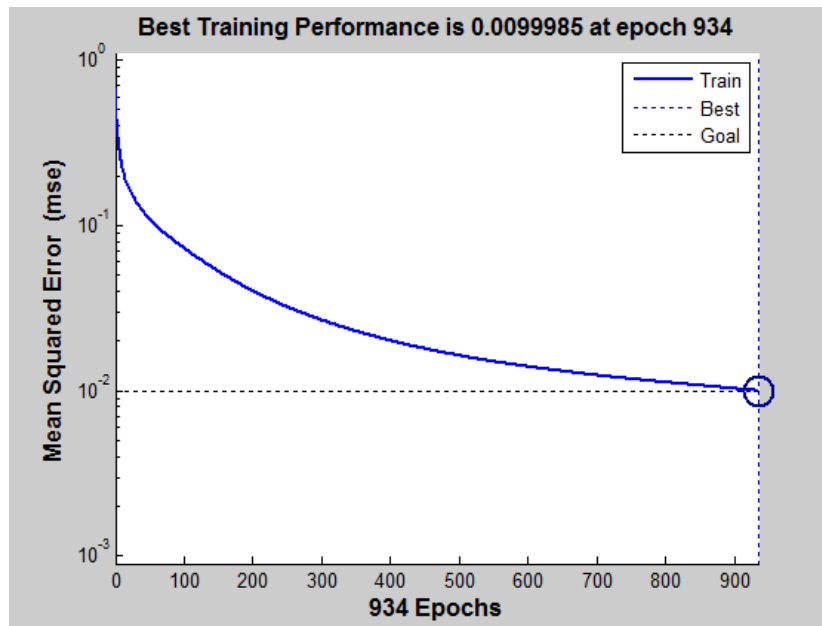
Tanulási algoritmusok összehasonlítása

A 7-7 neuronból álló 2 rejtett rétegű neurális hálózat tanítását többféle tanulási algoritmussal elvégeztem. Megvizsgáltam, hogy a különböző tanulási algoritmusokkal mennyi idő és hány epochszám alatt lehet elérni 0,01 értékű átlagos négyzetes hibát (5. táblázat).

Tanulási algoritmus	epochszám	Tanulási idő
Backpropagation algoritmus tanulási paraméter 0,1	896,6	9,9896
Backpropagation algoritmus tanulási paraméter 0,15	687,2	8,472
Backpropagation algoritmus tanulási paraméter 0,5	219,2	2,4778
Backpropagation algoritmus momentummal tanulási paraméter 1 momentum konstans 0,7	185,2	2,085
Levenberg-Marquardt algoritmus	7	8,9872
Fletcher-Reeves konjugált gradiens módszer	44	1,9586

5. táblázat a tanulási algoritmusok eredményei

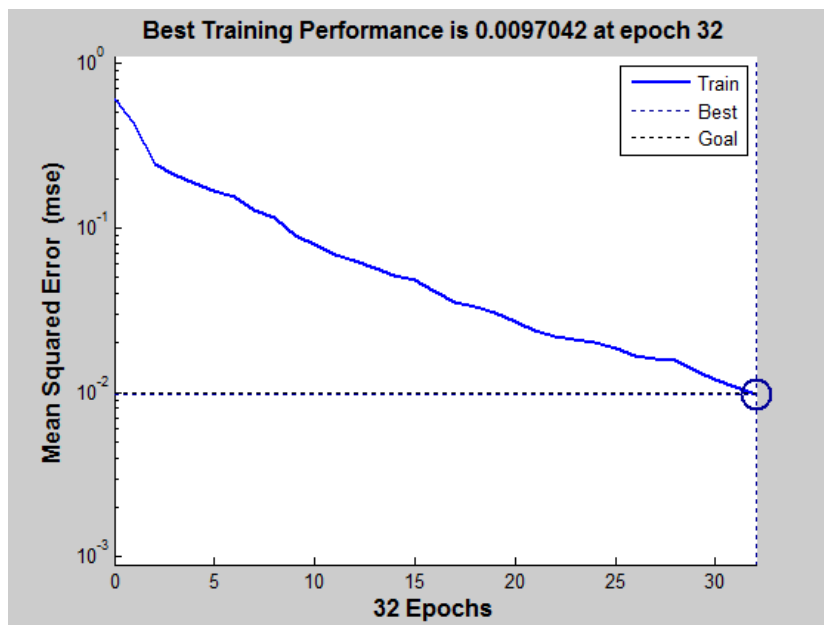
A backpropagation algoritmusnál a tanulási paraméter értéke nagy mértékben befolyásolja az algoritmus sebességét. Ha a tanulási paraméter optimális értékét próbálgatással lehet meghatározni. Ha a tanulási paraméter túl kicsire választjuk, akkor a súlyok kis mértékben fognak változni, lassú lesz a konvergencia. Nagy tanulási paraméter esetén gyorsabb lesz a konvergencia, de instabilitás, oszcillációs léphet fel. Momentum konstans hozzáadásával növelni lehet a tanulás gyorsaságát, mivel a momentum aluláteresztő szűrőként viselkedve csökkenni az oszcillációt, így nagyobb tanulási paraméterekre is stabilitást biztosít. Az eredményekből látható, hogy a backpropagation algoritmus nagy epochszám után éri el a 0,01 értékű átlagos négyzetes hibát, hiszen a hibafüggvény elsőrendű deriváltjait használja a hálózat súlyainak beállítására.(29. Ábra)



29. ábra Átlagos négyzetes hiba alakulása backpropagation algoritmussal (tanulási paraméter 0,1)

A Levenberg-Marquardt algoritmusnál lehet a legkisebb epochszám alatt elérni a 0,01 értékű átlagos négyzetes hibát. Ez annak köszönhető, hogy ez az algoritmus a másodrendű deriváltakat használja a tanítás során. A hálózat tanulási ideje magas, mivel az algoritmus során számítandó Jacobi-mátrix nagy, 75x500-as méretű. (75 a tanító minták száma, 500 a hálózat súlyainak száma)

A konjugált gradiens módszer a backpropagation algoritmushoz hasonlóan elsőrendű gradiens módszer, viszont a keresés nem a negatív gradiens irányában történik, hanem a konjugált gradiens módszer szerint meghatározott irányokban. A backpropagation algoritmushoz képest a tanulás gyorsabb lesz, valamint kisebb epochszám alatt éri el a hibafüggvény megfelelő mértékű csökkenését (30. ábra). A vonalmenti minimumkeresésre az aranymetszés keresést választottam. Ez az eljárás egyike a leggyakrabban használtaknak. Előnye, hogy minden lépésben csak egy olyan új pont adódik, ahol ki kell kiszámolni a függvényértéket.



30. ábra Átlagos négyzetes hiba alakulása Fletcher-Reeves konjugált gradiens módszer

A neurális hálózat osztályozásának pontossága

A hálózat tanítására a 75 képből álló tanító mintahalmazt használtam fel. A tanító halmazt 2 részre osztottam, 80%-át a hálózat tanítására, 20 %-át validációra használtam. A hálózat tanítása le fog állni, ha a validációs minta hibája 10 egymást követő epochban növekedni fog. A hálózat tanítására Fletcher-Reeves konjugált gradiens módszert alkalmaztam, mert ez biztosította a leggyorsabb tanulást.

Az osztályozás pontosságát az SVM-nél is használt 75 tesztkép segítségével határoztam meg. A kapott eredményeket a 6. táblázat tartalmazza, amiket 10 tanítás során kapott eredmények átlagaként számítottam.

Osztályozás pontossága	59-7-7-3 struktúrájú neurális hálózat
fű	96,4%
homok	83,6%
aszfalt	87,2%

6. táblázat az osztályozás pontossága 59-7-7-3 struktúrájú neurális hálózattal

5. Összefoglalás, továbbfejlesztési lehetőségek

Munkám során betekintést nyertem a képfeldolgozás egy fontos területébe, a talajfelismerésbe. A talajvizsgáló rendszer tervezését egy alapos irodalomkutatás előzte meg, áttekintettem a képi információ kinyerésére használható módszereket, valamint az osztályozásra használható eljárásokat. Dolgozatom első fele ezen módszerek elméleti bemutatásával foglalkozik. Ezt követően munkámat a szakirodalomban található talajfelismerő rendszerek vizsgálatával folytattam. Végezetül egy saját talajfelismerő rendszert terveztem, amelynek programrészleteit a Matlab-ban készítettem el.

Az elkészült alkalmazás továbbfejlesztésére számos lehetőség kínálkozik:

- A jelenlegi rendszer 3 talajtípus azonosítására képes. Az osztályok számát növelve a rendszert alkalmassá tehetjük további talajtípusok felismerésére is.
- Egyéb textúra leíró módszerek, valamint klasszifikációs módszerek alkalmazásával is elvégezhetjük az osztályozást, hiszen előfordulhat, hogy más módszerekkel még jobb felismerési arányt lehet elérni.

Köszönetnyilvánítás

Szeretnék köszönetet mondani belső konzulensemnek, Tormási Alexnek, hogy hasznos tanácsokkal és észrevételekkel segítette munkámat.

Külön köszönöm szüleimnek, hogy minden körülmények között mellettem álltak és segítették tanulmányaimat.

Irodalomjegyzék

- [1] Fazekas Gábor, Hajdu András: Képfeldolgozási módszerek egyetemi jegyzet, Debreceni Egyetem, Informatikai Intézet, 2004, 2-4. old.
- [2] P. Mohanaiah, P. Sathyanarayana, L. GuruKumar: Image Texture Feature Extraction Using GLCM Approach, International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013
- [3] R. Jain, R. Kasturi, B. G. Schunck: Machine Vision, McGraw-Hill, 1995, 234-239. old.
- [4] M. Pietikäinen, A. Hadid, G. Zhao, T. Ahonen: Computer Vision Using Local Binary Patterns, Springer-Verlag, 2011 3-19. old, 32-33. old.
- [5] Random Forest http://en.wikipedia.org/wiki/Random_forest letöltés ideje: 2014.05.10.
- [6] Dobó András: Mesterséges Intelligencia I. gyakorlat http://www.inf.u-szeged.hu/~dobo/MI1_gyak_jegyzet_2013-2014_I_felev.pdf letöltés ideje: 2014.05.11.
- [7] Dr. Bodon Ferenc: Adatbányászati algoritmusok, elektronikus jegyzet 116-118. old.
- [8] SVM https://dms.sztaki.hu/sites/dms.sztaki.hu/files/file/2013/svm_ea.pdf letöltés ideje: 2014.05.12.
- [9] Ispány Márton, Jeszenszky Péter: Tartóvektorgépek <http://www.inf.unideb.hu/valseg/dolgozok/ispány/DataMine/Alkalmazasok/svm.pdf> letöltés ideje: 2014.05.12.
- [10] Dr. Nagy Tamás: Lagrange dualitás, Miskolci Egyetem, Alkalmazott Matematikai Tanszék <http://www.uni-miskolc.hu/~matente/oktatasi%20tananyagok/LAGRANGE%20DUALITAS.pdf> letöltés ideje: 2014.05.13.
- [11] Fazekas István: Neurális hálózatok, Debreceni Egyetem, Informatikai Kar, 2013, 15-44. old.
- [12] M. T. Hagan, H. B. Demuth, M. Beale: Neural Network Design, PWS Publishing Company, 1996 Chapter 2, 4, 11, 12
- [13] S. Laible, Y. N. Khan, K. Bohlmann, A. Zell: 3D LIDAR- and Camera-Based Terrain Classification Under Different Lighting Conditions, Chair of Cognitive Systems, Computer Science Department, University of Tübingen
- [14] Y. N. Khan, F. Komma, K. Bohlmann, A. Zell: Grid-based Terrain Classification for Outdoor Robots using Local Features
- [15] M. Demuth, M. Beale: Neural Network Toolbox for Use with Matlab 5.14-5.30

[16] John C. Platt: Fast Training of Support Vector Machines using Sequential Minimal Optimization, 2000

Mellékletek

I. számú melléklet: Uniform LBP kódja:

```
I=imread('E:/kep1.jpg');
m=size(I,1);
n=size(I,2);
I2= double(rgb2gray(I));

for i=2:m-1
    for j=2:n-1
        J0=I2(i,j);                %gc
        d0=I2(i,j+1)>J0;            %g0
        dx1=I2(i,j+1)-I2(i,j);
        dy1=I2(i+1,j)-I2(i,j);
        dxy1=I2(i,j)-I2(i,j+1)-I2(i+1,j)+I2(i+1,j+1);
        d1=(J0+0.5*sqrt(2)*dx1+0.5*sqrt(2)*dy1+0.5*dxy1)>J0;    %g1
        d2=I2(i+1,j)>J0;            %g2
        dx3=I2(i,j)-I2(i,j-1);
        dy3=I2(i+1,j-1)-I2(i,j-1);
        dxy3=I2(i,j-1)-I2(i,j)-I2(i+1,j-1)+I2(i+1,j);
        d3=(I2(i,j-1)+(1-0.5*sqrt(2))*dx3+0.5*sqrt(2)*dy3+(1-
0.5*sqrt(2))*0.5*sqrt(2)*dxy3)>J0;    %g3
        d4=I2(i,j-1)>J0;            %g4
        dx5=I2(i-1,j)-I2(i-1,j-1);
        dy5=I2(i,j-1)-I2(i-1,j-1);
        dxy5=I2(i-1,j-1)-I2(i-1,j)-I2(i,j-1)+I2(i,j);
        d5=(I2(i-1,j-1)+(1-0.5*sqrt(2))*dx5+(1-0.5*sqrt(2))*dy5+(1-0.5*sqrt(2))*
0.5*sqrt(2))*dxy5)>J0;    %g5
        d6=I2(i-1,j)>J0;
        dx7=I2(i-1,j+1)-I2(i-1,j);
        dy7=I2(i,j)-I2(i-1,j);
        dxy7=I2(i-1,j)-I2(i-1,j+1)-I2(i,j)+I2(i,j+1);    %g6
        d7=(I2(i-1,j)+0.5*sqrt(2)*dx7+(1-0.5*sqrt(2))*dy7+0.5*sqrt(2)*dxy7)>J0;    %g7
```

```

LBP(i-1,j-1)=d0*2^0+d1*2^1+d2*2^2+d3*2^3+d4*2^4+d5*2^5+d6*2^6+d7*2^7;

end

end

LBP=uint8(LBP);
LBPv=reshape(LBP,1,size(LBP,1)*size(LBP,2));
Hist=hist(LBPv,0:255);
imshow(LBP);
plot(Hist);

%table létrehozása
table = zeros(1, 256);

nextLabel = 1;
for k = 1:256
    bits=de2bi(k-1,8);
    changes=sum(abs(diff(bits)));
    if bits(8)-bits(1)~=0
        changes=changes+1;
    end
    if changes <= 2
        table(k) = nextLabel;
        nextLabel = nextLabel + 1;

    else
        table(k) = 0;

    end
end
end

```

```
%LBP-ból LBP uniform

for i=1:1:256
    x=table(i);
    if x==0
        LBPu(1)=LBPu(1)+Hist(i);
    else
        LBPu(x+1)=Hist(i);
    end
end
end
```

II. számú melléklet: SVM osztályozó GUI kódja:

```
function varargout = LBPhistogram1(varargin)
%LBPHISTOGRAM1 M-file for LBPhistogram1.fig
%   LBPHISTOGRAM1, by itself, creates a new LBPHISTOGRAM1 or raises the existing
%   singleton*.
%
%   H = LBPHISTOGRAM1 returns the handle to a new LBPHISTOGRAM1 or the handle
%   to
%   the existing singleton*.
%
%   LBPHISTOGRAM1('Property','Value',...) creates a new LBPHISTOGRAM1 using the
%   given property value pairs. Unrecognized properties are passed via
%   varargin to LBPhistogram1_OpeningFcn. This calling syntax produces a
%   warning when there is an existing singleton*.
%
%   LBPHISTOGRAM1('CALLBACK') and
%   LBPHISTOGRAM1('CALLBACK',hObject,...) call the
%   local function named CALLBACK in LBPHISTOGRAM1.M with the given input
%   arguments.
%
%   *See GUI Options on GUIDE's Tools menu. Choose "GUI allows only one
%   instance to run (singleton)".
%
% See also: GUIDE, GUIDATA, GUIHANDLES

% Edit the above text to modify the response to help LBPhistogram1

% Last Modified by GUIDE v2.5 02-Nov-2014 09:37:05

% Begin initialization code - DO NOT EDIT
gui_Singleton = 1;
gui_State = struct('gui_Name',    mfilename, ...
    'gui_Singleton', gui_Singleton, ...
    'gui_OpeningFcn', @LBPhistogram1_OpeningFcn, ...
    'gui_OutputFcn', @LBPhistogram1_OutputFcn, ...
    'gui_LayoutFcn', [], ...
    'gui_Callback', []);
if nargin && ischar(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end

if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:});
else
    gui_mainfcn(gui_State, varargin{:});
end
% End initialization code - DO NOT EDIT
```

```

% --- Executes just before LBPhistogram1 is made visible.
function LBPhistogram1_OpeningFcn(hObject, eventdata, handles, varargin)
% This function has no output args, see OutputFcn.
% hObject    handle to figure
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
% varargin   unrecognized PropertyName/PropertyValue pairs from the
%            command line (see VARARGIN)

% Choose default command line output for LBPhistogram1
handles.output = hObject;

% Update handles structure
guidata(hObject, handles);

% UIWAIT makes LBPhistogram1 wait for user response (see UIRESUME)
% uiwait(handles.figure1);

% --- Outputs from this function are returned to the command line.
function varargout = LBPhistogram1_OutputFcn(hObject, eventdata, handles)
% varargout  cell array for returning output args (see VARARGOUT);
% hObject    handle to figure
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Get default command line output from handles structure
varargout{1} = handles.output;

% --- Executes on button press in pushbutton1.
function pushbutton1_Callback(hObject, eventdata, handles)
handles.output = hObject;
[fn pn] = uigetfile('*.jpg','select jpg file');
complete = strcat(pn,fn);
set(handles.edit1,'string',complete);
I = imread(complete);
axes(handles.axes1);
imshow(I,[]);
guidata(hObject, handles);
set(hObject,'UserData',I);

% --- Executes on button press in pushbutton2.
function pushbutton2_Callback(hObject, eventdata, handles)
h = findobj('Tag','pushbutton1');
I= get(h,'UserData');

m=size(I,1);
n=size(I,2);

```

```

I2= double(rgb2gray(I));

switch get(handles.popupmenu1,'Value')
case 1

for i=2:m-1
    for j=2:n-1
        J0=I2(i,j); %gc
        d0=I2(i,j+1)>J0; %g0
        dx1=I2(i,j+1)-I2(i,j);
        dy1=I2(i+1,j)-I2(i,j);
        dxy1=I2(i,j)-I2(i,j+1)-I2(i+1,j)+I2(i+1,j+1);
        d1=(J0+0.5*sqrt(2)*dx1+0.5*sqrt(2)*dy1+0.5*dxy1)>J0; %g1
        d2=I2(i+1,j)>J0; %g2
        dx3=I2(i,j)-I2(i,j-1);
        dy3=I2(i+1,j-1)-I2(i,j-1);
        dxy3=I2(i,j-1)-I2(i,j)-I2(i+1,j-1)+I2(i+1,j);
        d3=(I2(i,j-1)+(1-0.5*sqrt(2))*dx3+0.5*sqrt(2)*dy3+(1-0.5*sqrt(2))*0.5*sqrt(2)*dxy3)>J0; %g3
        d4=I2(i,j-1)>J0; %g4
        dx5=I2(i-1,j)-I2(i-1,j-1);
        dy5=I2(i,j-1)-I2(i-1,j-1);
        dxy5=I2(i-1,j-1)-I2(i-1,j)-I2(i,j-1)+I2(i,j);
        d5=(I2(i-1,j-1)+(1-0.5*sqrt(2))*dx5+(1-0.5*sqrt(2))*dy5+(1-0.5*sqrt(2))*0.5*sqrt(2)*dxy5)>J0; %g5
        d6=I2(i-1,j)>J0;
        dx7=I2(i-1,j+1)-I2(i-1,j);
        dy7=I2(i,j)-I2(i-1,j);
        dxy7=I2(i-1,j)-I2(i-1,j+1)-I2(i,j)+I2(i,j+1); %g6
        d7=(I2(i-1,j)+0.5*sqrt(2)*dx7+(1-0.5*sqrt(2))*dy7+0.5*sqrt(2)*dxy7)>J0; %g7
        LBP(i-1,j-1)=d0*2^0+d1*2^1+d2*2^2+d3*2^3+d4*2^4+d5*2^5+d6*2^6+d7*2^7;

    end
end
LBP=uint8(LBP);
LBPv=reshape(LBP,1,size(LBP,1)*size(LBP,2));
Hist=hist(LBPv,0:255);
Hist=Hist./sum(Hist);
axes(handles.axes2);
plot(Hist);
set(hObject,'UserData',Hist);

case 2

for i=2:m-1
    for j=2:n-1

```

```

J0=I2(i,j); %gc
d0=I2(i,j+1)>J0; %g0
dx1=I2(i,j+1)-I2(i,j);
dy1=I2(i+1,j)-I2(i,j);
dxy1=I2(i,j)-I2(i,j+1)-I2(i+1,j)+I2(i+1,j+1);
d1=(J0+0.5*sqrt(2)*dx1+0.5*sqrt(2)*dy1+0.5*dxy1)>J0; %g1
d2=I2(i+1,j)>J0; %g2
dx3=I2(i,j)-I2(i,j-1);
dy3=I2(i+1,j-1)-I2(i,j-1);
dxy3=I2(i,j-1)-I2(i,j)-I2(i+1,j-1)+I2(i+1,j);
d3=(I2(i,j-1)+(1-0.5*sqrt(2))*dx3+0.5*sqrt(2)*dy3+(1-0.5*sqrt(2))*0.5*sqrt(2)*dxy3)>J0;
%g3
d4=I2(i,j-1)>J0; %g4
dx5=I2(i-1,j)-I2(i-1,j-1);
dy5=I2(i,j-1)-I2(i-1,j-1);
dxy5=I2(i-1,j-1)-I2(i-1,j)-I2(i,j-1)+I2(i,j);
d5=(I2(i-1,j-1)+(1-0.5*sqrt(2))*dx5+(1-0.5*sqrt(2))*dy5+(1-0.5*sqrt(2))*0.5*sqrt(2)*dxy5)>J0; %g5
d6=I2(i-1,j)>J0;
dx7=I2(i-1,j+1)-I2(i-1,j);
dy7=I2(i,j)-I2(i-1,j);
dxy7=I2(i-1,j)-I2(i-1,j+1)-I2(i,j)+I2(i,j+1);
%g6
d7=(I2(i-1,j)+0.5*sqrt(2)*dx7+(1-0.5*sqrt(2))*dy7+0.5*sqrt(2)*dxy7)>J0;
%g7
LBP(i-1,j-1)=d0*2^0+d1*2^1+d2*2^2+d3*2^3+d4*2^4+d5*2^5+d6*2^6+d7*2^7;

end
end
LBP=uint8(LBP);
LBPv=reshape(LBP,1,size(LBP,1)*size(LBP,2));
Hist=hist(LBPv,0:255);

%table létrehozása
table = zeros(1, 256);
ch=zeros(1,256);
nextLabel = 1;
for k = 1:256
    bits=de2bi(k-1,8);
    changes=sum(abs(diff(bits)));
    if bits(8)-bits(1)~=0
        changes=changes+1;
    end
    if changes <= 2
        table(k) = nextLabel;
        nextLabel = nextLabel + 1;
        ch(k)=changes;
    else
        table(k) = 0;
        ch(k)=changes;
    end
end

```

```

        end
    end
    %LBP-böl LBP uniform

    for i=1:1:256
        x=table(i);
        if x==0
            LBPu(1)=LBPu(1)+Hist(i);
        else
            LBPu(x+1)=Hist(i);
        end
    end
end

axes(handles.axes2);
LBPu=LBPu./sum(LBPu);
set(hObject,'UserData',LBPu);
plot(LBPu);

end

% --- Executes on selection change in popupmenu1.
function popupmenu1_Callback(hObject, eventdata, handles)
% hObject    handle to popupmenu1 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hints: contents = cellstr(get(hObject,'String')) returns popupmenu1 contents as cell array
%        contents{get(hObject,'Value')} returns selected item from popupmenu1

% --- Executes during object creation, after setting all properties.
function popupmenu1_CreateFcn(hObject, eventdata, handles)
% hObject    handle to popupmenu1 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called

% Hint: popupmenu controls usually have a white background on Windows.
%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function edit1_Callback(hObject, eventdata, handles)
% hObject    handle to edit1 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

```

```
% Hints: get(hObject,'String') returns contents of edit1 as text
%      str2double(get(hObject,'String')) returns contents of edit1 as a double
```

```
% --- Executes during object creation, after setting all properties.
```

```
function edit1_CreateFcn(hObject, eventdata, handles)
% hObject    handle to edit1 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called
```

```
% Hint: edit controls usually have a white background on Windows.
```

```
%      See ISPC and COMPUTER.
```

```
if ispc && isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end
```

```
% --- Executes on button press in pushbutton3.
```

```
function pushbutton3_Callback(hObject, eventdata, handles)
% hObject    handle to pushbutton3 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
t= cputime;
```

```
switch get(handles.popupmenu1,'Value')
case 1
    TrainingSet = xlsread('E:/circularLBP256norm.xls');
    h = findobj('Tag','pushbutton2');
    Hist= get(h,'UserData');
    TestSet=zeros(1,256);
    TestSet(1,:)=Hist ;
    GroupTrain= xlsread('E:/GroupTrain.xls');
    u=unique(GroupTrain);
    numClasses=length(u);
    result = zeros(length(TestSet(:,1)),1);
```

```
%build models
```

```
for k=1:numClasses
    %Vectorized statement that binarizes Group
    %where 1 is the current class and 0 is all other classes
    G1vAll=(GroupTrain==u(k));
    models(k) = svmtrain(TrainingSet,G1vAll,'boxconstraint',0.1);
end
```

```
x=zeros(length(TestSet(:,1)),numClasses);
y = zeros(length(TestSet(:,1)),1);
```

```

%classify test cases
for j=1:size(TestSet,1)
    for k=1:numClasses

        x(j,k)=svmclassify(models(k),TestSet(j,:));

        if sum(x(j,:))==1
            if x(j,k)==1;
                result(j)=k;

            end

        else

            result(j)=0;

        end

    end

end

disp(Hist(1));
disp(x);

if result==1
    disp('grass');
elseif result==2
    disp('sand');
elseif result==3
    disp('asphalt');
else
    disp('nonclassified');
end

e= cputime-t;
disp(e);

case 2
TrainingSet = xlsread('E:/uniformLBPnorm.xls');
h = findobj('Tag','pushbutton2');
LBPu= get(h,'UserData');
TestSet=zeros(1,59);

```

```

TestSet(1,:)=LBPu;
GroupTrain= xlsread('E:/GroupTrain.xls');
u=unique(GroupTrain);
numClasses=length(u);
result = zeros(length(TestSet(:,1)),1);

%build models
for k=1:numClasses
    %Vectorized statement that binarizes Group
    %where 1 is the current class and 0 is all other classes
    G1vAll=(GroupTrain==u(k));
    models(k) = svmtrain(TrainingSet,G1vAll,'boxconstraint',0.1);
end

x=zeros(length(TestSet(:,1)),numClasses);
y = zeros(length(TestSet(:,1)),1);

%classify test cases
for j=1:size(TestSet,1)
    for k=1:numClasses

        x(j,k)=svmclassify(models(k),TestSet(j,:));

        if sum(x(j,:))==1
            if x(j,k)==1;
                result(j)=k;

            end
        else

            result(j)=0;

        end
    end
end

disp(TestSet);
disp(x);

if result==1
    disp('grass');
elseif result==2

```

```
    disp('sand');  
elseif result==3  
    disp('asphalt');  
else  
    disp('nonclassified');  
end  
  
e= cputime-t;  
disp(e);  
  
end
```

III. számú melléklet: neurális hálózat tanítása backpropagation algoritmussal:

```
net = feedforwardnet([7 7]);
net.trainFcn='traingd';
net.trainParam.show = 50;
net.trainParam.epochs = 1000;
net.trainParam.lr=0.1;
net.trainParam.goal = 0.001;
net.trainparam.max_fail=10;
net.divideparam.trainratio=80;
net.divideparam.valratio=20;

net.layers{1}.transferFcn = 'tansig';
net.layers{2}.transferFcn = 'tansig';
net.layers{3}.transferFcn = 'purelin';

net=configure(net,p,t);
[net,tr] = train(net,p,t);
```

IV. számú melléklet: neurális hálózat tanítása Levenberg-Marquardt

algoritmussal:

```
net = feedforwardnet([7 7]);
net.trainFcn='trainlm';
net.trainParam.show = 50;
net.trainParam.epochs = 1000;
net.trainParam.goal = 0.001;
net.trainparam.max_fail=10;
net.divideparam.trainratio=80;
net.divideparam.valratio=20;
net.trainparam.mu_inc=4;
net.trainparam.mu_dec=0.5;

net.layers{1}.transferFcn = 'tansig';
net.layers{2}.transferFcn = 'tansig';
net.layers{3}.transferFcn = 'purelin';

net=configure(net,p,t);
[net,tr] = train(net,p,t);
```

V. számú melléklet: neurális hálózat tanítása Flecher-Reeves konjugált gradiens

módszerrel:

```
net = feedforwardnet([7 7]);
net.trainFcn='traincgf';
net.trainParam.show = 50;
net.trainParam.epochs = 1000;
net.trainParam.goal = 0.001;
net.trainparam.max_fail=10;
net.divideparam.trainratio=80;
net.divideparam.valratio=20
net.trainParam.searchFcn='srchgol';

net.layers{1}.transferFcn = 'tansig';
net.layers{2}.transferFcn = 'tansig';
net.layers{3}.transferFcn = 'purelin';

net=configure(net,p,t);
[net,tr] = train(net,p,t);
```