

BABEŞ-BOLYAI UNIVERSITY CLUJ-NAPOCA
FACULTY OF MATHEMATICS AND INFORMATICS
SPECIALIZATION: COMPUTER SCIENCE

License Thesis

Road sign recognition

Abstract

There are numerous methods for road sign recognition but the robust and cost-effective solution is still an active research subject. The recognisability of traffic signs is affected by many variables, like luminous intensity, shading, partial coverage and other obstacles. The developed system is recognising road signs using machine learning techniques taking in consideration the above mentioned factors. The most important factor is accuracy and speed. The application displays the examined image and analyses it on background threads. The program extracts random rectangles from the original picture and decides if it contains a traffic sign or not, and if it does it categorises them.

This work is the result of my own activity. I have neither given nor received unauthorized assistance on this work.

JULY 2015

SZABO ÁGNES-TERÉZ

ADVISOR:
CSATÓ LEHEL BABEŞ-BOLYAI UNIVERSITY,
FACULTY OF MATHEMATICS AND INFORMATICS

BABEŞ-BOLYAI UNIVERSITY CLUJ-NAPOCA
FACULTY OF MATHEMATICS AND INFORMATICS
SPECIALIZATION: COMPUTER SCIENCE

License Thesis

Road sign recognition



SCIENTIFIC SUPERVISOR:

CSATÓ LEHEL BABEŞ-BOLYAI UNIVERSITY,
FACULTY OF MATHEMATICS AND INFORMATICS

STUDENT:

SZABO ÁGNES-TERÉZ

JULY 2015

UNIVERSITATEA BABEŞ-BOLYAI, CLUJ-NAPOCA
FACULTATEA DE MATEMATICĂ ŞI INFORMATICĂ
SPECIALIZAREA INFORMATICĂ

Lucrare de licență

Recunoașterea semnelor de circulație



CONDUCĂTOR ȘTIINȚIFIC:

CSATÓ LEHEL
UNIVERSITATEA BABEŞ-BOLYAI,
FACULTATEA DE MATEMATICĂ ŞI INFORMATICĂ

ABSOLVENT:

SZABO ÁGNES-TERÉZ

IULIE 2015

BABEŞ-BOLYAI TUDOMÁNYEGYETEM KOLOZSVÁR
MATEMATIKA ÉS INFORMATIKA KAR
INFORMATIKA SZAK

Licensz-dolgozat

Forgalmi táblák felismerése



TÉMAVEZETŐ:

CSATÓ LEHEL
BABEŞ-BOLYAI TUDOMÁNYEGYETEM,
MATEMATIKA ÉS INFORMATIKA KAR

SZERZŐ:

SZABO ÁGNES-TERÉZ

2015 JÚLIUS

Tartalomjegyzék

1. Bevezető	4
1.1. A dolgozat célja	4
1.2. Megvalósítás	4
2. Eddigi eredmények	5
2.1. Szín és formai jellemzők kivonásával	5
3. Forgalmi táblák	6
3.1. Mik a forgalmi táblák?	6
3.2. Adathalmaz	6
3.2.1. Tanulási adatok	6
3.2.2. Tesztelési adatok	7
4. Darabolás	8
5. Képek előfeldolgozása	10
5.1. Színterek	10
5.1.1. RGB színtér	10
5.1.2. HSV színtér	11
5.2. Megvalósítás	11
5.3. Átméretezés	12
6. Neurális hálózatok	13
6.1. Bevezető	13
6.1.1. Biológiai modell	13
6.2. Mesterséges neuronok	13
6.2.1. Sigmoid neuron	14
6.3. Neurális hálók struktúrája	15
6.4. Sztochasztikus gradiens csökkentés	15
6.5. Backpropagation	16
7. Az algoritmus	17
8. A kimenet	19
9. Eredmények és következtetések	20
9.1. Rétegek számának meghatározása	20
9.2. Tanulási ráta meghatározása	20
9.3. Batch méretének meghatározása	21
9.4. Ciklusok számának meghatározása	21
10. Továbbfejlesztési lehetőségek	24

TARTALOMJEGYZÉK

11. Fontosabb programkódok listája

25

1. fejezet

Bevezető

1.1. A dolgozat célja

Számos olyan rendszer létezik ami megtalál és felismer forgalmi táblákat valós környezetből vett képeken, de a robusztus és költséghatékony megoldás mai napig aktív kutatási téma. Mivel több különböző típusú forgalmi tábla létezik és ezek között sok a hasonlóság, a megkülönböztetésük a mai napig kihívásnak bizonyul. A táblák felismerhetőségét befolyásolja több tényező, mint a fényerősség, árnyékolás, parciális fedés és egyéb akadályok.

Egyes autóvezetők számára nehézséget jelenthet a figyelmük megosztása, ami bizonytalansághoz, balesethez vezethet. Ha rendelkezésükre állna egy rendszer, ami képes figyelmeztetni őket a forgalmi táblákról és az általuk megszabott korlátról vagy információról, magabiztosabbá válhatnának.

A céлом egy gyors, hatékony és megbízható rendszert készítése forgalmi táblák felismerésére valós környezetben készített képeken.

1.2. Megvalósítás

A rendszer először előfeldolgozást végez a képeken. Mivel a tesztadatok nem azonos méretűek, átméretezi őket majd szín szerinti szűrést végez, hogy egyszerűsítse a feldolgozandó adatokat. Ezt követően neurális háló segítségével elkezdődik a tanulási folyamat. Miután a rendszer megtanulta a különböző osztályok jellemzőit, a rendszer használható forgalmi táblák felismerésére tetszőlegesen kiválasztott képeken.

A felhasználó által kiválasztott képen első sorban meg kell találni a forgalmi tábla helyét. Ennek érdekében a képet kisebb darabokra vágja és ezeken is végrehajtja az fentebb említett előfeldolgozást. Az így kapott képeket próbálja azonosítani a forgalmi táblákat.

Mivel a forgalmi táblák idővel változhatnak, lényeges szempont, hogy a tervezett rendszer könnyen alkalmazkodjon a változásokhoz, anélkül, hogy nagymértékben módosítani kelljen azt. A neurális hálók tökéletesek erre a feladatra, hiszen ha az adatok változnak is, elég a tanulási folyamatot megismételni; nincs szükség a kód módosítására. [8, 2010]

2. fejezet

Eddigi eredmények

Összefoglaló: Forgalmi táblák felismerésére több módszert fejlesztettek ki, ennek egyike a neurális hálókkal történő tanulás. Ebben a fejezetben bemutatok néhány alternatív módszert:

1. Szín és formai jellemzők kivonásával
2. Szín és formai jellemzők kivonásával
3. Szín és formai jellemzők kivonásával

2.1. Szín és formai jellemzők kivonásával

A forgalmi táblákat színük és formájuk segítségével lehet kategorizálni, teljes felismerést viszont nem lehet biztosítani. Az [Gao et al., 2006] hivatkozásban a forgalmi táblák belsejében levő szimbólum figyelmen kívül hagyásával végeztek kategorizálást, pusztán a forma és szín alapján.

Az [Hatzidimos, 2004] hivatkozásban két részre osztható a javasolt megoldás:

1. A tábla megtalálása a képen
2. A tábla felismerése

A [Hatzidimos, 2004] hivatkozásban a tábla behatárolásához szín szerinti szűrést használ. Az így létrehozott bináris képből kiválasztja a fekete pixelek sűrűsége alapján azt a területet ami potenciálisan tartalmazza a táblát. Ezután forma alapján végez egy szűrést a képen. Amennyiben úgy véli, hogy megtalálta a forgalmi tábla helyét, elkezdi a felismerést. Pixelenként összehasonlítja a piros, zöld és kék komponensét a kinyert képnek egy mintaképpel mindegyik osztályból és ezekből számol egy átlagot. Abba az osztályba sorolja a képet, amelyik mintaképre a legnagyobb átlag jön ki.

3. fejezet

Forgalmi táblák

3.1. Mik a forgalmi táblák?

A forgalmi táblák a forgalom irányításáért felelősek; figyelmeztetik, irányítják és informálják a gépkocsivezetőket. Három fő típusú forgalmi tábla van és mindegyik különböző formával rendelkezik:

1. utasító táblák (kör)
2. figyelmeztető táblák (háromszög)
3. informáló táblák (téglalap)

Ez alól kivétel a STOP tábla, ami nyolcszög alakú.

További információval szolgál a forgalmi tábláról a színe. A forgalmi táblák figyelemfelkeltőek kell legyenek, ennek megfelelően alkalmazzák a színeket is. A megjelenő színek: piros, fehér, fekete, sárga és kék. [8, 2010]

A forgalmi táblák többsége két részből tevődik össze: egy külső telített sáv és egy belső szimbólum vagy szám. A belső szimbólum általában fekete.

3.2. Adathalmaz

Az adathalmazhoz egy forgalmi táblák felismerésére kitűzött németországi verseny adatait használtam fel. [?]

3.2.1. Tanulási adatok

Az adathalmaz 1213 tanulási adatot tartalmaz. A képek mérete 16x16 pixeltől 128x128 pixelig változik, különböző szemszögből és változatos fényhatások mellett. A képek 43 kategóriába vannak sorolva, a tartalmazott forgalmi tábla alapján.

Habár nagy mennyiségű tanulási adat áll rendelkezésre, ezek nincsenek egyenlően elosztva az osztályok között. Ez megnehezíti a tanulást a neurális háló számára. Azok az osztályok amelyekben kevés tanulási adat van, kevésbé vannak befolyással a neurális háló kimenetére.

A 3.1 ábrán látható példa négy tanítási adatra. Az első három képen levő táblák könnyedén felismerhetőek, míg a az utolsó tábla nehezen azonosítható. Nehéz megmondani, hogy 30-as, de 70-es vagy

3. FEJEZET: FORGALMI TÁBLÁK

80-as sebességkorlát. Az adathalmazban kis számban ugyan, de előfordulnak ehhez hasonló táblák. Ezek megnehezítik a tanulási folyamatot.



3.1. ábra. Példa tanulási adatokra

forrás:http://benchmark.ini.rub.de/?section=gt_sdb&subsection=dataset

3.2.2. Tesztelési adatok

900 teszt adat állt a rendelkezésemre, amiken nullától hatig akárhány forgalmi tábla megjelenhetett. Ezek jóval nagyobb méretű képek a tanításra szánt képeknél, 1360x800 pixel nagyságúak. A 3.2 ábrán látható egy példa ilyen képre. Ezen a képen két forgalmi tábla található.



3.2. ábra. Példa teszt adatra

forrás:http://benchmark.ini.rub.de/?section=gt_sdb&subsection=dataset

4. fejezet

Darabolás

Annak érdekében, hogy megtaláljunk egy forgalmi táblát egy képen, szegmentációt alkalmaztam. A szegmentáció az a folyamat, amikor kisebb képeket vágunk ki egy nagyobb képből [Ohlander et al., 1978].

Több különböző módszer létezik a szegmentálás végrehajtására. Az általam alkalmazott véletlenszerű kiválasztáson alapszik. Véletlenszerű téglalapokat generálok és vágom ezeket ki az eredeti képből.

Mivel a valós környezetből vett képek több forgalmi táblát is tartalmazhatnak, fontos hogy az egész képet bejárják ezek a darabok. Azáltal hogy nagy számban generálok ilyen darabokat, biztosítom, hogy nagy valószínűséggel megtaláljam a képen levő forgalmi táblákat. A táblák mérete is változik attól függően, hogy milyen távolságra van a fényképezőtől, ezért arra is figyelnem kellett, hogy a kivágott darabok mérete széles tartományban mozogjon.

Azáltal, hogy véletlenszerűen választom ki a vizsgálandó részeket az eredeti képből, ezeknek nagy része olyan kép lesz, ami nem tartalmaz forgalmi táblát vagy csak részlegesen fogja azt tartalmazni. Ez az oka annak, hogy a neurális háló akkor is megfelelően kell osztályozza a táblát ha csak 70-80% látszik belőle.

A 4.1 ábrán látható kép darabolásából nyert képekre látunk példákat a 4.2 ábrán. Az így létrejött képek között megtalálható a két forgalmi tábla (első és negyedik kép) de ezeken kívül olyan darabok is létrejönnek, amik nem tartalmaznak forgalmi táblát. Azok a darabok amik nem tartalmaznak forgalmi táblát más információt tartalmaznak, mint például emberek, fák, járművek, esetleg hirdető táblák.

4. FEJEZET: DARABOLÁS



4.1. ábra. Szegmentálás előtti kép

forrás:<http://benchmark.ini.rub.de/?section=gtsdb&subsection=dataset>



4.2. ábra. Példa szegmentálás utáni képekre

forrás:<http://benchmark.ini.rub.de/?section=gtsdb&subsection=dataset>

5. fejezet

Képek előfeldolgozása

Összefoglaló: Annak érdekében, hogy egyszerűsítsem a bevitt adatokat, előfeldolgozást végzek a képeken a színük alapján. A neurális háló a pixelek szürke árnyalatával dolgozik. A képek különböző fényviszonyok mellett készültek, ennek köszönhetően némelyik nagyon világos, mások pedig nagyon sötétek.

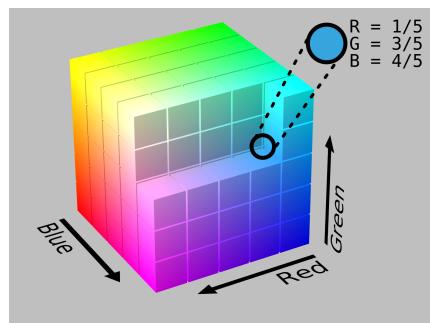
Megjegyzés: A képek valós környezetből származnak, így más objektumokat is tartalmazhatnak, mint például faágak.

5.1. Színterek

A színterek arra szolgálnak, hogy könnyebben tudjuk meghatározni, létrehozni és felfogni a színeket. Egy színt általában három paraméter határozza meg. Ezek fejezik ki az adott szín helyét az illető színtérben. Különböző színtereket ismerünk, attól függően válasszuk ki a megfelelőt, hogy mire akarunk alkalmazni. [Ford and Roberts, 1998]

5.1.1. RGB színtér

Az RGB színtér egy additív színrendszer, mely a tri-kromatikus elméletben alapszik. A tri-kromatikus elmélet lényege, hogy 3 különböző szín segítségével bármilyen szín kikeverhető. Az RGB színtér könnyen implementálható, de nem párhuzamos az emberi látással. Az RGB színteret gyakran felhasználják számítógépes alkalmazásokhoz, mert nincs szükség transzformációra ahhoz, hogy megjelenítsük az információt. A színtér legegyszerűbb ábrázolása egy kocka, ahol a 3 tengely a piros, zöld és kék színekkel egyenértékű.



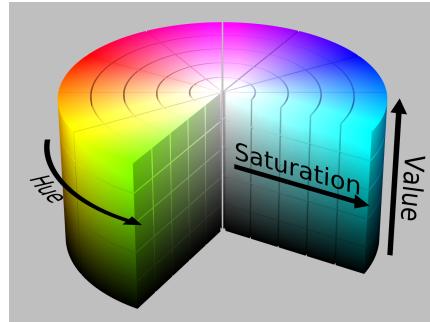
5.1. ábra. RGB színtér

forrás:http://upload.wikimedia.org/wikipedia/commons/8/83/RGB_Cube_Show_lowgamma_cutout_b.png

5. FEJEZET: KÉPEK ELŐFELDOLGOZÁSA

5.1.2. HSV színtér

Ellentétben az RGB ábrázolással, a HSV színtér henger-koordinátákat alkalmaz. Ez az egyik leggyakoribb ábrázolási módszer. Az árnyalat(hue) az az emberi érzet, amely szerint egy szín egy másik színre vagy több szín keverékére hasonlít a piros, sárga, zöld és kék színek közül. A szín telítettsége(saturation) egy szín színessége és a fényessége közti arány. A fényesség (value) az az emberi érzet, hogy egy terület több vagy kevesebb fényt áraszt magából.



5.2. ábra. HSV színtér

forrás:http://upload.wikimedia.org/wikipedia/commons/0/0d/HSV_color_solid_cylinder_alpha_lowgamma.png

5.2. Megvalósítás

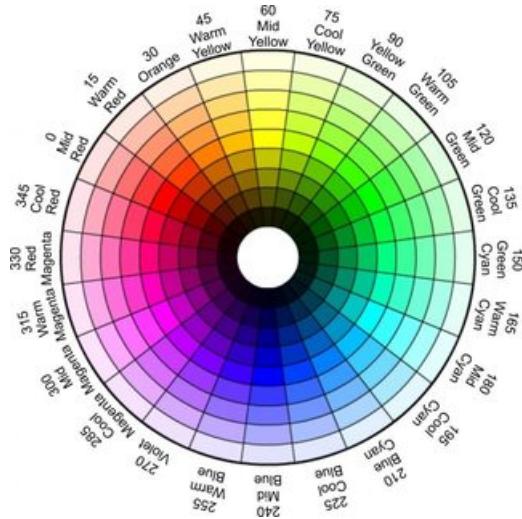
A forgalmi táblák figyelemfelkeltőek kell legyenek, ezért jól látható színeket használnak, mint például piros, kék, sárga, fekete és fehér.

A bemeneti képeket szín alapján szűröm, ezáltal felépíték egy új, fekete-fehér képet. Ha az árnyalat, a telítettség és a fényesség egy bizonyos intervallumban helyezkedik el, az új kép megfelelő pixelje fekete lesz, különben fehér. Ezeket az intervallumokat kísérletezéssel és a 5.3 ábra segítségével határoztam meg.

Szín	Árnyalat	Telítettség	Fényesség
Piros	0-20, 340-360	>0.3	>0.2
Sárga	25-60	>0.3	>0.2
Kék	180-270	>0.3	>0.2
Fekete	-	-	<0.3
Fehér	-	-	>0.85

5.1. táblázat. Elfogadási tartomány

5. FEJEZET: KÉPEK ELŐFELDOLGOZÁSA



5.3. ábra. Színkör

forrás:<http://stackoverflow.com/questions/21737913/image-of-hsv-color-wheel-for-opencv>

5.3. Átméretezés

Mivel a tanulási adatok és a teszt adatból kivágott képek mérete is változó, a képeket átméretezem 25x25 pixelre és ezzel dolgozom tovább.

6. fejezet

Neurális hálózatok

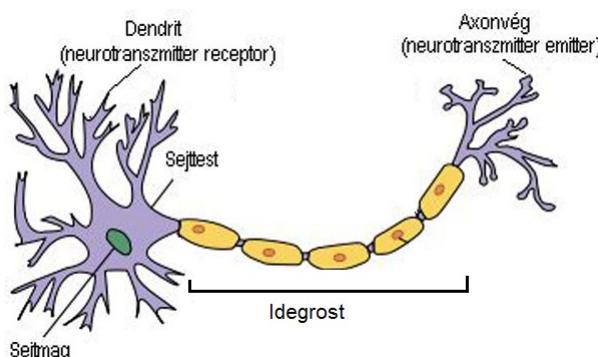
Összefoglaló: Ebben a fejezetben megnézzük a biológiai modellt ami alapján a mesterséges neuronok létrejöttek, megismerkedünk a mesterséges neuronokkal

6.1. Bevezető

6.1.1. Biológiai modell

Az idegsejtek (neuronok) az idegrendszer alkotóelemei. A neuron jeleket fogad és ezekre megfelelő választ generál. Az emberi idegrendszer nagyon komplex műveletek elvégzésére képes rövid idő alatt. Mivel ennek analógiára hozták létre a mesterséges neurális hálókat, ezért fontos a biológiai alapok megismerése.

A 6.1 ábrán látható egy idegsejt és annak részei. A dendritek más neuronokkal állnak kapcsolatban és a bejövő információt szállítják. A beérkezett információ feldolgozását a sejtmag hajtja végre. Az idegrost egy szigetelt vezető aminek az a feladata, hogy továbbszállítsa az információt. Az axonvégek más neuronok dendritjeivel vannak kapcsolatban, a választ szállítják tovább a velük kapcsolatban levő idegsejteknek.



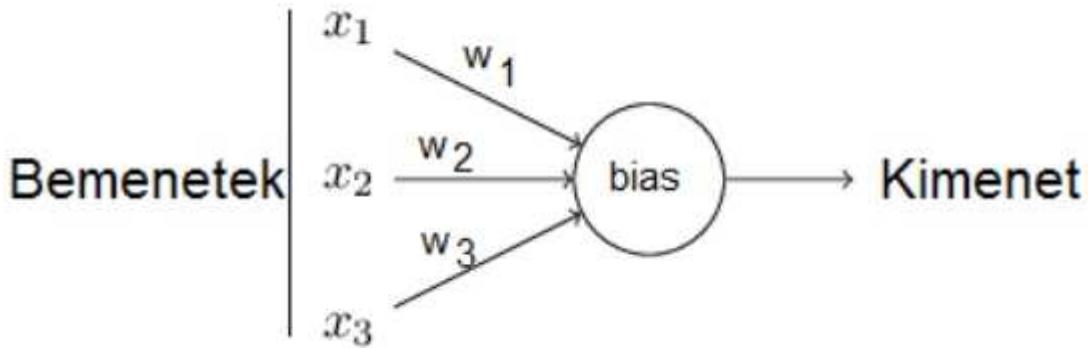
6.1. ábra. Idegsejt

6.2. Mesterséges neuronok

A neuronok a neurális hálók építőkövei.

6.2.1. Sigmoid neuron

A sigmoid neuron 0 és 1 közötti értékeket kap bemeneti paraméterként és létrehoz egy kimenetet ugyan-ebben az intervallumban.



6.2. ábra. Sigmoid neuron

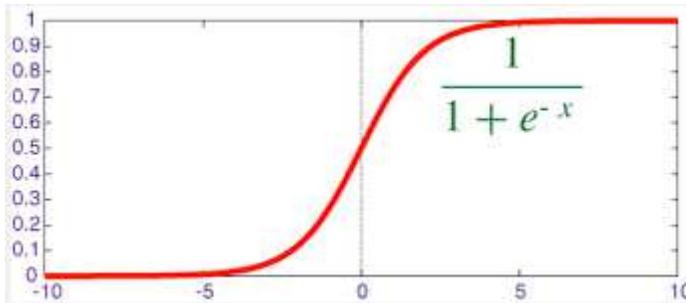
Mindegyik bemenethez súlyok vannak rendelve, ami azt jelképezi, hogy az illető bemenet mennyire játszik fontos szerepet a neuron kimenetében. minden neuronhoz tartozik továbbá egy küszöbérték ami a kisebb rendellenességek javítására szolgál.

A sigmoid neuron kimenete a 6.1 képlettel számítható ki.

$$\sigma\left(\sum_{i=1}^{n-1} w_i x_i + b\right) \quad (6.1)$$

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (6.2)$$

A 6.2 függvényt aktiválási függvénynek nevezik. Az aktiválási függvény szerepe, hogy a kimenetet egy adott intervallumba szorítsa. Ahogy ez a 6.3 ábrán is látszik, a sigmoid neuron esetében ez az intervallum 0 és 1 között van.

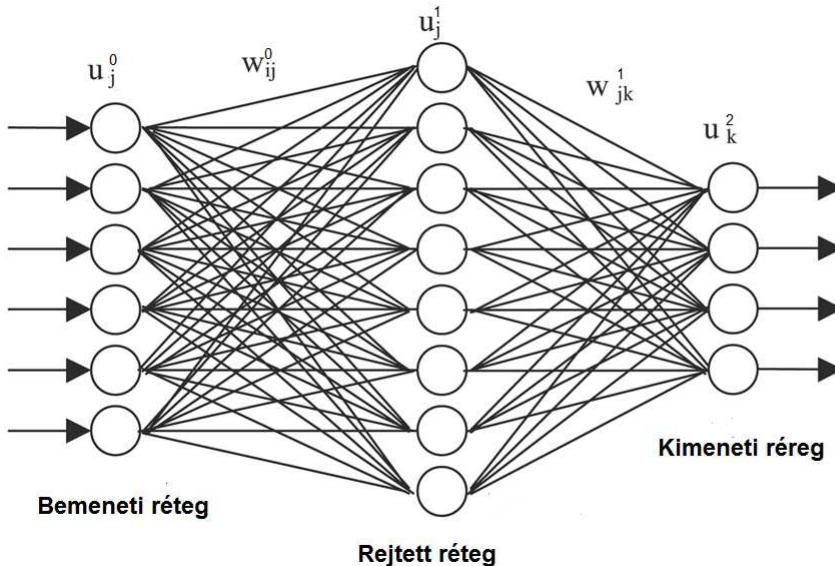


6.3. ábra. Sigmoid függvény grafikus képe

A sigmoid neuron egyik tulajdonsága, hogy kis változás a súlyokban kis változást eredményez a kimenetben. Ez legalább annyira előny is, mint hátrány. Előny, mert amikor a kimenet közel van az

6. FEJEZET: NEURÁLIS HÁLÓZATOK

elvárt kimenethez, a módosítások nem eredményezik azt, hogy a kimenet túlságosan eltávolodjon az elvárttól. Hátrány, mert amikor a kimenet nagyon helytelen, mint például a tanulás elején, hosszú időbe telik a javítása.



6.4. ábra. Egyszerű neurális háló

A 6.4 ábrán egy egyszerű szigmoid neuronokból álló neurális háló látható. A súlyvektorok mérete már ennél a kis példánál is viszonylag nagy, ezért egy nagy hálónál ezek az értékek jelentősen megnőnek.

6.3. Neurális hálók struktúrája

Olyan neurális hálót használtam, amelynek 625 bemenete és 43 kimenete van. A bemeneti neuronok a kép egy pixelének szürke árnyalatát tartalmazzák. A kimenetek száma megegyezik a forgalmi táblák osztályainak a számával.

6.4. Sztochasztikus gradiens csökkentés

Annak érdekében, hogy a súlyokat és küszöbértékeket a kimenetek az elvárt eredménytől való eltérése függvényében tudjam módosítani, a sztochasztikus gradiens csökkentést alkalmaztam. Ez a gradiens csökkentés egyszerűsített változata. Gradiens csökkentéssel a teljes tanítási adathalmazra számolja a hiba értékét, míg a sztochasztikus gradiens csökkentés közelíti a hibát egy bizonyos számú véletlenszerűen kiválasztott adat alapján. Minél nagyobb ez a szám, a közeliítés pontossága annyival jobb lesz, de mivel nem a teljes adathalmazzal dolgozik, ezért gyorsabb lesz.

A 6.3 egyenlet a négyzetes hibafüggvény. n a bemeneti tanítási adatok száma, $y(x)$ az x bemenetre a háló által generált kimenet és a az elvárt kimenet az x bemenetre. Amikor az elvárt és a kiszámított kimenet közel van egymáshoz, a hiba kicsi lesz. Mivel a célom, hogy a kiszámított kimenet az elvárt

6. FEJEZET: NEURÁLIS HÁLÓZATOK

kimenetet közelítse meg minél jobban, ezt a hibát kell minimalizálnom. Ennek érdekében bevezetem a ΔC -vel jelölt gradiensét a hibafüggvénynek. A gradiens értékét a 6.4 képlettel lehet kiszámítani.

$$C = \frac{1}{2n} \sum_x \|y(x) - a(x)\|^2 \quad (6.3)$$

$$\Delta C = \left(\frac{\partial C}{\partial w_i}, \frac{\partial C}{\partial b_j} \right)^T \quad (6.4)$$

A hiba csökkentése érdekében ismételtem alkalmaztam a 6.5 és a 6.6 szabályokat a súlyokra és küszöbértékekre. η a tanítási ráta. Ez egy kicsi érték kell legyen, de nem túl kicsi, mert akkor a tanulási folyamat túl lassú lenne.

$$w_i \rightarrow w_i - \eta \frac{\partial C}{\partial w_i} \quad (6.5)$$

$$b_j \rightarrow b_j - \eta \frac{\partial C}{\partial b_j} \quad (6.6)$$

6.5. Backpropagation

A backpropagation algoritmus arra szolgál, hogy a háló súlyait és küszöbértékeit úgy állítsa be, hogy a kimenet megközelítse minél jobban az elvárt kimenetet.

A 6.7 egyenletben z^l a súlyozott bemeneti vektor az l -edik rétegen. Ahogy a 6.8 egyenletben látható, a kimeneti vektor egy rétegen az előző réteg kimenetének használatával számolható ki, vagyis az információ továbbításával. A háló kimenetének kiszámítása után a 6.9-es egyenlet segítségével lehet kiszámolni az utolsó réteg hibáját, ahol L a rétegek számát jelöli, $\Delta_a C$ pedig a $\frac{\partial C}{\partial a_j^L}$ parciális deriváltakat tartalmazó vektor.

Maga a backpropagation a 6.10 képlet segítségével történik minden l rétegre, $L-1$ -től 2-ig.

$$z^l = (w^l a^{l-1} + b^l) \quad (6.7)$$

$$a^l = \sigma(z^l) \quad (6.8)$$

$$\delta^L = \Delta_a C \circ \sigma'(z^L) \quad (6.9)$$

$$\delta^l = ((w^{l+1})^T \delta^{l+1}) \circ \sigma'(z^l) \quad (6.10)$$

7. fejezet

Az algoritmus

1. Súlyok és küszöbértékek inicializálása véletlenszerű adatokkal
2. minden ciklusra végezd:
 - (a) Tanulási adatok összekeverése
 - (b) minden batch-re végezd:
 - i. minden tanulási adatra a batch-ból végezd:
 - A. Mindegyik neuron kimenetének kiszámítása a bemenettől a kimenet felé
 - B. Kimenet hibájának kiszámítása
 - C. minden neuron hibájának kiszámítása a kimenettől a bemenet felé haladva
 - ii. Gradiens csökkentés: súlyok és küszöbértékek újraszámolása

Algorithm 1 Bla

```
1: Súlyok és küszöbértékek inicializálása véletlenszerű adatokkal
2: for minden ciklus  $i$  do
3:   Tanulási adatok összekeverése
4:   for minden batch  $j$  do
5:     for minden  $d$ -re  $j$  batch-ból do
6:       minden réteg kimenetének kiszámítása a 6.2 egyenlet segítségével a bemenettől a kimenet felé
7:       Evaluate the fitness  $f(\vec{X}_i)$ 
8:       if  $f(\vec{X}_i) < f(\vec{P}_i)$  then
9:          $\vec{P}_i \leftarrow \vec{X}_i$ 
10:      end if
11:      if  $f(\vec{X}_i) < f(\vec{P}_g)$  then
12:         $\vec{P}_g \leftarrow \vec{X}_i$ 
13:      end if
14:    end for
15:  end for
16: end for
```

Ahhoz, hogy a háló tanulni tudjon, fontos az, hogy a súlyokat és küszöbértékeket megfelelően kicsi számokkal inicializáljuk. Ezt úgy valósítottam meg, hogy a 0 és 1 között véletlenszerűen generált számot elosztottam 3000-el.

7. FEJEZET: AZ ALGORITMUS

A batch-ben egyidőben jelen levő adatok minél változatosabbak kell legyenek, annak érdekében hogy tanulás közben egy időben figyelembe vegye a különböző osztályok tulajdonságait. Ezért van szükség az adatok összekeverésére. A beolvasás sorban történik, vagyis az azonos osztályba tartozók egymás után fognak szerepelni. Ha ebben a sorrendben kerülnének be a batch-be, akkor a háló először megpróbálná megtanulni az egyik osztály tulajdonságait, utána pedig sorban a következőkét. Az adatok keverésével ezt kerülöm el. [Nielsen, 2014]

8. fejezet

A kimenet

A neurális háló kimenete 43 neuront tartalmaz. Ez megegyezik a forgalmi táblák kategóriáinak számával.

Azt, hogy a háló melyik kategóriába sorolta az adott táblát a legnagyobb kimenetű neuron indexe dönti el. E mellett bevezettem egy olyan feltételt, hogy a legnagyobb és a második legnagyobb érték közötti különbség nagyobb kell legyen mint 0,2. Erre azért volt szükség, mert megtörténhet olyan, hogy a háló nem képes eldönteni, hogy melyik osztályba tartozik egy kép. Ekkor a kimenet több neuronjának is hasonló értéke lesz.

A kimenet alapján tehát három különböző lehetőség van: helyesen osztályoz egy képet, helytelenül osztályoz egy képet, nem képes osztályozni egy képet. Amikor nem képes osztályozni egy képet, azt úgy tekintem, hogy a kép nem tartalmaz forgalmi táblát.

9. fejezet

Eredmények és következtetések

A háló teszteléséhez a tesztelési adatokból kivágott forgalmi táblákat használtam.

Különböző paraméterekre (rétegek száma, rejtett réteg/rétegek mérete, batch méret, tanulási ráta) folytattam tesztelést az ideális körülmények megtalálása érdekében. Eleinte nem a teljes adathalmazra, hanem csak pár osztályra futtattam, a gyors eredmény érdekében. Ezek a tesztek gyorsan lefutottak és viszonylagos eredményt adtak a hálóról.

Amikor kis adathalmazzal teszteltem, kevés osztállyal, azt tapasztaltam, hogy ha az osztályok olyan táblákat tartalmaztak amik jelentősen különböznek egymástól, akkor 100%-ban felismeri a tesztadatokat. Ilyen osztályokra példa a STOP tábla és a 30-as sebességkorlátozó tábla. Abban az esetben, ha az osztályokba tartozó táblák hasonlítanak egymásra, mint például a 30-as és 50-es sebességkorlátozó tábla, a felismert táblák száma lecsökken, kb. 85%-ra.

9.1. Rétegek számának meghatározása

A rendszer a legjobb eredményt akkor adta, ha nem volt jelen rejtett réteg. Az által, hogy nem volt rejtett réteg a számítások egyszerűsödtek és a folyamat felgyorsult.

Abban az esetben, ha egy rejtett réteg volt, a tanulás kb. 100 ciklus után kezdődött el. A tanulási folyamat lassú volt és a végeredmény nem mutatott jobb eredményt mint amikor hiányzott a rejtett réteg.

Az elkövetkezendő teszteredmények olyan rendszerben készültek, ahol nem volt jelen rejtett réteg.

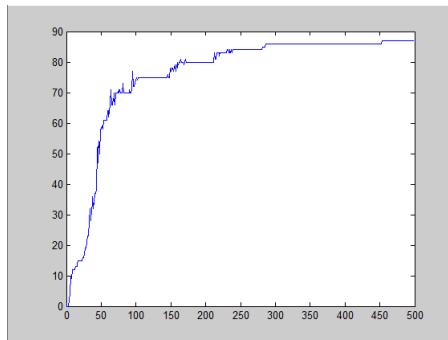
9.2. Tanulási ráta meghatározása

A tanulási ráta behatárolására először három, négy osztályra futtattam teszteket. Ez által gyorsan kaptam egy viszonyítási pontot, amit majd alkalmazni tudtam a több osztályra történő futáskor.

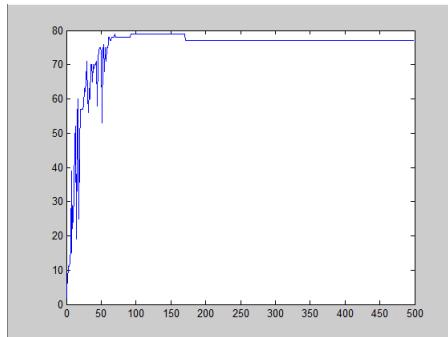
Elsőnek 1-es tanulási rátával próbálkoztam. Az eredmények elég pozitívak voltak, a 9.1 ábrán látható a helyesen kategorizált képek száma minden ciklusban.

Ezek után a tanulási ráta növelésével próbálkoztam. Habár a helyesen kategorizált képek száma minimálisan csökkent, ebben az esetben a háló sokkal hamarabb kezdte megtanulni az osztályok jellemzőit. A 9.2 ábrán jól látszik, hogy sokkal meredekebb az elején, vagyis gyorsabban tanul az elején és utána normalizálódik. A maximum elérése után a kezdeti oszcilláció is megszűnik.

9. FEJEZET: EREDMÉNYEK ÉS KÖVETKEZTETÉSEK



9.1. ábra. 1-es tanulási ráta



9.2. ábra. 3-as tanulási ráta

Végeztem teszteket más értékű tanulási rátával is, de ha nem változtattam nagy mértékben az értéket, akkor hasonló eredményt adtak. Ha túlságosan megnövelte az értékét, az oszcilláció ami a 9.2 ábra elején látható megnőtt és nagy mértékben lecsökkent a felismert táblák száma. Ha túl kicsi értéket választottam, a tanulási folyamat nagyon lassan indult be. Ebből kiindulva a további tesztekhez a tanulási rátát háromnak választottam meg.

9.3. Batch méretének meghatározása

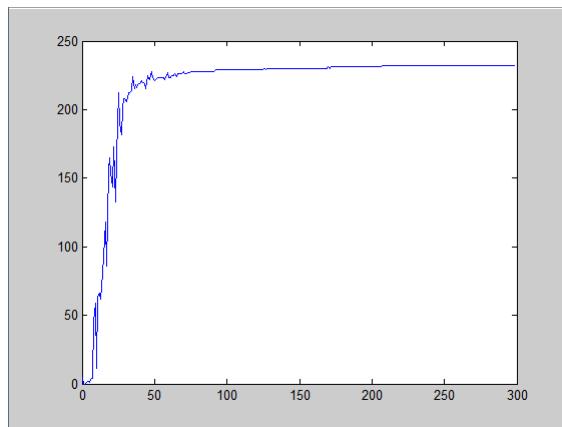
A batch méretének meghatározásához is több tesztet futtattam különböző értékekre. A behatárolást ebben az esetben is kis méretű adathalmazra végeztem.

A batch méretének változtatása nem volt nagy hatással a helyesen kategorizált képek számára. A 9.3, 9.4 és 9.5 ábrákon látható a 10-re, 20-ra, illetve 30-ra kapott eredmények. Nincs nagy különbség közöttük, de a 30-as értékre let legnagyobb a felismert táblák száma, ezért a továbbiakban ezzel az értékkel dolgoztam. Nagyobb mennyiségi tesztadatra megpróbáltam később a bath méretét 10-re állítani, de a tanulás jelentősen nehezebben indult be, ezért nem végeztem vele több tesztet.

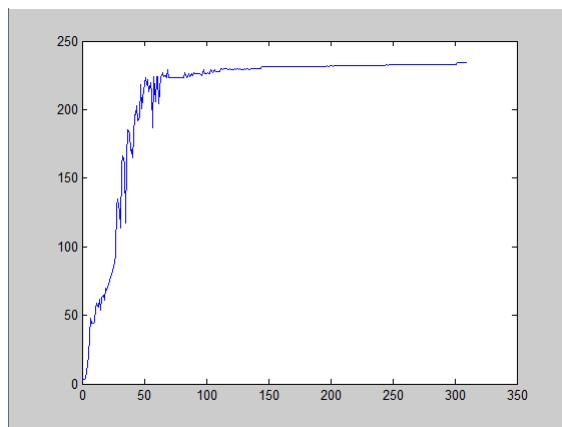
9.4. Ciklusok számának meghatározása

A ciklusok száma a tanulási folyamatban nem játszik olyan fontos szerepet mint a tanulási rata, viszont ha túl kicsinek választjuk meg a hálónak nincs elég ideje tanulni. A nagy ciklusszám nincs negatív hatással

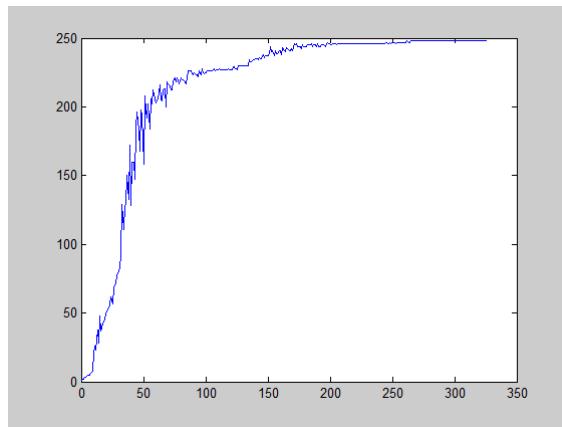
9. FEJEZET: EREDMÉNYEK ÉS KÖVETKEZTETÉSEK



9.3. ábra. 10-es bath méret eredményei



9.4. ábra. 20-as bath méret eredményei



9.5. ábra. 30-as bath méret eredményei

9. FEJEZET: EREDMÉNYEK ÉS KÖVETKEZTETÉSEK

a felismert táblák számára, de feleslegesen növeli a futási időt. Miután a háló elérte azt a pontot, hogy nem képes tovább fejlődni, felesleges további ciklusokat futtatni. Lehetne egy olyan feltételt beiktatni, hogy ha egy bizonyos százalékban felismeri a háló a táblákat akkor állítsuk le a tanulást meg mielőtt az összes ciklust végrehajtanánk, de ha ezt a százalékot túl nagyra állítjuk és ez által elérhetetlenné válik, akkor felesleges, ha pedig túl kicsire, a futási időt ugyan lecsökkent, de a háló még képes lett volna tanulni tovább.

100 ciklus ez esetek többségében elégsgesnek bizonyul, de néha 50 ciklus is ugyanazt eredményezte.

10. fejezet

Továbbfejlesztési lehetőségek

Az alkalmazás továbbfejlesztése szempontjából a pontosság és hatékonyság az elsődleges szempontok amiket szem előtt kell tartani. Ennek érdekében lehetne a képek előfeldolgozásán módosítani, a szín szerinti szűrést javítani, forma szerinti szűrést bevezetni.

A neurális háló szempontjából is több fejlesztési lehetőség áll rendelkezésre: más típusú neuronokat használni(nem szigmoid, hanem perceptron vagy más), konvolúciós neurális hálót alkalmazni, más hibafüggvényt használni, különböző típusú neuronokat beépíteni ugyanabba a hálóba.

11. fejezet

Fontosabb programkódok listája

```
1 #include "network.h"
2
3 //konstruktor
4 Network::Network(std::vector<int> sizes)
5 {
6     //retegek beallitasa
7     _num_layers = sizes.size();
8
9     _sizes = sizes;
10
11    //biasok inicializalasa
12    _biases = Matrix2D(_sizes);
13    _biases.randInit();
14
15    //sulyok inicializalasa
16    _weights = Matrix3D(_sizes);
17    _weights.randInit();
18 }
19
20 //az elozo reteg kimenete segitsegevel kiszamolja a kovetkezo reteg kimenetet
21 Vector Network::feedforward(Vector a)
22 {
23     for(unsigned int i = 0; i < _weights.GetMatrix2DCount(); i++) {
24         Matrix2D w_i = _weights.GetMatrix2D(i);
25         int s = _biases.GetVector(i).GetCount();
26         Vector new_a(s);
27         for (int j = 0; j < s; j++)
28             new_a.SetItem(j, w_i.dot(a).GetItem(j) + _biases.GetData() [i].GetData() [j]);
29         a = sigmoid(new_a);
30     }
31     return a;
32 }
33
34 //a helyesen kategorizat kepek szama
35 unsigned int Network::evaluate(DataSet testDataSet)
36 {
37     unsigned int correct = 0;
38     for(unsigned int i = 0; i < testDataSet.GetDataCount(); i++) {
39         Data testData = testDataSet.GetData() [i];
40         Vector tv = feedforward(testData.GetImageData());
41         double imgClass = tv.GetMax();
42         if (imgClass == testData.GetImageClass())
43             correct++;
44     }
45     return correct;
46 }
47
48 //szochasztkus gradiens csokkentes
49 void Network::SGD(DataSet dataset, int epochs, int mini_batch_size, double eta, DataSet
50 testDataSet)
51 {
52     unsigned int n = dataset.GetDataCount();
53
54     for (unsigned int j = 0; j < epochs; j++) {
55         //az adatok osszekeverese
56         dataset.Shuffle();
57         for (int i = 0; i <= n-mini_batch_size; i+=mini_batch_size) {
58             DataSet mini_batch = DataSet(dataset.GenerateMiniBatch(i,mini_batch_size));
59             updateMiniBatch(eta,mini_batch);
60         }
61     }
62
63     //a vegso biasok kiirasa hogy felhasznalhato legyen kesobb
64     struct stat buffer;
65     std::string name = "biases.txt";
66     if (stat(name.c_str(), &buffer) == 0) {
67         remove(name.c_str());
68     }
69 }
```

11. FEJEZET: FONTOSABB PROGRAMKÓDOK LISTÁJA

```

    }
    _biases.writeData(name);

    //a vegso sulyok kiirasa hogy felhasznalhato legyen kesobb
71     struct stat buffer2;
72     std::string name2 = "weights.txt";
73     if (stat(name2.c_str(), &buffer2) == 0) {
74         remove(name2.c_str());
75     }
76     _weights.writeData(name2);
77 }

//backpropagation alkalmazasa a batch-ekre
void Network::updateMiniBatch(double eta, DataSet mini_batch)
81 {
82     Matrix2D nabla_b(_sizes);
83     nabla_b.zeros();
84     Matrix3D nabla_w(_sizes);
85     nabla_w.zeros();
86     for(unsigned int i = 0; i < mini_batch.GetDataCount(); i++) {
87         Data mini_batch_row = mini_batch.GetData()[i];

88         Matrix2D nabla_b2(_sizes);
89         nabla_b2.zeros();
90         Matrix3D nabla_w2(_sizes);
91         nabla_w2.zeros();

92         Vector activation = mini_batch_row.GetImageData();
93         Matrix2D activations(_sizes, 0);
94         activations.SetDataRow(0, activation);
95         Matrix2D zs(_sizes);

96         //a kimenetek kiszamitása előre haladva
100        for(unsigned int j = 0; j < _biases.GetVectorCount(); j++) {
101            Vector z(_biases.GetVector(j).GetCount());
102            z.SetData(_weights.GetMatrix2D(j).dot(activation).add(_biases.GetVector(j)));

103            zs.SetDataRow(j, z);
104            activation = sigmoid_prime(z);
105            activations.SetDataRow(j+1, activation);
106        }

107        Vector outActivation = activations.GetVector(activations.GetVectorCount() - 1);
108        Vector preferedOutput(outActivation.GetCount());
109        //elvárt kimenet felepítése
110        preferedOutput.setItem(mini_batch_row.GetImageClass(), 1.0);

111        //a kimeneti reteg hibajának kiszamitása
112        Vector delta = cost_derivative(outActivation, preferedOutput).HadamardProd(
113            sigmoid_prime(zs.GetVector(zs.GetVectorCount() - 1)));
114
115        nabla_b2.SetDataRow(nabla_b2.GetVectorCount() - 1, delta);
116
117        nabla_w2.SetData(nabla_w2.GetMatrix2DCount() - 1, dot(delta, activations.GetVector(
118            activations.GetVectorCount() - 2)));
119
120        //backpropagation alkalmazva visszafele a retegek hibainak kiszamitasara
121        for (unsigned int j = 2; j < num_layers; j++) {
122            Vector z = zs.GetVector(zs.GetVectorCount() - j);
123            Vector spv = sigmoid_prime(z);
124            delta = dot(_weights.GetMatrix2D(_weights.GetMatrix2DCount() - j + 1), delta) .
125                HadamardProd(spv);

126            nabla_b2.SetDataRow(nabla_b2.GetVectorCount() - j, delta);
127            nabla_w2.SetData(nabla_w2.GetMatrix2DCount() - j, dot(delta, activations.GetVector(
128                activations.GetVectorCount() - j - 1)));
129        }
130        Nablas ns(nabla_b2, nabla_w2);

131        Matrix2D delta_nabla_b = ns._nabla_b;
132        Matrix3D delta_nabla_w = ns._nabla_w;
133        nabla_b = nabla_b.add(delta_nabla_b);
134        nabla_w = nabla_w.add(delta_nabla_w);
135    }

136    //sulyok es biasok megfelelo modosítása
137    _weights = _weights.sub(nabla_w.multiplyByScalar(eta / (double)mini_batch.GetDataCount()));
138    _biases = _biases.sub(nabla_b.multiplyByScalar(eta / (double)mini_batch.GetDataCount()));
139}

```

Irodalomjegyzék

Know your traffic signs official edition. London Department for Transport, 2010.

X. W. Gao, L. Padladchikova, D. Shaposhnikov, K. Hong, and N. Shevtsova, „Recognition of traffic signs based on their colour and shape features extracted using human vision models,” *Journal of Visual Communication and Image Representation*, vol. 17, no. 4, pp. 675–685, 2006.

J. Hatzidimos, „Automatic traffic sign recognition in digital images,” 2004.

R. Ohlander, K. Price, and D. R. Reddy, „Picture segmentation using a recursive region splitting method,” *Computer graphics and image processing*, vol. 8, pp. 313–333, 1978.

A. Ford and A. Roberts, „Colour space conversions,” 1998.

M. A. Nielsen, „Neural networks and deep learning,” 2014.

S. Escalera, X. Baró, O. Pujon, J. Vitria, and P. Radeva, „Traffic-sign recognition systems,” *Journal of Visual Communication and Image Representation*, vol. VI, 2011.

K. Brkic, „An overview of traffic sign detection methods,” 2010.