

KÖZLEKEDÉSI TÁBLÁK FELISMERÉSE NEURÁLIS HÁLÓZAT ALKALMAZÁSÁVAL

APPLYING NEURAL NETWORK FOR THE RECOGNITION OF TRAFFIC SIGNS

Neurohr Zsolt Barnabás¹, Tóthné Laufer Edit²

Óbudai Egyetem, Bánki Donát Gépész és Biztonságtechnikai Mérnöki Kar, Budapest, Magyarország

¹ zsoltbarni@gmail.com

² laufer.edit@bgk.uni-obuda.hu

Abstract

Artificial intelligence is one of the most dynamically developing areas of science today. Although it is not yet an integral part of our lives to use artificial intelligence solutions, it can be seen in terms of development, that it will become available to everyone in the coming decades, and not be exclusive to the richest. An important part of artificial intelligence research are the so-called soft computing methods, the most important of which are fuzzy logic, genetic algorithms and neural networks. In this article, the authors present a method of identifying certain traffic signs with the help of the neural networks.

Keywords: *image processing, neural network.*

Összefoglalás

Napjaink egyik legdinamikusabban fejlődő a tudomány területe a mesterséges intelligencia. Habár egyelőre nem szerves része az életünknek a mesterséges intelligencián alapuló megoldások alkalmazása, fejlődését tekintve megállapítható, hogy a következő évtizedekben mindenki számára elérhetővé válik és nem csak egy szűk réteg fogja tudni megengedni magának. A mesterséges intelligencia kutatás fontos részét képezik az úgynevezett lágy számítási módszerek, melyek közül a legfontosabbak a fuzzy logika, a genetikus algoritmusok és a neurális hálózatok. Jelen cikkünkben az szeretnénk bemutatni, hogy neurális hálózat segítségével, hogyan lehet egyes közlekedési táblákat felismerni.

Kulcsszavak: *képfeldolgozás, neurális hálózat.*

1. Bevezetés

A mindennapi forgalomban fontos szerepe van a KRESZ betartásának a balesetmentes közlekedés érdekében. Ehhez a közlekedési táblák felismerése és értelmezése elengedhetetlen.

Számos alkalmazással tudunk különböző objektumokat kimutatni. Ezek között megemlíthetjük a szegmentálást, a Kálmán-filtert vagy éppen a Moravec sarokpont detektort is.

A képfelismerő módszerek legtöbbje a lágy számítási módszereket, ezek közül is leginkább a neurális hálózatokat alkalmazza, azok taníthatósága miatt.

2. Neurális hálózatok működése

2.1. A Neurális hálózat felépítése

A mesterséges neurális hálózat felépítése az emberi agy szerveződését követi. Az emberi agy egyes elkülönült részei felelnek a látásért, hallásért, szaglásért és a különböző működési területekért, amik egymást nem tudják kiváltani. Ebből következik, hogy a mesterséges neurális hálózatot egy adott probléma megoldása céljából hozunk létre.

A hálózat párhuzamos felépítésének, és az eggy-mással kapcsolatban lévő sok kis számláló egy-

ségének köszönhetően nagy működési sebesség érhető el. A neurális hálózat egy bemeneti, egy kimeneti, és a kettő közötti rejtett rétegekből áll, ahogy az **1. ábrán** is látható. A rétegek egymáshoz neuronokon keresztül kapcsolódnak. Minden egyes összeköttetést súlytényezővel láttunk el, amihez egy értéket rendelünk hatékonyság alapján **[1]**.

A hálózatot nem programozzuk, hanem nagyon sok minta által tanítjuk. Ezek a tárolt adatok elosztottan szerepelnek, ezért hibátűrőek, ha egy-egy elemet kiveszünk vagy megváltoztatunk, a rendszer akkor is kiválóan fog működni.

Egy hálózat létrehozásához három feltétel szükséges. Az elsőhöz tartoznak a számítási egységek, az úgy nevezett neuronok, a második a hálózat összeköttetésének topológiája és végül a tanító szabályokat alkalmazó algoritmusok, amik a súlytényezők beállításáért felelősek.

2.2. Az alap neuron felépítése és működése

A neuron a többi neuronokból érkező bemeneteket súlyozott összegként számolja. Az eredményt a következő lépésben a transzfer függvények használatával értelmezzük, amint az a **2. ábrán** is látható, ahol x_i az i -edik bemenet, w_i a súlytényező, $f(x)$ a transzfer függvény.

Ilyen függvény lehet például az egység ugrás függvény. Ha az érték kisebb, mint nulla vagy egyenlő vele, akkor a kimenet nulla lesz, ekkor nem váltódik ki reakció. Viszont, ha magasabb, mint nulla akkor a kimenet az egyes értéket fogja felvenni.

További jellemző átviteli függvények lehet például a lineáris és a logisztikus függvények **[1]**.

2.3. A hálózat tanításának menete

A hálózat egy fekete doboznak tekinthető; a bemenethez egy adott kimenet, célérték tartozik. A

hálózat minden rétege az előző rétegtől kapott adatokat veszi figyelembe és alakítja át, majd továbbítja a következő rétegeknek.

A rétegtől rétegre tartó tanulás közben, a hálózat növeli az adatok összetettségét és részletességét. A háló közvetlenül az adatokból tanul, arra nincs befolyásunk, hogy az adott adatnak mely részleteit fogja megtanulni.

Eszerint megadunk egy tanító mintát a hálózat bemenetének és kiszámítjuk, a kimenetét, majd ezt összehasonlítjuk a célértékkel. Ha nem megfelelő eredményt értünk el, akkor a súlytényezőket addig kell hangolni, amíg nem érjük el a megfelelő célértéket. Ezt nevezzük ellenőrzött vagy felügyelt tanításnak, amikor a háló megtanulja a mintapárok által előhívott be- és kimenet leképvezését.

A hálózat akkor tekinthető működőképesnek, ha a tanító mintákat nem tartalmazó teszt mintákhoz is a megfelelő célértéket fogjuk kapni **[1]**.

2.3.1. A tanítási szabályok

Két tanítási szabály ismert az egyik a Hebb-szabály, ami azt mondja ki, hogy két összekötött neuron között akkor növekszik meg a súlytényező, ha egyidejűleg aktívak, illetve inaktívak. Ekkor megerősödik a kapcsolat, más esetben pedig leépül.

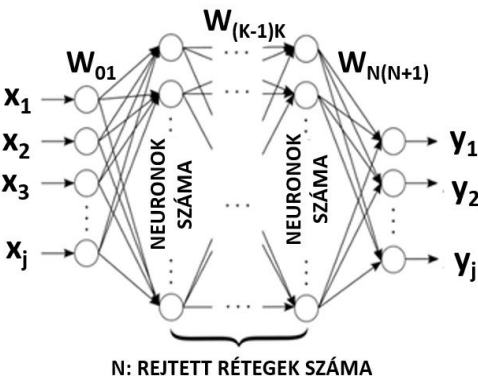
A másik a delta szabály, ekkor az algoritmus a súlyértéket úgy módosítja az optimális érték felé, hogy közben minimalizálja a különbséget a keletkező kimenet és a kívánt célérték között.

3. Esettanulmány

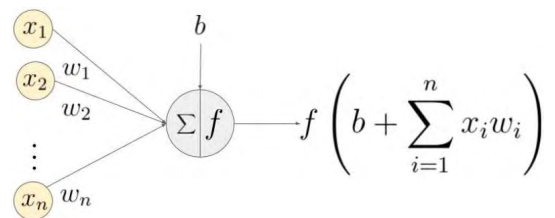
3.1. Tesztkörnyezet

A tanulmány célja négy különböző közlekedési tábla (a stop, az elsőbbségadás kötelező, behajtani tilos valamint a gyalogos átkelő hely) felismerése.

A feladat megoldása során a MATLAB program által nyújtott eszköztárakat vesszük segítségül. A hálózat létrehozásához és tanításához a Deep Learning Toolbox-ot használtuk. A hálózat tanítása a felügyelt tanítás módszerével valósul meg.



1. ábra. A neurális hálózat topológiája



2. ábra. Az alap neuron felépítése

3.2. Adatbázis létrehozása

Első lépésként egy adatbázist hoztunk létre, melynek képeit közlekedési tábla típus szerint almappákba rendeztük. Minden egyes kép saját mappájának nevét kapta. Ezután a képeket véletlenszerűen két adatbázisba osztottuk. Az egyik adatbázist tanító mintaként használtuk, a másik adatbázis a tesztelésre szolgált, ami alapján ellenőriztük a hálózat képfelismerési hatékonyságát.

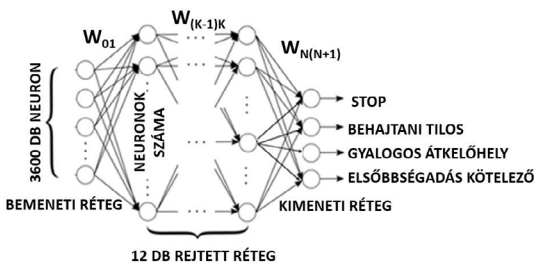
3.3. A neurális hálózat létrehozása

A létrehozott hálózat egy bemeneti, egy kimeneti, valamint a kettő közötti jellemzők felismerésére szolgáló 14 rejtett rétegből áll. A hálózat felépítése a 3. ábrán látható.

A hálózat bemeneti rétege 3600 neuronból épül fel, melyek mindegyike a 60x60-os felbontású bemeneti kép egy-egy képpontjának felel meg. A vizsgálat során azt tapasztaltuk, hogy nagyobb felbontású képméret alkalmazása nem járult hozzá nagyobb pontosságú kép felismeréshez, azonban lassította a hálózat működését.

Az első rejtett réteg az úgynevezett konvolúciós réteg, ahol a neuronok a bemeneti képnek egyszerre csak egy kis részletét dolgozzák fel. A neuronok által végzett művelet nagyon hasonlít a képfeldolgozásból ismert szűrők hatására. A konvolúciós neuronoktól azt várjuk, hogy valamilyen többé-kevésbé absztrakt fogalmat tanuljanak meg (pl. egy él detektálása). Ehhez általában egyetlen neuron nem elég, tehát párhuzamosan több neuront tanítunk ugyanazon az input pozíción. Ebben az esetben 40 neuront használunk, ezért 40 kimeneti mátrixot fogunk kapni [2]. A kimeneten az úgynevezett ReLu aktivációs függvényt alkalmaztuk. (4. ábra) Ennek működése részben biológia eredetű és azzal analóg, ahogy a neuronok aktiválódnak vagy sem.

Ha a kimenet értéke meghaladja a küszöbértéket, akkor a függvény saját magát adja vissza, ha a küszöbértéket nem haladja meg, akkor a kimenet egyértelműen nulla lesz [3].



3. ábra. A létrehozott neurális hálózat topológiája

Ahogy egyre mélyebbre haladunk, úgy válik egyre inkább szükségtelemmé a kis finom részletek megőrzése. Például ha az „él-detektáló” jelzett, hogy talált egy élet, akkor annak a pontos pozíciójára már nem lesz szükségünk a mélyebb rétegekben. Ezért a következő rétegben végezzük el az aktivált kimenetek egyszerűsítést. A fentebb említett rejtett rétegek egymásra építésével egyre absztraktabb fogalmak nyerhetők ki. Jelen esetben elégnek bizonyult háromszor megismételni az előbbi műveletet ahhoz, hogy megfelelő hatékonyságot kapjunk. Így a hálózat 14 rejtett réteget tartalmaz.

Légvégül a kimeneti réteg megadása következik, ami az osztályozást végzi és négy neuronból áll a négy osztálynak megfelelően.

A hálózat tanítása gradiens alapú szélső érték-kereső eljárással és felügyelt tanítással valósult meg.

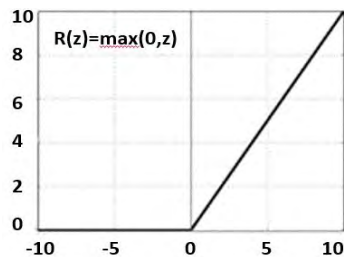
A jelen módszerek használatával másfél perces tanulással 85-90 százalék körüli hatékonysággal képes a létrehozott hálózat felismerni az előbbiekben felsorolt táblákat.

3.4. Alexnet

Összehasonlításképpen a tanítás folyamatát az Alex Krizhevsky által fejlesztett Alexnet nevezetű hálón is lefolytattuk, amivel már 99%-os hatékonyságot lehetett elérni hét perces tanulással.

A különbség abból adódhat, hogy míg az általunk fejlesztett háló tanítása kevesebb, mint háromszáz képpel történt, addig az Alexnet kiképzéséhez több millió kép állt rendelkezésre, valamint az Alexnet felépítése 25 rétegből áll ellentétben az általunk létrehozott 14 rétegű hálóval [4].

Az alábbi mátrixokkal jól szemléltethető a hálózatok hatékonysága, amelyben a késsel jelzett átlók a helyes megoldások számát, halványpirossal a hibák számát jelöli. A baloldali mátrixból világosan látszik, hogy az Alexnet segítségével szinte az összes tesztkép osztályát sikerült felismerni, míg az általunk fejlesztett háló több esetben is mellé detektálta a kép osztályát.



4. ábra. A ReLu függvény képe és képlete

