

Hollósi János

Gépi látás és mesterséges intelligencia eszközök
alkalmazása az autonómjármű fejlesztés területén

doktori tézisek

Prof. Dr. Pozna Claudiu Radu

Egyetemi tanár

Széchenyi István Egyetem

Dr. Ballagi Áron

Egyetemi docens

Széchenyi István Egyetem

Infrastrukturális Rendszerek Modellezése és Fejlesztése
Multidiszciplináris Műszaki Tudományi Doktori Iskola

Motiváció és célkitűzés

A mesterséges intelligencia tudományterületen belül az egyik legdinamikusabban fejlődő, és napjainkban egyre szélesebb körben hangsúlyt kapó terület a mesterséges neurális hálózatok világa. A terület fejlődése során egyre összetettebb és egyre komplexebb feladatok megoldására alkalmas hálózatok jelentek meg, melyek egyre nagyobb hatásfokkal képesek teljesíteni az adott feladat tekintetében. Az idő haladtával az elérhető számítási teljesítmény is fokozatosan növekszik. Ennek köszönhetően már nem jelent problémát sok rétegű, mély szerkezetű, magas paraméterszámú neurális hálózatok akár valós idejű használata sem. Ma már különféle gyakorlati alkalmazásokra, ipari felhasználási lehetőségekre talált a tudományterület. Ide sorolható az önvezető járművek világa is, amely napjaink egyik legnépszerűbb alkalmazási területe a mesterséges intelligenciának és számos más tudományterületnek egyaránt.

Az önvezető járművek fejlesztésének tekintetében a mesterséges intelligencia és a neurális hálózatok gyakorlati alkalmazása több fronton megmutatkozik. Annak ellenére, hogy a neurális hálózatok, és e területen belül is elsődlegesen a konvolúciós neurális hálózat nagy népszerűségnek örvendenek, biztonsági kockázatot is magukban rejtenek. A konvolúciós neurális hálózatok különféle rétegekből épülnek fel, ahol minden rétegnek más-más szerepe van. Feladattól függetlenül általában a cél az, hogy a bemeneti, nagy mennyiségű adat alapján a hálózat képes legyen alacsonyabb szintű absztrakciót végrehajtani azokon. Osztályozási feladatok esetén teszi ezt olyan szinten, hogy a kimeneti oldalon már csak az egyes osztálykategóriákhoz tartozó neuronok szerepeljenek. Ebből látszik, hogy ilyen esetben akár meglehetősen magas szintű adategyszerűsítés történik. Az ilyen jellegű absztrakciónak mindig az a célja, hogy a bemeneti adat számottevő, lényegi részeit reprezentáló jellemzőket legyen képes felismerni, kiemelni a hálózat. Bevált megoldás úgynevezett összevonó rétegek (pooling layer) alkalmazása ilyen jellegű

feladatok esetén. [1][2] Az összevonó réteg csökkenti a bemeneti jellemzőterkép méretét, az által, hogy a vizsgált területen a típusának megfelelő műveletet hajt végre. Több oka is van annak, hogy ilyen jellegű absztrakciós rétegeket alkalmazunk. A hálózatot transláció invariánsá teszi. Ennek köszönhetően például objektum detektálási feladat esetén nem jelent problémát a hálózat számára, ha a felismerendő objektum helyzete változik a bemeneti képsíkon. Másrészt, a jellemzőterkép méretének drasztikus csökkentésével arányosan csökken a hálózat paramétereinek száma, ezáltal a számítási komplexitása is. Ugyanakkor az összevonó réteg természetéből adódóan a bemeneti jellemzők jelentős részét elveszti, ami akár a hatékonyság és a megbízhatóság rovására is mehet. Ez pedig a konvolúciós neurális hálózatok egyik legjelentősebb problémája.

A konvolúciós neurális hálózatok másik aktuális problémája, hogy az alacsony és magas szintű jellemzők közötti kapcsolatot nehezen képes felfedezni és egy rendszerben kezelni. Egyszerű példa erre, hogy ha egy osztályozó hálózatnak egy emberi arcot kell felismernie. Az alacsonyabb szintű jellemzőkért felelős rétegek képesek realizálni az emberi arc különböző komponenseit, úgymint szemek, fülek, orr vagy száj. Viszont a magasabb szintű jellemzőkért felelős rétegekben az egyes elemek közötti reláció nem alakul ki megfelelő módon. Azaz nincs tisztában például azzal, hogy az arc részei hogyan helyezkednek el egymáshoz képest. Emiatt lehetséges egy ilyen hálózat megtévesztése olyan, úgynevezett támadó mintával, ami ugyan tartalmazza az emberi arc tipikus összetevőit, de azok olyan kompozícióban jelennek meg a bemeneti képen, hogy emberi résztvevő számára egyértelműen meghatározható, hogy az nem arcot tartalmaz, csupán részleteit lehet nyomokban felfedezni. Ezt a jelenséget nevezzük Picasso-problémának. [3]

Az olyan területeken, mint az önvezető járművek világa, ahol a rendszer meglehetősen magas biztonságkritikus szintet képvisel, nem megengedhető az, hogy a neurális hálózat teljesítménye ne érjen el egy elvárt minimális szintet, hiszen az

emberi, környezeti vagy gazdasági károkhoz vezethet. Ezt az elvárást azonban fenyegeti az összevonó rétegek alkalmazásából eredő adatredukció ténye, valamint az alacsony és a magas szintű jellemzők közötti kapcsolat kialakításának a hiányából fakadó megtéveszthetőség. A konvolúciós neurális hálózatok tárgyalt problémáinak kiküszöbölése céljából vezette be Geoffrey Hinton és kutatócsapata az úgynevezett kapszula hálózatok elméletét. A kapszula hálózat elméletéhez köthető első publikáció 2011-ben jelent meg. [4] Az elméletben rejlő lehetőségek bemutatásra kerültek, azonban ekkor még nem született meg az ilyen jellegű hálózatok tanításához szükséges mechanizmus. Erre 2017-ig kellett várni, amikor a dinamikus útválasztás algoritmus (dynamic routing algorithm) megjelenése lehetővé tette kapszula alapú hálózatok kialakítását és tanítását. [5][6] Azóta meglehetősen nagy érdeklődésnek örvend a terület és egyre szélesebb körben kerül alkalmazásra, folyamatosan gazdagszik új megoldásokkal. [7-12]

Munkám során a kapszula hálózatok területén végeztem kutatásokat, a gépi látás témakörén történő alkalmazhatóság jegyében. A feldolgozandó adatok és a körüljárt problémakörök jellegét tekintve fontos szempont volt az önvezető járművek területén történő alkalmazhatóság szem előtt tartása. Ezért elsődlegesen vizuális adatok osztályozása, szegmentálása során, másodsorban térbeli pontfelhők rekonstrukciója tekintetében végeztem a vizsgálatokat, és igyekeztem új eredményekkel gazdagítani ezt a friss területet.

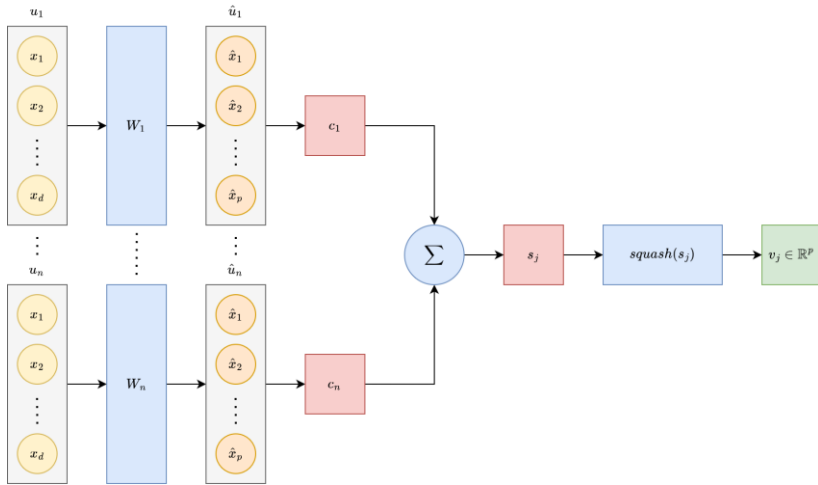
Kutatás során alkalmazott módszerek

A kutatásom során a neurális hálózatok egy új alterületével, a kapszula hálózatok elméletével foglalkoztam. Munkám során az új elmélet gyakorlati alkalmazhatóságát vizsgáltam, elsődlegesen vizuális és térbeli adatok feldolgozása, úgymint osztályozása, szegmentálása vagy rekonstrukciója esetén. A kutatási eredményeimet klasszikus konvolúciós neurális hálózatok alapján készült módszerekkel, state-of-the-art megoldásokkal hasonlítottam össze, ami jól rávilágít a kapszula hálózatok jelenlegi helyzetére, teljesítményére és a terület nyújtotta lehetőségekre, aktuális kérdéseire.

1. táblázat A kapszula és a neuron közötti főbb különbségek összehasonlítása

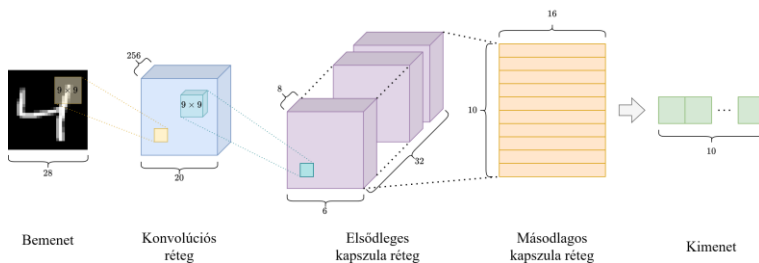
		Kapszula	Neuron
Műveletek	Bemenet	vektor: u_i	skalár: x_i
	Affin transzformáció	$\hat{u}_{ji} = W_{ij}u_i$	-
	Súlyozás	$s_j = \sum_i c_{ij} \hat{u}_{ji}$	$a_j = \sum_i W_i x_i + b$
	Nem-lineáris aktiváció	$v_j = \frac{\ s_j\ ^2}{1 + \ s_j\ ^2} \frac{s_j}{\ s_j\ }$	$y = f(a_j)$
	Kimenet	vektor: v_j	skalár: y

A kapszulát úgy tudjuk elképzelni, mint szorosan összetartozó neuronok csoportját, ahol egy adott jellemzőt, tulajdonságot a csoportba tartozó neuronok együttesen képesek definiálni. Tehát a fő különbség a neurális hálózathoz képest, hogy a számítás alapegysége skalárról vektorra változik, illetve ennek megfelelően a szükséges műveletek is módosulnak. Ezt foglalja össze az 1. táblázat, valamint egy kapszula felépítését az 1. ábra szemlélteti.



1. ábra Kapszula felépítése

Geoffrey Hinton és kutatócsapata a kapszula hálózat elméletének kidolgozása során elsődlegesen vizuális osztályozási feladatok esetén vizsgálták a kapszula hálózatok hatékonyságát, ahol sikerült megmutatni, hogy az új elméletnek köszönhetően jobb hatásfok érhető el, mint az addigi legjobb konvolúciós neurális hálózat alapú megoldások esetén. Az általuk publikált hálózat felépítését szemlélteti a 2. ábra.



2. ábra Kapszula hálózati architektúra

Annak ellenére, hogy a kapszula alapú hálózatok elmélete magasabb hatékonysági fokot ígér és jobban adaptálódik a biológiai neurális hálózatok működéséhez, ennek

az elméletnek is léteznek árnyoldalai, valamint jelenlegi nyitott kérdései, még feltáratlan részterületei.

A kapszula hálózatok jellemzően sokkal lassabbak, mint a konvolúciós neurális hálózatok, köszönhető ez az erős komplexitás-beli különbségnek. Ennek oka egyrészt, hogy az alapegység nem skalár, hanem tetszőleges dimenzió-számú vektor, ahol a vektortér mérete, valós gyakorlati feladatok esetén akár 16, 32 vagy még több dimenziós is lehet. Ez egyértelműen bonyolítja a számítást és növeli a számítás végrehajtásához szükséges időt vagy hardver kapacitást.

A hagyományos konvolúciós neurális hálózatok regularizációs technikáiból eredő információ vesztes kivédését célzó dinamikus útválasztási algoritmus a másik tényezője annak, hogy a kapszula alapú hálózatok miért lassabbak. Ugyan az útválasztási algoritmus által meghatározott kapcsolati súlyértékeknek köszönhetően magas hatásfok érhető el, de a számítás jellege végett, mivel a pontosítás több iteráción keresztül és komplexebb adatokon zajlik, a számítási idő is jelentősen megnövekszik.

A kapszula hálózatok, jellegüknél adódóan vektorokat szolgáltatnak kimenetként. Ezek a kimeneti vektorok jellemzően magas dimenzió számmal rendelkeznek. Az elmélet létrehozói például 16-dimenziós kimeneti vektorokat definiáltak, tehát minden osztály esetén 16 tulajdonság realizálja az osztályhoz tartozás valószínűségét. Ez a gyakorlatban azt jelenti, hogy egy konvolúciós neurális hálózathoz képesta kimenet mérete a 16-szorosára növekszik. Ha egy osztályozási feladat esetén 10 előre definiált osztály közül kell eldönteni, hogy a bemeneti vizsgált kép melyikhez tartozik, akkor konvolúciós neurális hálózat esetén a kimeneten 10 darab valós szám jelenik meg, melyek egy-egy valószínűségi értéket takarnak. Ellenben, a kapszula alapú hálózat esetén a kimenet 160 darab valós számot fog tartalmazni, tehát 16-szorosára nőtt a kimenet mérete. Emiatt olyan feladatok esetén, ahol a lehetséges osztályok száma egyébként is magas, nagy

mértékben megbonyolíthatja a kimenet generálását a kapszula alapú megoldás, ahol a végrehajtási idő elfogadhatatlan érték fölé emelkedhet.

Az elmélet kidolgozói csak egy szűk területen, egyszerű osztályozási feladatok esetén vizsgálták a kapszula alapú hálózatok teljesítőképességét. Adódik a kérdés, hogy más jellegű feladatok esetén, például detektálás vagy szegmentálás, milyen hatékonysággal alkalmazhatók az új elmélet nyújtotta lehetőséget. Ezen felül célszerű megvizsgálni, hogy más összetételű adathalmazok esetén mekkora határfok érhető el a konvolúciós neurális hálózatokhoz képest. Ide érhető magasabb osztályszámú adatkészletek, vagy olyan alacsony osztályszámú minták, ahol a különféle osztályokba tartozó objektumok közötti eltérés szűk tartományon mozog. Célszerű megvizsgálni magasabb dimenziószámú adatokon is a kapszula hálózatok működését is. Érdekes lehet annak a vizsgálata, hogy 3-dimenziós adatokon, például térbeli pontfelhők esetén hogyan adaptálható a kapszula hálózatok elmélete és milyen eredmény érhető el más megoldásokhoz viszonyítva.

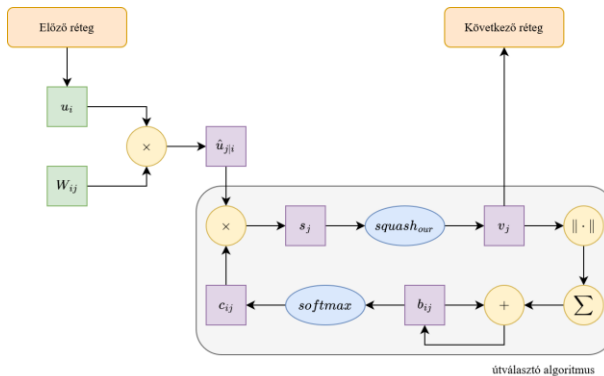
Tézisek

Első tézis

Kapszula alapú hálózatok esetén az egyes kapszula rétegek közötti, az információ áramlást szabályozó útválasztás együttható súlyértékeinek optimalizálásához új egyszerűsített iteratív dinamikus útválasztási algoritmust és hozzá kapcsolódó, a vektorok terén értelmezett aktivációs függvényt vezettem be. A vizsgált esetek tekintetében az általam javasolt megoldás hatékonyabb és gyorsabb módszernek bizonyult, mint az elmélet szerzői által bevezetett dinamikus útválasztás algoritmus.

Az első tézist alátámasztó saját publikációk: [HJ1], [HJ2], [HJ3], [HJ4], [HJ5]

A kapszula alapú hálózatok hatékonyságának egyik kulcskérdése az alacsonyabb és a magasabb rétegekben található kapszulák közötti kapcsolat súlyértékeinek a realizálása. Ennek köszönhetően az alacsonyabb szinten megjelenő jellemzők és a magasabb szintű jellemzők közötti kapcsolat, összetartozás kialakul a tanítási folyamat során.



3. ábra Javasolt útválasztó algoritmus

Bár az elmélet szerzői által javasolt dinamikus útválasztás algoritmus nagyon hatékony és jó eredménnyel szolgál, a sebessége elmarad a konvolúciós neurális hálózatok esetén tapasztalhatókhöz képest, ami gátolja a kapszula hálózatok gyakorlati felhasználhatóságát, elterjedését. Ezért egy új, egyszerűsített iteratív útválasztó algoritmust javasoltam, valamint hozzá tartozó, a vektorok terén értelmezett aktivációs függvényt írtam fel, amely paramétereizhetőségének köszönhetően jobban adaptálódik a kívánt feladathoz. A javasolt algoritmus blokkvázlatát a 3. ábra mutatja.

2. táblázat Mérési hibák összehasonlítása

	MNIST	F-MNIST	S.NORB	CIFAR10	SVHN	GTSRB
Sabour és társai	0,45%	9,00%	11,15%	31,49%	8,05%	2,73%
Saját megoldás	0,41%	8,35%	8,81%	28,26%	6,90%	2,22%

3. táblázat Visszahívás (recall) pontszám

	MNIST	F-MNIST	S.NORB	CIFAR10	SVHN	GTSRB
Sabour és társai	0,9939	0,9159	0,9056	0,6790	0,8958	0,9513
Saját megoldás	0,9948	0,9124	0,9138	0,6925	0,9012	0,9519

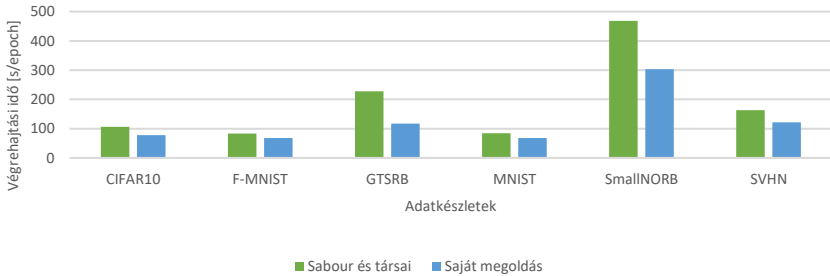
4. táblázat Dice pontszám

	MNIST	F-MNIST	S.NORB	CIFAR10	SVHN	GTSRB
Sabour és társai	0,9940	0,9153	0,9052	0,6748	0,9012	0,9576
Saját megoldás	0,9949	0,9118	0,9131	0,6882	0,9046	0,9576

5. táblázat F1-pontszám

	MNIST	F-MNIST	S.NORB	CIFAR10	SVHN	GTSRB
Sabour és társai	0,9940	0,9159	0,9056	0,6790	0,9084	0,9711
Saját megoldás	0,9949	0,9124	0,9138	0,6925	0,9122	0,9727

A bevezetett útválasztó algoritmus hatékonyságát több különféle és jellegében eltérő adatkészlet esetén hasonlítottam össze a dinamikus útválasztás algoritmusával. A mérési eredményeimet a 2-5. táblázatok foglalják össze, ahol a tesztelés során mért különféle mérőszámok eredménye láthatók.



4. ábra Végrehajtási idő a különféle adatkészleteken

Fontos szempont volt, hogy az algoritmus ne csak hatékonyságát tekintve teljesítsen jól, de képes legyen jobb sebesség elérésére. Ezért a méréseket garantáltan ugyanolyan körülmények között végeztem el, így a számítási idők alapján jól látható a komplexitás-beli különbség, melyet a 4. ábra szemléltet.

Második tézis

Neurális hálózatok tanító adatkészletét célzó, gradiens alapú zajterheléses támadási módszerek nagy hatásfokkal képesek rontani a konvolúciós neurális hálózatok hatékonyságán. Megmutattam, hogy különféle korszerű támadási módszerek alkalmazása esetén a kapszula hálózatok osztályozási hatékonysága kisebb mértékben romlik, mint a vizsgált state-of-the-art konvolúciós neurális hálózatok esetén.

A második tézist alátámasztó saját publikációk: [HJ6], [HJ7], [HJ8], [HJ9]

A konvolúciós neurális hálózatok egyik fontos problémája, a hálózatok megtéveszthetősége, becsaphatósága. Napjainkban egyre szélesebb körben jelenik meg a neurális hálózatok valós gyakorlatban történő alkalmazása, ahol bizonyos esetekben a rendszer tévedésének minimalizálása kritikus, vagy a megtévesztése nem megengedhető következményekkel járna. Ezért különösen fontos a hálózatok tévedhetőségének a vizsgálata. Kutatásaim során megmutattam, hogy a kapszula hálózatok elmélete szerint kialakított rendszerek nagyságrenddel jobb ellenállóságot mutatnak a gradiens alapú zajterheléses támadási megoldásokkal szemben.

Kutatásom során state-of-the-art jellegű konvolúciós neurális hálózatok hatékonysága és a kapszula hálózatok hatékonysága közötti különbségre világítottam rá, különféle támadási módszerek alkalmazása során. A kapszula hálózatok területén az elmélet kidolgozói által javasolt dinamikus útválasztás algoritmus, valamint az általam bevezetett egyszerűsített iteratív útválasztás közötti különbség is látható. Ezeket, különféle mérőszámok mentén, a 6-9. táblázatok összegzik.

6. táblázat Mért hatékonyságok a támadási módszerek esetén

	-	FGSM	FFGSM	TPGD
SimpleCNN	0,9999	0,2521 (-74,79%)	0,2494 (-75,05%)	0,5743 (-42,56%)
Wide-ResNet	1,0000	0,2294 (-77,06%)	0,1763 (-82,37%)	0,4057 (-59,43%)
CapsNet	0,9991	0,7324 (-26,70%)	0,6504 (-34,90%)	0,6822 (-31,72%)
CapsNet (saját)	0,9993	0,7941 (-20,53%)	0,6280 (-37,16%)	0,8346 (-16,48%)

7. táblázat *F1-pontszám*

	-	FGSM	FFGSM	TPGD
SimpleCNN	0,9994	0,4183	0,1683	0,6191
Wide-ResNet	0,9987	0,2738	0,1981	0,5280
CapsNet	0,9973	0,6086	0,7626	0,4793
CapsNet (saját)	0,9955	0,7587	0,7969	0,8385

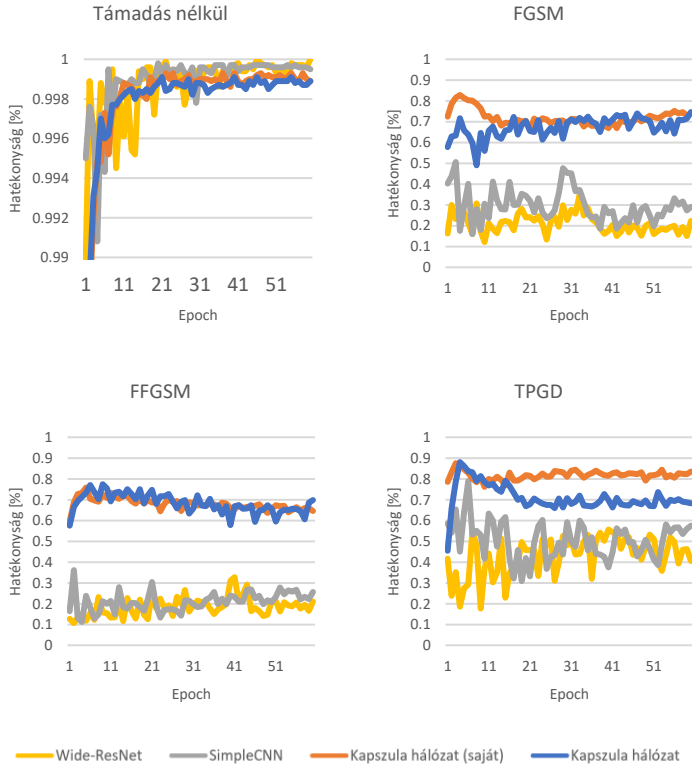
8. táblázat *Pontosság (precision) pontszám*

	-	FGSM	FFGSM	TPGD
SimpleCNN	0,9987	0,3536	0,1390	0,5541
Wide-ResNet	0,9985	0,1757	0,1264	0,3793
CapsNet	0,9956	0,5153	0,7135	0,4114
CapsNet (saját)	0,9944	0,7062	0,7668	0,8138

9. táblázat *Visszahívás (recall) pontszám*

	-	FGSM	FFGSM	TPGD
SimpleCNN	0,9989	0,3858	0,1549	0,5841
Wide-ResNet	0,9986	0,2538	0,1962	0,5045
CapsNet	0,9956	0,5711	0,7273	0,4568
CapsNet (saját)	0,9944	0,7268	0,7694	0,8146

Az 5. ábra a tanítási folyamat során mért hatékonyságokat mutatja, ahol a különbség szemmel látható. Támadás nélkül minden megoldás jól teljesít, a különbségek elenyészőek. A támadási módszerek esetén a neurális hálózat alapú megoldások hatékonysága elmarad a kapszula alapú megoldással szemben.



5. ábra Hatékonyságok a támadási módszerek esetén

Harmadik tézis

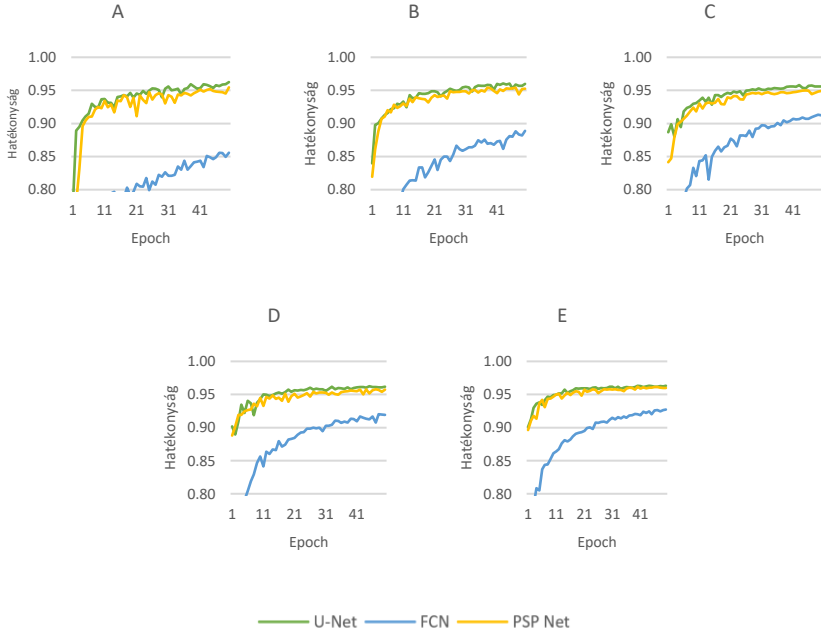
Kapszula hálózatok elméletén alapuló szegmentáló hálózatokat valósítottam meg. Megmutattam, hogy számítógéppel generált virtuális tanítóminták esetén a U-Net és PSP Net kapszula hálózatok a vizsgált esetekben több, mint 95%-os hatékonyságot képesek elérni, ami lehetővé teszi a gyakorlati alkalmazásukat. Rávilágítottam, hogy FCN alapú kapszula hálózat hatékonysága a virtuális minták növelésével javítható. Így a kapszula hálózatok elmélete alkalmazható olyan esetben, amikor rendelkezésre álló valós tanítóminták száma korlátozott.

A harmadik tézist alátámasztó saját publikációk: [HJ10], [HJ11], [HJ12], [HJ13], [HJ14]

A kapszula hálózatok elméletét elsődlegesen objektum osztályozási feladatokra hozták létre és vizsgálták meg a hatékonyságát ilyen jellegű konvolúciós neurális hálózatokhoz viszonyítva. Kutatásom során az elméletet szegmentálási feladat esetén alkalmaztam, ahol egy speciális esetben, úgynevezett virtuális adatkészletek alkalmazása során mutattam meg a kapszula hálózatokkal elérhető teljesítményt. Ebben az megközelítésben a tanítás során nem csak valós, hanem számítógéppel generált virtuális mintákon történik a hálózatok kiképzése. Mindez hasznos olyan feladatok esetén, amikor nem áll rendelkezésre kellő mennyiségű tanítóminta. A kialakított adatkészletek összetételét a 10. táblázat foglalja össze. Három eltérő és jól bevált neurális hálózati architektúrát alakítottam át a kapszula hálózatok elmélete alapján, majd végeztem méréseket a létrehozott adatkészletek felhasználásával.

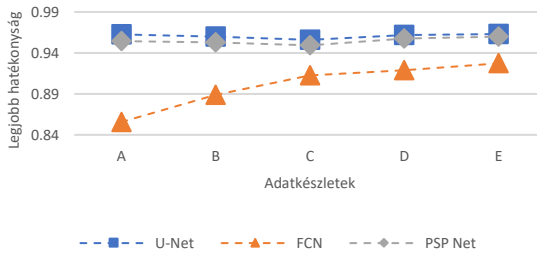
10. táblázat Létrehozott adatkészletek összetétele

Adatkészlet neve	Tanítókészlet		Tesztkészlet	
	Virtuális minták	Valós minták	Virtuális minták	Valós minták
<i>A</i>	0	500	0	125
<i>B</i>	500	500	0	250
<i>C</i>	1500	500	0	500
<i>D</i>	1500	1000	0	625
<i>E</i>	1500	1500	0	750



6. ábra Szegmentálás hatékonysága a különféle adatkészleteken

A tanítás során elért hatékonyságot a 6. ábra, az adatkészletenként elért legjobb eredményeket a 7. ábra szemlélteti. Látható, hogy több esetben sikerült megmutatni, hogy a kapszula hálózatok elmélete alkalmas ilyen feladatok végrehajtására.



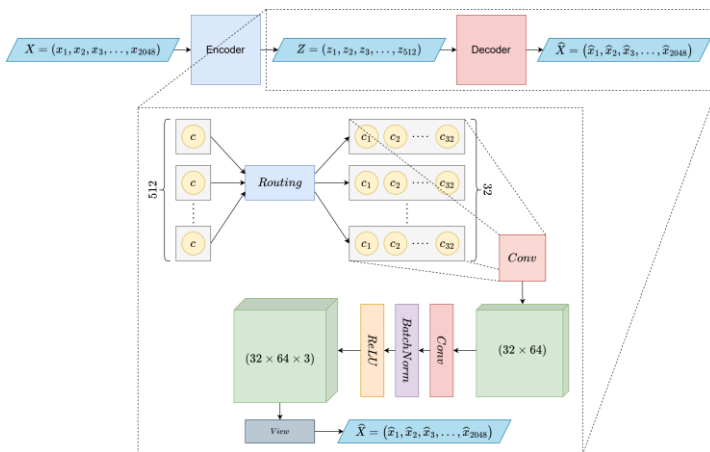
7. ábra Legjobb hatékonyságok változása az adatkészleteken

Negyedik tézis

Pontfelhők rekonstrukciójára képes kapszula alapú autoenkóder hálózatot valósítottam meg és hasonlítottam össze konvolúciós neurális hálózat alapú megoldásokkal. Megmutattam, hogy a kapszula hálózatok elmélete alkalmazható térbeli adatok feldolgozása során. A vizsgálatok döntő többségében a kapszula alapú megoldás hatékonyabb, mint a konvolúciós neurális hálózatok, valamint az esetek több, mint harmada esetén a kapszula alapú megoldással jobb hatékonyság érhető el kevesebb tanítási idővel.

A negyedik tézist alátámasztó saját publikációk: [HJ15], [HJ16], [HJ17], [HJ18]

A kapszula hálózatok elméletét alkalmaztam térbeli adatok, pontfelhők feldolgozása során, ahol azt vizsgáltam a kutatásom során, hogy ilyen jellegű megoldással milyen hatások érhető el 3-dimenziós adatok rekonstrukciója esetén. Ehhez javasoltam egy kapszula alapú dekóder hálózatot, ami tetszőleges enkóder hálózattal ötvözve képes rekonstrukciós feladatok végrehajtására. Az általam kialakított kapszula alapú dekóder hálózatot a 8. ábra mutatja be.



8. ábra Javasolt kapszula dekóder hálózat

11. táblázat Rekonstrukció mérési hibája (forgatás nélkül)

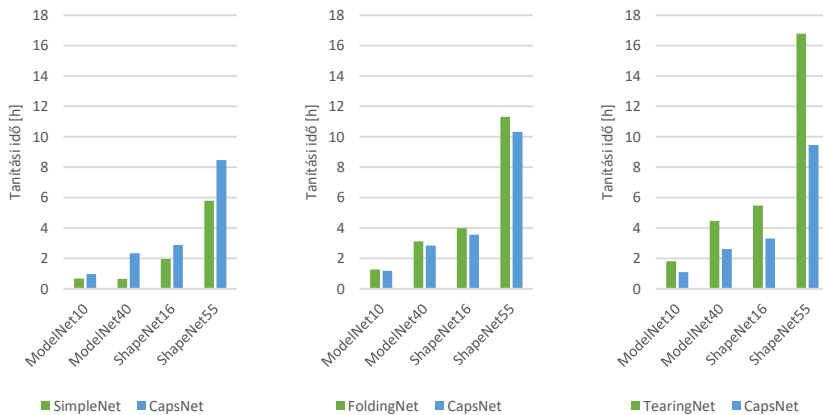
	ModelNet10	ModelNet40	Shapenet16	ShapeNet55
Achlioptas és tsai	3,7346	3,2378	2,7896	3,0954
Saját megoldás	2,4936	2,7500	1,9490	2,2465
Yang és tsai	2,3230	2,5099	1,6566	2,0187
Saját megoldás	1,7244	2,2486	1,6014	1,9531
Pang és tsai	1,8104	1,2012	1,4819	1,6082
Saját megoldás	1,7625	1,4882	1,2009	1,7708

A mérések során több state-of-the-art autoenkóder hálózattal hasonlítottam össze az általam javasolt kapszula alapú megoldást. A mérési eredményeket az 11. és a 12. táblázatok foglalják össze. Előbbi esetben a bemeneti pontfelhőn transzformációt nem hajtottam végre, utóbbi során véletlenszerű forgatásokat eszközöltem, ezzel komplexebb esetben is feltárva a hatékonyságokat. A mérési eredmények is alátámasztják a megoldásom hatékonyságát, ahol a vizsgált esetek több, mint 70%-a esetén nagyobb hatékonyságot mutatott a kapszula alapú megoldásom.

12. táblázat Rekonstrukció mérési hibája (forgatással)

	ModelNet10	ModelNet40	Shapenet16	ShapeNet55
Achlioptas és tsai	6,2518	5,3737	3,7729	3,8842
Saját megoldás	4,7726	4,3178	3,0158	3,5781
Yang és tsai	4,0988	3,3240	2,5575	2,7561
Saját megoldás	4,0979	3,8986	3,0966	2,5966
Pang és tsai	3,5793	2,2440	1,8911	1,7647
Saját megoldás	3,5935	2,5450	1,7760	2,2420

A kutatás során megvizsgáltam a hálózatok komplexitását is, amire jó mérőszám a tanítási folyamat során mért végrehajtási idő. Ezeket a 9. ábra mutatja, ahol látható, hogy a három esetből kettőnél, a két legkomplexebb neurális hálózat esetén, a kapszula alapú megvalósításom gyorsabban teljesített. Így az esetek 37,5%-a esetén az általam bevezetett megoldás nem csak jobb, de rövidebb tanulási idővel képes azt elérni.



9. ábra Tanítási idő a különféle hálózatok esetén

Konklúzió

Doktori értekezésemben a kapszula hálózatok elméletének kutatásával, gyakorlati alkalmazhatóságával foglalkoztam. Összefoglaltam a konvolúciós neurális hálózatok aktuális problémáit, melynek nyomán a kapszula hálózatok elmélete megszülethetett. Részleteztem az új elméleti megközelítés megértéséhez szükséges tágabb háttérismeretet. Kutatómunkám eredményeként 4 tézist fogalmaztam meg, melyekkel a kapszula hálózatok területét szeretném gazdagítani. Javaslatot tettem egy egyszerűsített iteratív útválasztó algoritmus és a vektorok terén értelmezett paramétereztető aktivációs függvényre. A bevezetett módszerrel, a mérési eredmények alapján, sikerült hatékonyabb és gyorsabb megoldást javasolnom, mint az elmélet szerzői. Megmutattam, hogy gradiens alapú zajterheléses támadási módszerek esetén kapszula alapú megoldások jobban teljesítenek, mint a korszerű konvolúciós neurális hálózatok. Így biztonságkritikus rendszerek esetén célszerű kapszula rétegek alkalmazásával növelni a rendszer robusztusságát. Rávilágítottam, hogy kapszula hálózatok elmélete alapján módosított szegmentáló neurális hálózatok kellő hatékonyságot képesek elérni virtuális adatkészleteken történő tanítás során. Így jól alkalmazhatók olyan szituációban, amikor a feladat jellegéből adódóan tanítóminta csak korlátozottan férhető hozzá. Végül megkonstruáltam egy kapszula alapú dekóder hálózatot, amely képes térbeli pontfelhők esetén hatékony rekonstrukciós feladatot végrehajtani, a bemutatott esetek jelentős részében jobb teljesítménnyel és gyorsabban, mint a vizsgált korszerű neurális hálózatok. A téziseimben megfogalmazott, a kutatómunkám során elért eredményeket korszerű, több esetben state-of-the-art megvalósításokkal hasonlítottam össze, különféle szempontok szerint. A mérési eredmények alátámasztják, hogy a kapszula hálózatok sok potenciált rejtnek magukban és érdemes kiemelt hangsúlyt fektetni a területre. Az általam javasolt megoldások, a mérési eredményeken alapuló összehasonlító elemzések alapján, a vizsgált esetek tekintetében hatékony és több esetben jobb megoldást jelentenek, mint a konvolúciós hálózat alapú megvalósítások.

Irodalomjegyzék

A téziszüzetben hivatkozott saját publikációk

- [HJ1] **J. Hollósi**, Á. Ballagi & C. R. Pozna, “Simplified Routing Mechanism for Capsule Networks,” Algorithms, vol. 16, no. 7, p. 336, Jul. 2023, doi: 10.3390/a16070336.

Független hivatkozások száma: 0

- [HJ2] **J. Hollósi**, “Útválasztó algoritmusok hatékonysági vizsgálata a kapszula hálózatok területén,” Digitális Járműipari Kutatások a Széchenyi István Egyetemen - Konferenciakiadvány 2021, Győr, Magyarország, p. 64, 2021.

Független hivatkozások száma: 0

- [HJ3] **J. Hollósi** & Á. Ballagi, “Training Capsule Networks with Various Parameters,” 2019 IEEE 13th International Symposium on Applied Computational Intelligence and Informatics (SACI), May 2019, pp. 191–196. doi: 10.1109/SACI46893.2019.9111574.

Független hivatkozások száma: 1

- [HJ4] **J. Hollósi**, “Kapszula hálózatok útválasztási algoritmusának hatékonysági vizsgálata,” Tavasz Szél 2019 Konferencia. Nemzetközi Multidiszciplináris Konferencia - Konferenciakiadvány, Debrecen, Magyarország.

Független hivatkozások száma: 0

- [HJ5] **J. Hollósi**, Á. Ballagi, G. Kovács, S. Fischer & V. Nagy, “Face Detection Using a Capsule Network for Driver Monitoring Application,” Computers, vol. 12, no. 8, p. 161, Aug. 2023, doi: 10.3390/computers12080161.

Független hivatkozások száma: 0

- [HJ6] **J. Hollósi**, Á. Ballagi & C. R. Pozna, “Capsule Network based 3D Object Orientation Estimation,” in 2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Jul. 2023, pp. 1–6. doi: 10.1109/ICECCME57830.2023.10252762

Független hivatkozások száma: 0

- [HJ7] **J. Hollósi**, Á. Ballagi & C. R. Pozna, “Objektum orientáció becslés kapszula hálózatok felhasználásával,” JKK Digitális Járműipari Kutatások a Széchenyi István Egyetemen – Mesterséges Intelligencia a mobilitásban konferencia 2022, Győr, Magyarország, 2022.

Független hivatkozások száma: 0

- [HJ8] **J. Hollósi**, Á. Ballagi & C. R. Pozna, “Kapszula hálózatok hatékonysági vizsgálata ellenséges minták esetén,” AUTONÓM JÁRMŰVEK - Jövőformáló járműipari kutatások Konferenciakiadvány, Győr, Magyarország, 2021, pp. 77–88.

Független hivatkozások száma: 0

- [HJ9] **J. Hollósi**, E. Horváth & C. R. Pozna, “Two-Stage Racetrack Segmentation Method Using Color Feature Filtering and Superpixel-Based Convolutional Neural Network,” 2018 IEEE 12th International Symposium on Applied Computational Intelligence and Informatics (SACI), May 2018, pp. 000131–000136. doi: 10.1109/SACI.2018.8440968.

Független hivatkozások száma: 0

- [HJ10] **J. Hollósi**, R. Krecht, N. Markó & Á. Ballagi, “Improving the efficiency of neural networks with virtual training data,” Hungarian Journal of Industry and Chemistry, vol. 48, no. 1, pp. 3–10, Jul. 2020, doi: 10.33927/hjic-2020-02.

Független hivatkozások száma: 0

- [HJ11] **J. Hollósi** & Á. Ballagi, “Capsule Networks for Object Segmentation Using Virtual World Dataset,” Sensors & Transducers Journal, vol. 244, no. 5, pp. 20–27, 2020.

Független hivatkozások száma: 0

- [HJ12] **J. Hollósi** & Á. Ballagi, “Training Capsule Networks with Virtual Dataset for Object Segmentation,” Advances in Signal Processing and Artificial Intelligence, Proceedings of the 2nd ASPAI’ 2020 Conference, Berlin, Németország: International Frequency Sensor Association (IFSA) Publishing, S. L., 2020, pp. 145–149.

Független hivatkozások száma: 0

- [HJ13] **J. Hollósi** & Á. Ballagi, “Training Neural Networks with Computer Generated Images,” 2019 IEEE 15th International Scientific Conference on Informatics, Poprád, Szlovákia, 2019, pp. 155–160. doi: 10.1109/Informatics47936.2019.9119273.

Független hivatkozások száma: 2

- [HJ14] **J. Hollósi**, R. Krecht & N. Markó, “Neurális hálózatok hatékonyságának növelése számítógéppel generált tanítási mintákkal,” AUTONÓM JÁRMŰVEK – A JÖVŐ KÖZLEKEDÉSE : Konferenciakiadvány 2019, Győr, Magyarország, 2020, pp. 48–59.

Független hivatkozások száma: 0

- [HJ15] **J. Hollósi**, Á. Ballagi & C. R. Pozna, “Capsule-based Autoencoder Network for Pointcloud Reconstruction,” 2023 IEEE 21st World Symposium on Applied Machine Intelligence and Informatics (SAMI), Jan. 2023, pp. 121–126. doi: 10.1109/SAMI58000.2023.10044532.

Független hivatkozások száma: 0

- [HJ16] **J. Hollósi**, C. R. Pozna & Á. Ballagi, “Kapszula alapú autoenkóder hálózatok rekonstrukciós hatékonysági vizsgálata pontfelhők esetén,” Mobilitás és környezet - Jövőformáló járműipari kutatások konferencia, Győr, Magyarország, 2022.

Független hivatkozások száma: 0

- [HJ17] **J. Hollósi**, “Kapszula alapú mesterséges neurális hálózatok vizsgálata az önvezető járművek tekintetében,” AUTONÓM JÁRMŰVEK – Intelligens közlekedési rendszerek a fenntarthatóságért Konferenciakiadvány 2020, Győr, Magyarország, 2020, pp. 7–14.

Független hivatkozások száma: 0

- [HJ18] **J. Hollósi** & C. R. Pozna, “Improve the Accuracy of Neural Networks using Capsule Layers,” 2018 IEEE 18th International Symposium on Computational Intelligence and Informatics (CINTI), Nov. 2018, pp. 000015–000018. doi: 10.1109/CINTI.2018.8928194.

Független hivatkozások száma: 7

További saját publikációk

- [HJ19] G. Teschner, C. Hajdu, **J. Hollósi**, N. Boros, A. Kovács & Á. Ballagi, “Digital Twin of Drone-based Protection of Agricultural Areas,” 2022 IEEE 1st International Conference on Internet of Digital Reality (IoD), Jun. 2022, pp. 99–104. doi: 10.1109/IoD55468.2022.9986763.
- [HJ20] C. Hajdu, **J. Hollósi**, G. Teschner, Á. Ballagi & A. Kovács, “Protection of Agricultural Areas from Wild Animals with Drone-based Intelligent System,” 12th IEEE International Conference on Cognitive Infocommunications (CogInfoCom 2021) : Proceedings, 2021, pp. 651–656.
- [HJ21] C. Hajdu, **J. Hollósi**, R. Krecht, Á. Ballagi & C. R. Pozna, “Economical Mobile Robot Design Prototype and Simulation for Industry 4.0 Applications,” 2020 IEEE 3rd International Conference and Workshop in Óbuda on Electrical and Power Engineering (CANDO-EPE), Nov. 2020, pp. 155–160. doi: 10.1109/CANDO-EPE51100.2020.9337786
- [HJ22] T. Dobay, **J. Hollósi**, & N. Markó, “Objektum detektáló neurális hálózatok alkalmazási lehetőségei az autonóm járművek területén,” AUTONÓM JÁRMŰVEK - Jövőformáló járműipari kutatások Konferenciakiadvány, Győr, Magyarország, 2021, pp. 27–32.
- [HJ23] Á. Titrik, I. Lakatos & **J. Hollósi**, “Hangvezérlő rendszer integrálása a járműbe a kényelmi és biztonsági szint emelése érdekében,” XVI. Innováció és Fenntartható Felszíni Közlekedés Konferencia (IFFK 2022), Budapest, Magyarország, 2022.
- [HJ24] I. Lakatos, Á. Titrik & **J. Hollósi**, “Demonstrating the Options for Automated Advanced Selective Waste Gathering,” 2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY), Sep. 2018, pp. 167–172. doi: 10.1109/SISY.2018.8524707.
- [HJ25] **J. Hollósi**, N. Markó & Á. Ballagi, “Forgalomtechnikai terelő bója észlelésen alapuló pozíció és orientáció meghatározása,” Digitális Járműipari Kutatások a Széchenyi István Egyetemen - Konferenciakiadvány 2021, Győr, Magyarország, 2021, pp. 56–63.
- [HJ26] N. Markó & **J. Hollósi**, “Környezetérzékelés közös gerinchálózattal rendelkező neurális hálózatgyüttes segítségével,” AUTONÓM JÁRMŰVEK – Intelligens közlekedési rendszerek a fenntarthatóságért Konferenciakiadvány 2020, Győr, Magyarország, 2020, pp. 75–80.

- [HJ27] N. Markó & **J. Hollósi**, “Háromdimenziós objektumdetektálás implementációja,” AUTONÓM JÁRMŰVEK – WORKSHOP-SOROZAT, Győr, Magyarország, 2021, pp. 78–84.
- [HJ28] **J. Hollósi** & Á. Ballagi, “Kamera-Lidar Kalibráció,” AUTONÓM JÁRMŰVEK – WORKSHOP-SOROZAT, Győr, Magyarország, 2021, pp. 123–130.
- [HJ29] C. Hajdu, E. Horváth, P. Körös, Á. Ballagi & **J. Hollósi**, “Esettanulmány autonóm tesztjármű fejlesztésére,” AUTONÓM JÁRMŰVEK – A JÖVŐ KÖZLEKEDÉSE : Konferenciakiadvány 2019, Győr, Magyarország, 2020, pp. 38–47.
- [HJ30] N. Markó & **J. Hollósi**, “Gépi látás mélytanulás segítségével,” AUTONÓM JÁRMŰVEK – A JÖVŐ KÖZLEKEDÉSE : Konferenciakiadvány 2019, Győr, Magyarország, 2019, pp. 25–29.
- [HJ31] I. Lakatos, Á. Titrik, & **J. Hollósi**, “Emelt szintű szelektív hulladékgyűjtés automatizálási lehetőségének bemutatása,” IFFK 2018: XII. Innováció és fenntartható felszíni közlekedés, Budapest, Magyarország, 2018, pp. 383–388.
- [HJ32] E. Horváth, C. R. Pozna, C. Hajdu & **J. Hollósi**, “A use case of the simulation-based approach to mobile robot algorithm development,” 2016 IEEE 14th International Symposium on Applied Machine Intelligence and Informatics (SAMi), Jan. 2016, pp. 311–314. doi: 10.1109/SAMI.2016.7423026.
- [HJ33] R. Krecht, C. Hajdu & **J. Hollósi**, “Önvezető funkciók megvalósítására alkalmas jármű Unreal Engine 4 alapú szimulációja,” XXVIII. Nemzetközi Gépészeti Konferencia – OGÉT 2020 : 28th International Conference on Mechanical Engineering, Kolozsvár, Románia, 2020, pp. 227–230.
- [HJ34] I. Lakatos, Á. Titrik & **J. Hollósi**, “Integration of the voice control system into the vehicle to increase comfort and safety levels,” Proceedings in Conference of Informatics and Management Sciences : The 7th International Virtual Conference, Zsolna, Szlovákia, 2018, pp. 120–123.
- [HJ35] C. R. Pozna, E. Horváth & **J. Hollósi**, “The inverse kinematics problem, a heuristical approach,” 2016 IEEE 14th International Symposium on Applied Machine Intelligence and Informatics (SAMi), Jan. 2016, pp. 299–304. doi: 10.1109/SAMI.2016.7423024.

A téziszűzetben hivatkozott egyéb publikációk

- [1] A.Zafar, M.Aamir, N. M. Nawi, A. Arshad, S. Riaz, A. Alruban, A. K. Dutta & S. Almotairi, “A Comparison of Pooling Methods for Convolutional Neural Networks,” *Applied Sciences*, vol. 12, no. 17, p. 8643, Aug. 2022, doi: 10.3390/app12178643.
- [2] H. Gholamalinezhad & H. Khosravi, “Pooling Methods in Deep Neural Networks, a Review.” *arXiv*, Sep. 16, 2020. doi: 10.48550/arXiv.2009.07485.
- [3] V. Gliozzi, G. L. Pozzato & A. Valese, “Combining neural and symbolic approaches to solve the Picasso problem: A first step,” *Displays*, vol. 74, p. 102203, Sep. 2022, doi: 10.1016/j.displa.2022.102203.
- [4] G. E. Hinton, A. Krizhevsky & S. D. Wang, “Transforming Auto-Encoders,” in *Artificial Neural Networks and Machine Learning – ICANN 2011*, T. Honkela, W. Duch, M. Girolami & S. Kaski, Eds., in *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer, 2011, pp. 44–51. doi: 10.1007/978-3-642-21735-7_6.
- [5] S. Sabour, N. Frosst & G. E. Hinton, “Dynamic routing between capsules,” in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, in *NIPS’17*. Red Hook, NY, USA: Curran Associates Inc., Dec. 2017, pp. 3859–3869.
- [6] S. Sabour & G. E. Hinton & N. Frosst, “Matrix capsules with EM routing,” presented at the *Sixth International Conference on Learning Representations*, Vancouver, Canada, Apr. 2018.
- [7] T. Vijayakumar, “Comparative Study of Capsule Neural Network in Various Applications,” *JAICN*, vol. 01, no. 01, pp. 19–27, Sep. 2019, doi: 10.36548/jaicn.2019.1.003.
- [8] R. Shi & L. Niu, “A brief survey on Capsule Network,” in *2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, Dec. 2020, pp. 682–686. doi: 10.1109/WIIAT50758.2020.00103.
- [9] J. Li, Q. Zhao, N. Li, L. Ma, X. Xia, X. Zhang, N. Ding & N. Li, “A Survey on Capsule Networks: Evolution, Application, and Future Development,” in *2021 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS)*, Dec. 2021, pp. 177–185. doi: 10.1109/HPBDIS53214.2021.9658349.

-
- [10] O. El Alaoui-Elfels & T. Gadi, “From Auto-encoders to Capsule Networks: A Survey,” *E3S Web Conf.*, vol. 229, p. 01003, 2021, doi: 10.1051/e3sconf/202122901003.
- [11] M. Kwabena Patrick, A. Felix Adekoya, A. Abra Mighty & B. Y. Edward, “Capsule Networks – A survey,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 1, pp. 1295–1310, 2022, doi: 10.1016/j.jksuci.2019.09.014.
- [12] F. D. S. Ribeiro, K. Duarte, M. Everett, G. Leontidis & M. Shah, “Learning with Capsules: A Survey.” *arXiv*, Jun. 06, 2022. doi: 10.48550/arXiv.2206.02664.