

Doktori (PhD) értekezéstézisei

Hatékony Neurális Háló Alapú Architektúrák Fejlesztése Orvosi Adatok Feldolgozásához

Bogacsovics Gergő

Témavezető: Dr. Hajdu András



DEBRECENI EGYETEM

Informatikai Tudományok Doktori Iskola

Debrecen, 2024

Tartalomjegyzék

1. Bevezetés	1
2. Kétlépcsős architektúra tervezése a betegségek pontos előrejelzésére	5
3. A hagyományos képfeldolgozó módszerekkel és a CNN-ek által kinyert jellemzők kombinálása	9
4. Teljesen konvolúciós együttes modell automatikus sejtsegmentációhoz	13
5. Diverz együttes modellek építése a tagmodellek közötti hasonlóság büntetésével	16
Hivatkozások	20
Publikációs lista	26

1. Bevezetés

Az elmúlt évtizedekben az orvostudomány jelentős átalakulásnak és paradigmaváltásnak volt kitéve. Az új képalkotó technikák és a korszerűbb orvosi berendezések megjelenésével a különböző területek a napi szinten keletkező adatmennyiség ugrásszerű növekedését tapasztalták. A mélytanuló algoritmusok megjelenése lehetővé tette ezen adatok feldolgozását valamint megbízható megoldások kidolgozását. A disszertációban több újszerű módszer kerül bemutatásra a mélytanuló modellek pontosságának és megbízhatóságának további növelésére, és használjuk őket különböző problémák megoldására a klinikai területen.

A disszertációban szereplő megoldások két csoportra oszthatók. Az első csoportba tartozó eljárások célja a neurális hálók pontosságának és használhatóságának javítása a hagyományos módszerek és a mélytanuló technikák kombinálásával. Megmutatjuk, hogy elő tudunk állítani olyan modelleket, amelyek mindkét megközelítés előnyeit magukban hordozzák. Az ilyen módon létrehozott modellek ugyanis nem csak kisebb adathalmazokon képesek működni, de robusztusabbak illetve rugalmasabbak is, mint a hagyományos módszerek. A disszertációban bemutatott módszerek második csoportja olyan hatékony együttes modellek (ensemble) kifejlesztésére összpontosít, amelyek a hagyományos ensemble megközelítéseknel pontosabbak és hatékonyabbak.

Először az elméleti modellek és neurális hálók hatékony fúziójára összpontosítottunk. Megmutattuk, hogy mindkét megközelítésnek megvannak a maga előnyei és hiányosságai. Ebből a célból egy új, kétlépcsős architektúrát fejlesztettünk ki, amely egyesíti a két megközelítés előnyeit, és olyan be-

tegségek terjedésének előrejelzésére használtuk, mint az influenza és a COVID-19. Megmutattuk, hogy az architektúra először egy neurális hálót tanít be egy Kermack-McKendrick-modell [1] közelítésére, amely egy SIR-modell [2], majd a hálót valós adatokon tanítjuk tovább. E lépéseket követve a háló először megtanulhatja az elméleti modell közelítését, majd a második lépésben a modell valós adatokon történő tanítása következik. Az első lépés során a hálózatnak lehetősége van arra, hogy a szabályozottabb elméleti modell segítségével megtanulja a probléma legfontosabb aspektusait. Ezt követően, a második lépés során a súlyok finomhangolása történik a pontosság maximalizálása érdekében.

A disszertáció következő részében részletesen bemutatunk a hagyományos képfeldolgozással kinyert jellemzők fontosságát, amelyeket a neurális hálók és a mélytanulás térnyerésével nagyon kevés kutatás használt fel az elmúlt években. Kitértünk arra, hogy ezek kinyerése szakmai személyzetet igényel, valamint általában nehézkes és időigényes. Megjegyeztük azonban, hogy olyan releváns jellemzőket is megragadhatnak, amelyek felismerésével a mélytanuló technikáknak gondjaik lehetnek. Ezen okból kifolyólag ezeknek a hagyományos módszerekkel kinyert jellemzőknek a konvolúciós neurális hálók (CNN-ek) által kinyert jellemzőkkel történő hatékony kombinációjára összpontosítottunk. Megmutattuk, hogy ez a kombináció korántsem triviális, ezért három különböző változatot mutattunk be: egy sekély (egyrétegű), egy mélyebb (többrétegű) megközelítést, valamint egy olyan architektúrát, amely két osztályozót tanít be, mindkettőt az adott jellemzőkre. Ez utóbbi architektúra izoláltan tanítja be a két osztályozót, majd egyszerű átlagolással aggregálja kimeneteiket, mint egyfajta "mini-ensemble".

Ezután bemutattunk egy adathalmazt és egy új együttes keretrendszert a sejtek hatékony szegmentálására digitalizált Pap kenet képeken. Bemutattuk, hogy a keretrendszerünk, ellentétben a hagyományos módszerekkel, saját súlyozással rendelkezik. A keretrendszer egy FCN-32 [3] architektúrát használ gerinceként, és az eredeti bemeneti képpel együtt néhány előzetesen betanított FCN-32, FCN-16 és FCN-8 hálózat kimeneteit is megkapja. Ezen keretrendszer több változatát is kiértékeltek, ahol mindegyik változat a három FCN-modell különböző részalalmazát használta bemenetként. Megmutattuk, hogy teljesítmény tekintetében mind az eredeti modelleket, mind más módszereket is felül tudtunk múlni.

A disszertáció utolsó részében a diverz együttes modellek építésére összpontosítottunk. Megemlítettük, hogy a diverzitás a legújabb tanulmányok szerint [4, 5] az ensemble általános teljesítménye szempontjából fontos metrika. Ezt követően részleteztük, hogy ezzel ellentétben a diverzitást általában a legtöbb kutatás nem veszi figyelembe az együttesek építése során. Ezután több olyan újszerű keretrendszert mutattunk be, amelyek az összes megfelelő tagmodellt egyszerre tanítják be. A keretrendszerek a költségfüggvénybe egy további regularizációs kifejezést építenek be, amely a tagmodellek diverzitásának méréséért felelős. Megmutattuk, hogy a modellek hasonlóságának mérésének egyik lehetséges módja az egyes modellek által kinyert jellemzővektorok felhasználása. Ennek oka, hogy két azonos vagy nagyon hasonló jellemzőkkel működő modell redundánsnak tekinthető, míg a nagyon eltérő jellemzőket használó modellek nagy valószínűséggel eltérő módon működnek, ami értékesebbé teszi őket az együttes építése során. Ezután két fő keretrendszert mutattunk be, amelyek a

koszinusz hasonlóságra és a hisztogramveszteségre [6] támaszkodtak a különböző tagmodellek által kinyert vektorok hasonlóságának mérésére. Ezenfelül ugyancsak bemutattuk a kereprendszerek két kiterjesztését, amelyek lehetővé teszik azok alkalmazását súlyozott költségfüggvény vagy olyan tagmodellek használata esetében is, amelyek különböző hosszúságú jellemzővektorokat állítanak elő.

2. Kétlépcsős architektúra tervezése a betegségek pontos előrejelzésére

Az elméleti modelleket gyakran használják a kutatásban, mivel viszonylag egyszerű, de hatékony és kifejező módját nyújtják a különböző jelenségek modellezésének. Az elméleti modellek ráadásul akkor is használhatók, ha a rendelkezésre álló adathalmaz kevés elemből áll, amennyiben a probléma lényege ismert, mint például egy adott betegség természetes terjedését leíró képlet. A fő probléma azonban az, hogy összetettebb modellek használatakor a matematikai leírás is bonyolultabbá válik. Emiatt az elméleti modellek általában kerülnek a nagyfokú komplexitást, és a lehető legegyszerűbb definícióra törekednek. Ez matematikailag kezelhetőbb modellt eredményez, ugyanakkor általában nem optimális teljesítményhez is vezet. Az ilyen típusú modellek másik hátránya, hogy az elméleti modellek általában túl merevek, ami még nehezebbé teszi alkalmazásukat az adott problémára.

Másrészt a mélytanulás alapú megoldások a rugalmasságukról valamint arról ismertek, hogy a nagyobb számú paraméterek miatt hatékonyabban illeszkednek a tanító adathalmazhoz. Például kimutatták [7, 8], hogy a neurális hálókkal olyan modelleket hozhatunk létre, amelyek képesek matematikai függvények közelítésére. Ez azt jelenti, hogy a neurális hálók a betanítás után képesek úgy viselkedni, mint az elméleti modellek, miközben megtartják legnagyobb előnyüket, azaz a rugalmasságukat, mely összességében jobb teljesítményt eredményez a bonyolultabb problémák esetében.

Kutatásunk célja egy olyan keretrendszer megalkotása volt, amely lehetővé teszi, hogy olyan mélytanuló modelleket épít-

sünk, amelyek hatékonyan ötvözik mind az elméleti modellek, mind a hagyományos neurális hálók főbb előnyeit. Nevezetesen olyan modellek létrehozására törekedtünk, amelyek kisebb adathalmazok használata esetén is alkalmazhatók, matematikailag robusztusabb alapokkal rendelkeznek, ugyanakkor rugalmasabbak, mint az egyszerű elméleti modellek, és végső soron jobb teljesítményt érnek el. Ehhez egy újszerű, kétlépcsős megközelítést mutattunk be, amely először a neurális háló alapú modellt az eredeti elméleti modell közelítésére tanítja be. Ezt követően, a második lépésben a valós adathalmazon finomhangoltuk a súlyait.

1.1. Tézis *Kétlépcsős architektúrát dolgoztam ki az elméleti modellek és a mélytanuló módszerek hatékony kombinálására. Az architektúra először egy neurális hálót tanít az elméleti modell által generált szintetikus adatokon. A következő lépésben a modell tovább tanul a valós adatokon, mely során az némileg eltérhet az eredeti elméleti modelltől annak érdekében, hogy maximalizálja teljesítményét az eredeti adathalmazon.*

A kétlépcsős architektúra alkalmazhatóságának és hatékonyságának validálása érdekében két betegség, az influenza és a COVID-19 terjedésének előrejelzésével mértük a teljesítményét. Elméleti modellként egy egyszerű SIR-modellt [2], az úgynevezett Kermack-McKendrick-modellt [1], mélytanuló modellként pedig egy egyszerű, három rejtett réteggel rendelkező neurális hálót használtunk. Azért választottuk ezt a konkrét elméleti modellt, mert bár a modell viszonylag egyszerű, mégis matematikailag megalapozott, hiszen differenciálegyenleteken alapul. A neurális háló alapú modell esetében is főként egyszerűbb architektúrákra összpontosítottunk a kif-

nomultabbak, például a rekurrens hálók (RNN) vagy a hosszú-rövid távú memória (LSTM) használata helyett. Ezt azért tettük, hogy megmutassuk, hogy az általunk javasolt architektúra nem specializált, és számos problémára és területen alkalmazható. Kísérleteztünk az időablak növelésével is a modell bemenetei és kimenetei esetében, hogy még tovább növeljük a modell teljesítményét.

1.2. Tézis *A javasolt kétlépcsős architektúra több változatát dolgoztam ki különböző méretű bemeneti és kimeneti ablakok használatával, és értékeltem őket ki az influenza és a COVID-19 esetek terjedésének előrejelzésére.*

A kísérleti eredmények alapján a modell felülmúlta mind az eredeti elméleti modellt, mind a kizárólag az eredeti adathalmazon betanított neurális hálókat. Kísérleteink után azt is megállapítottuk, hogy eredményeink és megfigyeléseink szerint a kétlépcsős megközelítésünk szerinti tanítás után kapott modell rugalmasabb volt, mint az eredeti elméleti modell, és pontosabb, mint a kizárólag az eredeti adathalmazon tanított neurális hálók. Megmutattuk ugyanis, hogy a modellt csak a COVID-19 első hullámán tanítva a kivont jellemzők ténylegesen segíthettek a modellnek a második hullám előrejelzésében, jelentősen felülmúlva az eredeti elméleti modellt, amely csak egyetlen hullámot tudott megjósolni.

Végül arra is kitértünk, hogy ez a módszer a transzfer tanulás (transfer learning) egy másik típusának tekinthető. Nevezetesen ahelyett, hogy nagy, előre betanított neurális hálókat használnánk, egy elméleti modellt választhatunk, amely nagyjából hasonlít a valós adatokra, és a neurális hálót ezen a függvényen taníthatjuk be. Azt is megjegyeztük, hogy a java-

solt keretrendszer általános jellegéből adódóan nincsenek korlátozások az elméleti modellel kapcsolatban, mivel az bármilyen modell lehet, amennyiben a kimenetei összehasonlíthatók egy neurális háló kimeneteivel. Ez azt jelenti, hogy az architektúra potenciálisan számos tudományágban használható különböző problémák megoldására. Végül azt is kiemeltük, hogy az elméleti modellnek nem kell tökéletesen illeszkednie a valós adatokhoz. Ehelyett elegendő, ha a valós adatok legfontosabb jellemzői (terjedés, jelleg stb.) kódolva vannak az elméleti modellben. A részletes kísérleti eredmények és megállapításaink folyóiratcikk formájában kerültek publikálásra [9].

3. A hagyományos képfeldolgozó módszerekkel és a CNN-ek által kinyert jellemzők kombinálása

Az olyan szembetegségek, mint a diabéteszes retinopátia (DR) és a diabéteszes makulaödéma (DME) komoly veszélyt jelentenek napjainkban. A világ népességének jelentős részét érintik. Ezért rendkívül fontos olyan megbízható megoldások kifejlesztése, amelyek pontosan felismerik ezeket a betegségeket. 2019-ben a becslések szerint 1,5 millió halálesetet okozott közvetlenül a cukorbetegség, és az Egészségügyi Világszervezet jelentése szerint világszerte 422 millió ember szenved a betegségben [10]. Így az olyan megbízható rendszerek kifejlesztése, amelyek ezeket a betegségeket korai stádiumukban képesek felismerni, a látáskárosodás előrehaladásának lelassításával az életminőség javulásához vezethet.

A jelenlegi módszerek némelyike a szakértők által, hagyományos képfeldolgozó módszerekkel kinyert jellemzőket használja a rendellenességek felismerésére. Ezeknek a megoldásoknak a kifejlesztése azonban gyakran időigényes, és általában nem elég pontosak. Emiatt számos olyan megoldást javasoltak az utóbbi időkben, amelyek mélytanulás alapú technikákat használnak, ahol nincs szükség szakértőkre a jellemzők kinyeréséhez, így a modellek fejlesztése egyszerűbbé és gyorsabbá válik. Másrészt ezeknek a módszereknek is vannak hátrányai. Nevezetesen, általában figyelmen kívül hagyják a rendkívül értékes, hagyományos módszerekkel kinyert jellemzőket, amelyek potenciálisan hozzájárulhatnak a pontosabb előrejelzésekhez.

[11] egy olyan megoldást javasolt, amely a konvolúciós neu-

rális háló (CNN) által kinyert jellemzőket kombinálja a hagyományos módszerekkel kinyert jellemzőkkel, hogy javítsa az osztályozási teljesítményt. Az ötlet [12]-ből származik, ahol a szerzők azt állították, hogy egy CNN képes automatikusan számos fontos lokális jellemzőt kinyerni a képekből egy konvolúciós szűrővel. A helyi jellemzők mellett azonban a globális jellemzők, például a kontúrjellelmzők és a strukturális jellemzők is fontosak, és számos képfeldolgozási feladatban meghatározó szerepet játszanak. Ráadásul a DR és a DME betegségekre utaló jelek közül sok jól leírható a kontúrjellelmzőkkel, ami azt jelenti, hogy e jellemzők figyelembevételével javíthatjuk a CNN-alapú szűrőrendszer végső pontosságát.

Ebben a disszertációban több újszerű megoldást mutatunk be a [11]-ben közölt módszer kiterjesztésével és néhány korszerű neurális háló alapú architektúra alkalmazhatóságának vizsgálatával. Szisztematikusan megvizsgáltuk a javasolt módszertan előnyeit és hátrányait. Megmutattuk, hogy az eredeti keretrendszer rendelkezett néhány figyelemre méltó korláttal, és bemutattunk több olyan módosítást, amelyek megoldást nyújtottak ezekre a problémákra. A [11]-ben javasolt architektúra a hagyományos módszerekkel kinyert jellemzőket és az AlexNet modell által kinyert jellemzőket úgy kapcsolta össze, hogy csak egyetlen súlykészsletet használt, hogy az összekapcsolt vektort közvetlenül a modell előrejelzéseivé alakítsa át. Megjegyeztük, hogy a jellemzőknek ez az azonnali redukciója előrejelzésekké nem ideális, mivel a keretrendszernek nincs lehetősége összetettebb minták megtanulására.

Ennek megfelelően kísérleteinket a jellemzők mélyebb kombinációjával kezdtük. Megállapítottuk, hogy ennek egyik lehetséges módja az, hogy nagyobb számítási kapacitást biz-

tosítunk a keretrendszer azon részének, amely a hagyományos képfeldolgozó módszerekkel kinyert jellemzőkkel dolgozik, ami úgy valósítható meg, hogy ezen jellemzőket több rétegen keresztül áramoltatjuk át. Ily módon a keretrendszer a megnövekedett számítási kapacitás-nak köszönhetően több, a hagyományos módszerekkel kinyert jellemzőkben jelen lévő jellemzőt és mintát lehet képes megragadni.

2.1. Tézis *Kifejlesztettem a [11]-ben közölt módszer egy továbbfejlesztett változatát. Ugyanazon bemeneti képet használva a hagyományos módszerekkel kinyert jellemzők kiszámításra kerülnek és sűrű rétegek egy sorozatán haladnak keresztül míg a modell egy másik ága az eredeti konvolúciós neurális háló alapú architektúrát használja a kép feldolgozására. Mindegyik ág előállítja a saját kimenetét, amelyeket aztán átlagolva kapjuk meg a végső predikciót.*

Az új keretrendszer bevezetése után megvitattuk, hogy bár megoldja a jellemzővektorok sekély kombinációjának problémáját, a jellemzők most külön-külön kerülnek feldolgozásra. Megjegyeztük, hogy ez nem feltétlenül jelent problémát, mivel a továbbfejlesztett keretrendszer most egyfajta "mini-együttesként" működik, ahol minden modell külön-külön működik, és a kimeneteket egyszerűen átlagoljuk. Ugyanakkor azt is kiemeltük, hogy így a keretrendszernek nincs esélye olyan minták felismerésére, amelyekhez egyszerre van szükség néhány hagyományos módszerekkel kinyert és néhány CNN-jellemző jelenlétére. Ennek megoldására a keretrendszer egy másik változata került megvalósításra.

2.2. Tézis *Kifejlesztettem a [11]-ben közölt módszer egy olyan továbbfejlesztett változatát, amely az utolsó konvolúciós*

réteg által és a hagyományos módszerekkel kinyert jellemzőket egyaránt felhasználja. Mindkettőt együttesen, egyszerre áramoltatja át a következő rétegeken, lehetőséget adva a neurális hálónak összetettebb minták felismerésére, ahol a hagyományos módszerekkel és a CNN által kinyert jellemzők egyidejű jelenlétére van szükség.

Végezetül több egyéb eljárást is implementáltunk, amelyekkel összehasonlítottuk megoldásaink teljesítményét. Kiértékeljük a neurális hálók teljesítményét a hagyományos módszerekkel kinyert jellemzők nélkül, különféle gépi tanulási modellekét, amelyek csak a hagyományos módszerekkel kinyert jellemzőket használják, valamint más korszerű megoldásokat is megvizsgáltunk. Az átfogó összehasonlítás érdekében több különböző metrikát is használtunk. A javasolt architektúrák kiértékelése a diabéteszes makulaödéma és a diabéteszes retinopátia felismerésére több forrásból származó adathalmazok felhasználásával történt. A javasolt architektúrák mindegyike felülmúlta a hagyományos módszerekkel kinyert jellemzőket nem használó neurális hálókat, a csak hagyományos módszerekkel kinyert jellemzőkre támaszkodó megoldásokat és egyéb korszerű megközelítéseket is.

A bemutatott módszereket, valamint a részletes kísérleti eredményeket folyóiratcikkben publikáltuk [13].

4. Teljesen konvolúciós együttes modell a- utomatikus sejtszegmentációhoz

Az automatizált szűrőrendszerek létfontosságú szerepet játszanak a betegellátás színvonalának és az orvosi munkafolyamatok minőségének növelésében. A legszélesebb körben használt automatikus megoldások közé tartozik a Hologic ThinPrep Imaging System [14] és a Focal Point Slide Profiler [15], amelyeket az Egyesült Államok Élelmiszer- és Gyógyszerügyi Hivatala (FDA) is jóváhagyott. Mindkét megoldásnak azonban van néhány jelentős hátránya. A Hologic ThinPrep képalakító rendszer ugyanis csak a ThinPrep Pap-teszt keneteket tudja elemezni, melyeknek jóval magasabbak a költségei, mint a leggyakrabban alkalmazott Papanicolaou kenetvizsgálatnak [16], míg a Focal Point Slide Profiler csak a legalacsonyabb kockázatú kenetek 25%-át képes kizárni. Kutatásunk végső célja egy olyan teljesen automatikus szűrőrendszer kifejlesztése volt, amely képes felismerni a rákos sejteket a digitalizált Pap kenet képeken. A rendszer kifejlesztése folyamán arra törekedtünk, hogy leküzdjük e két megoldás korlátait. Azaz egy olyan rendszer kifejlesztése volt a célunk, amely a leggyakrabban alkalmazott Pap kenet képeket használja, és pontosabban rangsorolja azokat a kockázati szint szerint. Automatikus szűrőrendszerünk célja az egyes sejtek megtalálása és szegmentálása a nagy felbontású bemeneti képen. Ebben a disszertációban bemutattuk ennek a rendszernek egy olyan komponensét, amely a bemeneti képeken lévő sejtek pontos és megbízható szegmentálásáért felelős, valamint egy saját adathalmazt.

Az adathalmaz a Debreceni Egyetem Klinikai Központ Pa-

tológiai Osztályával folytatott együttműködés keretében készült. A disszertációban ugyancsak kifejtettük, hogy a rendelkezésre álló adathalmazok kis méretűek, és részleteztük, hogy nincs tudomásunk más hasonló, nyilvánosan elérhető adathalmazról a sejtek szegmentálásának feladatára. A manuális annotáció hosszadalmas és időigényes volta miatt megjegyeztük, hogy minden kísérlet csak az adathalmaz jelenleg elérhető részéhez férhetett hozzá. Adathalmazunk végleges, átfogóbb változatát, amely összesen 3565 darab 2000×2000 pixeles képet és a hozzájuk tartozó szegmentációs maszkokat tartalmazta [17]-ben tettük közzé.

Ez a disszertáció egy egyedülálló módszert is bemutatott a pontos és megbízható együttes modellek létrehozására, amelyek egy automatikus szűrőrendszer részeként használhatók a sejtek szegmentálására. A kapcsolódó elméleti háttér bemutatását a hagyományos ensemble megközelítések legnagyobb problémáinak bemutatásával kezdtük. Nevezetesen kiemeltünk néhány példát, ahol az egyes módszerek különböző sejteket ismertek fel. Taglaltuk, hogy ezekben az esetekben a hagyományos megközelítések nem lennének képesek hatékonyan kombinálni ezeket a kimeneteket, mivel nincs átfedés a generált szegmentációs maszkok között. Ezután bemutattuk a saját módszerünket, amely egy teljesen konvolúciós neurális háló (FCN) [3] architektúrát használ gerincként. Megmutattuk, hogy ennek a megközelítésnek a fő előnye, hogy a hagyományos együttes modellekkel ellentétben nem csak a tagmodellek kimeneteit képes aggregálni, hanem azok előrejelzéseit figyelmen kívül hagyva akár saját szegmentációs maszkot is képes generálni.

3.1. Tézis *Kidolgoztam egy egyedülálló együttes modellt, amely egy módosított FCN-32 architektúrát használ a gerincéként. A keretrendszer néhány előre betanított FCN-architektúra kimeneteit kapja meg az adott bemeneti képpel együtt, hogy predikciókat végezzen. A keretrendszer több változatát is megvalósítottam, amelyek különböző számú FCN-modellt használnak bemenetként.*

Megmutattuk, hogy az architektúra FCN-32 gerincének és saját súlykészletének köszönhetően a keretrendszer képes a jellemzők és minták felismerésére mind a bemeneti képet, mind pedig a tagmodellek előrejelzéseit használva. Ezen minták segítségével automatikusan el tudja dönteni, hogy milyen súlyt rendeljen az egyes tagokhoz, vagy akár azt is felismerheti, ha egy adott tagmodell alulteljesít a bemeneti kép egy adott jellemzőjének jelenlétében. A javasolt architektúrák különböző változatainak a saját adatállományunkon történő kiértékelése után megjegyeztük, hogy a javasolt keretrendszer jelentősen felülmúlta mind az egyes tagmodelleket, mind a többi korszerű megoldást is. Az előállított szegmentálási maszkok minősége megközelítette, sőt egyes esetekben meg is haladta az emberi teljesítményt.

Az ismertetett, átfogó adatállományhoz tartozó folyóirat-cikket közlésre benyújtottuk [17]. Az előre betanított FCN-hálózatok és a bemeneti kép kombinált kimeneteit használó ensemble keretrendszer konferenciacikk formájában került publikálásra [18].

5. Diverz együttes modellek építése a tag-modellek közötti hasonlóság büntetésével

Bár az elmúlt évtizedben számos mélytanulás alapú technikát javasoltak különböző orvosi feladatok megoldására, mint például emlőrák [19, 20] és agydaganat [21, 22] felismerésére vagy bőrrák [23] azonosítására, az egyes modellek és architektúrák teljesítménye alkalmazásonként nagyon eltérő. Gyakori példaül, hogy egy adott modell jól teljesít egy feladat esetében, míg egy másik alkalmazás során visszaesik a teljesítménye. Ez a jelenség népszerűsítette az együttes modellek használatát [24, 25, 26], amelyek több különböző modellt egyesítenek.

Ezen algoritmusok fő előnye, hogy a különbözőképpen működő modellek integrálásával az ensemble potenciálisan jobban teljesíthet, mint az egyes modellek. Ennek oka, hogy azokban az esetekben, amikor egy modell rosszul teljesít, a többi modell, amennyiben azok kellően pontosak és többségben vannak, a megfelelő irányba billentheti az együttes modell kimeneteit. Az ilyen technikák alkalmazásával olyan modelleket fejleszthetünk ki, amelyek ellenállóbbak a kiugró értékekkel vagy a bemenetekben fellelhető szabálytalanságokkal szemben, és jobb általánosítási képességekkel rendelkeznek. A hagyományos technikák azonban általában nem veszik figyelembe a diverzitást, vagyis azt, hogy az ensemble modelljei mennyire különbözőek. Ez komoly hiányosság, hiszen többször kimutatták [4, 5], hogy a diverzitás jótékony hatással van a pontos és jól teljesítő ensemble modellek összeállítására.

Ebben a disszertációban egy új, ensemble alapú módszer mutattunk be az agydaganatok megbízható és pontos osztályo-

zására T1-súlyozott, kontraszt javított, mágneses rezonancia képalkotó (MRI) eszközzel készített felvételek alapján. Több olyan technikát is bemutattunk, amelyek a hagyományos ensemble megközelítésekkel ellentétben közvetlenül mérik és optimalizálják a tagmodellek diverzitását a tanítás során. Az ensemble modellek felépítéséhez az együttesek tagjaként a leggyakrabban használt korszerű CNN architektúrákat vettük figyelembe, mint például az AlexNet [27, 28], MobileNetv2 [29, 30, 31], EfficientNet [32, 33], és ShuffleNet v2 [31, 34].

A bemutatott módszer két részre E és D osztja a CNN architektúrákat, ahol E a konvolúciós rétegeket, D pedig a modell teljesen kapcsolt rétegeit jelöli. Bevezettük azt az elképzelésünket, hogy bármely CNN-modell esetében az E kimeneteit (más szóval a jellemzővektorokat) egy magas dimenziójú látens vektortér néhány látens vektoraként kezelhetjük. Azt is részleteztük, hogy amennyiben ezek a jellemzővektorok különbözőek (azaz a különböző tagok különböző jellemzőhalmazokon operálnak), úgy az együttes diverzitása várhatóan megnő. Ezután bemutattuk a keretrendszerünket, amely a tagmodellek által kinyert jellemzővektorok (E kimenetei) közötti különbségek maximalizálására törekszik.

4.1. Tézis *Kifejlesztettem egy újszerű együttes keretrendszert, amely minden tagját egyszerre tanítja és a költségfüggvényben egy egyedi regularizációs kifejezést használ, hogy maximalizálja a tagmodellek által kinyert jellemzővektorok közötti különbséget. A tagmodellek esetében több különböző neurális hálóval kísérleteztem és a koszinusz hasonlóságot használtam a jellemzővektorok közötti hasonlóság mérésére.*

A disszertációban felvázoltunk néhány problémát is a ko-

szinusz hasonlóság használatával kapcsolatban a jellemzővektorok összehasonlítására a keretrendszerünkben. Megmutattuk ugyanis, hogy nem kezeli jól a sorrendhez kapcsolódó hasonlóságokat, például amikor két jellemzővektor pontosan ugyanazokat az értékeket tartalmazza, csupán kissé permutált módon. Ilyen esetekben a koszinusz hasonlóság nem érzékeli a jellemzők közötti teljes hasonlóságot, ami nem optimális eredményekhez vezethet. Azt is megmutattuk, hogy a probléma egyik lehetséges megoldása a hisztogram alapú megközelítés használata, amely nem támaszkodik a jellemzővektorok elemeinek sorrendjére.

4.2. Tézis *Kidolgoztam a keretrendszer továbbfejlesztett változatát, amely a koszinusz hasonlóság helyett a hisztogramvesztést használja a kivont jellemzővektorok közötti hasonlóság mérésére. A keretrendszer ezen változata minimalizálja annak a valószínűségét, hogy két különböző modellből származó jellemzővektorok jobban hasonlítanak egymásra, mint az azonos modell által előállított vektorok.*

Azt is kiemeltük, hogy a keretrendszerek egyik fő hátránya az volt, hogy csak olyan tagmodellekkel tudtak működni, amelyek azonos dimenziójú jellemzővektorokat hoztak létre. Részleteztük, hogy ez a hasonlósági függvények korlátja volt, mivel definíció szerint a belső szorzat, amely mindkét hasonlóság kiszámításában fontos lépés, két különböző méretű vektor esetén nem definiált.

4.3. Tézis *Kifejlesztettem egy olyan kiterjesztést, amely lehetővé teszi mindkét keretrendszert számára, hogy azok különböző hosszúságú jellemzővektorokkal rendelkező tagmodelleket használjanak azáltal, hogy a vektorokat egy közös dimen-*

zióra hoztam, mielőtt megmértem a köztük lévő hasonlóságot. A megközelítés érvényességéről az adathalmazon végzett gondos kiértékeléssel győződtem meg és megállapítottam, hogy a kiterjesztés semmilyen negatív hatást nem gyakorolt a keretrendszerre.

A disszertációban azt is említettük, hogy az orvosi adathalmazok jellegéből adódóan rendkívül fontos, hogy egy technika kiegyensúlyozatlan adathalmazok esetében is alkalmazható legyen. Ilyenkor egy súlyozott költségfüggvényt használunk, mely során minden egyes mintára kiszámítunk egy-egy együtthatót a hozzájuk tartozó osztálycímkék ritkaságától függően. Ezután megmutattuk, hogy a súlyozott költségfüggvény használata nincs negatív hatással a javasolt keretrendszerekre, mivel a bevezetett regularizációs kifejezés egyenlő hangsúlyt fektet az egyes tagmodellekre és az azok által előállított jellemzővektorokra, figyelmen kívül hagyva az osztálycímkét. Mindkét keretrendszert és azok kiterjesztéseit egy agydaganatok MRI-felvételeit tartalmazó adathalmazon értékeltük ki. Megállapítottuk, hogy kísérleteink során a keretrendszerek felülmúlták az összes többi korszerű megközelítést, a tagmodelleket és a hagyományos ensemble módszereket is.

Az eredeti keretrendszert, amely a koszinusz hasonlóságot használta, konferenciatickk formájában publikáltuk [35]. Továbbá a hisztogram veszteséget és a bemutatott kiterjesztéseket alkalmazó továbbfejlesztett keretrendszerről készített cikkünket közlésre benyújtottuk [36].

Hivatkozások

- [1] Kermack, W. O. és McKendrick, A. G., „A contribution to the mathematical theory of epidemics,” *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, vol. 115, no. 772, pp. 700–721, 1927.
- [2] Harko, T., Lobo, F. S., és Mak, M., „Exact analytical solutions of the susceptible-infected-recovered (sir) epidemic model and of the sir model with equal death and birth rates,” *Applied Mathematics and Computation*, vol. 236, pp. 184–194, 2014.
- [3] Long, J., Shelhamer, E., és Darrell, T., „Fully convolutional networks for semantic segmentation,” *IEEE conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3431–3440, 2015.
- [4] Liu, L., Wei, W., Chow, K.-H., Loper, M., Gursoy, E., Truex, S., és Wu, Y., „Deep neural network ensembles against deception: Ensemble diversity, accuracy and robustness,” *2019 IEEE 16th international conference on mobile ad hoc and sensor systems (MASS)*, pp. 274–282, IEEE, 2019.
- [5] Zhang, S., Liu, M., és Yan, J., „The diversified ensemble neural network,” *Advances in Neural Information Processing Systems*, vol. 33, pp. 16001–16011, 2020.
- [6] Ustinova, E. és Lempitsky, V., „Learning deep embeddings with histogram loss,” *Advances in neural information processing systems*, vol. 29, 2016.

- [7] Zainuddin, Z. és Pauline, O., „Function approximation using artificial neural networks,” *WSEAS Transactions on Mathematics*, vol. 7, no. 6, pp. 333–338, 2008.
- [8] Hornik, K., „Approximation capabilities of multilayer feedforward networks,” *Neural networks*, vol. 4, no. 2, pp. 251–257, 1991.
- [9] Bogacsovics, G., Hajdu, A., Lakatos, R., Beregi-Kovács, M., Tiba, A., és Tomán, H., „Replacing the sir epidemic model with a neural network and training it further to increase prediction accuracy,” *Annales Mathematicae et Informaticae*, vol. 53, pp. 73–91, Eszterházy Károly Egyetem Líceum Kiadó, 2021.
- [10] World Health Organization, „*Global report on diabetes*. World Health Organization, 2016.
- [11] Harangi, B., Toth, J., Baran, A., és Hajdu, A., „Automatic screening of fundus images using a combination of convolutional neural network and hand-crafted features,” *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2699–2702, IEEE, 2019.
- [12] Zhang, T., Zeng, Y., és Xu, B., „Hcnn: A neural network model for combining local and global features towards human-like classification,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 30, no. 01, p. 1655004, 2016.
- [13] Bogacsovics, G., Toth, J., Hajdu, A., és Harangi, B., „Enhancing cnns through the use of hand-crafted features in

- automated fundus image classification,” *Biomedical Signal Processing and Control*, vol. 76, p. 103685, 2022.
- [14] Biscotti, C. V., Dawson, A. E., Dziura, B., Galup, L., Darragh, T., Rahemtulla, A., és Wills-Frank, L., „Assisted primary screening using the automated thinprep imaging system,” *American journal of clinical pathology*, vol. 123, no. 2, pp. 281–287, 2005.
 - [15] Chute, D., Lim, H., és Kong, C. S., „BD focalpoint slide profiler performance with atypical glandular cells on surepath papanicolaou smears.,” *Cancer cytopathology*, vol. 118, no. 2, pp. 68–74, 2010.
 - [16] Papanicolaou, G. N. és Traut, H. F., „The diagnostic value of vaginal smears in carcinoma of the uterus,” *Am. J. of Obstet. and Gynec.*, vol. 42, no. 2, pp. 193 – 206, 1941.
 - [17] Harangi, B., Bogacsovics, G., Toth, J., Kovacs, I., Dani, E., és Hajdu, A., „Pixel-wise segmentation of cells in digitized pap smear images.” Beküldve: *Scientific Data*.
 - [18] Bogacsovics, G., Hajdu, A., és Harangi, B., „Cell segmentation in digitized pap smear images using an ensemble of fully convolutional networks,” *2021 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pp. 1–6, IEEE, 2021.
 - [19] Gao, F., Wu, T., Li, J., Zheng, B., Ruan, L., Shang, D., és Patel, B., „Sd-cnn: A shallow-deep cnn for improved breast cancer diagnosis,” *Computerized Medical Imaging and Graphics*, vol. 70, pp. 53–62, 2018.

- [20] Titoriya, A. és Sachdeva, S., „Breast cancer histopathology image classification using alexnet,” *2019 4th International conference on information systems and computer networks (ISCON)*, pp. 708–712, IEEE, 2019.
- [21] Khairandish, M. O., Sharma, M., Jain, V., Chatterjee, J. M., és Jhanjhi, N., „A hybrid cnn-svm threshold segmentation approach for tumor detection and classification of mri brain images,” *Irbm*, vol. 43, no. 4, pp. 290–299, 2022.
- [22] Kibriya, H., Masood, M., Nawaz, M., és Nazir, T., „Multiclass classification of brain tumors using a novel cnn architecture,” *Multimedia Tools and Applications*, vol. 81, no. 21, pp. 29847–29863, 2022.
- [23] Zhang, N., Cai, Y.-X., Wang, Y.-Y., Tian, Y.-T., Wang, X.-L., és Badami, B., „Skin cancer diagnosis based on optimized convolutional neural network,” *Artificial intelligence in medicine*, vol. 102, p. 101756, 2020.
- [24] Hekal, A. A., Moustafa, H. E.-D., és Elnakib, A., „Ensemble deep learning system for early breast cancer detection,” *Evolutionary Intelligence*, pp. 1–10, 2022.
- [25] Gupta, N., Bhatele, P., és Khanna, P., „Glioma detection on brain mris using texture and morphological features with ensemble learning,” *Biomedical Signal Processing and Control*, vol. 47, pp. 115–125, 2019.
- [26] Nguyen, Q. H., Do, T. T., Wang, Y., Heng, S. S., Chen, K., Ang, W. H. M., Philip, C. E., Singh, M., Pham, H. N., Nguyen, B. P., és others,, „Breast cancer prediction using

- feature selection and ensemble voting,” *2019 International Conference on System Science and Engineering (ICSSE)*, pp. 250–254, IEEE, 2019.
- [27] Krizhevsky, A., Sutskever, I., és Hinton, G. E., „Imagenet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [28] Lu, S., Wang, S.-H., és Zhang, Y.-D., „Detection of abnormal brain in mri via improved alexnet and elm optimized by chaotic bat algorithm,” *Neural Computing and Applications*, vol. 33, pp. 10799–10811, 2021.
- [29] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., és Chen, L.-C., „Mobilenetv2: Inverted residuals and linear bottlenecks,” *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4510–4520, 2018.
- [30] Arfan, T. H., Hayaty, M., és Hadinegoro, A., „Classification of brain tumours types based on mri images using mobilenet,” *2021 2nd International Conference on Innovative and Creative Information Technology (ICITech)*, pp. 69–73, IEEE, 2021.
- [31] Roslidar, R., Saddami, K., Arnia, F., Syukri, M., és Munadi, K., „A study of fine-tuning cnn models based on thermal imaging for breast cancer classification,” *2019 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)*, pp. 77–81, IEEE, 2019.

- [32] Tan, M. és Le, Q., „Efficientnet: Rethinking model scaling for convolutional neural networks,” *International conference on machine learning*, pp. 6105–6114, PMLR, 2019.
- [33] Shah, H. A., Saeed, F., Yun, S., Park, J.-H., Paul, A., és Kang, J.-M., „A robust approach for brain tumor detection in magnetic resonance images using finetuned efficientnet,” *IEEE Access*, vol. 10, pp. 65426–65438, 2022.
- [34] Ma, N., Zhang, X., Zheng, H.-T., és Sun, J., „Shufflenet v2: Practical guidelines for efficient cnn architecture design,” *Proceedings of the European conference on computer vision (ECCV)*, pp. 116–131, 2018.
- [35] Bogacsóvics, G., Harangi, B., és Hajdu, A., „Increasing the diversity of ensemble members for accurate brain tumor classification,” *2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 529–534, IEEE, 2023.
- [36] Bogacsóvics, G., Harangi, B., és Hajdu, A., „Developing diverse ensemble architectures for automatic brain tumor classification.” Beküldve: *Multimedia Tools and Applications*.



Nyilvántartási szám: DEENK/132/2024.PL
Tárgy: PhD Publikációs Lista

Jelölt: Bogacsovics Gergő

Doktori Iskola: Informatikai Tudományok Doktori Iskola

MTMT azonosító: 10078365

A PhD értekezés alapjául szolgáló közlemények

Idegen nyelvű tudományos közlemények hazai folyóiratban (1)

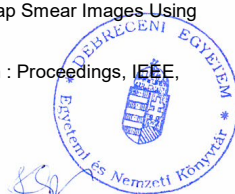
1. **Bogacsovics, G.**, Hajdu, A., Lakatos, R., Beregi-Kovács, M., Tiba, A., Tomán, H.: Replacing the SIR epidemic model with a neural network and training it further to increase prediction accuracy.
Ann. Math. Inform. 53, 73-91, 2021. ISSN: 1787-5021.
DOI: <http://dx.doi.org/10.33039/ami.2021.02.003>

Idegen nyelvű tudományos közlemények külföldi folyóiratban (1)

2. **Bogacsovics, G.**, Tóth, J., Hajdu, A., Harangi, B.: Enhancing CNNs through the use of hand-crafted features in automated fundus image classification.
Biomed. Signal Process. Control. 76, 1-10, 2022. ISSN: 1746-8094.
DOI: <http://dx.doi.org/10.1016/j.bspc.2022.103685>
IF: 5.1

Idegen nyelvű konferencia közlemények (2)

3. **Bogacsovics, G.**, Harangi, B., Hajdu, A.: Increasing the diversity of ensemble members for accurate brain tumor classification.
In: 36th IEEE International Symposium on Computer-Based Medical Systems, CBMS 2023, Institute of Electrical and Electronics Engineers Inc., [s.l.], 529-534, 2023. ISBN: 9798350312249
4. **Bogacsovics, G.**, Hajdu, A., Harangi, B.: Cell Segmentation in Digitized Pap Smear Images Using an Ensemble of Fully Convolutional Networks.
In: 2021 IEEE Signal Processing in Medicine and Biology Symposium : Proceedings, IEEE, Philadelphia, 1-6, 2021. ISBN: 9781665428972





További közlemények

Magyar nyelvű tudományos közlemények hazai folyóiratban (5)

5. **Bogacsovics, G.**, Hajdu, A., Harangi, B., Lakatos, I., Lakatos, R., Szabó, M., Tiba, A., Tóth, J., Tarcsi, Á.: Adatelemzési folyamat és keretrendszer a közigazgatás számára.
Közigazgatástudomány. 1 (2), 146-158, 2021. ISSN: 2786-1910.
DOI: <http://dx.doi.org/10.54200/kt.v1i2.24>
6. **Bogacsovics, G.**, Hajdu, A., Harangi, B., Lakatos, I., Lakatos, R., Szabó, M., Tiba, A., Tóth, J.: Napelemfarmok Magyarország területén történő elhelyezését segítő döntéstámogató rendszer fejlesztése.
Közigazgatástudomány. 1 (2), 134-145, 2021. ISSN: 2786-1910.
DOI: <http://dx.doi.org/10.54200/kt.v1i2.23>
7. Bátfai, N., Besenczi, R., Szabó, J., Jeszenszky, P., Buda, A., Jármí, L., Lovas, R. B., Pál, M. K., **Bogacsovics, G.**, Kovács, E.: DEAC-Hackers: játszó hackerek, hackelő játékosok.
Inform Társad. 18 (1), 132-146, 2018. ISSN: 1587-8694.
DOI: <http://dx.doi.org/10.22503/infars.XVIII.2018.1.9>
IF: 0.222
8. Bátfai, N., **Bogacsovics, G.**, Paszerbovics, R., Antal, A., Czevár, I., Kelemen, V., Besenczi, R.: E-sportlók mérése.
Inform. Társad. 18 (1), 147-155, 2018. ISSN: 1587-8694.
DOI: <http://dx.doi.org/10.22503/infars.XVIII.2018.1.10>
IF: 0.222
9. Bátfai, N., Bersenszki, M., Lukács, M., Besenczi, R., **Bogacsovics, G.**, Jeszenszky, P.: Az e-sport és a robotpszichológia közös jövője.
Inform. Társad. 16 (4), 26-39, 2017. ISSN: 1587-8694.
DOI: <http://dx.doi.org/10.22503/infars.XVI.2016.4.2>
IF: 0.023

Idegen nyelvű tudományos közlemények külföldi folyóiratban (3)

10. Lakatos, R., **Bogacsovics, G.**, Harangi, B., Lakatos, I., Tiba, A., Tóth, J., Szabó, M., Hajdu, A.: A Machine Learning-Based Pipeline for the Extraction of Insights from Customer Reviews.
Big Data Cogn. Comput. 8 (3), 1-24, 2024. EISSN: 2504-2289.
DOI: <http://dx.doi.org/10.3390/bdcc8030020>
IF: 3.7 (2022)
11. Bátfai, N., Papp, D., Besenczi, R., **Bogacsovics, G.**, Veres, D.: Benchmarking Cognitive Abilities of the Brain with the Event of Losing the Character in Computer Games.
Stud. Univ. Babeş-Bolyai. Inform. 64 (1), 15-25, 2019. ISSN: 1224-869X.
DOI: <http://dx.doi.org/10.24193/subbi.2019.1.02>





12. Bátfai, N., Papp, D., **Bogacsovics, G.**, Szabó, M., Simkó, V. S., Bersenszki, M., Szabó, G., Kovács, L., Kovács, F., Varga, E. S.: Object file system software experiments about the notion of number in humans and machines.
Cognition, Brain, Behavior. 23 (4), 257-280, 2019. ISSN: 2247-9228.
DOI: <http://dx.doi.org/10.24193/cbb.2019.23.15>

Időgen nyelvű konferencia közlemények (3)

13. **Bogacsovics, G.**, Harangi, B., Beregi-Kovács, M., Kupás, D., Lakatos, R., Serbán, N. D., Tiba, A., Tóth, J.: Assessing Conventional and Deep Learning-Based Approaches for Named Entity Recognition in Unstructured Hungarian Medical Reports.
In: 2024 IEEE 22nd World Symposium on Applied Machine Intelligence and Informatics (SAMI). Ed.: Kovács Levente, Liberios Vokorokos, IEEE, Piscataway, 77-82, 2024. ISBN: 9798350317206
14. Lakatos, R., **Bogacsovics, G.**, Hajdu, A.: Predicting the direction of the oil price trend using sentiment analysis.
In: IEEE 2nd Conference on Information Technology and Data Science (CITDS) : Proceedings. Ed.: Fazekas István, Institute of Electrical and Electronics Engineers (IEEE), Piscataway, 177-182, 2022.
15. Harangi, B., Tóth, J., **Bogacsovics, G.**, Kupás, D., Kovács, L., Hajdu, A.: Cell detection on digitized Pap smear images using ensemble of conventional image processing and deep learning techniques.
In: 11th International Symposium on Image and Signal Processing and Analysis (ISPA 2019). Eds.: S. Lončarić, R. Bregović, M. Carli, M. Subašić, Institute of Electrical and Electronics Engineers (IEEE), Piscataway, NJ, USA, 38-42, 2019. ISBN: 9781728131405

A közlő folyóiratok összesített impakt faktora: 9,267

A közlő folyóiratok összesített impakt faktora (az értekezés alapján szolgáló közleményekre): 5,1

A DEENK a Jelölt által az iDEa Tudóstérbe feltöltött adatok bibliográfiai és tudomány-metriai ellenőrzését a tudományos adatbázisok és a Journal Citation Reports Impact Factor lista alapján elvégezte.

Debrecen, 2024.04.09.



Short thesis for the degree of doctor of philosophy (PhD)

Developing Performant Neural Network Architectures for Processing Medical Data

by Gergő Bogacsovics

Supervisor: Dr. András Hajdu



UNIVERSITY OF DEBRECEN

Doctoral School of Informatics

Debrecen, 2024

Contents

1	Introduction	1
2	Designing a two-step architecture for the accurate prediction of diseases	4
3	Combining hand-crafted features with features extracted by CNNs	8
4	Fully convolutional ensemble model for automatic cell segmentation	12
5	Building diverse ensemble models by penalizing the similarity between the member models	15
	References	19
	List of publications	25

1 Introduction

In the last decades, the medical field was exposed to a major transformation phase and paradigm shift. With the appearance of new imaging techniques and more modern medical equipment, the various fields experienced a sharp increase in the amount of data generated each day. The appearance of deep learning algorithms made it possible to process this data and provide reliable solutions. In this dissertation, we present multiple novel methods to further increase the accuracy and reliability of deep learning models and use them to solve various problems in the clinical domain.

The solutions included in this dissertation can be divided into two groups. The first group focuses on improving the accuracy and usability of neural networks by combining traditional methods with deep learning techniques. We show that by doing so, we can train models that have the benefits of both approaches. Namely, models that can not only operate on smaller datasets but are also more robust and more flexible than regular traditional methods. The second group of techniques aims to develop efficient and performant ensemble methods that are more powerful than traditional ensemble approaches.

First, we focused on the effective fusion of theoretical models and neural networks. We showed that both of these approaches have their own advantages and shortcomings. For this purpose, we developed a novel two-step architecture that combines the benefits of both approaches and used it for predicting the spread of diseases such as influenza and COVID-19. We showed that the architecture first trains a neural network to approximate a Kermack-McKendrick model [1], which is a

SIR model [2], and then the network is trained further on real data. By following these steps, the network can first learn to approximate the theoretical model, and then, during the second step, the model is trained on the real data. During the first step, the network has the opportunity to learn the most important aspects of the problem by using the more regulated theoretical model. Then, during the second step, its weights are fine-tuned to maximize its accuracy.

In the next part of the dissertation, we detailed the importance of hand-crafted features, which, with the rise of neural networks and deep learning, very few research utilized in the recent years. We discussed that extracting them requires professional personnel and is usually cumbersome and time-consuming. However, we noted that they may also capture relevant features that deep learning techniques may have trouble recognizing. Thus, we focused on the effective combination of these hand-crafted features with the ones extracted by convolutional neural networks (CNNs). We showed that this combination is by no means trivial, and therefore presented three different variants: one shallow (one layer), one deeper (multi-layer) approach, and an architecture that trains two classifiers, each for the given features. This last architecture trains the two classifiers in an isolated manner and then aggregates their outputs using simple averaging, ultimately acting as a "mini-ensemble".

Next, we presented a dataset and a novel ensemble framework for the effective segmentation of cells on digitized Pap smear images. We showcased that our framework, in contrast to traditional methods, has its own set of weights. The framework utilizes an FCN-32 [3] architecture as its backbone and

receives the outputs of some pretrained FCN-32, FCN-16, and FCN-8 networks in conjunction with the original input image. We evaluated multiple versions of this framework, where each version used a different subset of the three FCN models as inputs. We showed that we were able to surpass both the original networks and other methods as well in terms of performance.

In the last part of the dissertation, we focused on building diverse ensemble models. We started by explaining that diversity, according to most recent studies [4, 5], is an important metric from the perspective of the overall performance of the ensemble. Then, we detailed that contrary to this, diversity is usually not considered when building ensembles. After this, we presented multiple novel frameworks that train all the corresponding member models at the same time. These frameworks include an additional regularization term in the cost function, which is responsible for measuring the diversity of the member models. We showed that one possible way of measuring the similarity of the models is by using the feature vectors extracted by each model. This is because two models operating on the same or highly similar features can be considered redundant, while models that operate on highly dissimilar features should operate in different manners, making them more valuable when building ensembles. Then, we proposed two main frameworks, which relied on the cosine similarity and the histogram loss [6] to measure the similarity of the vectors extracted by the different member models. We also presented two extensions of the frameworks that make it possible to use the frameworks in a weighted cost function or when using member models that produce feature vectors of different lengths.

2 Designing a two-step architecture for the accurate prediction of diseases

Theoretical models are often used in research, as they provide a relatively simple, yet effective and expressive way of modelling various phenomena. Moreover, theoretical models can be used even when the dataset at hand is small, given that the core nature of the problem is known, such as a formula describing the natural spread of a given disease. The main problem, however, is that when using more complex models, the mathematical description also becomes more complex. For this reason, theoretical models generally avoid large complexity and aim for the simplest possible definition. This results in a mathematically more manageable model, but at the same time it also generally leads to sub-optimal performance. Another drawback of these models is that theoretical models are usually too rigid, making their application to a given problem all the more tedious.

On the other hand, deep learning-based solutions are well known for their flexibility and ability to fit the training data more effectively due to the higher number of parameters. For example, it has been shown [7, 8] that with neural networks, we can build models that can approximate mathematical functions. This implies that once trained, artificial neural networks should be able to behave like theoretical models, while still retaining their biggest advantage, which is their flexibility, resulting in an overall better performance for more complex real-world problems.

The objective of our research was to construct a framework that makes it possible to build deep learning models that

can effectively combine the main benefits of both theoretical models and standard neural networks. Namely, we aimed to create models that can be applied even when using smaller datasets and have a mathematically more robust foundation, while also being more flexible than simple theoretical models, ultimately achieving better performance. For this, we proposed a novel two-step approach that first trains the neural network model to approximate the original theoretical model. Then, during the second step, we fine-tuned its weights on the real dataset.

Thesis 1.1. *I have developed a two-step architecture for the effective combination of theoretical models and deep learning methods. The architecture first trains a neural network on synthetic data generated by the theoretical model. In the next step, the model is further trained on the real data, during which it is allowed to slightly deviate from the original theoretical model to maximize its performance on the original dataset.*

In order to validate the viability and effectiveness of the two-step architecture, we evaluated its performance for predicting the spread of two diseases: influenza and COVID-19. We used a simple SIR model [2] called the Kermack-McKendrick model [1] as the theoretical model and a simple dense network with three hidden layers as the deep learning model. Our reasoning for choosing this particular theoretical model is that although the model is relatively simple, it is still mathematically founded, being based on differential equations. For the neural network model, we also mainly focused on simpler architectures instead of using more sophisticated ones like recurrent neural networks (RNNs) or long short-term memory

(LSTM). We did so to show that our proposed architecture is not a specialized one, and can be used for a variety of problems and areas. We also experimented with increasing the time window for both the inputs and outputs of the model in an attempt to increase its performance.

Thesis 1.2. *I have elaborated multiple variants of the proposed two-step architecture. I used input and output windows of different sizes and evaluated them for predicting the spread of influenza and COVID-19 cases.*

According to the experimental test results, the model outperformed both the original theoretical model and neural networks trained solely on the original dataset. After our experiments, we have also noted that, according to our results and observations, the model obtained after training it according to our two-step approach was more flexible than the original theoretical model and more accurate than neural networks trained solely on the original dataset. We showed that training the model only on the first wave of COVID-19, the extracted features could actually help the model in making predictions for the second wave, greatly surpassing the original theoretical model, which could only predict a single wave.

Finally, we also explained how this method could be viewed as another type of transfer learning. Namely, instead of using large pre-trained neural networks, we can instead choose a theoretical model that roughly resembles the real data and have the neural network learn on that function. We also noted that, due to the generic nature of the proposed framework, there are no restrictions regarding the theoretical model, as it can be any kind of model as long as its outputs can be compared

to the outputs of a neural network. This implies that the architecture can potentially be used in a variety of disciplines to solve a variety of problems. Lastly, we also highlighted that the theoretical model does not need to perfectly fit the real data. Instead, it is enough if the key features of the real data (spread, nature, etc.) are encoded in the theoretical model. The detailed experimental results and our findings have been published in [9].

3 Combining hand-crafted features with features extracted by CNNs

Eye diseases such as diabetic retinopathy (DR) and diabetic macular edema (DME) pose a major threat in today’s world. They affect a significant portion of the global population. Consequently, it is extremely important to develop reliable solutions that can accurately detect these diseases. In 2019, an estimated 1.5 million deaths were directly caused by diabetes and, according to the World Health Organization, 422 million people worldwide have the disease [10]. Therefore, developing reliable systems that could detect these diseases in their earlier stages could lead to increased life quality by slowing down the progression to vision impairment.

Some of the current methods use the hand-crafted features, extracted by experts to detect these diseases. However, developing these solutions is often time-consuming and they are usually not accurate enough. Due to this, several solutions have been proposed that use deep learning techniques, where experts are not required to extract these features, making the development of models easier and faster. On the other hand, these methods also have some disadvantages. Namely, they usually ignore the highly valuable hand-crafted features, which could potentially contribute to more accurate predictions.

[11] proposed a solution that combines the features extracted by a convolutional neural network (CNN) with traditional, hand-crafted features in an attempt to improve classification performance. The idea derived from [12], where the authors claimed that a CNN can automatically extract many important local, textual features from images by convolving

with a sliding window and forming a filter. However, besides the local features, global features, such as contour features and structural ones, are also important and have played a crucial role in many image processing tasks. Moreover, many of the signs implying DR and DME can be described well by their contour feature, which implies that by considering these features, we may be able to improve the final accuracy of a CNN-based screening system.

In this dissertation, we presented multiple novel solutions by extending [11] and investigating the applicability of some additional state-of-the-art neural networks. We systematically examined the advantages and disadvantages of the proposed methodology. We showed that the original framework had a few notable limitations and presented multiple modifications that addressed these problems. The architecture proposed in [11] concatenated the hand-crafted features and the features extracted by an AlexNet model by using only a single set of weights to directly transform the concatenated vector into the predictions of the model. We noted that this immediate reduction of the features into predictions is not ideal as the framework does not have the opportunity to learn more complex patterns.

We started by experimenting with a deeper combination of the features. We noted that one possible way to accomplish this is by giving more computational capacity to the part of the framework that operates on the hand-crafted features, which can be done by feeding the hand-crafted features through more layers. This way, the framework should be able to capture more features and patterns present in the hand-crafted features, due to the increased computational capacity.

Thesis 2.1. *I have developed an improved version of the model proposed in [11]. Using the same input image, the hand-crafted features are computed and passed through a series of dense layers, while another branch of the model uses the original convolutional neural network architecture to process the image. Each branch produces its own outputs, which are then averaged to get a final prediction.*

After the introduction of this new framework, we discussed that although it solves the problem of the shallow combination of the feature vectors, the features are now trained separately. We noted that this may not be a problem, as the improved framework now acts a sort of "mini-ensemble", where each model operates separately, and the outputs are simply averaged. However, we also highlighted that this way, the framework has no chance to recognize patterns that require the presence of both some hand-crafted and CNN features at the same time. To solve this, another version of the framework was implemented.

Thesis 2.2. *I have developed an improved version of the model proposed in [11] that uses the features extracted by the last convolutional layer in conjunction with the hand-crafted features. Both of them are fed through the subsequent dense layers together, at the same time, providing the opportunity for the network to recognize more complex patterns, where the simultaneous presence of both hand-crafted and CNN features are required.*

Finally, we implemented several baselines to compare the performance of our solutions with. We measured the performance of regular neural networks without the hand-crafted

features, machine learning models only using the hand-crafted features, and other state-of-the-art baselines. We used a variety of metrics to provide a fair comparison. All of the proposed architectures were evaluated for the task of detecting diabetic macular edema and diabetic retinopathy using a dataset constructed from multiple sources. The proposed architectures all outperformed the neural networks that did not use the hand-crafted features, solutions only relying on hand-crafted features, and other state-of-the-art approaches as well.

All of the presented methods, as well as the detailed experimental results have been published in [13].

4 Fully convolutional ensemble model for automatic cell segmentation

Automated screening systems play a vital role in increasing the level of patient care and quality of medical workflows. Some of the most widely used automatic solutions include the Hologic ThinPrep Imaging System [14] and the Focal Point Slide Profiler [15], which both have been approved by the U.S. Food and Drug Administration (FDA). However, both of these solutions have some significant drawbacks. Namely, the Hologic ThinPrep Imaging System can only analyze ThinPrep Pap Test slides which have much higher costs than the most commonly applied Papanicolaou smear test [16], while the Focal Point Slide Profiler can eliminate only up to 25% of the lowest-risk slides. The final objective of our research was to develop a fully automatic screening system that can recognize cancerous cells in digitized Pap smear images. While developing this system, we aimed to overcome the limitations of these two solutions. That is, our intent was to develop a system that uses the most commonly applied Pap smear test images and ranks the slides by the level of risk more accurately. The objective of our automatic screening system is to find and segment each cell from the high-resolution input image. In this dissertation, we presented a component of this system, responsible for the accurate and reliable segmentation of cells in the input images, as well as a dataset of our own.

The dataset was created in a collaboration with the University of Debrecen Clinical Center, Department of Pathology. In the dissertation, we explained that available datasets are small in size and detailed that there is currently no other compara-

ble, publicly available dataset for the task of cell segmentation. Due to the long and time-consuming nature of manual annotation, we noted that each of our experiments only had access to the currently available part of the dataset. We published the final, most comprehensive version of our dataset, containing 3 565 images of size 2000×2000 and their corresponding segmentation masks, in [17].

This dissertation also presented a unique method for building accurate and reliable ensemble models that could be used as a part of an automatic screening system for cell segmentation. We started by introducing the biggest problems of traditional ensemble approaches. Namely, we showed some examples where the individual methods recognized different cells. We explained that in these cases, traditional approaches would not be able to effectively combine these outputs, as there is no overlap between the generated segmentation masks. Then, we proposed our method, which utilizes a fully convolutional neural network (FCN) [3] architecture as its backbone. We showed that the main benefit of this approach is that, contrary to traditional ensembles, it can not only aggregate the outputs of the member models, but instead it can also generate its own segmentation mask, disregarding the predictions of the member models.

Thesis 3.1. *I have developed a unique ensemble approach that utilizes a modified FCN-32 architecture as its backbone. The framework receives the outputs of some pre-trained FCN architectures in conjunction with the given input image to generate its predictions. I have implemented multiple versions of the framework, which all use a different number of FCN models as their inputs.*

We showed that, due to the FCN-32 backbone of the architecture and its own set of weights, the framework has the ability to recognize features and patterns using both the input image and the predictions of the member models. Using these patterns, it can automatically decide what weight to assign to each member or even detect if a given member model underperforms in the presence of a given feature in the input image. After evaluating all of the different versions of the proposed architectures on our own dataset, we noted that the proposed framework greatly outperformed both the individual member models and other state-of-the-art approaches as well. The quality of the resulting segmentation masks was close to, or in some cases even surpassed, human performance.

The described, comprehensive dataset has been submitted for publication in [17]. The ensemble framework using the combined outputs of the pre-trained FCN networks and the input image has been published in [18].

5 Building diverse ensemble models by penalizing the similarity between the member models

Even though many deep learning-based techniques have been proposed in the last decade for solving a variety of medical tasks, like breast cancer [19, 20] and brain tumor [21, 22] detection or identifying skin cancer [23], the performances of the individual models and architectures vary greatly from application to application. For example, it is common that a specific model performs well in one task while its performance falls back in another scenario. This phenomenon has popularized the usage of ensemble models [24, 25, 26], which merge multiple different models into one.

The main benefit of these algorithms is that by integrating models that operate differently, the ensemble has the potential to perform better than the individual models. This is because, in cases when one model performs poorly, the other models, as long as they perform well and they are in the majority, can tilt the outputs of the ensemble in the right direction. By using such techniques, we can develop models that are more resistant to outliers or irregularities in the inputs and have better generalization capabilities. However, traditional techniques usually do not consider the diversity, in other words, how different the models of the ensemble are. This is a serious shortcoming, as it has been shown multiple times [4, 5] that diversity has a beneficial effect in building accurate and performant ensemble models.

In this dissertation, we presented a novel ensemble-based method for the reliable and accurate classification of brain tu-

mors from T1-weighted contrast-enhanced magnetic resonance imaging (MRI) images. We introduced multiple techniques that, in contrast to traditional ensemble approaches, directly measure and optimize the diversity of the member models during training. For building the ensemble models, we considered some of the most commonly used state-of-the-art CNN architectures as the members of the ensembles, such as AlexNet [27, 28], MobileNetv2 [29, 30, 31], EfficientNet [32, 33], and ShuffleNet v2 [31, 34].

The presented method divided any CNN architecture into two parts E and D , where E denotes the convolutional layers and D denotes the dense layers of the model. We introduced our idea that, for any CNN model, we can treat the outputs of E (in other words, the feature vectors) as some latent vectors in a high-dimensional latent vector space. We also detailed that as long as these feature vectors are different (i.e., the different members operate on different sets of features), the diversity of the ensemble should be increased. Then, we presented our framework that tries to maximize the differences between the feature vectors (the outputs of E) extracted by the member models.

Thesis 4.1. *I have developed a novel ensemble framework that trains all of its members at the same time using a custom regularization term in the cost function to maximize the difference between the feature vectors extracted by the member models. I have experimented with multiple different networks for the member models and used the cosine similarity to measure the similarity between the feature vectors.*

In this dissertation, we have also outlined some problems

with using the cosine similarity to compare the feature vectors in our framework. Namely, we showed that it does not handle out-of-order similarities well, for example when two feature vectors contain the exact same values, just in a slightly permuted manner. In such cases, the cosine similarity fails to capture the full similarity between the features, which may lead to sub-optimal results. We also showed that one possible solution to this problem is to use a histogram-based approach, which does not rely on the order of the elements of the feature vectors.

Thesis 4.2. *I have developed an improved version of the framework that uses the histogram loss instead of the cosine similarity to measure the similarity between the extracted feature vectors. This version of the framework minimizes the probability of feature vectors originating from two different models being more similar than vectors produced by the same model.*

We have also highlighted that one major disadvantage of the frameworks was that they could only operate with member models that produced feature vectors of the same dimension. We detailed that this was a limitation of the similarity functions, as, by definition, the inner product, which is a crucial step in the calculation of both of the similarities, is not defined for two vectors of different sizes.

Thesis 4.3. *I have developed an extension for both frameworks, which allows them to use member models with varying feature vector lengths by bringing the vectors to a common dimension before measuring the similarity between them. I have ascertained the validity of this approach by carefully evaluating*

it on the dataset and found that the extension did not impose any negative effects on the framework.

In the dissertation, we also mentioned that due to the nature of medical datasets, it is also extremely important for a technique to be able to be used when dealing with imbalanced datasets. In such cases, a weighted cost function is used, when a coefficient is calculated for each sample, depending on the rarity of their corresponding class labels. Then, we showed that using a weighted cost function does not have a negative effect on the proposed frameworks, since the introduced regularization term puts equal emphasis on each member model and their produced feature vectors, ignoring the class label. Both frameworks and all of their extensions were evaluated on a dataset containing MRI images of brain tumors. We noted that in our experiments, the frameworks outperformed all other state-of-the-art approaches, the member models, and traditional ensemble methods as well.

The original framework that used the cosine similarity has been published in [35]. We have also submitted a paper for publication in [36] about the improved framework that used the histogram loss and the presented extensions.

References

- [1] W. O. Kermack and A. G. McKendrick, “A contribution to the mathematical theory of epidemics,” *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, vol. 115, no. 772, pp. 700–721, 1927.
- [2] T. Harko, F. S. Lobo, and M. Mak, “Exact analytical solutions of the susceptible-infected-recovered (sir) epidemic model and of the sir model with equal death and birth rates,” *Applied Mathematics and Computation*, vol. 236, pp. 184–194, 2014.
- [3] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in *IEEE conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3431–3440, 2015.
- [4] L. Liu, W. Wei, K.-H. Chow, M. Loper, E. Gursoy, S. Truex, and Y. Wu, “Deep neural network ensembles against deception: Ensemble diversity, accuracy and robustness,” in *2019 IEEE 16th international conference on mobile ad hoc and sensor systems (MASS)*, pp. 274–282, IEEE, 2019.
- [5] S. Zhang, M. Liu, and J. Yan, “The diversified ensemble neural network,” *Advances in Neural Information Processing Systems*, vol. 33, pp. 16001–16011, 2020.
- [6] E. Ustinova and V. Lempitsky, “Learning deep embeddings with histogram loss,” *Advances in neural information processing systems*, vol. 29, 2016.

- [7] Z. Zainuddin and O. Pauline, “Function approximation using artificial neural networks,” *WSEAS Transactions on Mathematics*, vol. 7, no. 6, pp. 333–338, 2008.
- [8] K. Hornik, “Approximation capabilities of multilayer feedforward networks,” *Neural networks*, vol. 4, no. 2, pp. 251–257, 1991.
- [9] G. Bogacsovics, A. Hajdu, R. Lakatos, M. Beregi-Kovács, A. Tiba, and H. Tomán, “Replacing the sir epidemic model with a neural network and training it further to increase prediction accuracy,” in *Annales Mathematicae et Informaticae*, vol. 53, pp. 73–91, Eszterházy Károly Egyetem Líceum Kiadó, 2021.
- [10] World Health Organization, *Global report on diabetes*. World Health Organization, 2016.
- [11] B. Harangi, J. Toth, A. Baran, and A. Hajdu, “Automatic screening of fundus images using a combination of convolutional neural network and hand-crafted features,” in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2699–2702, IEEE, 2019.
- [12] T. Zhang, Y. Zeng, and B. Xu, “Hcnn: A neural network model for combining local and global features towards human-like classification,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 30, no. 01, p. 1655004, 2016.
- [13] G. Bogacsovics, J. Toth, A. Hajdu, and B. Harangi, “Enhancing cnns through the use of hand-crafted features in

- automated fundus image classification,” *Biomedical Signal Processing and Control*, vol. 76, p. 103685, 2022.
- [14] C. V. Biscotti, A. E. Dawson, B. Dziura, L. Galup, T. Darragh, A. Rahemtulla, and L. Wills-Frank, “Assisted primary screening using the automated thinprep imaging system,” *American journal of clinical pathology*, vol. 123, no. 2, pp. 281–287, 2005.
 - [15] D. Chute, H. Lim, and C. S. Kong, “BD focalpoint slide profiler performance with atypical glandular cells on surepath papanicolaou smears,” *Cancer cytopathology*, vol. 118, no. 2, pp. 68–74, 2010.
 - [16] G. N. Papanicolaou and H. F. Traut, “The diagnostic value of vaginal smears in carcinoma of the uterus,” *Am. J. of Obstet. and Gynec.*, vol. 42, no. 2, pp. 193 – 206, 1941.
 - [17] B. Harangi, G. Bogacsovics, J. Toth, I. Kovacs, E. Dani, and A. Hajdu, “Pixel-wise segmentation of cells in digitized pap smear images.” Submitted for publication to *Scientific Data*.
 - [18] G. Bogacsovics, A. Hajdu, and B. Harangi, “Cell segmentation in digitized pap smear images using an ensemble of fully convolutional networks,” in *2021 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pp. 1–6, IEEE, 2021.
 - [19] F. Gao, T. Wu, J. Li, B. Zheng, L. Ruan, D. Shang, and B. Patel, “Sd-cnn: A shallow-deep cnn for improved

- breast cancer diagnosis,” *Computerized Medical Imaging and Graphics*, vol. 70, pp. 53–62, 2018.
- [20] A. Titoriya and S. Sachdeva, “Breast cancer histopathology image classification using alexnet,” in *2019 4th International conference on information systems and computer networks (ISCON)*, pp. 708–712, IEEE, 2019.
 - [21] M. O. Khairandish, M. Sharma, V. Jain, J. M. Chatterjee, and N. Jhanjhi, “A hybrid cnn-svm threshold segmentation approach for tumor detection and classification of mri brain images,” *Irbm*, vol. 43, no. 4, pp. 290–299, 2022.
 - [22] H. Kibriya, M. Masood, M. Nawaz, and T. Nazir, “Multiclass classification of brain tumors using a novel cnn architecture,” *Multimedia Tools and Applications*, vol. 81, no. 21, pp. 29847–29863, 2022.
 - [23] N. Zhang, Y.-X. Cai, Y.-Y. Wang, Y.-T. Tian, X.-L. Wang, and B. Badami, “Skin cancer diagnosis based on optimized convolutional neural network,” *Artificial intelligence in medicine*, vol. 102, p. 101756, 2020.
 - [24] A. A. Hekal, H. E.-D. Moustafa, and A. Elnakib, “Ensemble deep learning system for early breast cancer detection,” *Evolutionary Intelligence*, pp. 1–10, 2022.
 - [25] N. Gupta, P. Bhatele, and P. Khanna, “Glioma detection on brain mris using texture and morphological features with ensemble learning,” *Biomedical Signal Processing and Control*, vol. 47, pp. 115–125, 2019.
 - [26] Q. H. Nguyen, T. T. Do, Y. Wang, S. S. Heng, K. Chen, W. H. M. Ang, C. E. Philip, M. Singh, H. N. Pham,

- B. P. Nguyen, *et al.*, “Breast cancer prediction using feature selection and ensemble voting,” in *2019 International Conference on System Science and Engineering (ICSSE)*, pp. 250–254, IEEE, 2019.
- [27] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [28] S. Lu, S.-H. Wang, and Y.-D. Zhang, “Detection of abnormal brain in mri via improved alexnet and elm optimized by chaotic bat algorithm,” *Neural Computing and Applications*, vol. 33, pp. 10799–10811, 2021.
- [29] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “Mobilenetv2: Inverted residuals and linear bottlenecks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4510–4520, 2018.
- [30] T. H. Arfan, M. Hayaty, and A. Hadinegoro, “Classification of brain tumours types based on mri images using mobilenet,” in *2021 2nd International Conference on Innovative and Creative Information Technology (ICITech)*, pp. 69–73, IEEE, 2021.
- [31] R. Roslidar, K. Saddami, F. Arnia, M. Syukri, and K. Munadi, “A study of fine-tuning cnn models based on thermal imaging for breast cancer classification,” in *2019 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)*, pp. 77–81, IEEE, 2019.

- [32] M. Tan and Q. Le, “Efficientnet: Rethinking model scaling for convolutional neural networks,” in *International conference on machine learning*, pp. 6105–6114, PMLR, 2019.
- [33] H. A. Shah, F. Saeed, S. Yun, J.-H. Park, A. Paul, and J.-M. Kang, “A robust approach for brain tumor detection in magnetic resonance images using finetuned efficientnet,” *IEEE Access*, vol. 10, pp. 65426–65438, 2022.
- [34] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, “Shufflenet v2: Practical guidelines for efficient cnn architecture design,” in *Proceedings of the European conference on computer vision (ECCV)*, pp. 116–131, 2018.
- [35] G. Bogacsovics, B. Harangi, and A. Hajdu, “Increasing the diversity of ensemble members for accurate brain tumor classification,” in *2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 529–534, IEEE, 2023.
- [36] G. Bogacsovics, B. Harangi, and A. Hajdu, “Developing diverse ensemble architectures for automatic brain tumor classification.” Submitted for publication to *Multimedia Tools and Applications*.



Registry number:
Subject:

DEENK/132/2024.PL
PhD Publication List

Candidate: Gergő Bogacsovics

Doctoral School: Doctoral School of Informatics

MTMT ID: 10078365

List of publications related to the dissertation

Foreign language scientific articles in Hungarian journals (1)

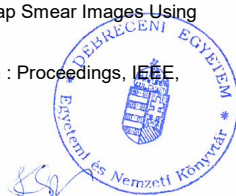
1. **Bogacsovics, G.**, Hajdu, A., Lakatos, R., Beregi-Kovács, M., Tiba, A., Tomán, H.: Replacing the SIR epidemic model with a neural network and training it further to increase prediction accuracy.
Ann. Math. Inform. 53, 73-91, 2021. ISSN: 1787-5021.
DOI: <http://dx.doi.org/10.33039/ami.2021.02.003>

Foreign language scientific articles in international journals (1)

2. **Bogacsovics, G.**, Tóth, J., Hajdu, A., Harangi, B.: Enhancing CNNs through the use of hand-crafted features in automated fundus image classification.
Biomed. Signal Process. Control. 76, 1-10, 2022. ISSN: 1746-8094.
DOI: <http://dx.doi.org/10.1016/j.bspc.2022.103685>
IF: 5.1

Foreign language conference proceedings (2)

3. **Bogacsovics, G.**, Harangi, B., Hajdu, A.: Increasing the diversity of ensemble members for accurate brain tumor classification.
In: 36th IEEE International Symposium on Computer-Based Medical Systems, CBMS 2023, Institute of Electrical and Electronics Engineers Inc., [s.l.], 529-534, 2023. ISBN: 9798350312249
4. **Bogacsovics, G.**, Hajdu, A., Harangi, B.: Cell Segmentation in Digitized Pap Smear Images Using an Ensemble of Fully Convolutional Networks.
In: 2021 IEEE Signal Processing in Medicine and Biology Symposium : Proceedings, IEEE, Philadelphia, 1-6, 2021. ISBN: 9781665428972





List of other publications

Hungarian scientific articles in Hungarian journals (5)

5. **Bogacsovics, G.**, Hajdu, A., Harangi, B., Lakatos, I., Lakatos, R., Szabó, M., Tiba, A., Tóth, J., Tarcsi, Á.: Adatelemzési folyamat és keretrendszer a közigazgatás számára.
Közigazgatástudomány. 1 (2), 146-158, 2021. ISSN: 2786-1910.
DOI: <http://dx.doi.org/10.54200/kt.v1i2.24>
6. **Bogacsovics, G.**, Hajdu, A., Harangi, B., Lakatos, I., Lakatos, R., Szabó, M., Tiba, A., Tóth, J.: Napelemfarmok Magyarország területén történő elhelyezését segítő döntéstámogató rendszer fejlesztése.
Közigazgatástudomány. 1 (2), 134-145, 2021. ISSN: 2786-1910.
DOI: <http://dx.doi.org/10.54200/kt.v1i2.23>
7. Bátfa, N., Besenczi, R., Szabó, J., Jeszenszky, P., Buda, A., Járm, L., Lovas, R. B., Pál, M. K., **Bogacsovics, G.**, Kovács, E.: DEAC-Hackers: játszó hackerek, hackelő játékosok.
Inform Társad. 18 (1), 132-146, 2018. ISSN: 1587-8694.
DOI: <http://dx.doi.org/10.22503/infars.XVIII.2018.1.9>
IF: 0.222
8. Bátfa, N., **Bogacsovics, G.**, Paszerbovics, R., Antal, A., Czevár, I., Kelemen, V., Besenczi, R.: E-sportlók mérése.
Inform. Társad. 18 (1), 147-155, 2018. ISSN: 1587-8694.
DOI: <http://dx.doi.org/10.22503/infars.XVIII.2018.1.10>
IF: 0.222
9. Bátfa, N., Bersenszki, M., Lukács, M., Besenczi, R., **Bogacsovics, G.**, Jeszenszky, P.: Az e-sport és a robotpszichológia közös jövője.
Inform. Társad. 16 (4), 26-39, 2017. ISSN: 1587-8694.
DOI: <http://dx.doi.org/10.22503/infars.XVI.2016.4.2>
IF: 0.023

Foreign language scientific articles in international journals (3)

10. Lakatos, R., **Bogacsovics, G.**, Harangi, B., Lakatos, I., Tiba, A., Tóth, J., Szabó, M., Hajdu, A.: A Machine Learning-Based Pipeline for the Extraction of Insights from Customer Reviews.
Big Data Cogn. Comput. 8 (3), 1-24, 2024. EISSN: 2504-2289.
DOI: <http://dx.doi.org/10.3390/bdcc8030020>
IF: 3.7 (2022)
11. Bátfa, N., Papp, D., Besenczi, R., **Bogacsovics, G.**, Veres, D.: Benchmarking Cognitive Abilities of the Brain with the Event of Losing the Character in Computer Games.
Stud. Univ. Babeş-Bolyai. Inform. 64 (1), 15-25, 2019. ISSN: 1224-869X.
DOI: <http://dx.doi.org/10.24193/subbi.2019.1.02>





12. Bátfai, N., Papp, D., **Bogacsovics, G.**, Szabó, M., Simkó, V. S., Bersenszki, M., Szabó, G., Kovács, L., Kovács, F., Varga, E. S.: Object file system software experiments about the notion of number in humans and machines.
Cognition, Brain, Behavior. 23 (4), 257-280, 2019. ISSN: 2247-9228.
DOI: <http://dx.doi.org/10.24193/cbb.2019.23.15>

Foreign language conference proceedings (3)

13. **Bogacsovics, G.**, Harangi, B., Beregi-Kovács, M., Kupás, D., Lakatos, R., Serbán, N. D., Tiba, A., Tóth, J.: Assessing Conventional and Deep Learning-Based Approaches for Named Entity Recognition in Unstructured Hungarian Medical Reports.
In: 2024 IEEE 22nd World Symposium on Applied Machine Intelligence and Informatics (SAMI). Ed.: Kovács Levente, Liberios Vokorokos, IEEE, Piscataway, 77-82, 2024. ISBN: 9798350317206
14. Lakatos, R., **Bogacsovics, G.**, Hajdu, A.: Predicting the direction of the oil price trend using sentiment analysis.
In: IEEE 2nd Conference on Information Technology and Data Science (CITDS) : Proceedings. Ed.: Fazekas István, Institute of Electrical and Electronics Engineers (IEEE), Piscataway, 177-182, 2022.
15. Harangi, B., Tóth, J., **Bogacsovics, G.**, Kupás, D., Kovács, L., Hajdu, A.: Cell detection on digitized Pap smear images using ensemble of conventional image processing and deep learning techniques.
In: 11th International Symposium on Image and Signal Processing and Analysis (ISPA 2019). Eds.: S. Lončarić, R. Bregović, M. Carli, M. Subašić, Institute of Electrical and Electronics Engineers (IEEE), Piscataway, NJ, USA, 38-42, 2019. ISBN: 9781728131405

Total IF of journals (all publications): 9,267

Total IF of journals (publications related to the dissertation): 5,1

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

09 April, 2024

