

Doktori (PhD) értekezés tézisei

**Stochastic and deterministic optimization methods and
their applications**

Tiba Attila

Témavezető: Dr. Hajdu András



DEBRECENI EGYETEM

Informatikai Tudományok Doktori Iskola

Debrecen, 2022

Tartalomjegyzék

1. Bevezetés	1
2. Egy sztochasztikus megközelítés az együttes rendszerek metszésére erőforráskorlátok mellett	3
3. Együttes módszerek alkalmazása az orvostudományban	6
3.1. A koronavírus (SARS-CoV-2) betegség járványgönbéjének előrejelzése mesterséges intelligencia segítségével: Egy alkalmazás az első és második hullámra	6
3.2. Kiugró és rossz minőségű orvosi képek detektálása összetett mélytanuló rendszerrel	7
4. Determinisztikus módszerek mintázatok szabályosságának mérésére	8
Hivatkozások	12
Publikációs lista	15

1. Bevezetés

Az együttes-alapú megközelítések nagyon hatékonyak bizonyulnak a különböző területeken az egyedi tagok pontosságának növelésében, ha valamilyen szavazási szabályt alkalmazunk az egyes döntések összesítésére.

A disszertáció első részében azt vizsgáljuk, hogy hogyan lehet megtalálni és jellemezni a legnagyobb pontossággal rendelkező együttes rendszereket, ha az együttesek tagjainak összköltsége korlátozott. Ez a kérdés nemlineáris és nem szeparálható célfüggvénnyel rendelkező hátizsák problémához vezet bináris és többsztályos osztályozási feladatok esetén, ha az aggregáláshoz a többségi szavazást választjuk. Mivel a hagyományos megoldási módszerek nem alkalmazhatók erre a feladatra, ezért egy új sztochasztikus megközelítést vezetünk be a bináris esetben, ahol az energiafüggvényt a tagok pontosságának együttes valószínűségi függvényeként tárgyaljuk. Továbbá bemutatunk néhány elméleti eredményt a várható együttes pontosság és annak varianciája tekintetében a többsztályos osztályozási feladatban, amelyek segíthetnek a hátizsák probléma hatékony megoldásában ebben az esetben is.

A disszertáció következő részében az együttes-alapú módszerek alkalmazásának néhány esetét mutatjuk be az orvostudomány területén. Először a COVID-19 világjárvánnyal kapcsolatos kutatásunkat ismertetjük, amelynek célja a COVID-19 járványgörbéinek (napi új esetek) előrejelzése hivatalos járványügyi adatok felhasználásával, egymással összekapcsolt alhálózatokból álló neurális hálózati archi-

tektúra segítségével, majd az előrejelzett modellek összehasonlítása és validálása a megfigyelt adatokkal.

Ezt követően megvizsgáljuk az együttes-alapú módszerek hatékonyságát konvolúciós neurális hálózatokon alapuló (CNN) modellek esetében. Bemutatunk egy összetett, SVM (Support Vector Machine) osztályozóval kombinált konvolúciós neurális hálózatokból (CNN) álló, a többségi szavazás koncepciója alapján döntést hozó és nagyon pontos szűrést eredményező kiugró érték detektálási módszert.

A disszertáció utolsó részében determinisztikus optimalizálási módszereket mutatunk be a képminták szabályosságának mérésére. A textúra elemzés évek óta nagy figyelmet kap a kutatók részéről. Számos esetben fontos annak meghatározása, hogy a textúra szabályos-e vagy, hogy a textúrán belül a mintázat valamilyen szabályosságot követ-e. Az együttes előfordulási mátrixokat, mint másodrendű statisztikai leírók forrásait gyakran használják a textúra osztályozási feladatokban. Egy ilyen mátrix létrehozásához szükségünk van egy helyzetvektorra, amelynek végpontjaiban ellenőrizni tudjuk a lehetséges intenzitás értékeket. Ebben a disszertációban egy hatékony algoritmust javasolunk olyan helyzetvektorok felkutatására, amelyek szerint a textúra mintázata ismétlődik, és így az együttes előfordulási mátrixból származtatott leírók (Haralick-jellemzők) alkalmasak a minta szabályosságának jellemzésére. Végezetül egy algoritmust adunk meg annak eldöntésére, hogy egy digitális képen valamilyen ismétlődően előforduló minta periodikus jellegűnek tekinthető-e vagy sem.

2. Egy sztochasztikus megközelítés az együttes rendszerek metszésére erőforráskorlátok mellett

Az együttes rendszerek számos alkalmazási területen meglehetősen népszerűek, és az egyedi megközelítések döntési pontosságának növelésére alkalmazzák őket.

A jelenlegi szakirodalomban az együttes-alapú megközelítéseket rendszeresen alkalmazzák a mintaosztályozók [3] vagy detektáló algoritmusok [2] kimeneteinek aggregálására, általában valamilyen többségi szavazáson alapuló szabály alapján. Ennek során az együttes rendszerek létrehozásának talán legnegatívabb tulajdonságával, az erőforrások növekvő igényével is szembesülhetünk. Ez a fajta költség az adott probléma jellegzetességeinek megfelelően jelentkezhet például végrehajtási/tanítási idő, vagy az együttes komponensek létrehozásához szükséges munkaidő formájában. Így a legpontosabb együttes rendszer összeállításának elsődleges célja mellett természetes kényszerként jelenik meg e célkitűzés költségkorlátozása.

A jelenlegi szakirodalom a hatékony együttes rendszer kiválasztását a lehetséges tagok egy csoportjából együttes metszésnek nevezi [18]. Még ha nem is alkalmazunk erőforráskorlátozást, a lehetséges együttes tagok egy részhalmaza jobb teljesítményhez vezethet, mint az összes tag kiválasztása. Sőt, a legjobb stratégia egy olyan együttes rendszer összeállítása, amely olyan jól teljesítő tagokból áll, amelyek viselkedése is változatos.

Ebben a munkában egy, az erőforrásokra vonatkozó, egyetlen korlátozással rendelkező problémát elemezünk annak érdekében, hogy a [9]-ban szereplő aggregációs szabályhoz hasonlóan a többségi szavazással kialakított energiafüggvény tekintetében a legpontosabb együttes rendszert állítsuk elő. Az általunk figyelembe vett korlátozó megkötés a tanítási időnek felel meg, azonban bármilyen más típusú erőforrás is figyelembe vehető.

Az alábbi tézispont összefoglalja az ezzel kapcsolatos megállapításokat.

1.1. Tézis. *Egy új sztochasztikus modellt hoztam létre az együttes-alapú bináris osztályozáshoz, amely az energiát a tagok pontosságának együttes valószínűségi függvényeként tekinti. Elméleti úton bebizonyítottam és kísérletileg igazoltam, hogy ez a modell hatékonyan beépíthető egy sztochasztikus keresési folyamatba, mint megállási szabály a pontos együttes rendszerek megtalálása érdekében.*

Alapvetően egy rendezésen alapuló, sztochasztikus mintavételezéssel kombinált megközelítést követünk az együttes rendszerek összeállításához, emellett azonban saját hozzájárulásként egy új heurisztikát is javasolunk ehhez [8]. Nevezetesen, az egyéni pontosság és költség mellett a kiválasztás során minden egyes lehetséges tag számára olyan hasznossági értéket számítunk, amely tükrözi annak közvetlen viselkedését a célfüggvény tekintetében, amely esetünkben a többségi szavazási szabályon alapul. Újszerű sztochasztikus keresési módszerünk bizonyítottan versenyképes a szimulált hűtés-sel (SA) [4], és más metszési módszerekkel [5, 13]. Emellett a ja-

vasolt heurisztika sikeresen beilleszthető ezekbe az általános sztochasztikus keresési stratégiákba is.

Az alábbi tézispont összefoglalja az ezzel kapcsolatos megállapításokat.

1.2. Tézis. *Egy újszerű és hatékony, az energiához jobban illeszkedő sztochasztikus keresési módszert adtam meg, amely más sztochasztikus stratégiákba is beépíthető. Nyilvánosan elérhető adatbázisok és szabványos metrikák segítségével kimutattam a módszer hatékonyságát más korszerű módszerekkel szemben.*

Egy objektumdetektálási problémán keresztül azt is bemutatjuk, hogy a javasolt modell hogyan általánosítható többosztályos osztályozási feladatokra [6]. Ehhez megadunk néhány elméleti eredményt a várható együttes pontosság és annak szórása tekintetében a többosztályos osztályozási feladatban [16].

Az alábbi tézispont összefoglalja az ezzel kapcsolatos megállapításokat.

1.3. Tézis. *Új elméleti eredményeket vezettem le a várható együttes pontosság és annak varianciájára vonatkozóan többosztályos osztályozási feladatra a hátizsák probléma kényszerfeltétele mellett. Az eredményeket egy speciális objektumdetektálási problémán értékeltem ki, amely célja a vakfolt megtalálása retinális képeken.*

3. Együttes módszerek alkalmazása az orvostudományban

3.1. A koronavírus (SARS-CoV-2) betegség járványgörbéjének előrejelzése mesterséges intelligencia segítségével: Egy alkalmazás az első és második hullámra

A COVID-19 világjárványt a globális közegészségügyet fenyegető legnagyobb veszélynek tekintik.

A szakirodalomban a különböző fertőző betegségek dinamikájának leírására és becslésére különféle matematikai modelleket használnak [10]. Ezek a modellek, amelyeket a fertőző betegségek dinamikájának szimulálására alkalmaznak, statisztikai, empirikus vagy gépi tanulási módszereken alapulnak [14]. A mesterséges neurális hálózatok egy formáját, a hosszú rövidtávú memóriával (LSTM) rendelkező rekurrens neurális hálózatokat (RNN) korábban az influenzajárvány modellezésére és előrejelzésére használták, kimagasló versenyképességgel és megbízható eredményekkel [17].

Tanulmányunk célja, hogy a hivatalos járványügyi adatok felhasználásával előrejelezzük a COVID-19 járványgörbéit (napi új esetek) rekurrens neurális hálózatok (RNN) segítségével, majd összehasonlítsuk és validáljuk a prediktált modelleket a megfigyelt adatokkal [11].

Az alábbi tézispont összefoglalja az ezzel kapcsolatos megállapításokat.

2.1. Tézis. *Új módszert dolgoztam ki a COVID-19 új eseteinek előrejelzésére rekurrens neurális hálózatok együttes rendszerének felhasználásával, extra jellemzők hozzáadásával, és a transzfer tanulást kombinálva az összekapcsolt alhálózatok komplex architektúrájával. A megközelítést a COVID-19 világjárvány első és második hullámának adatain teszteltem, és összehasonlítottam más korszerű módszerekkel. Az új modell realisabb előrejelzést biztosít hosszabb időszakra vonatkozóan.*

3.2. Kiugró és rossz minőségű orvosi képek detektálása összetett mélytanuló rendszerrel

A fejezet utolsó részében a kiugró értékek detektálásával kapcsolatos eredményeinket mutatjuk be, szintén egy együttes-alapú megközelítést alkalmazva [15].

Számos oka lehet annak, hogy egy adatbázisban nem megfelelő adatok fordulnak elő. A pontos eredmények és következtetések levonásához elengedhetetlen ezen elemek felderítése és kiküszöbölése. Az adatbázisban keletkező anomáliák kiszűrésének folyamatát kiugró érték detektálásnak nevezzük. A kiugró értékek, amelyek az adatok többi megfigyeléseitől eltérő szélsőséges értékek, egy mérés változékonyságára, kísérleti hibákra vagy újdonságra utalnak.

Célunk egy olyan eljárás kidolgozása volt, amely képes hatékonyan kiszűrni az adathalmazban a különböző típusú anomáliákat mutató képeket. A kiugró értékek detektálásának konkrét alkalmazásaként színes szemfenékképeket és bőrelváltozásokat tartalmazó bőrgyógyászati képeket vizsgáltunk.

Az alábbi tézispont összefoglalja a kapcsolódó eredményeket.

2.2. Tézis. *Létrehoztam egy együttes-alapú kiugró érték detektálási módszert, ahol az együttes tagjai konvolúciós neurális hálózatok (CNN) egy SVM (Support Vector Machine) osztályozóval kombinálva. Az aggregáláshoz többségi szavazást használtam, ami nagyon pontos kiugró érték szűrést eredményezett. A javasolt módszer teljesítményét retina-, illetve bőrelváltozásokat tartalmazó képekből álló adatbázisok szűrésére kiértékeltem. Az eredmények azt mutatják, hogy a javasolt együttes rendszer javította az egyedi tagok hatékonyságát a kiugró értékek felismerésében mind a retinális, mind a bőrelváltozásos képek esetében.*

4. Determinisztikus módszerek mintázatok szabályosságának mérésére

A textúrák vizsgálata évek óta nagy figyelmet kap a kutatók részéről. Számos esetben fontos annak meghatározása, hogy a textúra szabályos-e, vagy legalábbis annak megállapítása, hogy a textúrán belül a mintázat valamilyen szabályosságot mutat-e. A minták szabályosságának ellenőrzését széles körben használják például az orvostudományban [1, 7].

Az együttes előfordulási mátrixokat mint másodrendű statisztikai leírók forrásait gyakran használják a textúra osztályozási feladatokban. Egy ilyen mátrix létrehozásához szükségünk van egy helyzetvektorra, amelynek végpontjaiban ellenőrizni tudjuk a lehetséges

intenzitás értékeket. A helyzetvektort általában rövidnek tekintik, azonban a szakirodalomban nincs kifejezett ajánlás a hosszára vonatkozóan. A jelenlegi empirikus alkalmazási séma szerint a vektort 45, 90, 135 fokos szögben elforgatjuk, és minden egyes változathoz létrehozunk az együttes előfordulási mátrixokat. Szembesülhetünk azonban azzal a helyzettel, amikor nincs előzetes tudásunk az esetleg ismétlődő textúraelemek távolságáról és irányáról. Ezekben az esetekben a fő irányokban a rövid helyzetvektorok az együttes előfordulási mátrixból kinyert kevésbé karakterisztikus leíróhoz vezethetnek. Az ismétlődő komponenseket feltáró vektorok legegyszerűbb megtalálása időigényes keresésen (brute-force eljárás) alapulhat az összes lehetséges helyzetvektor ellenőrzésével. Ez a művelet azonban igen költséges, különösen nagyméretű képek elemzése esetén.

A megfelelő helyzetvektorok megtalálására olyan módszertant vezetünk be, amely könnyen és rövid idő alatt ajánl egy megfelelő vektort az együttes előfordulási mátrix létrehozásához. Megközelítésünk a Lenstra-Lenstra-Lovász (LLL) algoritmuson alapul [12], amely felhasználásával az eljárásunk képes a minta szabályosságának vizsgálatára a mintából kinyert pontfelhőre egy jól közelítő rácsot illeszteni. A közelítés hibájából a minta szabályosságára is lehet következtetni, és a közelítő rácsokat meghatározó bázisvektorok valamilyen ismétlődésre utalnak a mintában.

Az ennek megfelelő eredményeket a következő két tézispont tartalmazza.

3.1. Tézis. *Összeállítottam egy hatékony algoritmust olyan helyzetvektorok felkutatására, amelyek szerint a textúra mintázata ismétlődik, és így az együttes előfordulási mátrixból származtatott leírók (Haralick jellemzők) képesek jellemezni ezen textúra szabályosságát. Erre a célra jól közelítő rácsokat határoztam meg az LLL algoritmus segítségével, amely polinomiális futási idővel rendelkezik, így sokkal hatékonyabb megoldást nyújt, mint a brute-force algoritmuson alapuló keresés.*

3.2. Tézis. *Az előző tézispontban meghatározott módszert egy saját képészleten értékeltem ki. Az együttes előfordulási mátrixokat az általánosan használt szomszédsági vektorok, az LLL algoritmus által javasolt vektor, valamint a brute-force eljárással végzett kereséssel talált vektorok segítségével állítottam össze. Ezután a képeket Naive Bayes, Bayes-háló és 5 rejtett réteggel rendelkező több-rétegű perceptron osztályozó (MLPC) segítségével osztályoztam. A javasolt megközelítés az optimális (brute-force alapuló) megoldásokhoz nagyon közeli megoldásokat adott.*

A fenti megközelítésünkben csak azt követeltük meg, hogy egy közelítő rácspontra a bemeneti halmaz minden egyes pontjának közeli környezetébe essen. Ez a hibamérés azonban figyelmen kívül hagyja a bemeneti mintázat esetleges lyukait, mivel nem bünteti azokat a fordított eseteket, amikor nincsenek a közelítő rácsponthoz közeli bázisponthoz. Ezért e probléma megoldása érdekében hibamérésünket kiegészítjük egy kiegészítő ellenőrzéssel, azért hogy a közelítő rácsponthoz száma közel legyen a bemeneti ponthalmaz kardinalitá-

sához. A megfelelő kiegészítő feltétel úgy fogalmazható meg, hogy az eredeti ponthalmaz konvex burkában lévő rácspontokat megszámloljuk.

Az alábbi tézispont összefoglalja az ezzel kapcsolatos megállapításokat.

3.3. Tézis. *Egy új algoritmust adtam meg annak eldöntésére, hogy egy digitális képen ismétlődően előforduló minta periodikusnak tekinthető-e vagy sem. Ennek érdekében kivontam a textúra elemeket, és egyetlen képponttal ábrázoltam őket. A rácsméletet és az LLL algoritmust használtam arra, hogy rácsokat illesszek erre a ponthalmazra, valamint Barvinok hatékony számlálási módszerét a szabályosság meghatározására. Néhány megfelelő transzformáció után a konvex burkot vettem figyelembe a hiányzó mintakomponensekre illesztett rácspontok felderítésére és büntetésére.*

A módszer gyakorlati demonstrációjaként bemutatjuk, hogyan alkalmazható a bőrelváltozások képein a típusos/atípusos pigment-hálózatok szegmentációs hibáinak felismerésére.

Hivatkozások

- [1] Abásolo, D., Hornero, R., Espino, P., Poza, J., Sánchez, C.I., de la Rosa, R.: Analysis of regularity in the EEG background activity of Alzheimer’s disease patients with Approximate Entropy. *Clinical Neurophysiology* **116**(8), 1826–1834 (2005)
- [2] Antal, B., Hajdu, A.: An ensemble-based system for microaneurysm detection and diabetic retinopathy grading. *IEEE Trans. on Biomed. Eng.* **59**(6), 1720–1726 (2012). DOI 10.1109/TBME.2012.2193126
- [3] Antal, B., Hajdu, A.: An ensemble-based system for automatic screening of diabetic retinopathy. *Knowledge-Based Systems* **60**, 20–27 (2014). DOI <https://doi.org/10.1016/j.knosys.2013.12.023>
- [4] Du, K., Swamy, M.: *Search and Optimization by Metaheuristics: Techniques and Algorithms Inspired by Nature*. Springer International Publishing (2016)
- [5] Goldberg, D.: *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Publishing (1989)
- [6] Hajdu, A., Hajdu, L., Jónás, A., Kovács, L., Tomán, H.: Generalizing the majority voting scheme to spatially constrained voting. *IEEE Transactions on Image Processing* **22**(11), 4182–4194 (2013). DOI 10.1109/TIP.2013.2271116

- [7] Hajdu, A., Harangi, B., Besenczi, R., Lázár, I., Emri, G., Hajdu, L., Tijdeman, R.: Measuring regularity of network patterns by grid approximations using the LLL algorithm. In: 23rd International Conference on Pattern Recognition (ICPR 2016), pp. 1525–1530. Cancun, Mexico (2016)
- [8] Hajdu, A., Terdik, G., Tiba, A., Toman, H.: A stochastic approach to handle resource constraints as knapsack problems in ensemble pruning. *Machine Learning* **111**, 1551–1595 (2022). DOI [10.1007/s10994-021-06109-0](https://doi.org/10.1007/s10994-021-06109-0)
- [9] Hernández-Lobato, D., Martínez-Munoz, G., Suarez, A.: Statistical instance-based pruning in ensembles of independent classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31**(2), 364–369 (2009)
- [10] Hethcote, H.: Modeling heterogeneous mixing in infectious disease dynamics, p. 215–238. Publications of the Newton Institute. Cambridge University Press (1996). DOI [10.1017/CBO9780511662935.030](https://doi.org/10.1017/CBO9780511662935.030)
- [11] Kolozsvári, L.R., Bérczes, T., Hajdu, A., Gesztelyi, R., Tiba, A., Varga, I., Al-Tammemi, A.B., Szöllősi, G.J., Harsányi, S., Garbóczy, S., Zsuga, J.: Predicting the epidemic curve of the coronavirus (SARS-CoV-2) disease (COVID-19) using artificial intelligence: An application on the first and second waves. *Informatics in Medicine Unlocked* **25**, 100691 (2021). DOI <https://doi.org/10.1016/j.imu.2021.100691>

- [12] Lenstra, A.K., Lenstra, W., Lovász, L.: Factoring polynomials with rational coefficients. *Math. Ann* **261**, 515–534 (1982)
- [13] Martinez-Munoz, G., Suarez, A.: Using boosting to prune bagging ensembles. *Pattern Recognition Letters* **28**, 156–165 (2007). DOI <https://doi.org/10.1016/j.patrec.2006.06.018>
- [14] Siettos, C.I., Russo, L.: Mathematical modeling of infectious disease dynamics. *Virulence* **4**(4), 295–306 (2013). DOI [10.4161/viru.24041](https://doi.org/10.4161/viru.24041). PMID: 23552814
- [15] Tiba, A., Bartik, Z., Toman, H., Hajdu, A.: Detecting outlier and poor quality medical images with an ensemble-based deep learning system. In: 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA), pp. 99–104 (2019). DOI [10.1109/ISPA.2019.8868911](https://doi.org/10.1109/ISPA.2019.8868911)
- [16] Tiba, A., Hajdu, A., Terdik, G., Tomán, H.: Optimizing majority voting based systems under a resource constraint for multiclass problems. In: I. Faragó, F. Izsák, P.L. Simon (eds.) *Progress in Industrial Mathematics at ECMI 2018*, pp. 529–534. Springer International Publishing, Cham (2019)
- [17] Xianglei, Z., Fu, B., Yang, Y., Ma, Y., Hao, J., Chen, S., Liu, S., Li, T., Liu, S., Guo, W., Liao, Z.: Attention-based recurrent neural network for influenza epidemic prediction. *BMC Bioinformatics* **20**, 575 (2019). DOI [10.1186/s12859-019-3131-8](https://doi.org/10.1186/s12859-019-3131-8)
- [18] Zhou, Z.H.: *Ensemble Methods: Foundations and Algorithms*, 1st edn. Chapman & Hall/CRC (2012)



Nyilvántartási szám: DEENK/358/2022.PL
Tárgy: PhD Publikációs Lista

Jelölt: Tiba Attila

Doktori Iskola: Informatikai Tudományok Doktori Iskola

MTMT azonosító: 10068658

A PhD értekezés alapjául szolgáló közlemények

Idegen nyelvű tudományos közlemények külföldi folyóiratban (2)

1. Hajdu, A., Terdik, G., **Tiba, A.**, Tomán, H.: A stochastic approach to handle resource constraints as knapsack problems in ensemble pruning.
Mach. Learn. 111, 1551-1595, 2022. ISSN: 0885-6125.
DOI: <http://dx.doi.org/10.1007/s10994-021-06109-0>
IF: 5.414 (2021)
2. Kolozsvári, L. R., Bérczes, T., Hajdu, A., Gesztelyi, R., **Tiba, A.**, Varga, I., Al-Tammemi, A. B., Szöllősi, G. J., Kolozsváriné Harsányi, S., Garbóczy, S., Zsuga, J.: Predicting the epidemic curve of the coronavirus (SARS-CoV-2) disease (COVID-19) using artificial intelligence: an application on the first and second waves.
Informatics in Medicine Unlocked. 25, 1-13, 2021. ISSN: 2352-9148.
DOI: <http://dx.doi.org/10.1016/j.imu.2021.100691>

Idegen nyelvű konferencia közlemények (4)

3. **Tiba, A.**, Bartik, Z., Tomán, H., Hajdu, A.: Detecting outlier and poor quality medical images with an ensemble-based deep learning system.
In: 11th International Symposium on Image and Signal Processing and Analysis (ISPA), IEEE Inst Electrical Electronics Engineers Inc, Piscataway, 99-104, 2019. ISBN: 9781728131405
4. Hajdu, L., Harangi, B., **Tiba, A.**, Hajdu, A.: Detecting Periodicity in Digital Images by the LLL Algorithm.
In: Progress in Industrial Mathematics at ECMI 2018. Ed.: István Faragó, Ferenc Izsák, Péter L. Simon, Springer, Cham, 613-619, 2019, (Mathematics in Industry ; 30.) (The European Consortium for Mathematics in Industry ; 30.) ISBN: 9783030275495
5. **Tiba, A.**, Hajdu, A., Terdik, G., Tomán, H.: Optimizing Majority Voting Based Systems Under a Resource Constraint for Multiclass Problems.
In: Progress in Industrial Mathematics at ECMI 2018. Ed.: István Faragó, Ferenc Izsák, Péter L. Simon, Springer, Cham, 529-534, 2019, (Mathematics in Industry ; 30.) (The European Consortium for Mathematics in Industry ; 30.) ISBN: 9783030275495





6. **Tiba, A.**, Harangi, B., Hajdu, A.: Efficient Texture Regularity Estimation for Second Order Statistical Descriptors.
In: Proceedings of the 10th International Image and Signal Processing and Analysis (ISPA).
Ed.: Stanislav Kovačič, Sven Lončarić, Matej Kristan, Vitomir Štruc, Mladen Vučić, University of Zagreb, Zagreb, 90-94, 2017. ISBN: 9781509040117

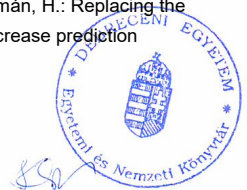
További közlemények

Magyar nyelvű tudományos közlemények hazai folyóiratban (2)

7. 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>
8. 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>

Idegen nyelvű tudományos közlemények hazai folyóiratban (3)

9. Lantang, O., Terdik, G., Hajdu, A., **Tiba, A.**: Comparison of single and ensemble-based convolutional neural networks for cancerous image classification.
Ann. Math. Inform. 54, 45-56, 2021. ISSN: 1787-5021.
DOI: <http://dx.doi.org/10.33039/ami.2021.03.013>
10. Lantang, O., Terdik, G., Hajdu, A., **Tiba, A.**: Investigation of the efficiency of an interconnected convolutional neural network by classifying medical images.
Ann. Math. Inform. 53, 219-234, 2021. ISSN: 1787-5021.
DOI: <http://dx.doi.org/10.33039/ami.2021.04.001>
11. 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)

12. Bankó, C., Nagy, Z. L., Nagy, M., Szemán-Nagy, G., Rebenku, I., Imre, L., **Tiba, A.**, Hajdu, A., Szöllösi, J., Kéki, S., Bacsó, Z.: Isocyanide Substitution in Acridine Orange Shifts DNA Damage-Mediated Phototoxicity to Permeabilization of the Lysosomal Membrane in Cancer Cells.
Cancers (Basel). 13 (22), 1-24, 2021. EISSN: 2072-6694.
DOI: <http://dx.doi.org/10.3390/cancers13225652>
IF: 6.575

Idegen nyelvű konferencia közlemények (2)

13. Lantang, O., **Tiba, A.**, Hajdu, A., Terdik, G.: Convolutional Neural Network For Predicting The Spread of Cancer.
In: Proceedings of the 10th IEEE International Conference on Cognitive Infocommunications : CogInfoCom 2019. Szerk.: Péter Baranyi, IEEE-Inst Electrical Electronics Engineers Inc, Piscataway, 175-180, 2019. ISBN: 9781728147932
14. Bérczes, A., Bérczes, T., Varga, I., **Tiba, A.**, Zsuga, J.: Using Laplacian spectrum to analyse the comorbidities network of hemorrhagic stroke.
In: Proceedings of the 10th IEEE International Conference on Cognitive Infocommunications : CogInfoCom 2019. Szerk.: Péter Baranyi, IEEE-Inst Electrical Electronics Engineers Inc, Piscataway, 53-60, 2019. ISBN: 9781728147932

A közlő folyóiratok összesített impact faktora: 11,989

A közlő folyóiratok összesített impact faktora (az értekezés alapján szolgáló közleményekre): 5,414

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

Debrecen, 2022.09.22.



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

**Stochastic and deterministic optimization methods and
their applications**

by Tiba Attila

Supervisor: Dr. Hajdu András



UNIVERSITY OF DEBRECEN

Doctoral School of Informatics

Debrecen, 2022

Contents

1	Introduction	1
2	A stochastic approach for ensemble pruning under resource constrains	3
3	Applications of ensemble methods in medicine	5
3.1	Predicting the Epidemic Curve of the Coronavirus (SARS-CoV-2) Disease Using Artificial Intelligence: An Application on the First and Second Waves . . .	5
3.2	Detecting outlier and poor quality medical images with an ensemble-based deep learning system . . .	6
4	Deterministic methods for measuring pattern regularity	7
	References	12
	List of publications	15

1 Introduction

Ensemble-based approaches are very effective in various fields in raising the accuracy of its individual members, when some voting rule is applied for aggregating the individual decisions.

In the first part of this dissertation, we investigate how to find and characterize the ensembles having the highest accuracy if the total cost of the ensemble members is bounded. This question leads to a Knapsack-problem with non-linear and non-separable objective function in binary and multiclass classification scenarios where majority voting is used for aggregation. As the conventional solving methods cannot be applied for this task, a novel stochastic approach was introduced in the binary case where the energy function is discussed as the joint probability function of the member accuracies. We show some theoretical results with respect to the expected ensemble accuracy and its variance also for the multiclass classification problem which can help us to solve the Knapsack-problem.

In the next part of the dissertation, we present some cases of the application of ensemble methods in the field of medicine. First, we show our research on the COVID-19 pandemic, which aimed to predict the COVID-19 epidemic curves (new cases per day) using official epidemiological data utilizing a neural network architecture consisting of interconnected subnetworks, and then compare and validate the predicted models with the observed data.

Furthermore, we investigate the effectiveness of ensemble methods for convolutional neural network (CNN) models. We present an

ensemble-based outlier detection method consisting of CNNs combined with a support vector machine (SVM) classifier. Experiments showed that it makes a majority voting-based decision very accurate in outlier filtering.

In the final part of the dissertation, we present deterministic optimization methods for measuring image pattern regularity. Texture analysis has received strong attention from researchers for many years. It is important in several cases to determine whether the texture is regular, or to determine at least the presence of some kind of regularity of the pattern within the texture. Co-occurrence matrices as sources of second-order statistical descriptors are commonly used in texture classification tasks. To generate such a matrix, we need a position vector to check possible intensity frequencies in its endpoints. In the dissertation, we propose an efficient algorithm to locate such position vectors according to which the pattern of the texture repeats and thus, the descriptors (Haralick features) derived from the co-occurrence matrix are capable to characterize the regularity of the pattern. Finally, we provide an algorithm to support the decision on whether some repeatedly occurring pattern in a digital image can be considered to have periodical nature or not.

2 A stochastic approach for ensemble pruning under resource constraints

Ensemble-based systems are rather popular in several application fields and are employed to increase the decision accuracy of individual approaches.

These approaches are also routinely considered in the current literature to aggregate the outputs of pattern classifiers [3] or detector algorithms [2] usually by some majority voting-based rule. In this process, we are also faced perhaps the most negative property of creating ensembles, that is, the increasing demand on resources. This type of cost may occur as the execution/training time and the working hours needed to create the ensemble components, etc., according to the characteristics of the given problem. Thus, in addition to the primary aim of the composition of the most accurate ensemble, a natural constraint emerges as a cost limitation for that endeavor.

The current literature mostly refers to the selection of an efficient ensemble from a pool of possible members as ensemble pruning [18]. Even if no resource constraints are applied, a subset of possible ensemble members may lead to better performance than selecting all the members. Moreover, the best strategy is to compose an ensemble having such good performing members which also have diverse behavior.

In this work, we analyze a single-constraint task on the resources to compose the most accurate ensemble regarding the energy formed by majority voting as the aggregation rule like in [9]. The constraint

we consider corresponds to the training time; however, any other type of resources could be considered.

The thesis point below summarizes my related findings.

Thesis 1.1. I have created a novel stochastic model for ensemble-based binary classification that considers the energy as the joint probability function of the member accuracies. I have theoretically proved and also validated experimentally that it can be efficiently incorporated in a stochastic search process as a stopping rule to find accurate ensembles.

Basically, we follow an ordering-based approach combined with stochastic sampling to compose the ensembles; however, additionally as a novel contribution we suggest a new heuristics for that [8]. Namely, besides its individual accuracy and cost, we calculate such a usefulness value for each possible member during the selection process that reflects its direct behavior according to the objective function, which is based on the majority voting rule in our case. Our novel stochastic search method is proven to be very competitive with respect to simulated annealing (SA) [4], and pruning methods [5, 13]. Also, the proposed heuristic can be successfully inserted into these general stochastic search strategies.

The thesis point below summarizes my related findings.

Thesis 1.2. I have provided a novel and efficient stochastic search method that better fits the given energy, which can be incorporated in other stochastic strategies as well. I have showed the dominance

of the method over other state-of-the-art ones using publicly available databases and standard metrics.

We also present how the proposed model is expected to be generalized to multiclass classification tasks with a demonstrative example on an object detection problem [6]. For this purpose, we provide some theoretical results with respect to the expected ensemble accuracy and its variance in the multiclass classification problem [16].

The thesis point below summarizes my related findings.

Thesis 1.3. *I have derived new theoretical results regarding the expected ensemble accuracy and its variance for the multiclass classification problem under a Knapsack constraint. I have verified the results in the optic disc detection problem in retinal images.*

3 Applications of ensemble methods in medicine

3.1 Predicting the Epidemic Curve of the Coronavirus (SARS-CoV-2) Disease Using Artificial Intelligence: An Application on the First and Second Waves

The COVID-19 pandemic is considered a major threat to global public health.

Various mathematical models may demonstrate and predict the dynamics of different infectious diseases [10]. These models, used to simulate the dynamics of infectious diseases, may be based on statistical, mathematical, empirical, or machine-learning methods [14].

One class of the artificial intelligence (AI) tools, the Recurrent Neural Networks (RNNs) with Long Short-Term Memory (LSTM) were previously used to model and forecast the influenza epidemic with strong competitiveness and reliable results [17].

The aim of our study is to use the official epidemiological data to forecast the epidemic curves (daily new cases) of the COVID-19 using RNNs, then to compare and validate the predicted models with the observed data [11].

The thesis point below summarizes my related findings.

Thesis 2.1. *I have developed a novel method to predict the new cases of COVID-19 using an ensemble of Recurrent Neural Networks, adding extra features, and combining transfer learning with a complex architecture of interconnected subnetworks. I have tested the approach on official data from the first and second waves of COVID-19 and compared it with other state-of-the-art methods. The new model provides a more realistic prediction over a longer period.*

3.2 Detecting outlier and poor quality medical images with an ensemble-based deep learning system

In the last part of this section, we present our results on outlier detection, also using an ensemble approach [15].

There are numerous reasons why inappropriate data can occur in a database. It is essential to detect and eliminate these elements for getting accurate results and conclusions. The process of filtering

anomalies generated in the database is called outlier detection. The outliers, that are extreme values deviating from other observations on data, indicate a variability in a measurement, experimental errors or a novelty.

Our aim was to develop a procedure that is able to efficiently filter images that exhibit different kind of anomalies in a dataset. As specific applications of outlier detection, we considered colour fundus images and skin lesions in dermatoscopic ones.

The thesis point blow summarizes my related findings.

Thesis 2.2. I have created an ensemble-based outlier detection method, where the members of the ensemble are Convolutional Neural Networks (CNNs) and a Support Vector Machine (SVM) is used as a classifier. I considered majority voting for aggregation which results in a very accurate outlier filtering. I have evaluated the performance of the proposed method for filtering databases consisting of retinal and skin lesion images, respectively. The results show that the proposed ensemble system improved the effectiveness of the member-level components in the outlier detection for both retinal and skin lesion images.

4 Deterministic methods for measuring pattern regularity

Examination of textures has received a strong attention form researchers for many years. It is important in several cases to determine whether the texture is regular or to determine at least the

presence of some kind of regularity of the pattern within the texture. Checking pattern regularity is widely used e.g. in medicine [1, 7].

Co-occurrence matrices as sources of second order statistical descriptors are commonly used in texture classification tasks. To generate such a matrix, we need a position vector to check possible intensity frequencies in its endpoints. The position vector is usually considered to be as a short one, however, no explicit recommendation for its length is available in the literature. The current empirical application scheme is to rotate the vector at angles of 45° , 90° , 135° , and create co-occurrence matrices for each variant. However, we can face the situation, when we do not have a priori knowledge regarding the distance and direction of the possibly repeating texture elements. In these scenarios, the short position vectors in the major directions might lead to a less characteristic descriptor extracted from the co-occurrence matrix. The simplest discovery of the vectors revealing the repeating components could be based on an exhaustive (brute force) search with checking all the possible position vectors. However, this operation is really expensive especially when large images are to be analyzed. Our goal is to provide a method for efficiently recommending position vectors tailored to the characteristics of a given texture.

To help with finding the appropriate position vectors, we introduce a methodology to recommend in a simple and fast way a suitable vector to generate the co-occurrence matrix. Our approach is based on the Lenstra-Lenstra-Lovász (LLL) algorithm [12] with making it capable to test the regularity of the pattern with automati-

cally fitting a grid on a point cloud extracted from the pattern. The LLL algorithm in its original form is a polynomial-time algorithm, which is able to perform basis reduction with providing a short vector in a new basis [12].

This algorithm becomes popular in several fields, mostly because of its polynomial running time and theoretical guarantees. With an appropriate modification of the original idea, an algorithm has been provided in [7] to find well-approximating grids on a set of points. From the error of the approximation one can also conclude to the regularity of the pattern, and the basis vectors spanning the approximating grids suggest some kind of repeat in the pattern.

My corresponding results are enclosed in the following two thesis points.

Thesis 3.1. *I have composed an novel algorithm on a theoretical basis to locate such position vectors according which the pattern of a texture repeats and thus, the descriptors (Haralick features) derived from the co-occurrence matrix are capable to characterize the regularity of it. For this purpose I have determined well-approximating grids by the LLL algorithm, which has a polynomial running time providing a much more efficient solution than brute force search.*

Thesis 3.2. *I have evaluated the method on an own image set. I have composed the co-occurrence matrices using the commonly used position vectors, the vector proposed by the LLL algorithm, and also the vectors found by brute force search. I have classified*

the images as regular and irregular ones by a Naive Bayes classifier, a Bayes Net and a multi-layer perceptron classifier (MLPC) with 5 hidden layers. The proposed approach has given the closest solutions to the optimal (brute force-based) ones.

The results show that the proposed approach is capable to suggest position vectors for an efficient co-occurrence matrix based texture analysis.

In our above approach we have required only that an approximating grid point should fall in a close environment of each point in the input set. However, this error measurement ignores possible holes in the input pattern since does not punish the reversed cases, when there are no base points close to the approximating grid points. Thus, to resolve this issue we complete our error measurement with a complementary check that the number of the approximating grid points should be close to that of the cardinality of the input point set. The proper extra condition can be formulated by counting lattice points in the convex hull of the original point set.

The thesis point below summarizes my related findings.

Thesis 3.3. I have provided a new algorithm to decide whether some repeatedly occurring pattern in a digital image has periodical nature. Accordingly, I have extracted textural elements and represented them by single pixels. I have used lattice theory and the LLL algorithm to fit lattices to this point set, and an efficient counting method of Barvinok to determine regularity. After some appropriate transformations I have considered the convex hull to

detect and punish lattice points fitted on missing pattern components.

As a practical demonstration of the method we have presented how it can be applied to recognize segmentation errors of atypical/typical pigmented networks in skin lesion images.

References

- [1] Abásolo, D., Hornero, R., Espino, P., Poza, J., Sánchez, C.I., de la Rosa, R.: Analysis of regularity in the EEG background activity of Alzheimer’s disease patients with Approximate Entropy. *Clinical Neurophysiology* **116**(8), 1826–1834 (2005)
- [2] Antal, B., Hajdu, A.: An ensemble-based system for microaneurysm detection and diabetic retinopathy grading. *IEEE Trans. on Biomed. Eng.* **59**(6), 1720–1726 (2012). DOI 10.1109/TBME.2012.2193126
- [3] Antal, B., Hajdu, A.: An ensemble-based system for automatic screening of diabetic retinopathy. *Knowledge-Based Systems* **60**, 20–27 (2014). DOI <https://doi.org/10.1016/j.knosys.2013.12.023>
- [4] Du, K., Swamy, M.: *Search and Optimization by Metaheuristics: Techniques and Algorithms Inspired by Nature*. Springer International Publishing (2016)
- [5] Goldberg, D.: *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Publishing (1989)
- [6] Hajdu, A., Hajdu, L., Jónás, A., Kovács, L., Tomán, H.: Generalizing the majority voting scheme to spatially constrained voting. *IEEE Transactions on Image Processing* **22**(11), 4182–4194 (2013). DOI 10.1109/TIP.2013.2271116

- [7] Hajdu, A., Harangi, B., Besenczi, R., Lázár, I., Emri, G., Hajdu, L., Tijdeman, R.: Measuring regularity of network patterns by grid approximations using the LLL algorithm. In: 23rd International Conference on Pattern Recognition (ICPR 2016), pp. 1525–1530. Cancun, Mexico (2016)
- [8] Hajdu, A., Terdik, G., Tiba, A., Toman, H.: A stochastic approach to handle resource constraints as knapsack problems in ensemble pruning. *Machine Learning* **111**, 1551–1595 (2022). DOI [10.1007/s10994-021-06109-0](https://doi.org/10.1007/s10994-021-06109-0)
- [9] Hernández-Lobato, D., Martínez-Munoz, G., Suarez, A.: Statistical instance-based pruning in ensembles of independent classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31**(2), 364–369 (2009)
- [10] Hethcote, H.: Modeling heterogeneous mixing in infectious disease dynamics, p. 215–238. Publications of the Newton Institute. Cambridge University Press (1996). DOI [10.1017/CBO9780511662935.030](https://doi.org/10.1017/CBO9780511662935.030)
- [11] Kolozsvári, L.R., Bérczes, T., Hajdu, A., Gesztelyi, R., Tiba, A., Varga, I., Al-Tammemi, A.B., Szöllősi, G.J., Harsányi, S., Garbóczy, S., Zsuga, J.: Predicting the epidemic curve of the coronavirus (SARS-CoV-2) disease (COVID-19) using artificial intelligence: An application on the first and second waves. *Informatics in Medicine Unlocked* **25**, 100691 (2021). DOI <https://doi.org/10.1016/j.imu.2021.100691>

- [12] Lenstra, A.K., Lenstra, W., Lovász, L.: Factoring polynomials with rational coefficients. *Math. Ann* **261**, 515–534 (1982)
- [13] Martinez-Munoz, G., Suarez, A.: Using boosting to prune bagging ensembles. *Pattern Recognition Letters* **28**, 156–165 (2007). DOI <https://doi.org/10.1016/j.patrec.2006.06.018>
- [14] Siettos, C.I., Russo, L.: Mathematical modeling of infectious disease dynamics. *Virulence* **4**(4), 295–306 (2013). DOI 10.4161/viru.24041. PMID: 23552814
- [15] Tiba, A., Bartik, Z., Toman, H., Hajdu, A.: Detecting outlier and poor quality medical images with an ensemble-based deep learning system. In: 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA), pp. 99–104 (2019). DOI 10.1109/ISPA.2019.8868911
- [16] Tiba, A., Hajdu, A., Terdik, G., Tomán, H.: Optimizing majority voting based systems under a resource constraint for multiclass problems. In: I. Faragó, F. Izsák, P.L. Simon (eds.) *Progress in Industrial Mathematics at ECMI 2018*, pp. 529–534. Springer International Publishing, Cham (2019)
- [17] Xianglei, Z., Fu, B., Yang, Y., Ma, Y., Hao, J., Chen, S., Liu, S., Li, T., Liu, S., Guo, W., Liao, Z.: Attention-based recurrent neural network for influenza epidemic prediction. *BMC Bioinformatics* **20**, 575 (2019). DOI 10.1186/s12859-019-3131-8
- [18] Zhou, Z.H.: *Ensemble Methods: Foundations and Algorithms*, 1st edn. Chapman & Hall/CRC (2012)



Registry number: DEENK/358/2022.PL
Subject: PhD Publication List

Candidate: Attila Tiba

Doctoral School: Doctoral School of Informatics

MTMT ID: 10068658

List of publications related to the dissertation

Foreign language scientific articles in international journals (2)

1. Hajdu, A., Terdik, G., **Tiba, A.**, Tomán, H.: A stochastic approach to handle resource constraints as knapsack problems in ensemble pruning.
Mach. Learn. 111, 1551-1595, 2022. ISSN: 0885-6125.
DOI: <http://dx.doi.org/10.1007/s10994-021-06109-0>
IF: 5.414 (2021)
2. Kolozsvári, L. R., Bérczes, T., Hajdu, A., Gesztelyi, R., **Tiba, A.**, Varga, I., Al-Tammemi, A. B., Szöllősi, G. J., Kolozsváriné Harsányi, S., Garbóczy, S., Zsuga, J.: Predicting the epidemic curve of the coronavirus (SARS-CoV-2) disease (COVID-19) using artificial intelligence: an application on the first and second waves.
Informatics in Medicine Unlocked. 25, 1-13, 2021. ISSN: 2352-9148.
DOI: <http://dx.doi.org/10.1016/j.imu.2021.100691>

Foreign language conference proceedings (4)

3. **Tiba, A.**, Bartik, Z., Tomán, H., Hajdu, A.: Detecting outlier and poor quality medical images with an ensemble-based deep learning system.
In: 11th International Symposium on Image and Signal Processing and Analysis (ISPA), IEEE Inst Electrical Electronics Engineers Inc, Piscataway, 99-104, 2019. ISBN: 9781728131405
4. Hajdu, L., Harangi, B., **Tiba, A.**, Hajdu, A.: Detecting Periodicity in Digital Images by the LLL Algorithm.
In: Progress in Industrial Mathematics at ECMI 2018. Ed.: István Faragó, Ferenc Izsák, Péter L. Simon, Springer, Cham, 613-619, 2019, (Mathematics in Industry ; 30.) (The European Consortium for Mathematics in Industry ; 30.) ISBN: 9783030275495
5. **Tiba, A.**, Hajdu, A., Terdik, G., Tomán, H.: Optimizing Majority Voting Based Systems Under a Resource Constraint for Multiclass Problems.
In: Progress in Industrial Mathematics at ECMI 2018. Ed.: István Faragó, Ferenc Izsák, Péter L. Simon, Springer, Cham, 529-534, 2019, (Mathematics in Industry ; 30.) (The European Consortium for Mathematics in Industry ; 30.) ISBN: 9783030275495





6. **Tiba, A.**, Harangi, B., Hajdu, A.: Efficient Texture Regularity Estimation for Second Order Statistical Descriptors.
In: Proceedings of the 10th International Image and Signal Processing and Analysis (ISPA).
Ed.: Stanislav Kovačič, Sven Lončarić, Matej Kristan, Vitomir Štruc, Mladen Vučić, University of Zagreb, Zagreb, 90-94, 2017. ISBN: 9781509040117

List of other publications

Hungarian scientific articles in Hungarian journals (2)

7. 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>
8. 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>

Foreign language scientific articles in Hungarian journals (3)

9. Lantang, O., Terdik, G., Hajdu, A., **Tiba, A.**: Comparison of single and ensemble-based convolutional neural networks for cancerous image classification.
Ann. Math. Inform. 54, 45-56, 2021. ISSN: 1787-5021.
DOI: <http://dx.doi.org/10.33039/ami.2021.03.013>
10. Lantang, O., Terdik, G., Hajdu, A., **Tiba, A.**: Investigation of the efficiency of an interconnected convolutional neural network by classifying medical images.
Ann. Math. Inform. 53, 219-234, 2021. ISSN: 1787-5021.
DOI: <http://dx.doi.org/10.33039/ami.2021.04.001>
11. 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)

12. Bankó, C., Nagy, Z. L., Nagy, M., Szemán-Nagy, G., Rebenku, I., Imre, L., **Tiba, A.**, Hajdu, A., Szöllősi, J., Kéki, S., Bacsó, Z.: Isocyanide Substitution in Acridine Orange Shifts DNA Damage-Mediated Phototoxicity to Permeabilization of the Lysosomal Membrane in Cancer Cells.
Cancers (Basel). 13 (22), 1-24, 2021. EISSN: 2072-6694.
DOI: <http://dx.doi.org/10.3390/cancers13225652>
IF: 6.575

Foreign language conference proceedings (2)

13. Lantang, O., **Tiba, A.**, Hajdu, A., Terdik, G.: Convolutional Neural Network For Predicting The Spread of Cancer.
In: Proceedings of the 10th IEEE International Conference on Cognitive Infocommunications : CogInfoCom 2019. Szerk.: Péter Baranyi, IEEE-Inst Electrical Electronics Engineers Inc, Piscataway, 175-180, 2019. ISBN: 9781728147932
14. Bérczes, A., Bérczes, T., Varga, I., **Tiba, A.**, Zsuga, J.: Using Laplacian spectrum to analyse the comorbidities network of hemorrhagic stroke.
In: Proceedings of the 10th IEEE International Conference on Cognitive Infocommunications : CogInfoCom 2019. Szerk.: Péter Baranyi, IEEE-Inst Electrical Electronics Engineers Inc, Piscataway, 53-60, 2019. ISBN: 9781728147932

Total IF of journals (all publications): 11,989

Total IF of journals (publications related to the dissertation): 5,414

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.

22 September, 2022

