

FELTÁRÓ KUTATÁS A HIPERPARAMÉTEREK SZEREPÉRŐL CNN-ALAPÚ ESZKÖZAZONOSÍTÁSBAN HPC ALKALMAZÁSÁVAL

Pap Martin Felicián, Kovács Ákos

Széchenyi István Egyetem Digitális Fejlesztési Központ

martin.pap@ddc.sze.hu, kovacs.akos@ddc.sze.hu

Absztrakt

A rádiófrekvenciás ujjlenyomat-alapú (RF fingerprinting) eszközazonosítás megbízható módszert kínál a vezeték nélküli kommunikációban részt vevő fizikai eszközök egyedi azonosítására, különösen olyan protokollok esetében, mint a Bluetooth. Kutatásunk célja egy mélytanuláson alapuló modell hiperparaméter-optimalizálása volt, amely képes felismerni az adóelektronikai eltérésekből származó jelsajátosságokat. A tanulmány során egy konvolúciós neurális háló (CNN) teljesítményét értékeltük többdimenziós hiperparaméter-tér feltárás révén, figyelembe véve a tanulási ráta, softmax hőmérséklet, optimalizáló algoritmus, osztályszám és mintaszám kombinációit.

Kulcsszavak: RF fingerprinting, Bluetooth, mélytanulás, HPC, open-set classification.

Abstract

Exploratory Research on the Role of Hyperparameters in CNN-based Device Identification Using HPC

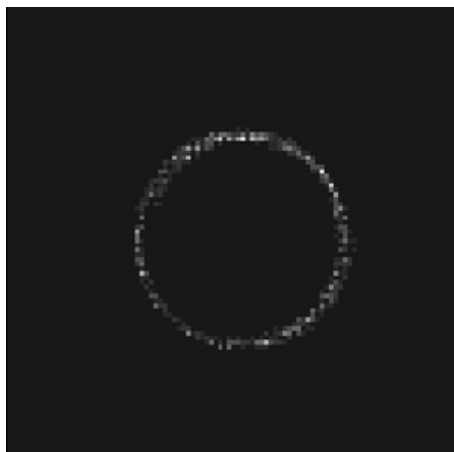
Radio frequency fingerprinting (RF fingerprinting) provides a reliable method for uniquely identifying physical devices participating in wireless communication, particularly in protocols such as Bluetooth. The aim of our research was to optimize the hyperparameters of a deep learning-based model capable of recognizing signal characteristics arising from transmitter hardware variations. In this study, we evaluated the performance of a convolutional neural network (CNN) through the exploration of a multidimensional hyperparameter space, taking into account combinations of learning rate, softmax temperature, optimization algorithm, the number of classes, and sample size.

Keywords: RF fingerprinting, Bluetooth, deep learning, HPC, open-set classification.

Bevezetés

A digitális eszközök megbízható azonosítása napjainkra a kiberbiztonság egyik kulcsfontosságú területévé vált, különösen a vezeték nélküli kommunikációs protokollokat (pl. Bluetooth) érintő környezetekben.

A rádiófrekvenciás ujjlenyomat-alapú (RF fingerprinting) eszközazonosítás lehetőséget nyújt arra, hogy egy-egy fizikai eszköz adóelektronikájának karakterisztikus eltéréseit kihasználva azonosítsuk annak kommunikációs mintázatát. Míg a korábbi módszerek jellemzően kézi jellemzőkivonásra és hagyományos klasszifikátorokra épültek, a közelmúltban a mélytanulási megközelítések számottevő előrelépést hoztak az eszközklasszifikáció pontosságában és általánosíthatóságában. Az 1. ábra szemlélteti az általunk használt Bluetooth-csomag alapján előállított ujjlenyomatot.



1. ábra Eszköz Bluetooth-csomag alapján előállított ujjlenyomat kép (saját ábra)

Kutatásunk a Bluetooth-csomagokra épülő rádiófrekvenciás ujjlenyomat-alapú (RF fingerprinting) eszközazonosításra fókuszált, zárt és nyílt osztályozási problémák kontextusában. A vizsgálat egyik központi eleme a mélytanuló modell átfogó tanítása volt, amely során széleskörű hiperparaméter-tér feltárást végeztünk.

A konvolúciós hálózat tanítása során számos kulcsfontosságú hiperparaméter (pl. tanulási ráta, optimalizáló algoritmus, softmax hőmérséklet, dropout arány, osztályszám, tanító elemek száma) értéke jelentős hatást gyakorol a modell prediktív teljesítményére és általánosíthatóságára. Az optimális konfiguráció kiválasztása nem végezhető megfigyelés vagy heurisztika alapján – különösen nem nyílt halmazos problémák esetén –, hanem szisztematikusan, többdimenziós paraméterter-feltárást igényel.

Ezért a kutatás során kiemelt szerepet kapott a nagy teljesítményű számítástechnikai (High Performance Computing, HPC) infrastruktúra alkalmazása, amely lehetővé tette

a több tanítási folyamat párhuzamos kezelését, a nagy méretű adatkészletek kezelését és a teljes kísérlet automatizált skálázását. A tanulmány különös figyelmet fordít annak bemutatására, hogy miként válik a HPC nem csupán elérhetővé, hanem nélkülözhetetlené a mélytanulási kísérletek hiperparaméter-optimalizálásának gyakorlatában, még mérsékelt modellméret mellett is.

A vizsgálathoz szükséges nagy számítási kapacitást a Digitális Kormányzati Fejlesztés és Projektmenedzsment Kft. (DKF) által biztosított Komondor HPC-rendszer tette elérhetővé.

Kapcsolódó kutatások és jelenlegi trendek

A rádiófrekvenciás (RF) ujjlenyomat-alapú eszközazonosítás olyan módszer, amely a vezetékek nélküli eszközök sugárzott jelébe kódolt, hardvereredetű torzításokat használja fel az egyedi eszközök felismerésére. Már korai kutatások igazolták, hogy például a Bluetooth-adók fizikai rétegű jelei egyedi mintázatot hordoznak: Hall és munkatársai (2006) analóg rádiójellemzők alapján detektáltak hálózatba illeszkedő *rogue* (illetéktelen) Bluetooth-eszközöket. [1] Később Ali et al. kézzel kigyűjtött jel feature-ök és hagyományos gépi tanulási algoritmusok (SVM, döntési fa) segítségével értékelték a Bluetooth fingerprinting hatékonyságát [2], míg Aghnaiya et al. variációs módusdekompozícióval (VMD) emelte ki a jelek stabil komponenseit a robusztusabb osztályozás érdekében [3]. Az alapötlet az, hogy minden adóberendezés RF-áramköre (oszillátorok, keverők, erősítők stb.) apró gyártási toleranciákból eredő eltéréseket visz a kisugárzott jelbe, amelyek az adott eszközre kvázi egyedi „ujjlenyomatként” szolgálnak. Ilyen jellemző például a vivőfrekvencia eltérése (CFO), az I/Q modulációs csatornák erősítés- és fáziskülönbsége vagy a tranziensek jelalakja, melyek eszközönként enyhén eltérnek, bár önmagukban nem lépik túl a szabványos toleranciákat. A korai tanulmányok többnyire egy-egy ilyen paraméter alapján kísérelték meg azonosítani a készülékeket, azonban újabb munkák rámutattak, hogy több különböző jel feature együttes használatával sokkal megbízhatóbb az eszközök megkülönböztetése. Peng és munkatársai például a differenciális konstelláció-nyomkép (Differential Constellation Trace Figure, DCTF) értékeit kombinálták CFO-val, modulációs hibákkal és I/Q eltolással, és így egyetlen jellemzőnél stabilabb, magasabb pontosságú azonosítást értek el. [4]

A mélytanulási algoritmusok számos olyan hiperparaméterrel rendelkeznek – például tanulási ráta, hálózati rétegek száma, neurális architektúra konfigurációk, regularizációs tényezők stb. – amelyek értékei kritikus módon befolyásolják a modell teljesítményét. Az optimális hiperparaméter-kombinációk szisztematikus felkutatása azonban igen erőforrás-igényes folyamat. Minden egyes kombináció esetén külön modellt kell betanítani és kiértékelni. A kombinációk száma a hiperparaméterek számával, valamint a lehetséges értékek halmazának méretével kombinatorikusan növekszik. Még viszonylag kis mélységű neurális hálók esetén is könnyen előállhat több tucatnyi vagy akár több száz, esetenként több ezer lehetséges konfiguráció, főleg, ha a modell teljesítményét alaposabban, például

keresztvalidációval értékeljük minden beállításnál. Ilyen átfogó kísérletek hagyományos egygépes környezetben gyakran rendkívül időigényesek vagy kivitelezhetetlenek.

A HPC-infrastruktúrák – tipikusan több csomópontba szervezett, nagy számú párhuzamos CPU-maggal és/vagy GPU-val rendelkező klaszterek – lehetővé teszik a hiperparaméter-keresés párhuzamosítását, jelentősen felgyorsítva a kísérletezést. [5] Egy HPC-klaszter erőforrásait kihasználva több modellt tudunk egyszerre, egymással párhuzamosan tanítani és tesztelni, így a hiperparaméter-tér különböző pontjain zajló kísérletek egyidejűleg futhatnak le. Ez a megközelítés drámai mértékben csökkenti a szükséges futási időt a szekvenciális (egymás utáni) kereséshez viszonyítva: egyes beszámolók szerint a nagymértékben párhuzamosított HPC-s hiperparaméter-optimalizálás az időigényt akár ~96%-kal is mérsékelheti. [6]

Egy orvosi gépi tanulási feladatban (Alzheimer-kór diagnosztikai modell) HPC-klaszteren valósítottak meg 100-szor ismételt, 5-fold keresztvalidáción alapuló SVM hiperparaméter-hangolást párhuzamos futtatással, ami a teljes számítási időt ~311 órától ~12 órára rövidítette – ~95–96%-os csökkenést eredményezve. [6]

Az általunk végzett kísérletben négyféle hőmérséklet-paraméter, háromféle tanulási ráta és háromféle optimalizáló algoritmus minden lehetséges kombinációját értékeltük ki, továbbá variálásra került a tréningben szereplő eszközkategóriák száma (2–11 különböző eszköz) és az egyes kategóriákhoz tartozó tanítóminták mennyisége (2 000-től 20 000-ig). Mindez összesen 1800 eltérő modell betanítását jelentette, ami a rendelkezésünkre álló, hagyományos környezetben szinte megoldhatatlan feladat lett volna.

Ez az eset rávilágít arra, hogy akár egy kisebb neurális háló architektúrateljesítményének szisztematikus feltérképezése is komoly számítási kapacitást igényelhet, ha egyszerre több tényezőt vizsgálunk meg alaposan.

HPC alkalmazása

A kutatás során a következő hiperparaméter-készletet vizsgáltuk szisztematikusan:

- softmax hőmérséklet: 1.0, 2.0, 5.0, 10.0,
- tanulási ráta: 0.01, 0.001, 0.0001,
- optimalizáló algoritmus: Adam, SGD, RMSprop,
- osztályszám: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11,
- osztályonkénti tanító halmaz méret: 2000, 5000, 10000, 15000, 20000.

Mindez

$$4 \times 3 \times 3 \times 10 \times 5 = 1800$$

modellkonfigurációt eredményezett, amelyeket külön-külön betanítottunk és kiértékelünk.

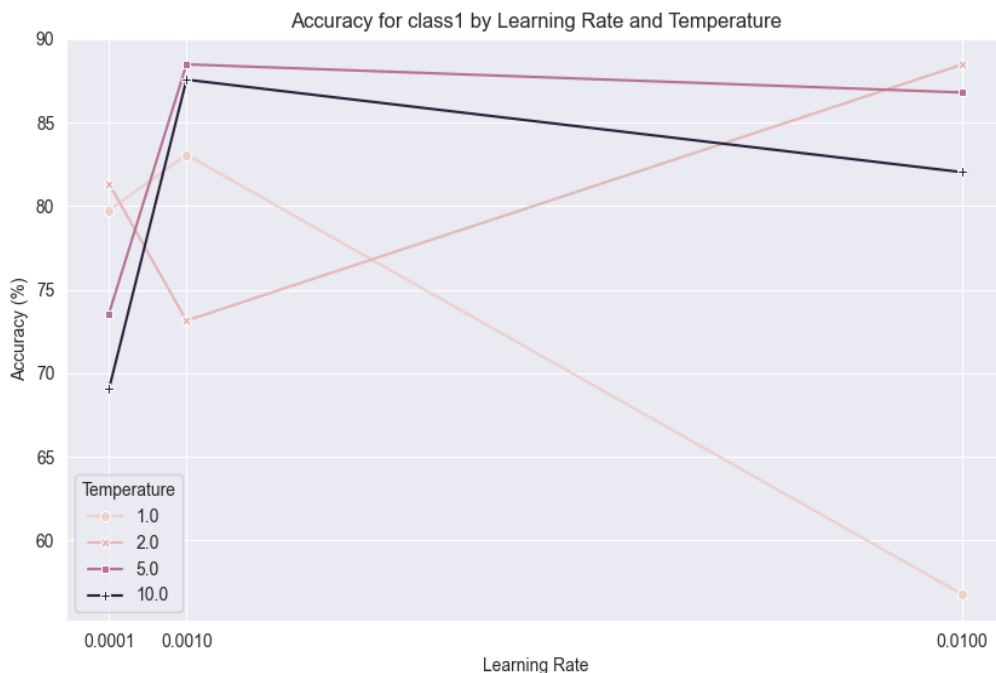
A HPC-használat indokoltsága

Az 1800 modell mindegyike külön tanítást és validációt igényelt; ez komoly idő- és számításkapacitás-igényt jelent, amelyet klasszikus egygépes környezetben nem lehetne hatékonyan kivitelezni.

A HPC lehetővé tette, hogy a 1800 modell párhuzamosan fusson. Az irodalomban hasonló hyperparaméteres SVM-futtatásoknál akár 96% időmegtakarítást mértek, ami egyértelműen igazolja ezt a megközelítést.[6]

A tanításhoz szükséges ~220 000 kép (11 eszköz × 20 000 kép) tárolása, előfeldolgozása és GPU-memóriába töltése csak HPC-n keresztül volt működtethető erőforráshatékonyan.

A HPC-infrastruktúra lehetővé tette, hogy többdimenziós paraméterter vizsgálatot hajtsunk végre: az egyes paraméterek (hőmérséklet, learning rate, optimizer, osztályok száma, adatméret) interakcióit feltérképezzük, ezáltal megalapozott következtetéseket tehetünk az általánosíthatóságról és teljesítményről.



2. ábra A learning rate és a temperature paraméter viszonya a modell pontosságához (saját ábra)

A 2. ábra egy példát mutat arra, hogy a HPC-alapú megközelítés milyen mélységű paraméterter-feltárást tett lehetővé. Jelen esetben a tanulási ráta (*learning rate*) és a hőmérséklet (*temperature*) különböző kombinációi mellett vizsgáltuk az osztályozási pontosságot.

Az ábrából jól látszik, hogy a paraméterek hatása nem triviális: egyes temperature-értékeknel a teljesítmény csökken a learning rate növelésével, míg más esetekben épp fordított tendencia figyelhető meg. Ezek az interakciók csak nagyszámú modellpárosítás tanítása és értékelése révén válhatnak láthatóvá – amit hagyományos környezetben nem lenne élet-szerű kivitelezni.

Megvalósítás

Egy 4× NVIDIA A100 GPU-s Komondor HPC node biztosította a párhuzamos modelltanításokat a komondor HPC-n, ahol SLURM volt a scheduler, amely hatékonyan ütemezte az 1800 modell tanítását. [9] 300 GPU órát vettünk igénybe a teljes tanítási és tesztelési folyamat futtatásához.

A checkpoint + early stopping mechanizmusok használata csak az ígéretes konfigurációk futását biztosította, ennek révén jelentős erőforrás-megtakarítást értünk el.

Minden futtatás eredményét külön teszthalmazon értékeltük ki, biztosítva, hogy nem csupán a tanítóadatokon, de valódi, korábban nem látott mintákon is mértünk teljesítményt – ez elengedhetetlen a valósághű általánosítás vizsgálatához. A teszthalmaz mérete osztályonként 1000 mintaelem volt. Tesztelés során a modell tanításában nem szereplő osztályokat összevontan az ismeretlen eszköz felismeréséhez használtuk.

A SLURM párhuzamos futtatása és checkpoint hibrid modell jelentős hatékonyságot biztosít komplex kísérleteknél a HPC-környezetben. [5]

A fejlesztések során világossá vált, hogy a laboratóriumi környezetre optimalizált TensorFlow/Keras-kód nem skálázódott megfelelően a Komondor párhuzamos környezetében. Ezzel szemben a PyTorch-alapú implementáció kiemelkedően stabilnak és hatékonyan futtathatónak bizonyult – különösen GPU-párhuzamosság esetén.

A laborból a valós HPC-környezetbe történő átjárás során alapvető szakmai előrelépést jelentett a PyTorch-ra való áttérés, a paramétergenerálás és SLURM-alapú szabályozás bevezetése. Ezek együtt lehetővé tették, hogy az 1800 konfiguráció párhuzamos futása hatékonyan és stabilan valósuljon meg – bizonyítva, hogy a helyes keretrendszer-választás és az erőforrás-particionálás kulcsfontosságú HPC-alapú kutatásokban.

További kutatási lehetőségek

A jelen kutatás eredményei rávilágítottak arra, hogy bár az RF fingerprinting módszerek hatékonyak, több kulcsfontosságú területen még jelentős eredmények érhetők el – különös tekintettel a több eszköz használatára és a valós környezetben előforduló zaj kezelésére.

Az RF fingerprinting jövőbeni fejlődése útközben egyre inkább a valós környezetek komplexitására reagál. A több eszköz egyidejű kezelése, a zaj- és csatornaváltozások beépítése, valamint az open-set skálázás mellett a HPC továbbra is alapvető támogatást nyújt a paraméterter és a kísérletsorozatok szisztematikus bővítéséhez. Ezek a kutatási irányok hozzájárulnak ahhoz, hogy a laboratóriumi eredményekből valódi, robusztus, gyakorlati alkalmazások váljanak.

Köszönetnyilvánítás

A kutatás/kutatómunka a Széchenyi István Egyetemért Alapítvány támogatásával valósult meg.

Köszönjük a Digitális Kormányzati Fejlesztési és Projektmenedzsment Kft.-nek, hogy hozzáférést biztosított számunkra a magyarországi Komondor HPC létesítményhez.

Irodalomjegyzék

- [1] J. Hall, M. Barbeau, and E. Kranakis, "Detecting rogue devices in Bluetooth networks using radio frequency fingerprinting," in Proc. 3rd IASTED Int. Conf. Commun. Comput. Netw. (CCN), 2006, pp. 108–113.
- [2] A. M. Ali, E. Uzundurukan, and A. Kara, "Assessment of features and classifiers for Bluetooth RF fingerprinting," IEEE Access, vol. 7, pp. 50524–50535, 2019, doi: [10.1109/ACCESS.2019.2911452](https://doi.org/10.1109/ACCESS.2019.2911452).
- [3] A. Aghnaiya, A. Ali, and A. Kara, "Variational mode decomposition based radio frequency fingerprinting of Bluetooth devices," IEEE Access, vol. PP, pp. 1–1, Oct. 2019, doi: [10.1109/ACCESS.2019.2945121](https://doi.org/10.1109/ACCESS.2019.2945121).
- [4] L. Peng, J. Zhang, M. Liu, and A. Hu, "Deep learning based RF fingerprint identification using differential constellation trace figure," IEEE Trans. Veh. Technol., vol. PP, pp. 1–1, Oct. 2019, doi: [10.1109/TVT.2019.2950670](https://doi.org/10.1109/TVT.2019.2950670).
- [5] L. Li, K. Jamieson, A. Rostamizadeh, E. Gonina, M. Hardt, B. Recht, and A. Talwalkar, "A system for massively parallel hyperparameter tuning," arXiv preprint arXiv:1810.05934, 2020. [Online]. Available: <https://arxiv.org/abs/1810.05934>
- [6] F. Zhang, M. Petersen, L. Johnson, J. Hall, and S. E. O'Bryant, "Accelerating hyperparameter tuning in machine learning for Alzheimer's disease with high performance computing," Frontiers in Artificial Intelligence, vol. 4, 2021, doi: [10.3389/frai.2021.798962](https://doi.org/10.3389/frai.2021.798962).
- [7] DKF, „Komondor rendszer SLURM alapú ütemezése,” docs.hpc.kifu.hu. [Online]. Available: <https://docs.hpc.kifu.hu>. [Accessed: Jun. 30, 2025].