Rádli Richárd Bence

"NOVEL DEEP LEARNING METHODS FOR OPTICAL RECOGNITION"

Című disszertációjának bírálata a nyilvános védésre

1. **Formai értékelés**

Bár a dolgozat első ránézésre terjedelmesnek tűnik a maga 166 oldalával, de a tényleges – a téziseket tárgyaló szakmai anyag ebből csak 103 oldal az irodalomjegyzéket is beleértve, azaz kimondható, hogy a dolgozat méretét tekintve megfelel egy PhD disszertáció elvárásainak. Formai oldalról nézve a dolgozat igényes. Az angolsága jó. A képletek az ábrák, a táblázatok rendezettek. Az egyes fejezeteket megfelelően bevezeti az irodalom áttekintésével.

2. **Tartalmi értékelés**

A dolgozat három témakört dolgoz fel, ezek eredményeit egy-egy téziscsoportban fogalmazza meg a Jelölt. Az első az autóenkoderekkel történő képi hibadetektálásra, a második a neurális hálózatok tanításának alternatív módszereire, a harmadik pedig metrikus hálózatokkal történő gyógyszer felismerésére fókuszál.

2.1 Második fejezet (első tézis)

Az első téziscsoport az autoecoder (AE) alapú képi hibadetekciót elemzi. A Jelölt három adathalmaz alkalmazásával vizsgál state-of-the-art irodalmi és általa javasolt AE variánsokat. Az első adatbázis hibás és jó textil mintákat tartalmaz. A második különböző állású ugyanolyan fajtájú csavarokat. A harmadik adatbázist a Jelölt maga készítette CPU-król. Ez az adatbázis - bár eléggé elrugaszkodik az ipari minőség-ellenőrzés körülményeitől, mert itt a hibák egy részét a képre utólag rárajzolt minták okozzák – mégis alkalmas arra, hogy egy harmadik képtípuson (részletgazdag elektronikai alkatrészekről felvett képeken) is be lehessen mutatni a kidolgozott módszert.

A Jelölt első körben két kiterjesztést alkalmaz az irodalmi autoencoderre. Egyrészt növeli a rétegszámot illetve a latens teret, másrészt kombinálja ezt denoising autoencoderrel. Az eredményekből az derül ki, hogy bár az autoencoder performanciája javul a model komplexitásának növelésével, de az eredmény hibadetekciós alkalmazása nem javul minden esetben. Ugyanis az egyik adatbázisra az egyik módszer teljesít jobban, a másikra a másik. Ennek megfelelően a Jelölt az 1.1 tézisében azt tudja csak megfogalmazni, hogy a jobb

minőséget biztosító autoencoder nem feltétlenül vezet jobb hibadetekcióhoz. A 2.3-as illetve a 2.4-es táblázatban mutatja be a jelölt az erre vonatkozó eredményeit. Ugyanakkor ebben a táblázatban egyes esetekben legjobbnak mondott eredmények **ROC AUC** értékei egyes esetekben csak a negyedik vagy az ötödik tizedesjegyben térnek el egymástól. Annak fényében, hogy a neurális hálózatok tanítása egy sztochasztikus módszer, biztosak lehetünk-e benne, hogy ezek a nagyon kis különbséggel első helyre befutó megoldásokból lehet-e következtetést levonni. Például, ha még egyszer lefuttatnánk a tanítást a Texture 2-es mintára, ahol a SSIM-AEe, z=500 és a SSIM-DAEe, z=500 ROC AUC értékek közötti különbsége annyira kicsi, hogy a négy tizedesjegyet tartalmazó táblázatban nem is látszik melyik a jobb, megint a SSIM-DAEe, z=500 jönne ki legjobbnak, vagy esetleg ezúttal a másik?

A Jelölt a 2.2-es tézisében alkalmazni kezd egy spatial transformert abból a célból, hogy az elforgatott mintákat visszaforgassa. Ez természetesen javít a hibadetekción, de ez igazából nem meglepő. Ugyanakkor a Jelölt is belefut abba a problémába, amibe a hasonló kísérletet elvégző cikk (bidart2019affine) is, nevezetesen, hogy a két háló szimultán tanítása beleragad lokális minimumokba. A Jelölt ennek megoldására előre betanította az STN hálózatot és csak ezt követően az autoencodert. Ez a megoldás minden alkalommal jelentős javulást eredményezett.

2.2 Harmadik fejezet (második tézis)

A harmadik fejezet a mély fully connected neurális hálózatok alternatív tanítási módszereivel foglalkozik. A Jelölt bemutat egy általa javasolt módszert, amelyik kiindul egy random súlyokkal generált egyrétegű hálózatból, majd csak mátrix inverziót használva, a leghasznosabb súlyokat kiválasztva – esetenként több párhuzamosan generált hálózatból- felépíti az egyre mélyebb hálózatot. A módszer meglepő módon akár 70-szeres tanítási sebesség növekedéshez is vezethet a mainstream backpropagation – ADAMS optimalizálóval szemben, ráadásul a hálózatok még jobban is teljesítenek.

A Jelölt a módszerét 13 publikus adatbázison mutatja be, összehasonlítva irodalmi más alternatív módszerekkel illetve a standard backpropagation megoldással. Az összehasonlításból az derül ki, hogy bár az irodalomban található gyorsabb megoldás is, de a Jelölt megoldása – hasonló nagyságrendű gyorsulást elérve, és a backpropagation–ADAMS párost egy nagyságrenddel megelőzve – egyes esetekben pontosabb eredményre vezet.

A kérdésem, hogy hol látja a módszer korlátját a Jelölt a mai neurális hálózat tanítási gyakorlatban? Mekkora komplexitású hálózatok taníthatóak hatékonyan ezzel a módszerrel?

2.3 Negyedik fejezet (harmadik téziscsoport)

A harmadik téziscsoportban a Jelölt bemutatott egy általa készített, metrikus módszerrel tanított 4 csatornás gyógyszerfelismerő hálózatot. A módszerben kifejezetten ügyes volt a szöveges gyógyszerleírások alkalmazása. A Jelölt a módszerét egy hozzáférhető adatbázison és egy saját maga által készített adatbázison tesztelte.

3. Összefoglalás

A dolgozat három téziscsoportja közül az utolsó két téziscsoportot tartom erős tudományos eredményeknek. Ezekről megállapítható, hogy a Jelölt sajátjai. Ezek alapján javaslom a nyilvános védés kitűzését és a fokozat odaítélését.

Budapest, 2026 június 26.

…….

Dr. Zarándy Ákos