

Horváth Péter válaszai a bírálók kérdéseire és megjegyzéseire

Szeretnék köszönetet mondani a bírálóknak, hogy elfogadták a felkérést, részletesen tanulmányozták a dolgozatomat, és megküldték bírálataikat, megjegyzéseiket és kérdéseiket. Az alábbiakban részletesen válaszolok ezekre. [KA.B] jelöli a feltett kérdést, míg [VA.B] az erre adott válaszomat.

Válaszok Prof. Dr. Csikász-Nagy Attila kérdéseire

[K1.1] Hány darab egyedi sejtet kell izolálni és karakterizálni ahhoz, hogy személyre szabott terápiás javaslatokat lehessen adni? Ez az analízis mennyi időt venne igénybe és mikorra várható ennek megjelenése a klinikumba?

[V1.1] Jelenlegi tudásunk alapján DNS-szekvenálás 1-5 sejt begyűjtése esetén lehetséges, mély transzkriptom analízist általában 100-1000 sejt közötti mennyiségből vagyunk képesek elvégezni (Brasko et al. 2018), lipidomikai elemzés 10-20 sejt esetében megbízható (Zsiros et al. 2023), és proteomikát jellemzően 100-200 sejtől tudunk végezni (Mund et al. 2022) (nem publikált adataink vannak sikeres 2-10 sejtől történt, mély proteomikai vizsgálatra egy hamarosan a piacra lépő készülékkel). Minden az előbbieken felsorolt szám humán daganatos mintákon értendő. Természetesen több rendelkezésre álló sejt növeli az analízis megbízhatóságát. A javasolt vagy ahhoz hasonló analízis klinikumba hozatala becslésünk szerint 5 éven belül megtörténhet, már jelenleg is vannak klinikai tesztek, pl. mi is futtatunk ilyet (lásd [V2.5]). Az analízis időtartama jelenleg a mintavételtől számított kb. két hét, melyet céleszközök bevezetésével jelentősen csökkenteni lehet majd, becsléseink szerint 1-2 napra.

[K1.2] A dolgozat benyújtásakor pár említett kézirat még bírálat alatt állt. Ezek közül valamelyikben sikerült újabb módszereket kifejleszteni, vagy inkább a módszerek használatára fókuszál a csoport?

[V1.2] Mindhárom javasolt tézis témakörében született saját és/vagy kollaborációs publikáció, melyek mind új módszereket, mind pedig a meglévők felhasználását írták le. Ezek közül a legfontosabbak:

Egysejtes képanalízis:

Fejlesztéseinket egysejt szegmentálásra koncentráltuk, olyan módszereket mutattunk be, melyek segítségével priori információt építettünk szegmentáló hálóba (Hirling et al., 2023a), kidolgoztuk az adatok intelligens augmentációját (Tasnádi et al., 2023), és a kiértékelés menetére adtunk ajánlást (Hirling et al., 2023b), valamint bemutattunk egy módszert különböző mikroszkópból származó felvételek regisztrációjára (Grexa et al., 2023).

Segmentation Metric Misinterpretations in Bioimage Analysis

D Hirling, ... P Horvath
Nature Methods, 2023

Structure preserving adversarial generation of labeled training samples for single cell segmentation

E Tasnadi, ... P Horvath
Cell Reports Methods, 2023

Unsupervised multimodal image registration with deep learning for biomedical microscopy

I Grexa, ... P Horvath
Briefings in Bioinformatics, 2023 (accepted in principal)

Cell segmentation and representation with shape priors

D Hirling, P Horvath
Computational and Structural Biotechnology Journal, 2023

Egysejt fenotipizálás:

Bemutattunk egy gyökeresen új látásmódot, mely segítségével az egysejt mikrokörnyezete hatékonyabban reprezentálható melytanuló algoritmusok számára (Tóth et al., 2022). Részt vettünk a Human Lung Atlas legfőbb publikációjában (Sikkema et al., 2023), valamint két fenotipizáló mélytanuló módszert írtunk le a Broad intézettel kollaborálva (Moshkov et al., 2023a, Moshkov et al., 2023b).

Fisheye transformation enhances deep-learning-based single-cell phenotyping by including cellular microenvironment

T Toth, D Bauer, F Sukosd, P Horvath
Cell Reports Methods 2022

An integrated cell atlas of the lung in health and disease

L Sikkema, C Ramirez-Suástegui, DC Strobl, TE Gillett, L Zappia, ...
Nature Medicine, 2023

Predicting compound activity from phenotypic profiles and chemical structures

N Moshkov, T Becker, K Yang, P Horvath, V Dancik, BK Wagner, ...
Nature Communications 2023

Learning representations for image-based profiling of perturbations

N Moshkov, M Bornholdt, S Benoit, M Smith, C McQuin, A Goodman, ...
Nature Communications 2023 (AIP)

Egysejt izoláció:

A módszer egy alkalmazását mutattuk be, mely egyben az első lipidomikai cikkünk lett (Zsíros et al., 2023). Szintén megjelent az első publikációnk egy új, több éve fejlesztett módszertanról, az egysejt mikroszkópia és Raman spektroszkópia korrelációjáról (Vörös et al., 2023).

Development of a Laser Microdissection-Coupled Quantitative Shotgun Lipidomic Method to Uncover Spatial Heterogeneity

V Varga-Zsíros, E Migh, A Marton, Z Kóta, C Vizler, L Tiszlavicz, P Horváth, ...
Cells 2023

Correlative Fluorescence and Raman Microscopy to Define Mitotic Stages at the Single-Cell Level: Opportunities and Limitations in the AI Era

C Voros, D Bauer, E Migh, I Grexa, AG Végh, B Szalontai, G Castellani, ... P Horvath
Biosensors 2023

[K1.3] A 2022-es Deep Visual Proteomics módszer igen nagy visszhangot kapott. A módszer továbbfejlesztése milyen irányban haladhat?

[V1.3] Három fő irányban folynak a sejt kiválasztási módszereink további technikai fejlesztései:

- 3D izoláció, hasonlóan a patch clamp-hez: Egy 10-15 µm nagyságú tűt vezérlünk 3D organoidokban a sejtek izolálásának céljából. Egyszerre 4-8 tű dolgozik majd.
- Dolgozunk egy galvó optikai lézeres kivágáson, mely a jelenlegihez képest kb. 10-100-szoros sebességgel képes egysejtek kivágására.
- Egy sokkal pontosabb regisztrációs módszert fejlesztünk, mely a modalitások közötti akár nemlineáris torzításokat valós időben korrigálja. Kézirat elfogadva (Grexá et al. 2023).

Válaszok Prof. Dr. Maurovich-Horvat Pál kérdéseire

[K2.1] A munkacsoport szinte minden vizsgálatukban támaszkodott a mesterséges intelligencia, a mélytanulás alkalmazására. Napjainkban gondok kezdenek felmerülni, hogy ezek a számítástechnikai módszerek nemcsak jó célokra, etikus módon használhatók, hanem lehetnek veszélyei, vadhajtásai is – amelyek jelenleg szabályozatlanok, pl. digitális mikroszkopiával készült felvételek módosítása, „hackelése” mesterséges intelligencia alkalmazásával. Hogy látja ezt a kérdést a Jelölt?

[V2.1] Köszönöm az érdekes kérdést. Természetes velejárója az egyre kifinomultabb mélytanulási módszereknek a hamisítás lehetősége. A GAN (Generative Adversarial Network) hálózatok alapvető tanítási módja is az, hogy egy generátor (csaló) és egy diszkriminátor (döntőbíró) verseng egymással. Sok más területen már köztudott, hogy a mélytanulás használható csalásra (lásd pl. a sok deepfake módszert). Már évek óta ismert az is, hogy voltak digitális patológiai esetek során mélytanulással történt csalások, jobb egészségügyi ellátás reményében, szintén több eset ismert a bíráló témakörében az orvosi képalkotásban. Véleményem szerint hasonló csalásoknak már most sincsenek technikai akadályai, és a jövőben ezek még kifinomultabbak lesznek, így szinte lehetetlen lesz ezek módszertani azonosítása. Azt gondolom, a mesterséges intelligencia nagyon erős jogi szabályozásának leszünk szemtanúi az elkövetkező években. Szintén érdekes kérdés, mennyire lehet meghamisítani kutatási eredményeket DL (Deep Learning) segítségével. A kiadók nagyon kifinomult módszereket használnak ennek kiszűrésére, de úgy vélem, a technológia itt is megállíthatatlan, és ezzel karöltve annak etikátlan felhasználása is.

[K2.2] Az értekezésben a daganatos sejtekkel kapcsolatos vizsgálatok érthető módon fókuszban álltak. Hogy látja a szerző, milyen további betegségtípusokban lehet még kiemelkedő szerepe az egysejt- vizsgálaton alapuló kórismének, ill. Terápiának?

[V2.2] Bár saját tapasztalatom legfőképpen a daganatterápiák egysejtes megközelítésében van, részt veszünk néhány más megbetegedés gyógyítására irányuló kutatásban is. Kiemelném ezek közül a (1) kardiovaszkuláris, (2) vírusfertőzéses, valamint (3) neurodegeneratív betegségeket. Mindhárom területen folytatunk kollaborációs kutatásokat.

- (1) A Semmelweis Egyetem Patológiai Intézetével közösen vizsgáljuk transzplantáció utáni esetleges gyulladások mértékét.
- (2) SARS-CoV-2 és influenza vírusok fertőzését, illetve azok sejtbe jutásának mechanizmusait vizsgáljuk egysejt szinten a Zürichi Műszaki Egyetemmel kollaborációban.
- (3) Alzheimer-modell egerek idegsejtjeinek morfológiai és molekuláris vizsgálatát végezzük további csoportokkal a Szegedi Biológiai Kutatóközpontban.

[K2.3] Néhány éven belül a mesterséges intelligencia támogatásával végzett kutatás jelentősen felgyorsíthatja az új tudományos felfedezéseket. Mi a jelölt véleménye hogyan alakítja át kutatók életét a mesterséges Intelligencia?

[V2.3] Megkülönböztetnék rövid és hosszú távú hatást. Úgy gondolom, a Bíráló itt a hosszú távú hatásokra gondolt, mégis megemlítenék néhány már jelenleg is meghatározó áttörést.

Rövid távon úgy gondolom, az adatok bányászatának gyorsabb és pontosabb megértésében segít az AI. Már most szemtanúi lehetünk világraszóló áttöréseknek, melyek nagyon hatékony adatfeldolgozást kínálnak, mint például az Alphafold, vagy a nyelvi modelleken alapuló ChatGPT, vagy a képfelismerésben használt konvolúciós hálózatok, melyek emberi pontossággal vagy még precízebben ismernek fel például melanomát. Ezek az adatbányász módszerek egyre kifinomultabbak, robosztusabbak és gyorsabbak lesznek, és jelenleg felfoghatatlan információáradattal szolgálnak, melyek már most is felgyorsították a tudományt; és ez csak a kezdet.

Feltételezésem szerint hosszú távon (10-20 év) a gépek már nem csak adatbányászatra, hanem adatok közötti összefüggések alapján hipotézisek felállítására és igazolására/elvetésére is képesek lesznek. Technológiai emberként nem vállalkoznék a jövő megjósolására, de úgy hiszem, a klasszikus értelemben vett adatok közötti összefüggések bizonyításán alapuló kutatást az AI magasabb szinten tudja majd művelni, mint az ember.

[K2.4] Milyen szerepe lehet a nagy nyelvi modelleken alapuló AI megoldások alkalmazásának a jelölt kutatási területén?

[V2.4] A tézis írásakor még csak kezdtek megjelenni a *transformer* alapú rendszerek, és elkezdődtek az első képi hálózatok fejlesztései, a *vision transformer*-ek. Mostanra ezek már szerves részét képezik mindennapjainknak. Csoportom esetében már most a használt modellek kb. 30-40%-a vision transformer alapú (2023 júliusi állapot). Meglátásom szerint ezen modellek

a nagy képi adatok esetében (pl. whole slide images, gigapixel méretben) jelentenek áttörést, mert képesek leszünk a kontextust jobban megérteni a finom részletek elvesztése nélkül.

Egy érdekes megközelítés lehet facebook által fejlesztett SAM (segment anything model)-je, melynek bemenete egy kép és egy szöveges input és az alapján szegmentálja a képen (pl "white cat" text input csak a fehér macskát szegmentálja a feketét nem). Ha rendelkezésünkre állna nagy szöveges adatbázis digitális patológiában/biológiában, akkor lehetne hasonló megoldásokat fejleszteni.

[K2.5] A doktori mű legvégén a jelölt az első klinikai sikertörténetet említi, amely nem publikált. Lehet tudni erről részleteket?

[V2.5] Az eset közlés alatt áll, néhány publikus részlet:

A páciens végstádiumban lévő 70 éves férfi (PatientX) primer melanómával (rapid progression melanoma) és nyirokcsomó, valamint tüdő áttéttel. Négy különböző daganatos sejtípust azonosítottunk AI segítségével a primer tumor metszetben, majd izoláltuk azokat. Az izolált mintákon proteomikai analízist végeztünk. A mintákban azonosítottunk egy mutációt, melyre létezik célzott terápia, és amelyet PatientX megkapott. PatientX állapotában jelentős javulást figyeltünk meg 14 hónapon keresztül, majd az alternatív terápia indítását követő 15. hónapban elhunyt. Tudomásunk szerint ez volt az első AI vezérelt egycell izolációs morfo-proteomikai alapon elvégzett személyre szabott kezelés.

Válaszok Prof. Dr. Ulbert István kérdéseire

A bíráló több esetben jelezte, hogy olyan állítások szerepelnek a tézisben, amik nem triviálisak, és nincs hozzájuk megfelelő irodalmi referencia. Ezekkel egyetértek, és köszönöm szépen a dolgozat ilyen precíz áttekintését. A feltett kérdésekre minden egyes esetben a forrás publikációkban bemutatott függelék dokumentumokban találunk választ. Ezeket részletesen idézem lentebb.

Szintén rámutat arra, hogy mindhárom tézispont esetében hiányzik a prior tudás és a módszerek esetleges limitációinak bemutatása. A stratégiám az volt, hogy a dolgozatom első fejezetében röviden áttekintem a szakma jelen állását. Egyetértek a limitációk bemutatásának hiányáról szóló megjegyzéssel, így ennek kifejtését itt részletesen megtettem.

K3.1 Záró aktusként következnek a véleményem szerint túlzottan tömör következtetések és a további tervek ismertetése. Jobb lett volna, ha a szerző például részletesebben diszkutálta volna ebben a műben is az általa kidolgozott módszerek határait és hibáit.

Limitációk, módszertani határok:

2. tézispont, képanalízis:

A képkorrektció területén a két legfontosabb limitáló tényező a megfelelő mennyiségű és eloszlású kép rendelkezésre állása és az esetleges nemlineáris torzító tagok. Bár meg kell jegyezni, hogy a módszert számos adathalmazon (12 adathalmaz) teszteltük és monitoroztuk a rendelkezésre álló képek függvényében elért rekonstrukciós pontosságot, azt tapasztaltuk, hogy

néhány esetben nagyobb számú képre van szükség. Ez legfőképp azon fluoreszcens képek esetében volt igaz, amik relatíve kevés vagy kis méretű objektumot tartalmaztak. Másodsorban, bár vizsgálataink során nem talákoztunk olyan adathalmazzal, ahol a javasolt lineáris modell nem lett volna megfelelő, elképzelhetőek olyan képalkotási technikák, melyek esetében magasabb fokú tagokat is érdemes kiszámítani. Ez technikailag lehetséges, de több felvételre lehet szükség.

Egysejt szegmentáció: Mindenképpen szétválasztanám a parametrikus modellekkel és energia minimalizációval történő szegmentációt a mélytanuló hálózatoktól.

Az előbbi esetében több limitáció ismert, melyek nagyon erősen korlátozzák felhasználhatóságukat. A legfontosabbnak ítélem az adattagtól való kitétséget. Ezen módszerek esetében nagyon nehéz általánosan jól működő adattagot készíteni, ezért nem tűz az az elvárás, hogy minden kísérlet futtatása előtt az adattag paramétereit újra kell hangolni. Másodsorban ezen módszerek erőssége a gyengéje is egyben, ugyanis minden esetben arra törekszünk, hogy a keresett alakzat leírására egy olyan parametrikus modellt készítsünk, mely matematikailag könnyen kezelhető, de a lehető legjobban ábrázolja a valós alakot. Ez nagyon gyakran nem triviális, például a bevezetett ellipszis modellünk (Molnar et al. 2016) nem általánosít jól minden sejtmag típusra.

Mélytanuló hálózatok esetében a legfőbb limitáció a tanítóhalmaz számossága, komplexitása és annak legmegfelelőbb felhasználása (gondolok itt pl. az augmentációra). Törekednünk kell a tanítóhalmaz leggondosabb elkészítésére, hiszen ennek meghatározó szerepe van a modell minőségében. A legfontosabb aspektusok a halmaz nagysága, a feldolgozandó adat lehető legátfogóbb reprezentációja a tanítóhalmazban, illetve a halmaz kiegyensúlyozottsága. Mindhárom szempont nem körültekintő megválasztása szignifikáns eredményromláshoz vezet, és a javasolt módszereket limitálja. Szintén figyelmet kell fordítani a tanítóhalmaz megfelelő augmentálására (mesterséges kiterjesztésére olyan példákkal amiket a meglévő példák közül különböző transzformációkkal állítunk elő). Egy jól megválasztott augmentáció kedvező hatással lehet a modellünk reprezentációs képességére, de nem körültekintően eljárva akár negatívan is befolyásolhatja a modellt. Szintén nagy körültekintést igényelnek a sokkal fejlettebb szintetizáló augmentációk, melyeket a Hollandi et al. 2022 és a Tasnadi et al. 2023 cikkekben mutattunk be. Ott több gyakorlati aspektust is leírtunk, melyek javíthatják a szegmentációt. Fontos limitáció, melyet érdemes megjegyezni, hogy a mélytanuló hálózatok sem képesek olyan adatra helyesen általánosítani, melyet még nem láttak, ezért olyan augmentációknak, melyek kiterjesztik az adat domain-t, fontos szerepük van (zaj, intenzitás, forgatás, új stílusok).

2. tézispont, fenotipizálás:

Fenotipizálás esetében két fontos limitációról kell beszélnünk, melyek (I.) a tanítóhalmazok mérete, valamint (II.) az *a priori* nem ismert fenotípusok kezelése.

Korábban részletes elemzéseket végeztünk arra vonatkozóan, hogy különböző sejtes minták esetében mekkora fenotipikus tanítóhalmaz szükséges (Smith et al. 2014). Ez a klasszikus gépi tanuló módszerek esetében 10-100 elem volt, míg a mélytanulók stabil működéséhez akár több ezer tanító minta is szükséges lehet. Ezt az utóbbi módszerek legfontosabb limitációjaként említeném.

Szintén érdekes és fontos megemlíteni a gépi döntés bizonytalanságát olyan sejtípusok esetében, melyek nem szerepeltek a tanítóhalmazban. Természetesen a gép itt is hoz egy matematikai alapú döntést, de semmi garancia, hogy ez orvosszakmailag helyes lesz. Számos módszer létezik outlier-ek kezelésére, de fontos a felügyelt módszerek ezen limitációjával tisztában lenni.

3. tézispont, egysejt izolálás:

A lézer mikrodiszekciónak három fontos limitációja van. Első és számukra a legfontosabb a vágási kontúr mentén keletkező égéstermék. Sajnos a lézer károsíthatja a fehérjét, a nukleinsavakat és a lipideket. Erre számos javító eljárást dolgoztak ki mások és mi is, melyeket alkalmazunk (pl. pulzáló lézerrel vágunk, ráhagyunk néhány μm -t a sejt körvonalára, ha lehetséges). Szintén érdemes tárgyalnunk a vágás sebességét. Sajnos ez egy nagyon erős limitáló faktor, főleg amíg az egysejt molekuláris analízis módszerek gyakran 100-1000 egysejtet is igényelnek. Jelenleg az áteresztő képességünk kb. 10 s/sejt. Végül megjegyezném a minta vastagságának kérdését. Állandó kompromisszumot kell kötnünk ebben, mivel a mikroszkópia a lehetséges legvékonyabb minták esetében működik a legjobban, míg a molekuláris analízisnek a lehető legnagyobb térfogatra van szüksége.

Az automatizált patch clamping sok limitációja közül két nehézségre hívnám fel a figyelmet. Első a képképzés, mely jelenleg oblique mikroszkópiával történik. Ezen képek a legnagyobb jóindulattal sem nevezhetők egyszerűnek és könnyedén interpretálhatónak. Ebben a környezetben nagyon komoly kihívás egysejteket megtalálni, azok közül az egészségeseket vagy még tovább menve adott fenotípusúakat. Másodsorban jegyezném meg a pipetta kiválasztott sejthez történő navigálását. Komoly kihívást jelentett algoritmizálni azt a folyamatot mely a pipetta útjába kerülő más sejtek, esetleg vérerek kikerülését vezérli.

[K3.2] Ami a tartalmi kérdéseket illeti, a disszertációban több helyen előfordulnak olyan indoklás, illetve referencia nélküli állítások, melyek nem triviálisak. Gondolok például a 7. oldalon a következőre: „...in most of the laboratories, the strategy of choice is retrospective multi-image correction function.”

[V3.2] A CIDRE módszerrel bemutató közleményünk (Smith et al. 2015) egyik fő előkészületi lépése egy átfogó kérdőív volt, melyet vezető mikroszkópos laborok töltöttek ki. Ebben felmértük a képjavítási szokásokat. A kérdőívet 170 kolléga töltötte ki világszerte. A Supplementary Note 2-ben található 14. kérdés konkrétan rákérdezett a jelenleg általuk alkalmazott képjavítási módszerre. Sajnálom, hogy a doktori értekezésemből kimaradt ennek idézése, köszönöm a Bírálónak, hogy erre felhívta a figyelmem.

[K3.3] A 14-15. oldalon megtudjuk, hogy a kalibráció mentes módszerek a kívánatosak, de nem tudjuk meg azt, hogy miért. Sejttem, de nem tudom, és ebben a szöveg sem segített. Kíváncsi lennék, miért hátrány a speciális kalibráció, ahogyan azt a szerző is sugallja.

[V3.3] Módszertanilag semmiképpen nem hátrány a kalibrációs slide-ok használata, azonban a gyakorlatban három limitáló tényezőjük van. Elsősorban az, hogy egy ilyen slide-ot be kell

szerezni, fluoreszcens mikroszkópia esetén minden csatornára külön-külön, egyéb esetekben pedig a felvenni kívánt mintával megegyező illuminációs tulajdonságút. Ezt a kalibrációs slide-ot optimális esetben minden kísérlet előtt/közben/után le kell fényképezni ami időigényes. Másodsorban, megfigyelésünk alapján a kalibrációs slide nem pontosan ugyanazokkal az optikai tulajdonságokkal rendelkezik, mint a minta. Megfigyeltük például, hogy fluoreszcens képek esetében a kalibrációs slide fluoreszkáló vastagsága nem azonos a mintáéval, ezért soha nem sikerült annyira pontos rekonstrukciót végeznünk, mint a javasolt CIDRE módszerrel. Végül, ha a kísérlet során elmarad a kalibrációs slide lefényképezése, akkor ezt pótolni nehézkes vagy gyakran lehetetlen. Köszönöm, hogy rávilágított arra, hogy ez kimaradt a téziszből.

[K3.4] Szintén hasonló csoportba tartoznak azok az állítások, melyek egy adott módszer pontosságára vonatkoznak, de sem magyarázat, sem idézet nem utal arra, miért van ez így, például a 17. oldal tetején: „... 10 images are sufficient for a good performance, while for datasets with low information content, approximately 100 images are necessary.”

[V3.4] Utalnék a Smith et al. 2015 Supplementary Figure 3-ra, ahol részletesen taglaltuk a rekonstrukció helyességét a rendelkezésre álló képek függvényében. Itt kiderült, hogy bizonyos modalitások és kísérleti típusok esetében különböző számú kép szükséges egy elfogadható pontosság eléréséhez. Köszönöm, hogy a bíráló felhívta a figyelmem, hogy ezen hivatkozás idézése kimaradt a disszertációból.

[K3.5] Hasonló kérdés vetődik fel például a gyakorló halmazok növelésének módszertanával kapcsolatban is, már ami a disszertációt illeti, nem a publikációkat. Kinyilatkoztatás szerű állításokat tartalmaz ezen a ponton is a mű, amelyek persze abszolút igazak is lehetnek, csak valahogy hiányoztak számomra a szövegben azok a mondatok, melyek indokolták volna az állítást. Például a 22. oldalon „We have augmented the training samples by creating realistic-looking artificial sample images with the texture, coloration and pattern elements from source images not included in the training set, using image style transfer (Fig 6).”

[V3.5] Köszönöm a Bíráló észrevételét. Bár értekezésem 22. oldalának idézett mondatát megelőző fél oldala az *image style transfer learning* augmentáció szerepét tárgyalja, úgy gondolom, ez valóban további részletezésre szorul. Az eredeti publikáció (Hollandi et al. 2020) Star*Methods szekciójában részletesen tárgyaljuk a "Training and Style Transfer Data Augmentation" alfejezetben, melynek citálása elmaradt a dolgozatban.

[K3.6] Ezzel kapcsolatban azért felmerült bennem az a kérdés is, ha mechanisztikus transzformációkkal növeljük a tanító halmaz számosságát, lásd 23. oldal: „Conventionally, a transformation (i.e. rotation, flipping, noise addition, etc.) or a series of transformations are applied on the original images.” akkor mire tanul rá gép - mert hogy okos - a transzformációkra, vagy arra, amire szerettük volna? Mennyire lehetünk bizonyosak abban, hogy ezeknek a dúsító módszereknek a segítségével valóban hasonló eredményeket érhetünk el, mintha ténylegesen bevittük volna az igazi, nem mesterségesen kreált adatmennyiséget a rendszerbe?

[V3.6] Ez egy nagyon érdekes kérdés, mert valójában tényleg a transzformációkat akarjuk megtanítani a gépnek, azaz, hogy milyen transzformáció utáni képeket tekintsen azonosnak/hasonlóknak. Például az intenzitás transzformációkkal azt tanítjuk meg, hogy a mikroszkóp különböző megvilágításai esetén invariáns maradjon a modellünk, vagy a forgatásokkal, hogy ugyanaz a sejt tetszőleges irányból identikusnak számítsen. Végül az augmentáció vs. több tanítóadat kérdésre a válasz, hogy egyáltalán nem lehetünk biztosak, sőt legtöbb esetben biztosan nem lesz jobb az augmentált adat, mint ugyanakkora, de eredeti annotált adat. A valós képek sokszínűsége reprezentálja az adat eloszlását a legjobban, így minél több eredeti tanító mintára kell törekedni. Másrészt az annotált adatok előállításuk sok esetben extrém drága, míg az augmentáció gyakorlatilag egy gombnyomás, mely szinte minden tanító eljárásba be van építve.

[K3.7] Mindhárom tézis kifejtésénél hiányoltam a disszertációban a módszertani határok legalább részleges bemutatását, és azoknak az a priori feltételezések rendszerének a definiálását, amelyek segítségével az adott módszertan felépíthető volt. Milyen feltételek mellett, milyen korlátokkal, milyen hibahatárral lehetett megvalósítani az adott feladatot? Bár a rövid tézises disszertáció terjedelme műfaji adottság, talán jobb lett volna kissé kiterjeszteni a kereteit.

[V3.7] Lásd [V3.1].

Szeged, 2023. 07. 27.

.....

Dr. Horváth Péter