

4. félévi beszámoló

Borkovits Bendegúz (borbende@phys-gs.elte.hu)

Statisztikus Fizika, Biológiai Fizika és Kvantumrendszerek Fizikája
PhD program

Témavezetők: Pollner Péter és Kontsek Endre

A dolgozat címe: Képdiaosztika adattudományi alapú fejlesztése

2025.05.21.

1. Bevezetés

Folytattam a Budapesti Műszaki és Gazdaságtudományi Egyetemen (BME), az Alkalmazott Biotechnológia és Élelmiszertudományi Tanszék laborjában közép-infravörös tartományban a spektrumok felvételét. A félév végére egy-két szövetminta kivételével teljesen elkészültem az adatfelvétellel, ez mennyiségre több, mint háromszáz szövetkorongot lemérésér, és így több, mint egymillió spektrum felvételét jelenti. Ugyanakkor, megkezdtuk a tömegspektrometriai mérések vastagság optimalizálását és rendkívül közel kerültünk az végleges adatfelvételhez.

Mindeközben, a sejt-, és hámdetekciós patológiai projekt során ezer nagy felbontású orvosi képen történt detekcióból készítettem statisztikákat és elkezdtem robusztusabb modelleket összeállítani.

2. Az aktuális félévben elvégzett kutatások ismertetése

A közép-infravörös (MIR) spektrumok felhasználásával a kezdeti gépi tanulási analízis folyamán kiderült, hogy sikeresen elkülöníthető az eddig vizsgált három szövet-típus (normál vastagbél, vastagbélrák primer carcinoma, májmetasztázis), habár további, a korongokon belüli strukturális információra van szükségünk ahhoz, hogy klinikai minőségű eredményeket produkáljanak a módszereink. Az adatbázisból, illetve az azt elemző módszerekből elkezdtem egy kéziratot írni, amelyet az év folyamán a Scientific Data folyóiratban tervezünk publikálni.

Összehasonlítottuk a MIR adatokon való predikciók által produkált eredményeket hasonló módon felhasznált tömegspektrumokon végrehajtott predikcióval. A különbség

nem számottevő, azonban ennek oka lehet az MS mérések során beállított metszetvastagság. A MIR spektrumok méréséhez 2 mikrométer vastag szövetmetszeteket használtunk. Az MS mérésekhez meg kell határoznunk az optimális vastagságot. Korábban ezt élelmiszerminták segítségével próbáltuk megvizsgálni, de a közelmúltban áttértünk az emberi szövetekből vett mintákra. Ugyanakkor, a felmerült a kérdés, hogy üveg vagy alumínium gőzölt lemezeken érdemes elvégezni a méréseket. A lemeztípus fontossága a szöveti és háttér spektrumok különbségének nagyságát befolyásolja. A projekt előzetes eredményei bemutatásra fognak kerülni június 12-én egy németországi konferencián, majd remélhetőleg rögtön utána publikálásra kerülnek egy kisebb impaktfaktorú folyóiratban.

A szövetmetszetek tárolására szolgáló lemezek alumínium gőzölési technikájának segítésére elkezdtünk dolgozni egy új kísérleti eszköz összeállításán és tesztelésén, melyet szabadalomra tervezünk előterjeszteni.

Emellett, folytattam a Semmelweis Egyetemhez tartozó sejt-, és hámdetekciós projektet. Több, mint ezer nagyfelbontású Whole Slide Image-en prediktáltam a két neurális háló felhasználásával, majd pedig a detektált sejt hámszövetben levő sűrűségéhez kapcsolódó statisztikát készítettem. Az objektum detekciós modellen még erősíteni kell, mert például egy a keresett sejthez hasonló hámsejtet is a keresett sejtnak prediktált. Kidolgozom egy VisionTransformer architektúra alapú modellt, mely előzetesen sokkal jobb eredményeket produkált az adott a sejdetekció feladatkörében, azonban nagyon nagyszámú fals pozitív jelölést generált. Ennek orvoslása, illetve még másik ezer képen való prediktálás a következő feladat, de így is a közelébe jutottunk az eddigi eredmények publikálásának.

3. Publikációk

Nem publikáltam ebben a félévben, viszont három különböző publikációhoz elegendő tartalmat sikerült elérni a projektek során. Reményeim szerint a következő félév elején be tudom küldeni a kéziratokat.

4. Tanulmányi tevékenység az aktuális félévben

A következő kurzusokat végeztem el ebben a félévben:

- Haladó gépi tanulás labor, (FIZ/3/093)
- A nyolcadik tantárgyamat az oktatási krediteimmal tervezem kiváltani, melyek száma (féléves felbontásban) $4+4+2+2=12$.

5. Konferenciák az aktuális félévben

Nem vettem részt konferencián ebben a félévben, mint előadó, viszont a tömegspektrometriai projektből készítettem konferencia anyagot, amit Kontsek Endre témavezetőm fog előadni június 12-én Németországban.

6. Oktatási tevékenység az aktuális félévben

- Haladó statisztika és modellezés (dsadvstatf20vm, FIZ/3/088, előadás, 6 kredit) előadás megtartása (3-szor 3 óra), tananyag kibővítése (~ 20 óra), ellenőrző zárt-helyi dolgozat felügyelése (2 óra) és javítása (3 óra).
- Tudományos modellezés számítógépes laboratórium (dsscimodf20lm, labor, 7 kredit), projektvezetés, konzultáció, előadások és beszámolók értékelése és javítása (heti 4 óra)

7. Elismerések

Továbbra is részesülök a DKÖP-2023 Doktori Kiválósági Ösztöndíj Program által nyújtott támogatásban.