

ŽILINSKÁ UNIVERZITA V ŽILINE
Fakulta riadenia a informatiky

AUTOREFERÁT
DIZERTAČNEJ PRÁCE

Žilinská univerzita v Žiline
Fakulta riadenia a informatiky

Ing. Dominika Petříková

Autoreferát dizertačnej práce

**Klasifikácia a predikcia histologických farbení pomocou umelých neurónových
sietí**

na získanie akademického titulu „**philosophiae doctor**“ (v skratke **PhD.**)
v študijnom programe doktorandského štúdia
aplikovaná informatika

v študijnom odbore:
informatika

Žilina, máj 2024

Dizertačná práca bola vypracovaná v dennej forme doktorandského štúdia na katedre Softvérových technológií na Fakulte riadenia a informatiky Žilinskej univerzity v Žiline

Predkladateľ: Ing. Dominika Petříková
Katedra softvérových technológií
Fakulta riadenia a informatiky
Žilinská univerzita v Žiline

Školiteľ: prof. Mgr. Ivan Cimrák, Dr.
Katedra softvérových technológií
Fakulta riadenia a informatiky
Žilinská univerzita v Žiline

Oponent: prof. Ing. Vanda Benešová, PhD.
Ústav počítačového inžinierstva a aplikovanej informatiky
Fakulta informatiky a informačných technológií
Slovenská Technická Univerzita Bratislava

Oponent: doc. Ing. Michal Gregor, PhD.
Ústav kompetencieschopnosti a inovácií
Žilinská univerzita v Žiline

Autoreferát bol rozoslaný dňa: 4. júla 2024

Obhajoba dizertačnej práce sa koná dňa 20.8.2024 o 9:00 h. pred komisiou pre obhajobu dizertačnej práce schválenou pracovnou skupinou odborovej komisie v študijnom odbore **informatika v študijnom programe aplikovaná informatika**, vymenovanou dekanom Fakulty riadenia a informatiky Žilinskej univerzity v Žiline.

prof. Ing. Ľudmila Jánošíková, PhD.
predsedníčka pracovnej skupiny odborovej komisie
v študijnom odbore **informatika**
v študijnom programe **aplikovaná informatika**

Fakulta riadenia a informatiky
Žilinská univerzita
Univerzitná 8215/1
010 26 Žilina

Abstrakt

DOMINIKA PETRÍKOVA: Klasifikácia a predikcia histopathologických ofarbení pomocou hlbokých neurónových sietí (dizertačná práca) - Žilinská univerzita v Žiline, Fakulta riadenia a informatiky, Katedra softvérových technológií - Školiteľ: prof. Mgr. Ivan Cimrák, Dr. - Stupeň odbornej kvalifikácie: philosophiae doctor v odbore Aplikovaná informatika - EDIS Žilina, apríl 2024 - 118 strán.

Dizertačná práca sa zaoberá aplikáciou hlbokého učenia v histológii pri klasifikácii nádorov s využitím polo-automaticky anotovaných dát. V prvej časti práce sa zaoberáme samotnými histologickými obrázkami a možnosťami automatickej tvorby anotácií z dvojíc histologických rezov s rôznymi ofarbeniami, kde prvý rez dostane anotáciu na základe vlastností zobrazenej na druhom reze toho istého tkaniva. S využitím optimalizačných metód a metód automatickej analýzy obrazu sa pokúšame tkanivo na skenoch najprv registrovať a následne kvantifikovať proliferčný biomarker Ki67. V druhej časti sa venujeme aplikácii hlbokých neurónových sietí na dataset anotovaný nami navrhnutou metódou za účelom klasifikácie hematoxylin-osinových snímok podľa intenzity Ki67 expresie v tkanive.

Kľúčové slová: neurónové siete, klasifikácia, hematoxylin-eosin, digitálna patológia, Ki67 proteín

1 Úvod

Umelé neurónové siete a hlboké učenie sú v súčasnosti jedným z popredných nástrojov strojového učenia vo viacerých oblastiach vrátane automatického rozpoznávania reči, spracovania jazyka, analýzy obrazu a diagnostiky porúch. V posledných desaťročiach došlo k obrovskému nárastu biomedicínskych dát, ako sú sekvencie genómov, štruktúry proteínov a lekárske snímky, vďaka pokroku v oblasti vysoko výkonných technológií. Preto je prirodzené, že sa uskutočnilo mnoho pokusov o preskúmanie potenciálu hlbokého učenia aj v tejto oblasti. Zjednodušením identifikácie nálezov, ktoré si vyžadujú liečbu, a podporou pracovného postupu experta analýzou obrazu a hlbokým učením sú kľúčovými nástrojmi na zlepšenie diagnostiky.

Napriek mnohým výhodám má hlboké učenie stále niekoľko významných nevýhod. Jednou z vlastností umelých neurónových sietí je, že sa správajú ako modely čiernej skrinky. To znamená, že poznáme len vstupy a výstupy. Zatiaľ čo v mnohých oblastiach a aplikáciách táto vlastnosť nie je prekážkou, v medicíne a biomedicíne si nemôžeme dovoliť robiť rozhodnutia bez jasného dôvodu, a preto je neprijateľné používať a spoliehať sa na rozhodnutie takéhoto modelu. Z dôvodu tejto obrovskej nevýhody hlbokého učenia je oblasť interpretovateľnosti neurónových sietí veľmi aktuálna.

V dizertačnej práci sme sa rozhodli zamerať na počítačom podporovanú diagnostiku rakoviny, konkrétne na klasifikáciu nádorov z histologických snímok. Analýza histologických obrazov sa považuje za zlatý štandard klinickej diagnostiky rakoviny. Histologická analýza sa vykonáva manuálnym skúmaním tenkého sklíčka (rezu) tkaniva pod optickým alebo elektrónovým mikroskopom. Pre diagnostiku rakoviny histopatológovia vizuálne skúmajú zákonitosti tvarov buniek a rozloženia tkaniva, rozhodujú, či sú oblasti tkaniva rakovinové, a určujú stupeň malignity [1].

V súčasnej klinickej praxi zohráva digitálna patológia kľúčovú úlohu a čoraz viac sa stáva technologickým imperatívom vo vedeckých laboratóriách. Nástup zobrazovania celých preparátov spolu s dostupnosťou rýchlejších sietí a nákladovo efektívnejších riešení na ukladanie dát, zjednodušili správu a zdieľanie digitálnych snímok pre klinické účely. Súčasne výrazné pokroky v strojovom učení uľahčili integráciu umelej inteligencie a digitálnej patológie, čím sa odomkli diagnostické možnosti založené na snímkach, ktoré boli predtým obmedzené na oblasti ako rádiológia a kardiológia. Množstvo výskumných prác, ktoré sa v poslednom čase venujú tejto oblasti, naznačujú, že existuje významný potenciál pre prelomové objavy v oblasti umelej inteligencie v doméne patológie [2].

Cieľom tejto dizertačnej práce je implementovať a aplikovať model neurónovej siete na problém klasifikácie so špecializáciou na histologické obrazy pri klasifikácii nádorov so zameraním na použitie rôznych farbení histologických rezov. Medzi čiastkové ciele vedúce k tomuto cieľu patria:

1. Vytvorenie datasetu

Veľké množstvo dobre anotovaných údajov je kľúčové pre aplikáciu hlbokého učenia. Získanie takýchto údajov v oblasti medicíny však zostáva výzvou, pretože si vyžaduje veľké množstvo práce zo strany expertov. Preto sa na začiatku zameriavame na vytvorenie správne anotovaného súboru údajov z dostupných údajov s minimálnou požiadavkou na manuálnu spoluprácu patológov.

2. Trénovanie neurónovej siete

Na takto vytvorenom súbore údajov sa uskutoční trénovanie, doladovanie a skúmanie viacerých architektúr s cieľom dosiahnuť čo najpresnejšie výsledky, ktorými sa dosiahne úroveň presnosti použiteľnú v praxi.

2 Súčasný stav

Tradične sa patologická diagnostika vykonáva tak, že patológ pozoruje zafarbenú vzorku z nádoru na sklenenom sklíčku pomocou mikroskopu, aby diagnostikoval rakovinu. V posledných rokoch sa hlboké učenie rýchlo rozvíja a medzitým sa čoraz viac celých tkanivových preparátov zachytáva digitálne pomocou skenerov a ukladá sa ako obrazy celých preparátov (WSI) [3]. Keďže sa digitalizuje veľké množstvo WSI, je prirodzené, že bolo vykonaných mnoho pokusov o preskúmanie potenciálu hlbokého učenia na analýzy histopatologických snímok.

Väčšina prístupov hlbokého učenia nepoužíva WSI ako vstup, pretože by bola výpočtovo nákladná (vysoká dimenzionalita). Namiesto toho sa extrahujú malé štvorcové políčka a priradujú im anotácie. Existujúce metódy možno rozdeliť do skupín podľa úrovne anotácií, ktoré používajú. Na základe typu anotácií použitých na tréning identifikujeme dve podkategórie: prístup so silnými anotáciami (anotácie na úrovni políčok) a prístup so slabými anotáciami (anotácie na úrovni snímok) [4]. Prvý prístup sa spolieha na identifikáciu oblastí záujmu a podrobnú lokalizáciu nádorov certifikovanými patológmi. Anotácie na úrovni políčok označujeme ako silné, čo znamená, že všetky extrahované políčka majú svoju vlastnú triedu označenia. Pre slabý prístup anotácií na úrovni snímok stačí priradiť konkrétnu triedu WSI. V dizertačnej práci sme sa rozhodli riešiť prístup silnej anotácie.

3 Definovanie problému

Hematoxylín-eozín (HE) sa považuje za zlatý štandard na hodnotenie mnohých nádorov a bežne sa používa na primárne diagnostikovanie na celom svete. Je to ľahko získateľný a cenovo výhodný. Obsahuje však len základné morfológické informácie, ako je tvar buniek a tkaniva zahŕňajúce parametre, ako je mitotická aktivita, invázia do susedných štruktúr alebo určovanie stupňa. Stupeň označuje mieru podobnosti medzi nádorovými bunkami a pôvodným zdravým tkanivom. Typicky dobre diferencované nádory (G1, G2) vykazujú viac priaznivejšiu prognózu, zatiaľ čo zle diferencované nádory (G3, resp. G4) majú tendenciu správať sa agresívnejšie a majú menej priaznivú prognózu [5]. V klinickej praxi sa na získanie molekulárnych informácií používaných pri diagnostike na potvrdenie a určenie podtypu, sa často používa imunohistochemické (IHC) farbenie. Takéto farbenie dokáže vizualizovať expresiu rôznych proteínov (napr. Ki67, estrogénový receptor) na bunkovej membráne alebo v jadre. Okrem toho je často potrebné vykonať niekoľko rôznych farbení IHC na získanie informácie o rôznych charakteristikách tkaniva, takže proces zahŕňa vytvorenie a kontrolu viacerých rôzne zafarbených tkanivových preparátov. Hoci farbenie IHC je v klinickej praxi štandardným postupom, má niekoľko obmedzení. IHC je časovo náročné, drahé a veľmi závislé na protokoloch manipulácie s tkanivom, keďže výstup je vyjadrený ako intenzita farbenia, alebo prítomnosť/neprítomnosť farbiva, alebo percento buniek, ktoré dosiahli určitú intenzitu farbiva. Okrem toho je interpretácia výsledkov vizuálna, a tak závisí od subjektívnej interpretácie patológov [6].

Mnohé nedávne štúdie ukázali, že existuje korelácia medzi HE a IHC farbenými preparátmi z rovnakej oblasti [7, 6, 8]. V takom prípade by malo byť možné modelovať vzťah medzi morfológickými informáciami v HE preparátoch a informáciami IHC a predpovedať expresiu špecifických proteínov priamo z preparátov farbených HE bez dodatočného farbenia IHC [9]. Tento prístup by mohol ušetriť čas aj peniaze alebo len poskytnúť druhý názor pri hodnotení farbenia IHC. Preto sme sa rozhodli použiť neuronové siete na úlohu predpovedania IHC markerov priamo z HE sklíčok. Po diskusii s patológmi sme sa zamerali na farbenie IHC s Ki67 antigénom v našom výskume.

Špecifické typy nádorov, ako sú neuroendokrinné neoplázie (NEN) postihujúce gastroenteropankreatobiliárny a respiračný trakt, si vyžadujú zahrnúť imunohistochemickú analýzu proliferatívnej aktivity nádoru do procesu klasifikácie. Táto analýza zahŕňa použitie IHC protilátky zameranej na proliferatívny faktor Ki67 [5]. Ki67 je proteín, ktorý sa nachádza v jadre nádorových buniek. Tento proteín sa nachádza len v bunkách, ktoré aktívne rastú a delia sa. Rakovinové bunky sa líšia od normálnych, zdravých telesných buniek. Jedna z vecí, ktoré naznačujú, že bunka zmutovala na rakovinovú bunku, je, že bunka neprestáva rásť a deliť sa. Pretože Ki67 sa spája s bunkami, ktoré sú rýchlo rastú a delia sa, niekedy sa považuje za dobrý marker proliferácie (rýchleho nárastu počtu buniek) [10]. Znalosť množstva Ki67-pozitívnych buniek je dôležité pre patológov pri liečebnom procese.

Vďaka spolupráci s Jesseniovou lekárskou fakultou v Martine Univerzity Komenského v Bratislave (JFMED UK) sa nám podarilo získať niekoľko súborov histologických snímok. Tie obsahovali neuroendokrinné nádory pľúc a seminómy, teda zhubné nádory zo zárodočných buniek semenníka.

Získané dáta neobsahujú žiadne dodatočné informácie o Ki67, ktoré by mohli byť použité ako anotácie, okrem samotných dvojíc skenov. Preto sme museli najprv navrhnúť metódu na odhad pomeru Ki67 pozitívnych buniek ku všetkým bunkám na políčkach zo skenov Ki67, ktoré sme potom mohli použiť ako anotácie pre políčka HE. Ak by sme však políčka vytvorili priamo zo skenov v obraze, neboli by sme schopní určiť, ktoré políčko z HE skenu sa zhoduje s políčkom z IHC skenu, a dvojice políčok z rovnakej oblasti obrazu by nezobrazovali rovnaké tkanivo. Z tohto dôvodu bolo potrebné urobiť otočenie a posunutie skenov tak, aby po prekrytí oboch snímok boli takmer identické. Na vytvorenie anotácií políčok je teda najprv potrebné získať zovňaznú dvojicu dvoch celých snímok tkaniva. Jeden s farbením HE a druhý s farbením Ki67. Proces anotácie si vyžadoval dva kroky: registráciu skenov a kvantifikáciu skóre Ki67.

4 HE tkanivo obsahuje informácie o Ki67 expresii – dôkaz koncepcie

Ako počiatočný dôkaz koncepcie pri klasifikácii markerov IHC z obrázkov HE, sme sa rozhodli natréňovať model na identifikáciu Ki67 pozitívnych a negatívnych buniek rovnako ako autori v článku [9]. Autori sa zaoberali problémom určenia počtu Ki67 pozitívnych buniek zo snímok HE pre liečbu viacerých typov rakoviny. Použitím zodpovedajúcich párov IHC a HE farbených obrázkov, modifikovaný ResNet-18 bol na úrovni buniek z HE obrázkov natréňovaný s presnosťou na validačnom súbore 0,9371. Na tréňovanie nášho modelu sme použili údaje uverejnené vo vyššie uvedenej štúdii. Dataset pozostával z políčok zobrazujúcich Ki67 pozitívne bunky, Ki67 negatívne bunky a pozadie. Súbor údajov sme náhodne rozdelili na tréňovaciu, testovaciu a validačnú množinu v pomere 8:1:1, hoci tento postup nie je úplne správny pre tento typ dát. Lepším riešením by bolo náhodné rozdelenie samotných WSI alebo dokonca pacientov, od ktorých sa údaje získali, ak je viac skenov od jednej osoby, a potom priradiť vygenerované políčka do množín na základe WSI, z ktorého pochádzajú. Keďže v súbore údajov boli zverejnené samotné políčka, nie WSI, a tiež tam neboli žiadne informácie o tom, ktoré políčka pochádzajú z rovnakých WSI, museli sme ich rozdeliť náhodne. Treba však poznamenať, že to môže viesť k skresleniu výsledkom, pretože sieť môže rozpoznať niektoré vlastnosti, ktoré sú jedinečné pre jednotlivé WSI a potom dosahovať lepšie výsledky na validačnom súbore. To je nežiaduci jav, pretože v aplikácii s reálnymi údajmi nie sú žiadne údaje o pacientoch súčasťou tréňovacej množiny.

V našom experimente sme použili dve modifikované architektúry NN, ResNet-18 a VGG-16. Modifikácia spočívala v nahradení poslednej plne prepojenej vrstvy novou vrstvou s počtom neurónov rovnajúcim sa trom, aby zodpovedal počtu tried, do ktorých chceme klasifikovať

obrázky. Váhy oboch NN boli inicializované na hodnoty predtrénovaných váh na datasete ImageNet, čo urýchlilo tréningovanie. Použili sme Adam optimalizátor so stratovou funkciou krížovej entropie a počet epôch bol 40. Koeficient učenia bol nastavený na $1e-3$ pre prvých 10 epôch a v ďalších epochách sme ju znížili na $1e-4$. Ako regularizáciu, aby sme zabránili pretrénovaniu modelu, sme použili penalizáciu L2 pridanú k stratovej funkcii s parametrom $1e-5$. Tento parameter je hodnota, ktorou sa násobí člen L2 pridaný k stratovej funkcii. Použili sme aj metódu skorého zastavenia, ktorá ukončí tréningovanie, ak chyba na validačnej množine začne rásť niekoľko epôch za sebou.

Architektúra ResNet dosiahla o niečo vyššiu presnosť ako VGG. Dokonca dosiahla vyššiu presnosť ako ten istý model, ktorý bol natrénovaný v článku [9] a dosiahol presnosť 0,95. Môže to byť spôsobené dlhším tréningom (viac epôch) a použitím metódy skorého zastavenia, ktorá zachováva model s najnižšou chybou na validačných údajoch.

5 Registrácia tkaniva

Predtým, ako sme mohli začať anotovať HE políčka s expresiou Ki67, bolo potrebné zabezpečiť, aby dvojice políčok z rovnakej oblasti na HE a IHC snímkach zodpovedali tomu istému tkanivu alebo tkanivu z tej istej oblasti sklíčka. Tkanivo na obrázkoch bolo rôzne otočené a posunuté, preto sme najprv hľadali spôsob, ako približne zarovnať obidva skeny. Treba zdôrazniť, že dvojice tkanív nie sú identické a preto nebolo možné vykonať porovnanie na úrovni buniek na žiadnom z uvedených súborov údajov. Dokonca aj v prípade, keď bolo zaručené, že HE a IHC rezy nasledovali bezprostredne po sebe a tkanivá sú susediace, neboli v oboch rezoch nájdené rovnaké bunky. Preto bolo potrebné vykonať porovnanie len na základe určitej podobnosti oblastí tkaniva napr. tvaru. Zosúladenie snímok pozostávalo z dvoch krokov, najprv z rotácie a potom posunutia, preto sme museli určiť tri transformačné parametre. Na základe informácií od lekárov by mal podobnú funkcionality poskytovať softvér SlideViewer v komerčnej verzii. Toto zarovnanie je však k dispozícii len na vizuálne preskúmanie v reálnom čase a nie je možné extrahovať zarovnané snímky alebo získať transformačné parametre.

V počiatočnom úsilí sa proces registrácie vykonával manuálne pod vedením skúsených lekárov. Táto metóda zahŕňala prekryvanie dvoch mierne priehľadných obrázkov vo vhodnom grafickom nástroji a ich zarovnanie na rovnaké súradnice. Následne bol jeden z nich podrobený rotačným a translačným úpravám, až kým nedošlo k približnému vizuálnemu zarovnaniu oboch tkanív. Tento postup bol časovo náročný, pretože okrem samotného zarovnania vyžadoval aj prepočítanie parametrov posunu z cm, v ktorých boli dané, na pixely. Taktiež výstupné parametre museli byť zaznamenané ručne bez možnosti automatického exportu. Aby bol tento prístup automatickejší, použili sme algoritmus SIFT na detekciu a priradovanie kľúčových bodov, ktoré možno opísať ako zaujímavé hlavné body alebo vlastnosti obrazu. Kľúčové body možno použiť na spoľahlivé porovnávanie medzi rôznymi pohľadmi na objekt alebo scénu. Na aplikovanie algoritmu SIFT sme použili implementáciu dostupnú v OpenCV. Všetky obrázky boli pred vstupom do algoritmu transformované do odtieňov sivej a výrazne zmenšené. Použité zmenšenie bolo približne 50-násobkom pôvodnej veľkosti, takže výsledné obrázky mali niekoľko pixelov. Najprv sme použili SIFT na obrázky s predspracovaním, ktoré zahŕňalo len zmenu veľkosti a škálovanie farieb na stupne sivej, čo malo minimálny úspech. Týmto spôsobom sa nám podarilo zaregistrovať iba jednu dvojicu obrázkov. Jednotlivé rezy sú napriek všeobecnej tvarovej podobnosti príliš odlišné. Na zníženie tejto rozdielnosti sme použili ďalšie metódy predspracovania. Najprv sa použilo binárne prahovanie, ktorého výsledkom boli čiernobiele obrázky. Samotné prahovanie neprinieslo žiadny úspech, preto sme obrázky rozšírili aj metódami binárnej morfológie, konkrétne eróziou a dilatáciou, ktoré sa použili v niekoľkých iteráciách s rôznymi veľkosťami jadier. Opäť to však nemalo požadovaný účinok, len pre jednu dvojicu sa nám podarilo nájsť

dobre zodpovedajúce kľúčové body. Zhoda kľúčových bodov sa zdala byť správna až po aplikácii rozmazania na oba obrázky. Celkovo sa podarilo urobiť zarovnanie v prípade 4 dvojíc obrázkov z 9. Teda sme aplikovanie SIFT algoritmu považovali za nedostatočné a rozhodli sme sa nepokračovať vo výskume tejto metódy.

Plne automatické vyhľadávanie kľúčových bodov a ich následné porovnávanie nebolo úspešné, preto sme sa rozhodli zvoliť iný prístup a zjednodušiť proces tým, že ho urobíme len poloautomatickým. Namiesto použitia automatického vyhľadávania kľúčových bodov sme ručne definovali dvojice kľúčových bodov, ktoré patria k sebe v každom HE a IHC skene, a potom sme ich len pomocou optimalizačnej metódy hľadali najlepšie transformačné parametre medzi nimi. Tento proces nebol taký pracný ako počiatočná manuálna registrácia, keďže sme mohli pracovať priamo v programe SlideViewer, ktorý umožňuje vytvárať anotácie a exportovať ich prostredníctvom programu SlideMaster do súborov xml. SlideViewer umožňuje aj otvorenie viacerých skenov súčasne vedľa seba, takže sme mohli mať naraz otvorené dvojice skenov a vytvárať anotácie kľúčových bodov v oboch. Pre každý snímok sme vytvorili 5 štvorcových anotácií kľúčových bodov. Okrem nich sme tiež vytvorili veľkú ohraničujúcu anotáciu celého tkaniva, na základe ktorej bol vytvorený výrez obrazu zo skenu. V tejto fáze boli kľúčové body definované ako štvorce, nie ako body, takže v ďalšom kroku sme vzali len stredu týchto štvorcov a prepočítali ich súradnice vzhľadom na anotáciu veľkého ohraničujúceho poľa, čím sme získali súradnice, ktoré slúžili ako vstup pre optimalizačný algoritmus. Pre optimalizačný problém sme použili algoritmus BFGS, patriaci do triedy kvázi-Newtonových metód, ktorý sa považuje za najpopulárnejšiu optimalizačnú metódu zo spomínanej triedy.

Po získaní transformačných parametrov sme najprv predspracovali oba obrázky do formátu png na základe ohraničujúcich anotácií. Potom sme porovnali ich rozmery a pridali biele pixely v oboch smeroch tak, aby mali oba obrázky rovnaké rozmery, pričom sme sa uistili, že ľavý horný roh pôvodného obrázka zostal na rovnakej pozícii v novom obrázku, a teda súradnice kľúčových bodov zostali nezmenené. Kvôli jednoduchšej manipulácii sme vykonali registračnú transformáciu s oboma obrázkami, pričom sme otočili obrázok IHC a posunuli obrázok HE. Celý postup by sme mohli zhrnúť do nasledujúcich krokov:

1. otočte obrázok IHC okolo jeho stredu s možnosťou rozšírenia, aby bol výsledný obraz dostatočne veľký pre celý otočený obraz IHC
2. vytvorte ďalší biely obrázok rovnakej veľkosti ako otočený obrázok IHC
3. vypočítajte vektor translácie obrazu HE vzhľadom na biely obraz tak, aby po umiestnení do bieleho obrázka bol v strede.
4. odčítajte parametre posunu získané optimalizáciou od translačného vektora z kroku 3
5. skopírujte každý pixel obrazu HE na súradnice posunuté o vektor z kroku 4 do bieleho obrazu.

Z 83 párov v datasete s nádormi semenníkov, ktoré nám boli poskytnuté, sa nám podarilo úspešne zaregistrovať 77 párov, čo predstavuje výrazné zlepšenie oproti predchádzajúcemu pokusu registrácie pomocou algoritmu SIFT. Zvyšných 6 párov bolo možné zaregistrovať ručne, ale vzhľadom na dostatočné množstvo údajov sme sa rozhodli ich vynechať.

6 Ki67 kvantifikácia

Zo zarovnaných a zaregistrovaných párov HE a IHC snímok bolo možné vytvoriť dvojice políčok. Odhad indexu Ki67 IHC políčok znamená vyhodnotiť pomer hnedých buniek ku všetkým bunkám. Pri našom počiatočnom skúmaní sme zvažovali možnosť prefarbenia obrazov IHC pomocou analýzy farieb pomocou histogramov a prahovania. Tento prístup bol zameraný na identifikáciu prevládajúcich farieb v obraze, pričom výstupom boli prahové hodnoty na prefarbenie všetkých pixelov. Avšak po skúmaní histogramov pre každý farebný kanál sme narazili na ťažkosti pri identifikovaní akýchkoľvek dominantných farieb okrem

bielej. V dôsledku toho sme zamietli túto metódu a namiesto toho sme zvážili metódu zhlukovania. Na vyhodnotenie pomeru Ki67 pozitívnych buniek je potrebné celý obraz rozdeliť na tri hlavné oblasti, ktoré predstavujú tieto tri hlavné farby:

- biela - pozadie
- modrá - Ki67 negatívne
- hnedá - Ki67 pozitívne.

Spomedzi rôznych dostupných algoritmov zhlukovania sme použili jeden z najznámejších, algoritmus zhlukovania K-means. Obmedzením, ktoré sa často uvádza pri algoritme K-means, je jeho nevyhnutná podmienka vopred definovať počet zhlukov. Vzhľadom na vysokú výpočtovú zložitosť zhlukovania na obrazoch generovaných z WSI sme sa pokúsili odhadnúť parameter k na základe experimentov vykonaných na súbore dát IHC obrazov s nižším rozlíšením pomocou takzvanej metódy lakt'a. Okrem nájdenia najlepšieho počtu zhlukov bolo potrebné zvoliť aj správny farebný priestor, nad ktorým sa má merať vzdialenosť. Preto sme spustili experimenty nad tromi priestormi: RGB, HSV (Hue, Saturation, Value) a Lab (Svetlosť, a - farebná zložka v rozsahu od zelenej po magenta, b - farebná zložka v rozsahu od modrej po žltú). Výsledky ukázali, že najvýraznejšie rozdiely vychádzali v priestore HSV pre počet zhlukov 6.

Zatiaľ čo snímky s NEN obsahovali veľké množstvo Ki67 pozitívnych buniek, obrázky so seminómami obsahovali prevažne Ki67 negatívne bunky, čo malo za následok, že sa v 6 výsledných centroidoch nenachádzali žiadne hnedé farby reprezentujúce pozitívne bunky hoci boli prítomné v tkanive. Preto sme zvýšili počet zhlukov na 12, pri ktorom sme boli schopní zachytiť aj hnedé bunky.

Výsledné centroidy zhlukov boli transformované z HSV do RGB priestoru a zobrazené v koláčovom grafe. Následne bolo potrebné vykonať vizuálnu analýzu a určiť, ktorý centroid bude patriť do ktorej kategórie (pozadie, Ki67 pozitívne, Ki67 negatívne). Nakoniec sa každý pixel podrobil prefarbeniu na základe jeho príslušnosti k určitému zhlukovému centroidu, pričom sa mu priradil jeden z troch vopred určených farieb.

V praxi sa výraz ki67 určuje ako pomer skutočného počtu pozitívnych buniek ku všetkým bunkám. Keďže by si to vyžadovalo nutnosť segmentovať jednotlivé bunky, rozhodli sme sa odhadnúť tento pomer podľa veľkosti plochy zaberanej hnedými pixelmi v pomere k celému povrchu buniek. Teda skóre Ki67 sa kvantifikuje ako počet pixelov zafarbených farbou zodpovedajúcim pozitívnym bunkám (hnedá), ku všetkým pixelom obsadeným akoukoľvek bunkou (hnedá a modrá). Aby sme sa vyhli generovaniu veľkého počtu prázdnych políčok, nastavili sme určité podmienky, na základe ktorých sme vybrali políčka do datasetu pre klasifikáciu. Tieto podmienky sa týkali pôvodných políčok HE a prefarbených políčok IHC, na základe ktorých bolo pridelené skóre. V prípade HE políčok sme hodnotili pomer pixelov, ktoré mali vo všetkých troch kanáloch RGB hodnotu na úrovni aspoň 230, čo zodpovedá bledým farbám pozadia. Keďže prefarbené IHC obrázky obsahovali len 3 farby, hodnotili sme priamo pomer bielych pixelov. Pri generovaní políčok pre klasifikačný dataset, sme nechávali len tie, ktoré obsahovali menej ako 40 % bledých pixelov v HE a menej ako 60 % bielych pixelov vo variante IHC. Pôvodne sme mali prahovú hodnotu nastavenú rovnako pre oba typy snímok, ale v prefarbených políčkach veľa bielych pixelov zaberajú aj medzibunkové priestory a to by odfiltrovalo zbytočne veľa vhodných políčok. Všetky ostatné políčka, ktoré nespĺňali tieto dve podmienky, boli vyradené.

Keďže sme nemali k dispozícii žiadne empirické anotácie pre poskytnuté skeny, prostredníctvom ktorých sme mohli overiť, či je táto metóda kvantifikácie správna, použili sme iné IHC snímky, pre ktoré sme mali presné manuálne kvantifikované anotácie. Tieto anotácie sme mohli porovnať s nami odhadnutými pomermi. Celkovo sme vyhodnotili pomery na 44 dostupných obrázkoch, čím sme získali priemernú chybu 5,22 %. Táto hodnota je považovaná za všeobecne prijateľnú vzhľadom na veľký rozsah intervalov, do ktorých sa

Ki67 klasifikuje. Hoci výsledok tejto validácie považujeme za prijateľný, je potrebné poukázať na niekoľko bodov:

- táto validácia je len hrubým odhadom skutočnej presnosti, vzhľadom na to, že použité dáta boli získané pomocou odlišných protokolov a farebne sa výrazne líšia od tých, ktoré chceme použiť na klasifikáciu,
- výsledný pomer je silne závislý od následného manuálneho zadelenia získaných centroidov do troch kategórií.

Lekársky softvér SlideViewer, ktorý patológovia používajú v praxi, poskytuje v komerčnej verzii aj mnoho ďalších metód analýzy tkanív, vrátane kvantifikácie IHC pozitivity tkaniva zobrazeného na obrazovke. To nám umožnilo získať určité množstvo relatívne dobre anotovaných údajov, ktoré možno použiť na overenie nami navrhovanej metódy kvantifikácie Ki67. Po vyladení parametrov potrebných na správny pomer kvantifikáciu odborníkmi, možno túto metódu použiť na skutočné tkanivo zobrazené na obrazovke a výsledky možno exportovať do programu Excel. Preto, sme požiadali patológov, aby anotovali ohraničujúce polia na celej obrazovke tkaniva a potom nad nimi spustili metódu kvantifikácie s vyexportovanými výsledkami. Vďaka zobrazeným anotáciám sme mohli urobiť výrez obrazu v tom istom reze a spustiť zhľukovanie nad tým istým rezom tkaniva, ktorého výsledok po odhadnutí pomeru Ki67 možno porovnať s výsledkom programu SlideViewer.

Získali sme 16 anotácií pochádzajúcich zo 7 WSI. Po extrakcii príslušných anotácií z našich prefarbených WSI sme porovnali výsledky a na základe rozdielu môžeme povedať, že sme výrazne podhodnotili hodnoty pozitivity s priemernou chybou 21,94, čo je veľmi veľká chyba vzhľadom na šírku intervalov, do ktorých chceme tkanivo klasifikovať. Rozptyl chyby bol tiež veľký a pohyboval sa od 1 do takmer 50%.

Na zníženie vysokej variability sme zmenili prístup k zhľukovaniu. Až doteraz sa zhľukovanie vykonávalo na každom snímku samostatne, takže chyba odhadu na každom snímku sa líšila. Pomocou dostupných anotovaných snímok sme sa zamerali na získanie presne vyladených centroidov, ktoré by boli spoločné pre všetky snímky. Doladené v zmysle udržania čo najnižšej chyby s nízkym rozptylom. V tomto prípade sa chyba vybalansovala, niektoré snímky sme podhodnotili a iné nadhodnotili a po bližšom skúmaní samotných snímok a odhadov lekárov sme začali mať pocit, že niekde musí byť koncepčná chyba, alebo že nám chýbajú nejaké informácie. Z našej diskusie sme zistili, že bunky nie sú klasifikované len na Ki67 pozitívne a Ki67 negatívne, ale v prípade Ki67 pozitívnych buniek sa rozlišujú 3 stupne intenzity pozitivity. Bunky, ktoré sú modré s menšou plochou hnedej farby, sa tiež považujú za slabo alebo stredne pozitívne. Na základe tohto zistenia je zrejmé, prečo náš anotačný systém nedokázal dosiahnuť nižšiu chybovosť. V tomto prípade je nahradenie buniek pixelmi nesprávne, pretože aj slabo pozitívna bunka, ktorá obsahuje veľa modrých pixelov a málo hnedých pixelov výrazne prispieje k zvýšeniu počtu Ki67 pozitivity. Je preto jasné, že pre budúcu kvantifikáciu bude potrebné navrhnúť iný prístup, pričom sa zohľadní, či je hnedá farba prítomná v bunke alebo nie.

7 Klasifikačné modely

V kapitole 4 sme ukázali, že sken tkaniva HE obsahuje aj molekulárne informácie o expresii Ki67. V tomto experimente však modely boli natréňované na rozlišovanie negatívnych a pozitívnych buniek z homogénnych oblastí, teda môžeme povedať, že model bol natréňovaný na klasifikáciu tkanív s pozitivitou 0 alebo 100. Naším cieľom je natréňovať model na klasifikáciu tkanív do rôznych intervalov pomeru pozitivity tkanív. To tiež znamená, že políčka zachytávajú heterogénne oblasti tkaniva s pozitívnymi aj negatívnymi bunkami. Schopnosť klasifikovať takéto obrazy pomocou neurónových sietí bola len našou hypotézou, preto sme v procese tréňovania modelov vyhodnocovali aj situácie, v ktorých sme z úplného rozsahu vynechali niektoré intervaly pomeru. Hoci takéto situácie nezodpovedajú reálnemu

scenáru pri používaní modelu, je veľmi dôležitá pre overenie hypotézy. Chyba anotácie môže spôsobiť, že snímky s rovnakým alebo podobným pomerom pozitivít budú v rôznych triedach, čím sa zhorší presnosť modelu. V takomto prípade nie sme schopní určiť, či je nízka presnosť spôsobená neschopnosťou modelu naučiť sa dané vlastnosti tkaniva alebo chybou anotácie. Vynechanie dát zo stredného intervalu nemá za následok miešanie údajov, a teda eliminuje tento problém.

Všetky experimenty s neurónovými sieťami boli vykonané na architektúre typu ResNet. Tieto architektúry napriek tomu, že sú staršie, dosahujú veľmi dobré výsledky na obrazových dátach a boli použité v mnohých histologických úlohách s dobrou presnosťou pri klasifikácii obrazu. Navyše v tejto časti nášho výskumu sme kladli väčší dôraz na automatické a poloautomatické generovanie dobre anotovaných dát, čo je nevyhnutné pre použitie metód hlbokého učenia.

Ako počiatočný dôkaz koncepcie na overenie presnosti a funkčnosti navrhovaného prístupu generovania automaticky anotovaných dát na účely klasifikácie sme extrahovali políčka z troch WSI NEN na trénovanie modelu neurónovej siete. Vygenerované políčka mali veľkosť 224x224 pixelov a boli rozdelené do dvoch kategórií na základe percentuálneho podielu Ki67-pozitívnych buniek. Prvú kategóriu tvorili políčka s menej ako 30 %, zatiaľ čo druhú zahŕňala políčka s viac ako 60 %. V tomto predbežnom experimente sme zámerne vylúčili hodnoty v strednom rozsahu, aby sme zjednodušili klasifikáciu. Náš dataset pozostával zo 1714 políčok. Vzhľadom na malý počet skenov použitých v datasete sme na vytvorenie testovacej sady nevyberali WSI, ktoré by sa mali oddeliť od ostatných a zahrnúť do testovacej sady. Namiesto toho sa aj do testovacej sady vyberali políčka zo všetkých skenov, rovnako ako v prípade trénovacej a validačnej sady.

Model dosahoval hodnoty presnosti 0,92 pre tréningovú, 0,94 pre validačnú a 0,82 pre testovaciu sadu. Podobný experiment sme vykonali aj na snímkach seminómov. Podobne ako v prípade dát NEN sme chceli vyskúšať binárnu klasifikáciu na najprv hraničné intervaly, takže sme dataset rozdelili do kategórií pod 10% a nad 60% hneď na začiatku. Okrem binárnej klasifikácie hraničných intervalov sme chceli otestovať aj to, ako dobre model zvláda binárnu klasifikáciu tried, ktoré od seba nie sú až tak vzdialené. Preto sme natrénovali binárny klasifikátor na rozdelenie políčok na triedy medzi 20 a 50% a nad 60%. Keďže oba uvedené modely dosiahli vysokú presnosť, rozšírili sme model na viactriednu klasifikáciu pridaním aj dolného intervalu pod 10%, čo nás priblížilo k želanej klasifikácii použiteľnej v praxi, ale stále sme zámerne vynechali stredné intervaly medzi susednými triedami. Hoci sa presnosť modelu znížila, bola stále dostatočne vysoká.

Pomocou postupu uvedeného v predchádzajúcich častiach sme vygenerovali celkovo takmer 340 000 anotovaných políčok s veľkosťou 224x224 zo 75 WSI. Aby sa validácia modelu čo najviac podobala praktickému používaniu modelu sme sa rozhodli rozdeliť súbor údajov na trénovaciu množinu a dve validačné množiny (VS). Prvá validačná množina bola vytvorená rozdelením samotných WSI, pričom sme vybrali 5, z ktorých sa v trénovacej množine nepoužijú žiadne údaje, a teda zodpovedajú reálnemu používaniu modelu v praxi, kde model musí predpovedať hodnoty pre nového pacienta. Druhý validačný set pozostával z 10% získaných políčok z trénovacej množiny na sledovanie priebehu tréningu a overenie modelu na dátach, ktoré neboli videné, ale pochádzajú zo známeho tkaniva modelu. Dôvod, prečo ich označujeme ako validačné súbory 1 a 2 namiesto validačnej a testovacej sady je ten, že vo všeobecnosti sa testovacia sada nepoužíva v procese doladovania hyperparametrov modelu. Testovacia množina sa používa len na konečné vyhodnotenie presnosti modelu, preto upúšťame od označovania prvého validačného súboru ako testovacieho.

Rozhodli sme sa používať triedy na základe odporúčaní patológov a preto sme súbor údajov rozdelili do troch kategórií: pod 20%, medzi 20% a 50% a nad 50%. Vzhľadom na značný objem dát sme sa rozhodli pre balansovanie datasetu prostredníctvom techniky

podvzorkovania, pri ktorej sa triedy s vyšším počtom údajov zredukovali. Najvyššia dosiahnutá presnosť na druhom validačnom sete bola 0,79, čo znamená pomerne mierny výkon. Analýza konfúznej matice pre najlepší model na druhom validačnom súbore ukázala, že model má najväčšie ťažkosti pri rozlišovaní obrázkov z kategórie pod 20 %. Vykazuje však zdatné rozlišovanie medzi ostatnými dvoma kategóriami. Na potvrdenie tohto pozorovania sme uskutočnili experiment, v ktorom sme natrénovali binárny model na rozlišovanie políčok z kategórií 20 - 50 % a nad 60 %. V tomto experimente dosiahol model presnosť 0,8595 na druhom VS. Následne po začlenení predtým vynechaného intervalu do nového modelu sa presnosť len nepatrne znížila na 0,8484 pri druhom VS a 0,5266 na prvom VS. Napriek tomu, že neurónové siete dosahujú celkom vysokú presnosť pri binárnej klasifikácii, majú potenciál na lepší výkon. Veľmi dôležité je aj to, že presnosť modelov na prvej validačnej množine bola veľmi nízka na úrovni nepoužiteľnosti pre všetky experimenty. To naznačuje, že sieť sa dokáže naučiť odhadovať pomer Ki67 pozitívnych buniek, ale len na tkanive, s ktorým má skúsenosti. Tento jav môže mať dve príčiny, ktoré bolo potrebné ďalej preskúmať.

Pri podrobnejšom skúmaní dát sme si všimli niekoľko nedostatkov. Dataset obsahoval aj políčka s čiernymi artefaktmi, rozmazané dáta, obrázky bez buniek alebo s iným typom tkaniva než nádorové. Pokúsili sme sa odstrániť všetky uvedené nedostatky z datasetu vytvorením oblastí záujmu (ROI), z ktorých sa budú vytvárať políčka na tréning siete. Po vygenerovaní políčok z ROI sa náš súbor údajov zmenšil a tiež sa zmenilo rozloženie údajov v jednotlivých triedach. Preto sme v prvom kroku vykonali binárnu klasifikáciu s kategóriami pod 20% a medzi 20% - 50%, okrem toho sme prvú triedu rozšírili a urobili sme rozdelenie pod 30% a medzi 30% - 50%. Presnosť modelu bola o 4 % vyššia, ako keď sme použili všetky údaje. Pri použití datasetu vytvoreného z dát s anotáciami na základe spoločných centroidov popisovaných v kapitole 6 sa presnosť modelu na druhej VS znížila a mierne vzrástla na prvej VS, avšak nárast nebol dostatočný, aby mohol byť model používaný v praxi.

8 Záver

V tejto práci sme sa zaoberali problémom klasifikácie histologického farbenia pomocou hlbokých neurónových sietí. Cieľom tejto práce bolo natrénovať klasifikačnú neurónovú sieť, ktorá by dokázala priamo predpovedať expresiu Ki67 zo snímok HE bez potreby dodatočného farbenia IHC. Je všeobecne známe, že hlboké neurónové siete vyžadujú veľké množstvo kvalitne anotovaných údajov, čo si vyžaduje časovo náročnú prácu expertov. V našej štúdii sme pracovali len s dvojicami zhodných skenov HE a Ki67 farbeného tkaniva bez akýchkoľvek ďalších informácií, ktoré by sa dali použiť ako anotácie, keďže sme mali k dispozícii len tento typ údajov. Z tohto dôvodu bolo nevyhnutné navrhnúť metódu na poloautomatické generovanie anotácie výlučne z dvojíc skenov HE a farbenia IHC.

V tejto práci sme sa snažili potvrdiť, že existuje korelácia medzi morfológickými a molekulárnymi informáciami zobrazenými v HE a Ki67 farbených tkanivách, a preto môžeme tento vzťah modelovať pomocou NN. Hoci sa táto korelácia preukázala v predchádzajúcich štúdiách, pracovali len s plne Ki67 pozitívnymi a negatívnymi vzorkami tkaniva. V našom výskume sme sa snažili otestovať hypotézu, že hlboké učenie dokáže klasifikovať aj tkanivové rezy, ktoré obsahujú heterogénne bunky, a teda majú rôznu expresiu Ki67. Okrem toho sme sa snažili dosiahnuť opísané ciele s použitím minimálnych až takmer žiadnych anotácií od lekárov pomocou navrhovanej vlastnej poloautomatickej metódy generovania anotácií z dvojíc HE a IHC skenov. Hoci naše modely nedosahujú použiteľný výkon na nových WSI, presnosť dosiahnutá na políčkach z tréningových WSI, ktoré sa nepoužili v tréningovom procese, dokazuje schopnosť NN rozpoznávať molekulárne vlastnosti tkaniva priamo zo snímok HE. Ak to zhrnieme, hlavné príspevky práce v oblasti aplikovanej informatiky a digitálnej patológie sú:

- návrh poloa automatizovanej metódy na registráciu a kvantifikáciu expresných údajov IHC pomocou metód analýzy obrazu, teda zjednodušenie procesu anotácie údajov expertmi
- potvrdzuje hypotézu, že NN modely dokážu predpovedať z HE molekulárne vlastnosti tkaniva bežne vyjadrené cez IHC farbenie, konkrétne index Ki67
- ukázali sme, že aj heterogénne tkanivo, ktoré neobsahuje čisto Ki67 pozitívne a negatívne bunky, možno klasifikovať pomocou neurónových sietí do tried pomernej pozitivita nazývaných stupne.

V posledných rokoch došlo k veľkému zlepšeniu v oblasti digitálnej patológie a s vynálezom skenerov, ktoré sú schopné zachytiť tkanivové rezy digitálne ako WSI, je prirodzené, že sa objavujú snahy o aplikáciu hlbokého učenia aj v tejto oblasti, keďže majú vysoký potenciál na obrazových dátach a boli úspešne aplikované v mnohých doménach na riešenie rôznych úloh. Napriek rastúcemu množstvu dostupných lekárskeho dát je množstvo dobre anotovaných dát stále nízke a ich získavanie je veľmi náročné. Použitím metód analýzy obrazu možno určiť expresiu proteínu Ki67 z IHC skenov, ktoré môžu slúžiť ako anotácia pre zodpovedajúce HE políčka z rovnakej oblasti tkaniva. Nami navrhovaná metóda kvantifikácie založenej na zhlukovaní a počítaní pixelov je použiteľná len na bunky, ktoré sú silne pozitívne a silne negatívne, t. j. čisto modré a čisto hnedé. V praxi sa však rozlišujú 4 kategórie buniek: negatívne (modré), slabo pozitívne, stredne pozitívne a silne pozitívne (hnedé). Slabo a stredne pozitívne bunky stále obsahujú veľké množstvo modrej farby, ale obsahujú aj stopy proteínu Ki67, t. j. obsahujú hnedé alebo hnedosivé škvrny. To sú informácie, ktoré sme na začiatku nášho výskumu nemali, a preto naša navrhovaná metóda kvantifikácie neposkytuje porovnateľné výsledky s lekárskeho softvérom. V budúcnosti bude potrebné upraviť náš postup tak, aby zohľadňoval túto skutočnosť a poskytoval tak spoľahlivejšie odhady pomeru Ki67. To by malo viesť k zníženiu chyby odhadu, ako aj zníženie rozptylu chýb na rôznych skenoch. S tým priamo súvisí ďalší krok výskumu, ktorým je zvýšenie presnosti modelov na nových WSI tak, aby modely boli použiteľné v reálnych situáciách a boli schopné klasifikovať aj tkanivo od nových pacientov. Predpokladom zvýšenia presnosti klasifikácie je však mať kvalitne anotované dáta. Iba v takom prípade môže byť následným krokom skúšať jemné doladenie modelov alebo iné architektúry, vrátane novších.

Hodnotenie proliferatívnej aktivity musí byť vykonávané len na oblastiach, ktoré obsahujú nádorové bunky. Imunitné bunky, ktoré sa v nádore môžu nachádzať nemajú byť súčasťou týchto odhadov. To znamená, že naša metóda bude v budúcnosti musieť vedieť rozlišovať medzi bunkami rôznych veľkostí a počítať len tie veľké (nádorové).

Všetky ciele uvedené v tejto časti spolu s problémami, ktoré sa budú riešiť neskôr, poukazujú aj na budúcu perspektívu tejto témy na potenciálne sľubné výsledky.

Zoznam použitej literatúry

- [1] L. He, L. R. Long, S. Antani, and G. R. Thoma, "Histology image analysis for carcinoma detection and grading," *Computer Methods and Programs in Biomedicine*, vol. 107, pp. 538–556, 2012.
- [2] M. K. K. Niazi, A. V. Parwani, and M. N. Gurcan, "Digital pathology and artificial intelligence," *Lancet Oncol.*, vol. 20, pp. e253–e261, May 2019
- [3] L. X. Pantanowitz, "Digital images and the future of digital pathology," *Journal of Pathology Informatics*, vol. 1, 2010.
- [4] N. Dimitriou, O. Arandjelović, and P. D. Caie, "Deep learning for whole slide image analysis: An overview," *Frontiers in Medicine*, vol. 6, 2019.
- [5] D. Petriková, I. Cimrák, K. Tobiášová, and L. Plank, "Ki67 expression classification from the images with semi-automated computer-generated annotations," in *BIOINFORMATICS*, 2024.
- [6] N. Naik, A. Madani, A. Esteva, N. S. Keskar, M. F. Press, D. Ruderman, D. B. Agus, and R. Socher, "Deep learning-enabled breast cancer hormonal receptor status determination from base-level h&e stains," *Nature Communications*, vol. 11, p. 5727, 2020.
- [7] Alexander, S. René, B. Michael, A. Maximilian, J. Philipp, K. Frederick, M. K.-R. S. Philipp, and Binder, *Interpretable Deep Neural Network to Predict Estrogen Receptor Status from Haematoxylin-Eosin Images*, pp. 16–37. Springer International Publishing, 2020
- [8] R. R. Rawat, I. Ortega, P. Roy, F. Sha, D. Shibata, D. Ruderman, and D. B. Agus, "Deep learned tissue "fingerprints" classify breast cancers by er/pr/her2 status from h&e images," *Scientific Reports*, vol. 10, p. 7275, 2020.
- [9] Y. Liu, X. Li, A. Zheng, X. Zhu, S. Liu, M. Hu, Q. Luo, H. Liao, M. Liu, Y. He, and Y. Chen, "Predict ki-67 positive cells in h&e-stained images using deep learning independently from ihc-stained images," *Frontiers in Molecular Biosciences*, vol. 7, 2020
- [10] Z. Kos and D. J. Dabbs, "Biomarker assessment and molecular testing for prognostication in breast cancer," *Histopathology*, vol. 68, pp. 70–85, 2016.