

Mitotikus sejtek azonosítása humán daganatos szövetekben mélytanuló algoritmusok használatával

Biológia TDK dolgozat, 2022. tavaszi forduló

Készítette:

Iván Zsanett Zsófia

Biológus MSc

Témavezetők:

Dr. Horváth Péter

ELKH-SZBK

Csapóné Dr. Miczán Vivien

ELKH-SZBK

Belső konzulens:

Dr. Huliák Ildikó



2022

Tartalom

Tartalom	2
1. Bevezetés	3
1.1 A mitózis szerepe a tumoros folyamatokban	3
1.2 Különböző mitotikus morfológiák bemutatása szövetmintákban	3
1.2.1 Profázis	4
1.2.2 Prometáfázis	5
1.2.3 Metafázis	5
1.2.4 Anafázis	5
1.2.5 Telofázis	6
1.2.6 Atípusos mitózis	6
1.3 Mitotikushoz hasonló morfológiák	7
1.4 Mélytanuló algoritmusok a diagnosztikában	8
1.5 Mitózis detekciós módszerek	9
1.5.1 Hagyományos gépi tanulási technikák	10
1.5.2 Mélytanuláson alapuló technikák	10
1.5.3 Kombinált technikák	10
1.6 A dolgozatban felhasznált módszerek elméleti áttekintése	10
1.6.1 A BIAS képfeldolgozó program	10
1.6.2 A MIDOG 21 kihívás adathalmaz	11
1.6.3 A YOLO detekciós algoritmus	11
1.6.4 A regressziós sík analízis	12
1.6.5 Lézer mikrodisszekció	13
2. Célkitűzés	14
3. Anyagok és módszerek	15
3.1 Emlőszövet MIDOG adathalmaz	15
3.2 Mellékvese tumor adathalmaz	15
3.3 Szövet előkészítés és festés	16
3.4 A képfeldolgozási munkafolyamat bemutatása	16
3.5 A tanító adat elkészítése	17
3.6 Mitotikus sejt detekció és klasszifikáció a YOLO detekciós algoritmussal	17
3.7 Regressziós sík analízis	18
3.8 Lézer mikrodisszekció	20
3.9 Ábrakészítés és a kódok elérhetősége	20
4. Eredmények	21
4.1 Mélytanuláson alapuló szegmentációs modell választása és optimalizálása mitotikus sejtekre	21
4.2 A kiválasztott mélytanuláson alapuló szegmentációs modell továbbfejlesztése	23
4.3 A mitotikus sejtek azonosítása a YOLO detekciós algoritmus segítségével	24
4.4 YOLO mitotikus fázis klasszifikáció	26
4.5 Mitotikus sejtek fázisának folytonos skálán való predikciója regressziós sík modell segítségével	28
4.6 Összetett munkafolyamat létrehozása mitotikus sejtek fázis-specifikus lézer-mikrodisszekciója érdekében	31
5. Diskusszió	33
5.1 A YOLO detekciós modell optimalizálása a MIDOG adathalmazon	34
5.2 Regressziós sík analízis fejlesztése és optimalizálása	35
5.3 Hisztológiai minták proteomikai és transzkriptomikai analízisének előkészítése	35
6. Köszönetnyilvánítás	36
7. Tételes szerzői hozzájárulás	37
Irodalomjegyzék	38

1. Bevezetés

1.1 A mitózis szerepe a tumoros folyamatokban

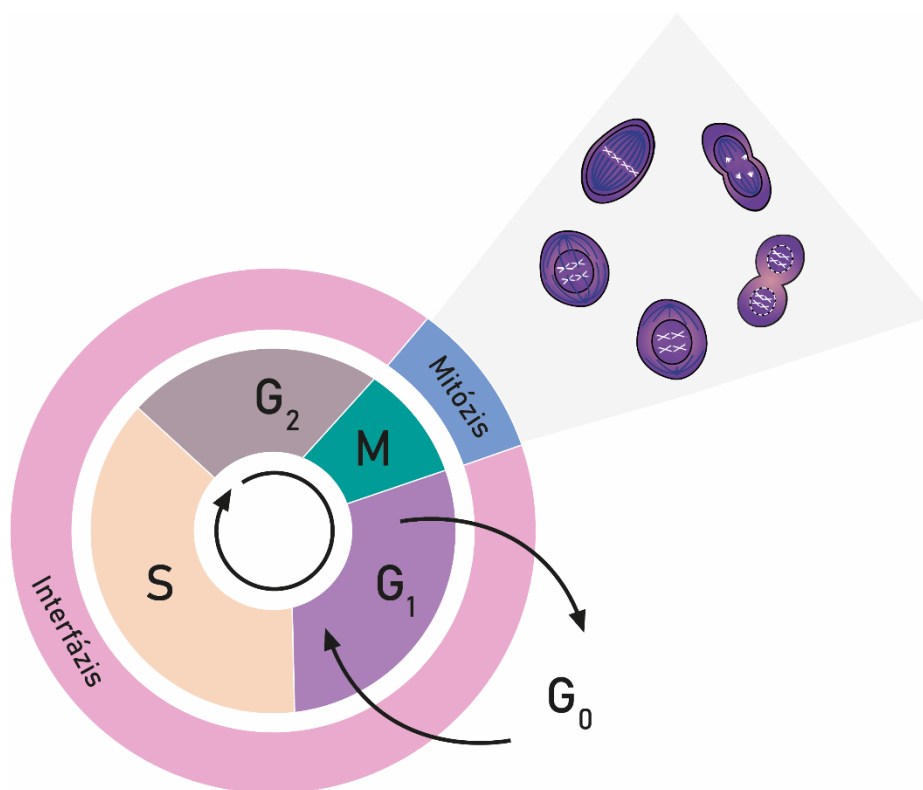
A normál mitózis alapvetően egy olyan folyamat, amely hozzájárul egy egészséges rendszer fenttartásához (például a szövetek kialakításával, a sebgyógyulással). Abban az esetben azonban, amikor ez a több ponton szabályozott rendszer felbomlik, a sejtek ellenőrizetlen és kontrollálatlan növekedésbe kezdenek, megnő a rák kialakulásának kockázata. Ennek a kontrollálatlan sejtszétválásnak a mértéke a tumor növekedésének üteme mellett annak agresszivitását, valamint áttétképzési potenciálját is befolyásolja, mindemellett fontos mérlegelési szempont a kezelés tekintetében (1). Erre alapozva több módszert is kifejlesztettek a szakemberek a sejtproliferáció aktivitásának mérésére (mitotikus index számolása, Ki67 és foszfohiszton H3 (PHH3) jelölés alkalmazása), azonban ezeknek a módszereknek is vannak hátrányai. Például az elterjedt Ki67 festés egy antitest alapú, immunhisztokémiai (IHC) festés, amely a sejtciklus összes aktív fázisát jelöli (G1, S, G2, mitózis), és az interfázisos sejtekben is mutat jelölést, ám az előbbiektől eltérő szubcelluláris kompartmentben és intenzitással, így nem mondható specifikusnak a mitotikus sejtekre (1). Habár a PHH3 egy specifikusabb IHC jelölés, mivel a foszforilált H3 hisztonfehérjét festi a mitotikus ciklusban, viszont ennek az általános alkalmazhatóságát rontja, hogy különböző tumortípusokban széles skálán mozog a mitotikus indexszel való korrelációja (2). Így a legszélesebb körben alkalmazott és elfogadott módszer jelenleg a patológusok körében a mitotikus index vizuális meghatározása hematoxin és eozin festett szövettani metszeteken, mivel ez egy egyszerű és viszonylag gyorsan kivitelezhető, könnyen nyomon követhető technika a tumoros folyamatok előrehaladtának és viselkedésének megfigyelésében (1, 3). Mindazonáltal ennek a technikának is vannak hátrányai. Az adatokat utólagosan korrigálni kell, mivel a különböző laboratóriumok különböző vastagságú metszeteket használnak a diagnosztikához (4), valamint általában a teljes tumorterület kiértékelése annak nagysága miatt pusztán emberi energia befektetésével hosszadalmas, és megnő a hibalehetőség esélye is (3). Ennek a feladatnak a megkönnyítésére számos digitális képfeldolgozási technikával próbálkoztak, azonban széles körben egyik sem terjedt el (3).

1.2 Különböző mitotikus morfológiák bemutatása szövetmintákban

Annak érdekében, hogy a jelenleg elérhető automatikus mitózis detekciós technikáknál jobb eredményeket lehessen elérni, a mitotikus index számolásának egységesítése mellett szükség van a különböző mitotikus morfológiák megismerésére, a jellemzők elkülönítésére (3). Ez különösen fontos abban az esetben, ha mesterséges intelligencián alapuló módszert szeretnénk

alkalmazni, hiszen ehhez szükség van egy körültekintően annotált mitotikus sejtekből álló adatszett használatára.

A mitózis a sejtciklus része, amelyet a különböző specifikus fehérjék térbeli és időbeli eloszlása és megjelenése befolyásol (5). A sejtben végbemenő számos folyamat végeredményeként pedig két utódsejt jön létre. A mitózis a sejtciklus viszonylag rövid szakasza. Nagyjából két órát tesz ki a hozzávetőlegesen 24 órás teljes sejtciklusból (1. ábra), és ezalatt a rövid idő alatt számos morfológiai változáson megy keresztül a sejt, amelyeknek ismerete fontos az atípusos mitózistól és a mitózishoz nagyon hasonló alakoktól való megkülönböztetés érdekében. Tehát általánosságban elmondható, hogy a mitotikus sejteket a különböző fázisspecifikus morfológiai jellemzők alapján határozzák meg és sorolják a nagy osztályok egyikébe (profázis, prometafázis, metafázis, anafázis és telofázis) (1).



1. ábra A sejtciklus különböző fázisainak sematikus ábrázolása

1.2.1 Profázis

A profázis során (2/A. ábra), a mitotikus fázis első lépéseként a kromoszómák elkezdnek tömörülni és ezáltal megvastagodni, feltekeredni, ezzel pedig kialakul egy hiperkromatikus struktúra. A sejtmaghártya felbomlik, és elkezd kialakulni a citoszkeletális orsó, amely a kromoszómákhoz rögzül, és mozgatja őket a mitózis további szakaszaiban. A profázis három alfázisra különíthető el: korai, középső és késői; azonban az első kettő esetében elmondható, hogy nagyon nehézkes az azonosításuk, hiszen a maghártya felbomlásakor a “sejtmagok” egy

ideig megtartják kerek szerkezetüket, ebből kifolyólag könnyen összetéveszthetők egy apoptotikus sejttel vagy limfocitával, amelyek szintén lekerekített körvonallal és kondenzált maganyaggal rendelkeznek. Ezen struktúrák egyértelmű azonosítása a patológusok számára is egy igen nehéz feladat, így a diagnosztikában legtöbbször csak a késői fázist veszik figyelembe, amely már egyértelműen elkülöníthető a többi sejttől, a szabálytalan, hullámos körvonalai miatt. A prometafázisba lépés kezdetén pedig szörszerű nyúlványok kezdenek el kialakulni, amelyek megfelelő felbontású mikroszkópos képeken könnyen azonosíthatók (1, 3).

1.2.2 Prometafázis

A prometafázis szakaszában (2/B. ábra) a sejt morfológiai változásai folytatódnak, erre a fázisra jellemző a „tüskeszerű”, „szörszerű” morfológia, amely nagyban megkönnyíti a felismerhetőséget. Ennek a fázisnak a végén a kromoszómák egy rövid időre gyűrű alakba rendeződnek, amely a prometafázis egyedi jellegzetessége (3).

1.2.3 Metafázis

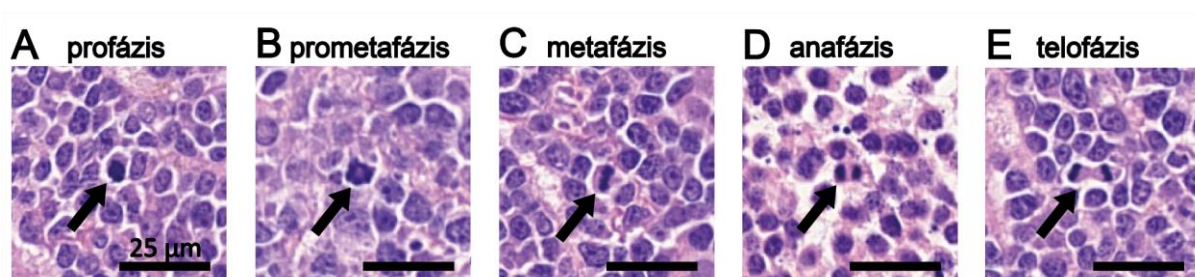
Amikor a sejtek eljutnak a metafázisba (2/C. ábra), addigra a magmembrán teljesen lebomlik, a kromoszómák az egyenlítői síkba vándorolnak. Hisztológiai mintákban egy lineáris sáv, ritkább esetekben gyűrű formájában látható a metszési síktól függően, amely nagyban megnehezíti a prometafázistól való elkülönítést az utóbbi esetben (1, 3). Mindezek mellett a metafázis is rendelkezik hozzá hasonló alakokkal, amelyek legtöbb esetben a rosszindulatú sejtek gyors növekedéséből fakadó struktúrák, esetlegesen rosszul fixálódott sztrómasejtek (1).

1.2.4 Anafázis

Az anafázis során (2/D. ábra) a testvérkromatidák elkezdnek elválni egymástól, és a sejt két ellentétes pólusára vándorolni a centromerek és húzófonalak segítségével. Egyes esetekben a testvérkromatidák „V” alakot formálhatnak a mozgítás következtében. Citológiailag ennek az eredménye két, egymással párhuzamos hiperkromatikus pálcaként jellemezhető, amely ugyanabban a sejtben található. A profázishoz hasonlóan az anafázis is részekre osztható: korai, amelyben a kromoszómák két csoportra oszlanak, és késői anafázisra, amelyben elkezdődik a kromoszómák szétválása. Az anafázis tekinthető a legkönnyebben azonosítható mitotikus fázisnak, hiszen nagyon egyedi tulajdonságokkal rendelkezik, annak ellenére is, hogy ez a legrövidebb és legdinamikusabb sejtosztódási fázis. A kromoszóma rendellenességgel rendelkező, hibás szétválásból fakadó alakjait is könnyen lehet azonosítani, ez nem befolyásolja a felismerhetőségét (1). Azonban az osztályozás során a telofázistól történő elkülönítése nehézkes, leggyakrabban a kromoszómacsoportok közötti távolság alapján történik (3).

1.2.5 Telofázis

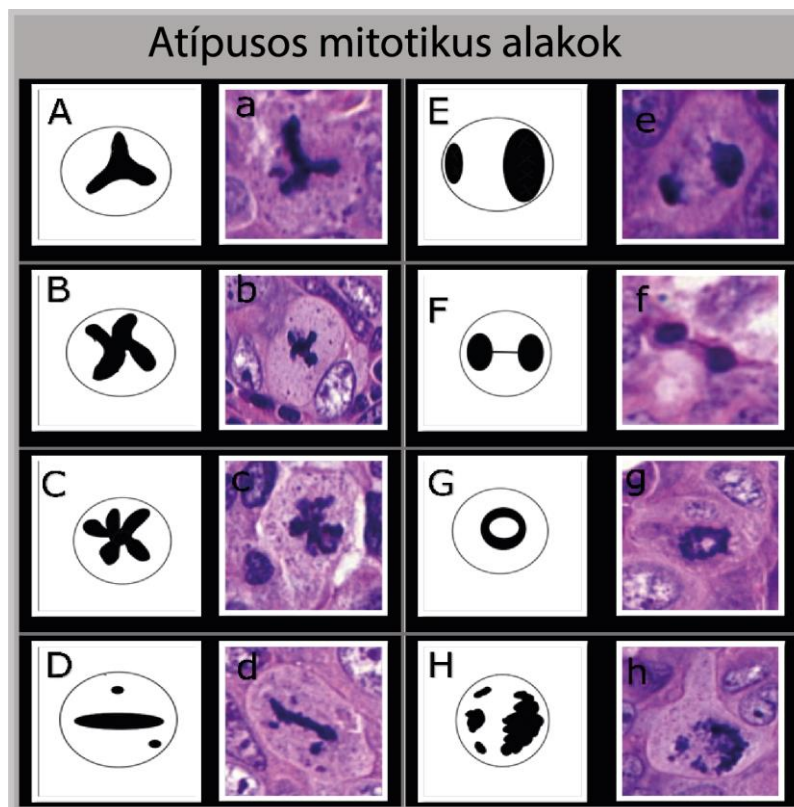
A telofázis során (2/E. ábra) a testvérkromatidák megérkeznek a sejt két egyenlítői pólusára, elkezd újra kialakulni a magmembrán és a nukleolusz. Az anyasejt citoplazmája szintén elkezd kétfelé osztódni, két leánysejt alakul ki azonos kromoszómakészlettel. Citológiai vizsgálatok során ebben a fázisban ritkán, de látható a kialakuló magmembrán, esetlegesen a két telofázisban lévő sejtet egy vékony citoplazmaszál kötheti össze (1, 3). A telofázist követően bekövetkezik a citokinézis, amely során fizikailag is elválík egymástól a két leánysejt. Amennyiben ez hibásan következik be, többmagvú sejtek alakulhatnak ki (3).



2. ábra: Mitotikus fázisok hisztológiai megjelenése hematoxin-eozinnal festett mellékvese tumor metszetben

1.2.6 Atípusos mitózis

A sejtciklus során minden fázist számos fehérje szabályoz és tart a megfelelő sejtosztódási folyamatban, ezzel megakadályozva a különböző mitotikus rendellenességek kialakulását. Azonban az ellenőrzőpontok hibái kontrollálatlan vagy rendellenes mitózist eredményezhetnek (1). A tumorok kialakulása során rengeteg genetikai hiba halmozódik fel, ehhez többek között a kromoszómák rendellenes szegregációja is hozzájárul (3, 6). Számos tanulmány alátámasztotta azt a feltételezést, miszerint a magas atípusos mitózisos sejtek (3. ábra) száma kedvezőtlen hatással van a tumor prognózisára, ebből kifolyólag az azonosításuk nagy jelentőséggel bír a jóindulatú és rosszindulatú tumorok megkülönböztetése során (1, 6).



3. ábra Atípusos mitotikus alakok sematikus és szövettani megjelenése

A, a Tripoláris mitózis; B, b Kvadripoláris mitózis; C, c Multipoláris mitózis; D, d Lag- típusú mitózis; E, e Poláris aszimmetria; F, f Anafázis híd; G, g Gyűrűs mitózis; H, h Szétszórt mitózis (A. Ibrahim és munkatársai alapján módosítva)

1.3 Mitotikushoz hasonló morfológiák

Az egyik legkritikusabb lépés a mitotikus index meghatározása során mind a kutatók, mint a mesterséges intelligencián alapuló mechanizmusok számára a mitotikus sejtek elkülönítése a mitotikushoz hasonló objektumoktól (1, 3). Általánosságban ezek közé sorolhatók az apoptotikus testek, hiperkromatikus sejtmagok, limfociták és egyéb gyulladásos sejtek, összezúzott sejtek, valamint a szöveti artefaktumok (1, 7). Megkülönböztetésükhöz alapos vizsgálatokra, szövettani ismeretekre, a képalkotás optimalizálására van szükség. A leggyakrabban előforduló, programozott sejthalálban résztvevő sejtek kisebb méretűek, sűrű, eozinofil citoplazmával rendelkeznek, a sejtmagok összezsugorodnak és kondenzálódnak a piknózis következtében, majd elkezdődik a fragmentálódásuk (kariorexis), végül a sejtmembrán leválik, és a magtörmelékek feloldódnak (kariolízis). Ezen folyamaton keresztülmenő sejtek a szövettani preparátumokban megtévesztően hasonlítanak a profázis kezdetleges szakaszához. A megkülönböztetésben a mag körvonalai (apoptotikus sejteknél sima, míg profázisú mitózisoknál hullámos), valamint a sejtek mérete (sejthalál során a sejt összezsugorodik, míg a sejtosztódás közben nem változik látványosan), ezen kívül a citoplazma eozinofil festődésének mértéke (apoptózisban erős, míg profázisban enyhe) nyújt segítséget.

Az apoptotikus sejtek azonosítása nemcsak a mitotikus index meghatározása szempontjából hasznos, hanem azért is, mert a nagy mennyiségű apoptotikus sejt legtöbb esetben a rák rossz prognózisával jár együtt (1). A limfociták, granulociták és egyéb gyulladásos sejtek szintén hiperkromatikus sejtmaggal rendelkeznek, amelyek optimális körülmények között könnyebben felismerhetők, azonban összetéveszthetők egy profázisos sejttel. A fő különbség a sejtípusok között a sejtmagban lévő kromatin eloszlása, amely a limfocitákban egyenletesebbnek és sűrűbbnek tűnik, valamint a plazmasejtek gyakran bazofil citoplazmát tartalmaznak, a neutrofil granulocitákban pedig a sejtmagok szegmentáltsága szembetűnő (1). Sötét hiperkromatikus sejtmagok gyakori velejárói a rosszindulatú tumorokban lévő sejteknek, és a kromatin kondenzáltsága miatt szintén a profázishoz hasonló morfológiával rendelkeznek. A mitózistól való megkülönböztetése a magmembrán hiánya alapján történik, a szabálytalan, „szőrös” sejt-körvonalak figyelembevételével, amely a prometafázisú osztódó sejtek egyik jellegzetessége, míg ezekre a sejtekre sima, érintetlen magmembrán jellemző (1). A rutinszerű formalinos fixálás alkalmazásából fakadóan a hibás vagy késleltetett fixáció következményeként artefaktumok keletkezhetnek, valamint a szöveti törmelékek olyan műtermékeket eredményezhetnek, amelyek megnehezítik a mitotikus alakzatok helyes azonosítását a hiperkromatikus festődésükből kifolyólag (1).

1.4 Mélytanuló algoritmusok a diagnosztikában

Manapság a gépi-és mélytanulási technikák egyre nagyobb népszerűségnek örvendenek az orvosi diagnosztika javításában és korszerűsítésében, hiszen lehetőséget nyújtanak a tökéletes reprodukálhatóságra, a pontosságra és a gyorsaságra, ezzel könnyítve az orvosok és kutatók dolgát (8). Azonban a diagnosztika területein a mesterséges intelligencia által vezérelt kísérletek az ígéretes eredmények ellenére a legtöbb esetben nem jutnak el a klinikai alkalmazhatóságig, mert ezek a folyamatok nagymértékben az orvosok által meghatározott adatokra támaszkodnak, amely egyének és laboratóriumok között változhatnak (8).

Kezdetben a képelemző algoritmusokat például a sejtek mennyiségi és minőségi (méret, darabszám, alak, kromatin eloszlása stb.) jellemzőinek meghatározására használták. Később azonban magukra a képekre fektették a hangsúlyt (kontraszt, orientáció stb.) abból a feltételezésből kiindulva, hogy ezek a típusú osztályozó algoritmusok megtalálják a különböző jellemzők közötti kapcsolatokat. Így alakult ki a modern digitális patológia, amely a legtöbb esetben felülmúlja a hagyományos képelemzési módszereket (8).

Annak köszönhetően, hogy a kutatók felfigyeltek a mesterséges intelligencia sikereire, az idő előrehaladtával nőtt a megoldani kívánt feladatok összetettsége. A digitális patológia fejlesztésére jött létre például a CAMELYON kihívás, amelyben az emlőrák metasztázisában

szerepet játszó sentinel nyirokcsomók kimutatásának fejlesztése volt a cél, elsőként lehetővé téve egy nagyobb elemszámú ($n = 1399$), nyílt hozzáférésű tanítóhalmaz meglétét. Hasonlóan a 2016-os TUPAC kihíváshoz, melynek során a mitotikus sejtek kimutatása volt a cél. A különböző kihívásokból kifolyólag a digitális patológia fejlődésnek indult, jelenleg pedig számos törekvés zajlik a rutinszerű alkalmazhatóságára, például a tumorok felismerésében és osztályozásában, a képszegmentálásban, a sejtek felismerésében és megszámlálásában, valamint a mitózis detektálásában (8).

1.5 Mitózis detekciós módszerek

A mitotikus sejtek számszerűsítése a szövettani metszeteken nagy jelentőséggel bír a diagnosztikában, mivel a tumor agresszivitásának értékelésére szolgál.

A WHO által előírt, az emlőrák osztályozására szolgáló „Nottingham Grading System” egyik elemeként is megjelenik a mitózisban lévő sejtek számának meghatározása (9). Azonban az osztódó sejtek kimutatása nehéz, hiszen egy összetett folyamatról van szó, amely során a sejtmag folyamatos átalakulásokon megy keresztül, ezeket a szakaszokat pedig változatos morfológia jellemzi (10).

A mitózis kimutatása a mai napig a legtöbb esetben manuális detekció útján történik, amelynek során a patológusok a fénymikroszkóp alatt figyelik meg és osztályozzák a szöveti metszeteket, azonban mindez egy viszonylag lassú folyamat, amely az egyéni tapasztalatok és tudás variabilitása alapján gyenge reprodukálhatóságot eredményez (9).

A számítógépes képfeldolgozási technológiák nagyban hozzájárulnak a diagnosztika pontosságának javításához, például a mitotikus sejtek felismeréséhez és klasszifikációjához. Ezeknek a programoknak a fő feladata, hogy a digitalizált patológiai metszetek segítségével megalkossanak egy modellt, amely elősegíti a mitotikus sejtek felismerését robosztusabb módon, a patológus által előre megtanított minták alapján (9). Azonban a legtöbb esetben ezek a rendszerek nem jutnak el a gyakorlati alkalmazhatóságig a korábban említett befolyásoló tényezők miatt (mitotikushoz hasonló alakok, mitotikus komplexitás), valamint a különböző laboratóriumokból származó hematoxilin-eozin festett metszetek változatossága szintén megnehezíti az automatizálást (9). Mindemellett fontos megemlíteni azt a tényt, miszerint a mitózis kis számban fordul elő egy látótéren belül a nem osztódó sejtekhez képest, amely a pozitív és negatív példák kiegyensúlyozatlanságához vezethet. Ezt a jelenséget tovább nehezíti a mesterséges intelligencián alapuló modellek azon tulajdonsága, miszerint a módszer a meglévő adathalmazokra támaszkodik, azonban a nyíltan hozzáférhető adatkészletek csak néhány képfeldolgozási kihívás kiírásából származnak, mely csak pár tumortípusra korlátozódik (főleg emlőrák), valamint előfordulhat, hogy a képminőség jobb a rutin

diagnosztikában használt képek minőségéhez képest, így a keletkező modellek alkalmazhatósága nem ideális (9).

1.5.1 Hagyományos gépi tanulási technikák

A hagyományos mitózis detekciós módszerek közvetlenül az eredeti RGB képeket dolgozzák fel úgy, hogy elsőként különböző képmanipulációs lépéseket hajtanak végre, majd azokról jellemző vektorokat nyernek ki, például a sejtek morfológiai és statisztikai paraméterei alapján, és ezekből építenek modelleket (9).

1.5.2 Mélytanuláson alapuló technikák

A mély neurális hálózatok nagyban hozzájárultak a számítógépes detektálási folyamatok javulásához, így előszeretettel alkalmazzák azokat az orvosi képfeldolgozás során is, például a sejtek szegmentálására, szövetek osztályozására vagy a mitózis detekciójára (9). Elsőként 2012-ben próbálkoztak ilyen módszerrel az ICPR (International Conference on Pattern Recognition) mitózis detektálási verseny keretében, amelynek során konvolúciós neurális hálót használtak (10). Mások pixelszintű osztályozókkal próbálkoztak, amelynek a lényege, hogy minden egyes pixel osztályozásra kerül, majd ebből számítják ki, hogy a pixelek milyen valószínűséggel tartoznak egy mitotikus sejthez (11). Egy következő megközelítés szegmentációs, detekciós és verifikációs mélytanulási algoritmusok kombinálásával közelítette meg a problémát (12).

1.5.3 Kombinált technikák

A hagyományos gépi tanuláson, és a mélytanulási algoritmusokon alapuló technikák különböző módszereken alapulnak, így az általuk kapott jellemzők is eltérnek. Ebből kifolyólag vannak olyan kutatók, akik megpróbálták kombinálni (9) a mélytanulási algoritmusokból kinyert jellemzőket a mesterségesen, a kutatók által felcímkezett jellemzőkkel, valamint figyelembe vették a szegmentált sejtmagok körüli területek jellemzőit is, hiszen a patológusok számára fontos tulajdonságokat hordoz, amelyet figyelembe vesznek a helyes azonosítás érdekében (13).

1.6 A dolgozatban felhasznált módszerek elméleti áttekintése

1.6.1 A BIAS képfeldolgozó program

Azért esett a képfeldolgozási lépések során a választásunk a kutatócsoportunk által kidolgozott Biological Image Analysis Software (BIAS) nevű programra, mert ez egy mesterséges intelligencián alapuló képfeldolgozó szoftver, amely lehetővé teszi a különböző kutatások során felhasznált két- és háromdimenziós minták digitális feldolgozását, ezzel hozzájárulva a kísérletek sikereihez.

A program megalkotói a tervezés során arra fektették a hangsúlyt, hogy a BIAS szoftver képes legyen a hagyományos fény- és fluoreszcens mikroszkópok, valamint a lézeres mikrodisszekción alapuló mikroszkópok közötti képfeldolgozási lépéseket összehangolni, amely segítségével zökkenőmentesen kombinálható a sejt kultúrák és hisztológiai minták nagyfelbontású képalkotása a mélytanuláson alapuló szegmentációval, valamint a fenotípus alapú osztályozással (14). A program fő profiljai közé tartozik az egysejt szinten történő lézer mikrodisszekció támogatása, amely a szoftver segítségével automatizálható, valamint a felhasználó igényeihez és a kísérlet tulajdonságaihoz igazítható. Választhatunk interaktív elemzést, fél-automatikus vagy teljesen automatikus analízist, ez nagyban lerövidítheti egy kísérlet idejét, így teret engedve a további vizsgálatoknak.

Egy munkafolyamat során sokszor szükséges riportokat és statisztikai adatokat generálni, amely szintén egy nehezebb és időigényesebb feladat, éppen ezért a BIAS-ban erre a célra is találunk beépített modulokat, valamint minden egyes munkafolyamat tartalmaz testre szabható modulokat.

1.6.2 A MIDOG 21 kihívás adathalmaz

Az elmúlt időszakban egyre nagyobb figyelmet kapott a rákkutatásban és a klinikai gyakorlatban a mitotikus index számolása, hiszen egyértelműen összefüggést mutat a tumorok proliferációjával. Jelenleg a legelterjedtebben használt technika a szöveti minták minősítése során egy képzett patológus által végzett vizuális kiértékelés. Ez viszonylag időigényes feladat, így a MICCAI (International Conference on Medical Image Computing and Computer Assisted Intervention) konferencia által kiírt kihívás a gépi tanuláson alapuló mitotikus sejt detektáló algoritmusok fejlődésének elősegítését célozta meg (15). A saját kísérleteink során mi is felhasználtuk az adathalmaz képeit, mind a regressziós sík alapú, mind pedig a detekción alapuló modellalkotáshoz. A MIDOG adathalmaz nagy előnye, hogy patológusok által validált annotációkat tartalmaz, amelyek kiegyensúlyozva tartalmazznak valódi mitotikus sejteket, és nehezen annotálható negatív példákat egyaránt.

1.6.3 A YOLO detekciós algoritmus

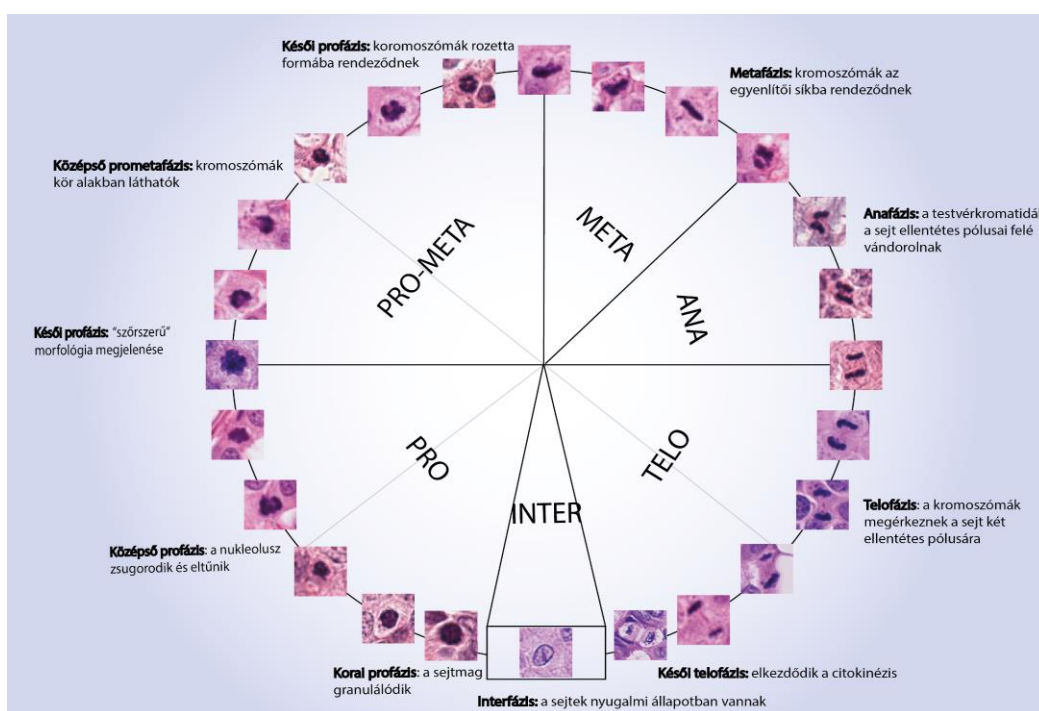
Jelenleg a leginkább elterjedt és használatban lévő objektum detekciós módszerek elve egy többlépéses folyamat: először potenciális befoglaló téglalapokat prediktálnak, majd azokon futtatják a klasszifikációs algoritmust, és egy későbbi lépésben finomítják a téglalapok helyét. Ez sok esetben egy lassú és több különböző komponens tanítását igénylő folyamat. A You Only Look Once (YOLO) módszer ezt a problémát igyekszik kiküszöbölni olyan módon, hogy az osztályozás helyett az objektumok (esetünkben a sejtek) detektálása befoglaló téglalapok regressziós problémájaként értelmezhető, amelyekhez hozzárendeljük az osztályok

valószínűségét. A modell meglehetősen gyors, ennek az oka pedig az, hogy a detektor képes a kívánt objektumok nagy hatékonyságú általánosítására (16).

1.6.4 A regressziós sík analízis

Számos biológiai folyamat, ideértve a sejtosztódást is, egy folytonos skálán változó jelenség, így a diszkrét osztályokra bontás csak korlátozott mértékben értelmezhető rajta, ezzel megnehezítve a gépi tanulás alkalmazhatóságát. Kutatócsoportunk tagjai ennek a problémának a kiküszöbölésére alkották meg a regressziós sík (RS) módszert (4. ábra), amely lehetővé teszi a morfológián alapuló gépi tanulást előosztályozás nélkül, ezzel valósághűbb módon elemezhetjük az adatainkat (17).

A regressziós sík az Advanced Cell Classifier (ACC) nevű hagyományos fenotípusos osztályozóba implementálva található meg az ACC v3.0 verziója óta. Hierarchikus módon minden egyes osztály egy külön regressziós síkkal bővíthető, ezzel lehetővé téve több regressziós sík bevonását egy adott projektbe (18). A regressziós síkra tetszőleges pozícióba helyezhetünk el különböző objektumokat, melyeket azután a pozíciójuk, egy egyedi azonosító és egy, a képek tulajdonságai alapján kiszámolt mátrix jellemez. A módszer távolságot mér, és ez alapján tudja a sejteket elhelyezni egymáshoz képest. Amennyiben egy új sejt helyét szeretnénk prediktálni a regressziós síkon, a modell kiszámolja, hogy mely értékek milyen mértékben korrelálnak a már meglévőkkel, és ezt súlyozza, ez alapján helyezi el a mitózis esetében egy olyan körlapon, melyben egy körcikket fenntartunk az interfázisban lévő sejteknek.



4. ábra: A regressziós sík illusztrációja az annotációs stratégiánkkal kiegészítve

1.6.5 Lézer mikrodisszekció

Miután sikerült azonosítani a sejteket, amelyekkel további kísérleteket szeretnénk végezni, a lézer mikrodisszekció egy gyors és hatékony módszer azok elkülönített kigyűjtésére. Segítségével a szöveti környezet ismeretében vagyunk képesek objektumokat (sejteket, de akár sejtmagokat vagy más sejtalkotókat) elkülöníteni a mintánkból. Ennek az elve a következő: a lézer a minták mikroszkópos vizualizációja mellett vágja ki a kísérlet szempontjából specifikus sejteket, akár egysejt szinten. Így egy olyan homogén mintát tudunk létrehozni, melyen olyan további vizsgálatok végezhetők el, mint például a minta proteomikai vagy transzkriptomikai összetételének meghatározása, ezáltal átfogóbb képet kaphatunk a vizsgált folyamatokról.

2. Célkitűzés

Dolgozatom célja az volt, hogy létrehozzak egy olyan munkafolyamatot, amellyel a világon elsőként formalin-fixált, paraffinba ágyazott, hematoxin-eozin-festett tumoros szövetekről készült mikroszkópos képeken azonosítani tudjuk a bennük található mitotikus sejteket, prediktálni tudjuk, hogy azok hol tartanak a mitózis folyamatában, és tetszőleges felbontással csoportokat tudjunk a sejtekből képezni, amelyek lézer mikrodisszekcióval történő kigyűjtésével és komplex omikai vizsgálatukkal új ismeretekre tehetünk szert a mitózis folyamatában résztvevő fehérjék dinamikájával kapcsolatban. Ehhez képfeldolgozási algoritmusok alkalmazására, tesztelésére és optimalizálására, valamint a kísérleti körülmények lézer mikrodisszekcióra való alkalmazására lesz szükség.

3. Anyagok és módszerek

3.1 Emlőszövet MIDOG adathalmaz

A szabadon hozzáférhető mitotikus sejt adathalmazok közül a 2021-es Mitosis Domain Generalization kihívás képeire esett a választásunk. A MICCAI (Medical Image Computing and Computer Assisted Intervention) konferenciához kapcsolódó kihívás középpontjában a digitális patológia fejlesztése állt. A kezdeményezés weboldaláról (<https://imig.science/midog2021/>) bárki ingyenesen letöltheti a patológusok által annotált adathalmazt, és szabadon felhasználhatja további kutatásokra. Az általunk felhasznált képek formalin-fixált, paraffinba ágyazott, hematoxin-eozinnal megfestett humán emlőtumor szövetről készültek három különböző mikroszkóp segítségével: Hamamatsu XR nanoszoomer 2.0, Hamamatsu S360, Aperio ScanScope C S2. Mikroszkóponként 50-50, teljes tárgylemezről készült kép került bele a válogatásba, melyekről egy képzett patológus egy-egy 2 mm²-es területet választott ki, melyet 10 látóterre osztva TIFF formátumban bocsátottak rendelkezésre. A mitotikus sejtek annotációját több szakértő közreműködésével, az összes osztódó sejt megtalálását célzó protokoll segítségével végezték el (19), emellett olyan nem osztó sejteket is megjelöltek, amelyek morfológiájukat tekintve nehezen elkülöníthetők a mitotikus sejtektől, hogy ezzel is segítsék az automatikus detekciót. Az így kialakított tanítóhalmaz 1721 mitotikus sejtet és 2714 nem osztódó sejtet tartalmazott.

3.2 Mellékvese tumor adathalmaz

A saját mitotikus adathalmaz kidolgozása során egy Epstein-Barr vírus asszociált Hodgkin-limfóma mellékvese tumor áttétjére esett a választásunk (20), amelynek a mikroszkópos képeit Dr. Sükösd Farkas patológus professzor bocsátotta rendelkezésünkre. A képek felhasználását a Szegedi Tudományegyetem Szent-Györgyi Albert Klinikai Központ Regionális Humán Orvosbiológiai Kutatásetika Bizottsága engedélyezte (engedélyszám: 5127, iktatószám: 17/2022-SZTE, protokollszám: DIGMAT-01). Ismert, hogy ennek a tumortípusnak az egyik legintenzívebb a mitotikus aktivitása, ebből kifolyólag sok osztódó sejt jelenlétére számítottunk. További pozitív tulajdonság, hogy a mellékvesében a szöveti struktúrák jól elkülöníthetők, ez megkönnyítette az annotálási folyamatot. Ebben a szövetben a limfociták és az apoptotikus sejtek megjelenése hasonlít a profázisban lévő osztódó sejtekéhez, így az előzőleg felállított annotációs stratégiánk alapján ezekre kellett különös figyelmet fordítanunk.

3.3 Szövet előkészítés és festés

A kiválasztott mellékvese minta feldolgozása a Szegedi Tudományegyetem Szent-Györgyi Albert Klinikai Központ Patológiai Intézetében történt a következő protokoll szerint:

A szövet eltávolítása sebészeti úton történt a páciens kezelésének részeként. A következő lépésben a minta egyenesen a 10%-os formaldehid oldatba került a műtőben. A minta szállítása a Patológia Intézetbe szobahőmérsékleten történt a fixáló oldatban, majd egy éjszakán tartó fixálás következett. Patológusi vizsgálat után a további feldolgozásra szánt részeket műanyag kazettákba, fixáló folyadékba helyezték. A paraffinálás délután kezdődött, az automatizált folyamat a következő lépéseket tartalmazta: 2 x 30 perc 70%-os etil-alkohol; 1 óra 90%-os etil-alkohol; 1,5 óra 96%-os etil-alkohol; 3 x 1,5 óra 2 rész izopropil-alkohol + 1 rész 96% etil-alkohol; 1,5 óra 1:1 arányú izopropil-alkohol+xilol; 2 x 1 óra xilol; 4 x 1 óra paraffin. Ezt követően a minta reggelig az utolsó paraffinos oldatban maradt. A paraffin blokkokból 2-10 µm vastag metszeteket készítettek mikrotóm segítségével. A metszetet üveg tárgylemezre rakták ki vízfürdő segítségével. A rutin szövettani hematoxilin-eozin festés festő automatában történt, a következő lépésekkel: 5 x 1,5 perc xilol; 4 x 1,5 perc etil-alkohol; 1 x 1,5 perc desztillált víz; 5 x 1,5 perc hematoxilin oldat (amely a következő összetevőkből áll 5 literre: hematoxilin 10 g; nátrium-jodát 1 g; alumínium-szulfát 88 g; desztillált víz 3750 ml; etilén-glikol 1250 ml; cc. ecetsav 100 ml); 1 x 1,5 perc desztillált víz; 1 x 1,5 perc 1% sósav 70%-os alkoholban oldva; 1 x 1,5 perc desztillált víz; 1 x 1,5 perc csapvíz; 1 x 1,5 perc eozin oldat (amely a következő összetevőkből áll 1000 ml-re: eozin Y (Sigma) 10 g, desztillált víz 1000 ml, kalcium-nitrát 1g); 1 x 1,5 perc desztillált víz; 3 x 1,5 perc etil-alkohol; 2x 1,5 perc acetone; 1 x 1,5 perc xilol. Leica Surgipath Micromount fedőanyaggal és fedőlemezzel való lefedés után a metszetről digitális képek készültek 3DHitech Panoramic MIDI tárgylemez szkennel segítségével (20 × objektív, 0.24 µm-es pixelméret).

3.4 A képfeldolgozási munkafolyamat bemutatása

Első lépésként az általunk választott mellékvese tumoros szövets képet a BIAS szoftverbe importáltuk. Ehhez segítséget nyújtott a külön applikációként rendelkezésre álló importer.

Ezt követte a BIAS programban történő előfeldolgozás, amely során kiválasztottuk a potenciálisan tumoros régiókat, mivel ott nagyobb arányban találhatóak osztódó sejtek.

Ezután elkezdődhetett a sejtek kontúrjainak kijelölése, vagyis a szegmentálás. A BIAS-ban számos beépített, különböző nagyítású és eltérő karakterisztikájú és festési eljárásokkal készült mintán tanított mélytanulás alapú szegmentációs modell áll rendelkezésre, valamint módosíthatjuk a paramétereket a nekünk leginkább megfelelő módon. Ezek közül mi a

következőt választottuk: “Histological Nucleus Segmentation v1” módszer, 4-es skálázási paraméter, 50% detekciós és 50% kontúr konfidencia.

3.5 A tanító adat elkészítése

Ezt követően pedig elkészítettük az osztályokat az előzetesen szegmentált objektumokból az annotációs stratégiánk alapján. Kezdeti manuális osztályozás után egy gépi tanuló algoritmus alkalmazásával automatikusan is gazdagíthattuk az osztályokat. Első lépésként mitotikus (n=793) és interfázisos (n=362) osztályokat készítettünk a mitózis detekciós feladat tanítóhalmazaként, később a mitotikus osztályt felbontottuk az öt nagy mitotikus fázisra (elemszámok: profázis= 76, prometafázis= 56, metafázis= 576, anafázis= 59, telofázis= 26).

3.6 Mitotikus sejt detekció és klasszifikáció a YOLO detekciós algoritmussal

Első lépésként kinyertük a mitotikus sejtek helyére vonatkozó információkat a MIDOG adathalmazhoz tartozó, annotációkat tartalmazó fájlból. Ezután kiszűrtük az osztódó sejtek közül az atípusos mitózis kritériumainak megfelelő sejteket, mivel ebben a kísérletben a normál mitotikus sejteket szerettük volna azonosítani. A következőkben három csoportra osztottuk az adathalmaz képeit: 70% került a tanítóhalmazba, amelyet közvetlenül a detekciós modell tanítására használtunk fel, 15% került a validációs halmazba, melyet a modell súlyainak optimalizálására használtunk, végül 15% került a teszt halmazba, melyekkel kizárólag a már kész modell hatékonyságát teszteltük.

Az eredeti képek viszonylag nagy méretűek, és kis sűrűségben találhatók meg rajtuk a mitotikus sejtek, így a tanításhoz az annotált mitotikus sejtek véletlenszerű 512x512 pixel környezetében készített képkivágásait használtunk fel a modellekészítés során, melyhez a YOLO módszer (16) 5. verzióját (<https://github.com/ultralytics/yolov5>) használtuk fel.

Kétféle modellt készítettünk: az első a mitotikus sejteket detektálja, így csak a mitotikus osztályt használtuk fel hozzá; a második pedig a mitózis nagy fázisai szerint klasszifikálja a sejteket. A klasszifikációs modell létrehozásához a regressziós sík alapjául szolgáló annotációt vettük figyelembe, majd a koordinátákat diszkrét osztályokká alakítottuk.

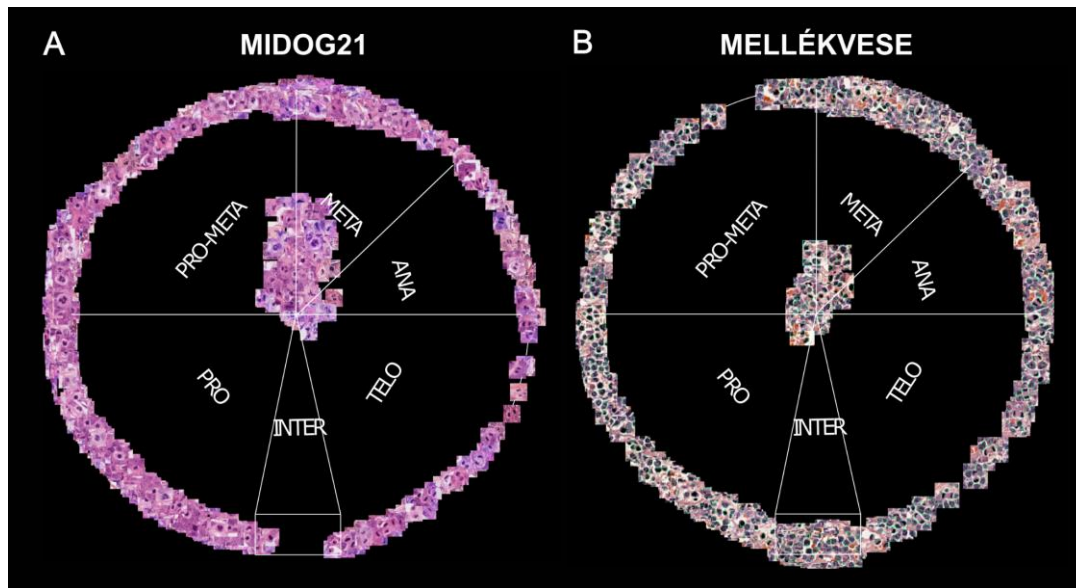
Öt különböző mélységű architektúrát teszteltünk, ezek közül a legjobbat választottuk ki a Jaccard-index hasonlósági metrika alapján. Az index úgy jellemzi az eredeti és a prediktált határoló téglalapok hasonlóságát, hogy veszi a téglalapok metszetének és uniójának hányadosát. Ez minél nagyobb szám, annál jobban átfednek a sejtek eredetileg meghatározott helyei a modell által prediktált helyekkel. Az architektúrákat 300 epochon keresztül és 4-es batch mérettel futtattuk, 0,01-es kezdeti tanulási rátát használva, amely 0,001-ig csökkent. A veszteségfüggvény optimalizációjához az Adam-módszert alkalmaztuk. A veszteségi értékek 120 epoch után konvergáltak, közülük a legalacsonyabb veszteségértékkel rendelkezőt

választottuk ki. Az eredmények kiértékeléséhez az AP50 (Average Precision, átlagos pontosság 0,5 metszet/ unió küszöbértéken), és az mAP (mean Average Precision, átlagos AP) értékeket számoltuk ki. A tanítást a következő specifikációkkal rendelkező számítógépen futtattuk: AMD Ryzen 7 5800X 8 magos processzor, @ 3.8GHz CPU, Nvidia A6000 48GB VRAM, 10752 CUDA magok @1410 Mhz, 336 tenzor mag, Ampere architektúra, 32 GB DDR4 of RAM.

3.7 Regressziós sík analízis

A regressziós modell elkészítéséhez első lépésként 150x150 pixeles képkivágásokat készítettünk a MIDOG adathalmaz teljes látóteret tartalmazó nyers képeiből a halmazhoz tartozó annotációs fájlok alapján. A képek közepén egy-egy mitotikusnak címkézett sejt volt látható. A képeket ezután importáltuk az Advanced Cell Classifier (ACC, 3.1 verzió, (18)) nevű programba, és az annotációs stratégiánkkal összhangban elhelyeztük az osztódó sejteket az általunk definiált regressziós síkon. A sík egy körlapon ábrázolva tartalmazta a mitózis fázisait. Az euklideszi tér 270 foknak megfelelő koordinátája jelenti itt a 0 fokot, és az óramutató járásának megfelelő irányban haladva a +/- 15 fok az interfázisnak felel meg, a fennmaradó 330 fokot pedig a következőképpen osztottuk fel: 15-90 fok - profázis, 90-180 fok- prometafázis, 180-225 fok- metafázis, 225-270 fok- anafázis, 270-345 fok- telofázis. A fázisokhoz tartozó szög nagyságát előkísérleteink alapján a folyamat során végbemenő morfológiai változásokhoz igazítottuk. Ez azt jelenti, hogy pl. a profázis során több olyan alfolyamat játszódik le a mitotikus sejtekben, amelynek megfelelő finomságú ábrázolásához nagyobb felbontásra van szükség, míg például a metafázis során a sejtek fenotípusa szempontjából kevésbé látványos változásokon megy keresztül a sejt, így fajlagosan kisebb felbontás is elegendő. A körlap sugara kb. 4500 pixel egy 10.000x10.000 pixel nagyságú négyzeten. Az annotáció során a körlap területére helyeztük a normál megjelenésű sejteket, míg a körlap közepe felé, kb. egy 3000 pixel sugarú belső körbe helyeztük az atípusos mitózist ábrázoló vagy nem egyértelműen annotálható sejteket (5/A. ábra).

A mellékvese tumor képeken hasonló módszerrel készítettük el a regressziós síkhoz tartozó reprezentációt (5/B. ábra), azzal a különbséggel, hogy a nyers képek kisebb felbontása miatt elegendő volt 75x75-ös képeket vágni ki a mitotikus sejtekről, valamint, azokat a BIAS program segítségével, saját annotációnk alapján készítettük. Összesen 1024 mitotikus és interfázisos sejtet választottunk ki az adathalmazból.



5. ábra: A regressziós sík annotációja a MIDOG (A) és a mellékvese (B) adathalmaz esetén

Az elkészült regressziós sík annotációs halmaz segítségével a következő lépésben regressziós modellt készítettünk a MIDOG adathalmaz képei alapján. Ennek bemenetétül a 16 bites nyers képekből normalizáció segítségével 8 bites képeket generáltunk. A tanításhoz a normál morfológiájú mitotikus sejteket használtuk fel, az atípusos sejteket elvetettük. A megmaradt 1718 sejt képét három csoportba osztottuk: 85% került a tanító halmazba, amelyet közvetlenül felhasználtunk a modell létrehozása során, 10% került a validációs halmazba, amely a tanítás során a súlyok ellenőrzésére szolgált, és 5% került a teszt halmazba, melyeket a modell kiértékelésére használtunk.

A tanító halmaz 1290 elemén kiterjedt augmentációs lépéseket hajtottunk végre a tanítás sikerességének és a modell általános alkalmazhatóságának elősegítése céljából, így több mint 11610 képen készülhetett el a tanítás. Az augmentáció során a következő módszerek kombinációját használtuk a generált képeken különböző valószínűséggel: hisztogram széthúzás, hisztogram kiegyenlítés, szín invertálás, a csatornák felcserélése, Gauss-féle zaj, impulzus zaj, elmosás, forgatás (-90 és +90 fok között), skálázás (0.9 és 1.1 között), tükrözés, transzláció (max. 3 pixel minden fő irányban).

A modell alapja egy ImageNet adatbázison előtanított Inception V3 (21) konvolúciós neurális hálózat volt, melyet úgy módosítottunk, hogy két értéket adjon eredményül, melyek a regressziós sík x és y koordinátájának felelnek meg.

A modell MATLAB (M2020b) segítségével készült egy alábbi tulajdonságokkal rendelkező számítógépen: Intel Xeon E5-2620 v4 @2.10GHz processzor, NVidia GeForce 2080 Ti 11 GB VRAM GDDR5x, 4352 CUDA mag @1350MHz Pascal architektúrával, 32 GB DDR4 RAM. A konstans 0.0001-es tanulási rátát előkísérletek segítségével állítottuk be. A nagymértékű

augmentáció miatt 5 epoch elegendő volt a tanítás során. A mini-batch méretet 16-ra állítottuk be. A tanítás kiértékeléséhez a négyzetes középhiba (root mean square error, RMSE) értéket választottuk. Ez a 10.000x10.000-es négyzetes regressziós síkon a prediktált és a valódi koordináták távolságaként is értelmezhető.

Ezután a mellékvese adathalmaz képeit használtuk bemenetként a MIDOG adathalmazon elkészült modellen, hogy prediktáljuk az egyes sejtek helyét a regressziós síkon.

3.8 Lézer mikrodisszekció

Mivel a BIAS program egyik fő profilja az egysejt kivágás optimalizálása, így ehhez a művelethez külön modulok tartoznak, amelyek segítségével megkaphatjuk a sejteink pontos koordinátáit, ezzel automatizálható a kivágó mikroszkóp. A lézer pontosságához igazodva állíthatjuk a sejtek kontúrját, ezzel elkerülhető, hogy fontos részekbe vágjunk bele. Ezt követően exportálhatjuk a kontúrokat, és átvihetjük őket az általunk használt Leica LMD6 mikrodisszekciós mikroszkópra. A vágás közben a minták gyűjtése PCR csövekbe történik.

3.9 Ábrakészítés és a kódok elérhetősége

A megjelenített diagramokat a MATLAB program R2020b verziójával készítettük, a többi ábrát az Adobe Illustrator (23.0.2 verzió) és Inkscape (1.1.1 verzió) szoftverekkel. Egy adott ábrán belül a különböző régiók intenzitás értékeit minden lépésben ugyanúgy módosítottuk.

A képfeldolgozási lépésekhez szükséges program kódokat MATLAB (M2020b) és Python (3.6.8 verzió) segítségével készítettük, azokat verziókövető rendszer segítségével privát repozitóriumban tároljuk a kísérletek publikálásáig.

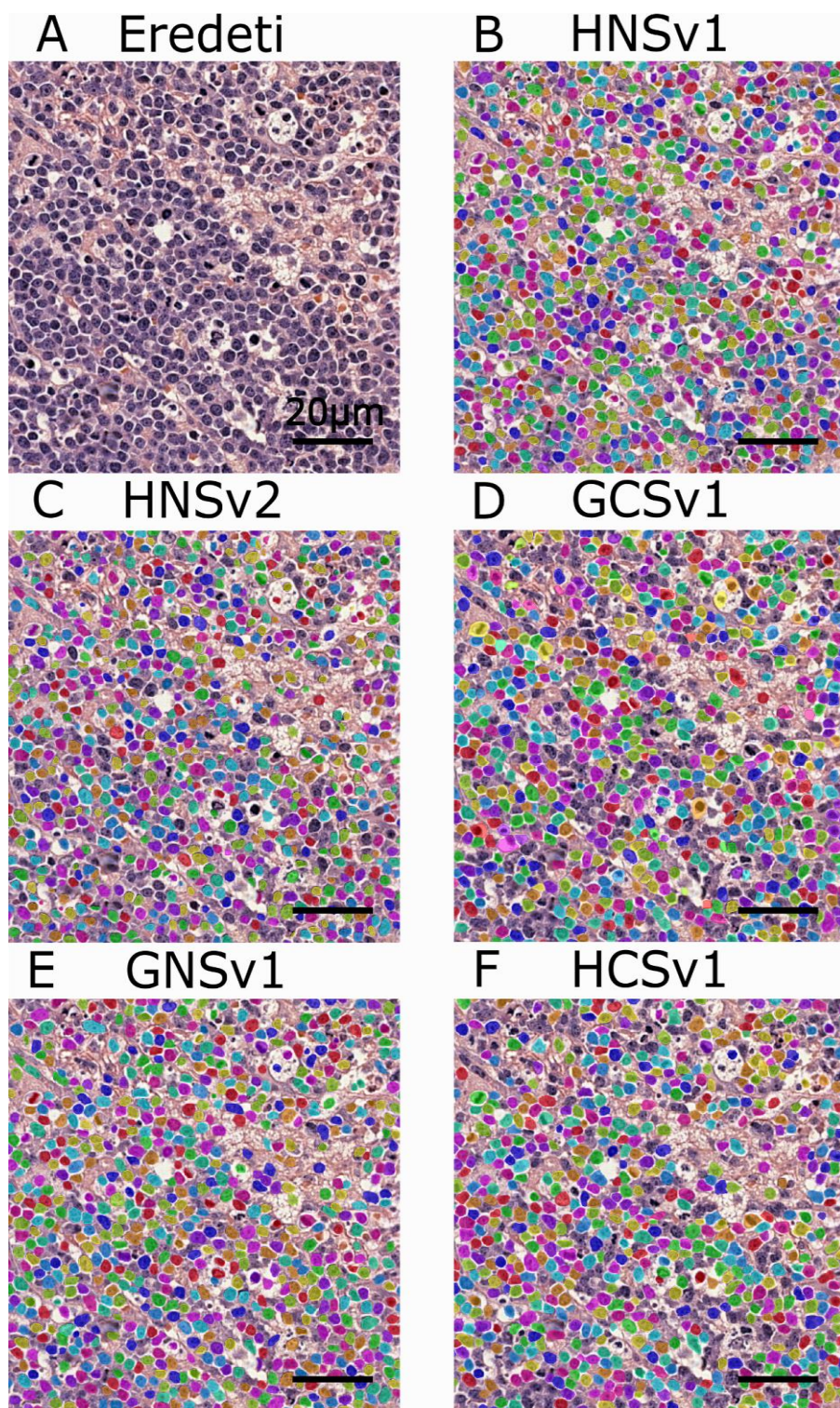
4. Eredmények

4.1 Mélytanuláson alapuló szegmentációs modell választása és optimalizálása mitotikus sejtekre

Annak érdekében, hogy képesek legyünk egy olyan munkafolyamat létrehozására, amely használható a mitotikus sejtek automatikus azonosítására és osztályozására a fázis-specifikus egysejt-kivágás megvalósítása céljából, az első feladatunk az volt, hogy az általunk választott hematoxilin-eozin festett humán mellékvese tumor adathalmaz képeiből összeállítsunk egy annotált tanító halmazt a mitotikus sejtekből. Az első megközelítésünk az volt, hogy a képeken szegmentáljuk a mellékvese sejtmagokat vagy sejttesteket, majd ez alapján címkézzük fel őket az annotáció során.

A képfeldolgozási folyamathoz a BIAS szoftvert választottuk, melyben számos hagyományos, illetve mélytanuláson alapuló szegmentációs módszer áll rendelkezésre a feladat elvégzéséhez. Előkísérleteink során és a feladat jellegéből adódóan úgy találtuk, hogy a mitotikus sejtek szegmentációjára a mélytanuló modellek előnyösebbek a hagyományos pl. küszöbölésen alapuló technikáknál, így ebben az irányban folytattuk a vizsgálódást. A legjobb mélytanuló szegmentációs modell kiválasztásához a meglévő modelleket hasonlítottuk össze kvalitatív és kvantitatív módszerekkel (6. ábra) (1. táblázat), figyelembe véve, hogy hány objektumot detektál a modell összesen, ebből mennyi a valódi sejt, valamint megtalálja-e az algoritmus a mitotikus sejteket, hiszen ez a legfontosabb és egyben legnehezebb feladat.

A mitotikus sejteket legnagyobb arányban a “Histological Nucleus Segmentation v1” (6/B. ábra) nevű modell azonosította, ezért ezt választottuk a további vizsgálatokhoz. Ennek a modellnek a tanítása hematoxilin-eozin festett hisztológiai képeken történt a CellProfiler program NucleAIzer keretrendszer segítségével (22).



6. ábra: A különböző mélytanuláson alapuló szegmentációs modellek eredményei mellékvese tumor metszeten

A: eredeti kép; Szegmentációs modellek: B: “Histological Nucleus Segmentation v1”; C: “Histological Nucleus Segmentation v2”; D: “Generic Cytoplasm Segmentation v1”; E: “Generic Nucleus Segmentation v1”; F: “Histological Cytoplasm Segmentation v1”. A különböző színű maszkok különböző szegmentált objektumokat (sejtmag, sejttest) jelölnek.

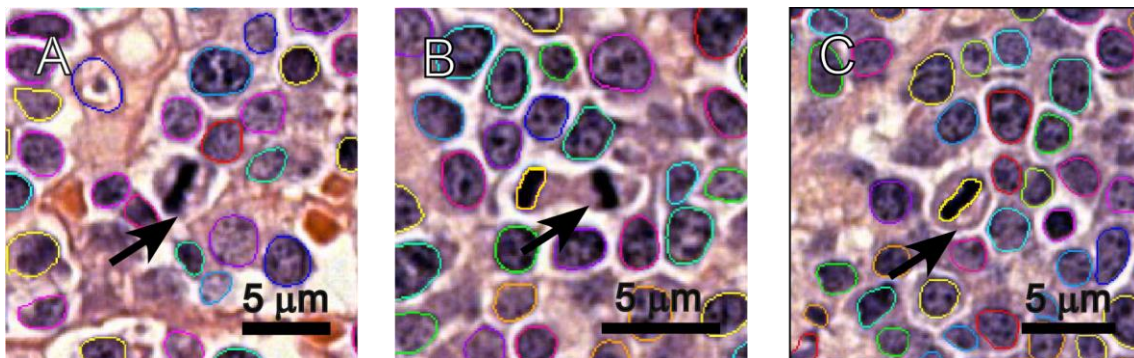
1. táblázat: A különböző mélytanuláson alapuló szegmentációs modellek összehasonlítása kvalitatív és kvantitatív módszerekkel

		HNSv1	HNSv2	GCSv1	GNSv1	HCSv1
		detektált objektumok száma				
Bemeneti skála	0,5	0	466	18	0	1
	1	7	873	532	19	76
	2	560	925	748	647	471
	3	829	949	438	872	711
	3,5	899	965	346	915	738
	4	949	958	301	931	755
Detekció konfidencia	5	950	960	302	931	756
	25%	949	925	748	872	738
	50%	949	925	748	872	738
Kontúr konfidencia	75%	895	889	371	753	666
	25%	949	925	748	872	738
	50%	949	925	748	872	738
	75%	949	925	748	872	738
Kvalitatív megfigyelések	80%	949	925	748	872	738
		Jó arányban detektálja a mitotikus sejteket, kis objektumok	Hiányos sejt szegmentáció, kis objektumok	Hiányos sejt szegmentáció, a mitotikusokat ~50%-ban szegmentálja	Hiányos sejt szegmentáció	Hiányos sejt szegmentáció, mitotikus sejtekre szuboptimális

Bemeneti skála: az eredeti kép mesterségesen nagyítható vagy kicsinyíthető a paraméterek változtatásával, amely a szegmentálandó kép és a tanítóhalmaz képeinek felbontásbeli különbségét, illetve a szegmentálni kívánt objektumok méretbeli különbségét küszöböli ki. **Detekció konfidencia:** annak a mérőszáma, hogy a detektált objektum milyen eséllyel felel meg a modell paramétereinek, növelésével például elvethetők a detektálni kívánt objektumoknál kisebb objektumok. **Kontúr konfidencia:** annak a mérőszáma, hogy az előző szűréseken keresztül jutott, detektált objektumok egyes pixeli, milyen eséllyel tartoznak a detektált objektumhoz. Növelésével például kissé nagyobb kontúrokat detektál a modell. A táblázatban félkövérrel emeltük ki az optimális paramétereket.

4.2 A kiválasztott mélytanuláson alapuló szegmentációs modell továbbfejlesztése

Az általunk kiválasztott, optimalizált paraméterekkel rendelkező módszer megbízhatóan képes volt szegmentálni a mellékvese szövet normál sejtjeit, és a mitotikus sejteket ez találta meg a legnagyobb hatásfokkal, ám ennek a modellnek is voltak hiányosságai az adathalmazunk képein (7. ábra). Például voltak olyan mitotikus sejtek, amelyeket nem ismert föl, sejtnek interpretált több olyan objektumot, amelyek nem voltak azok (például sejtközzöti állományt, sejttörmeléket), sokszor a szegmentálás során csak a kromatint vette figyelembe, nem pedig az egész sejtet.



7. ábra: A mitotikus sejtek szegmentációs hibáinak illusztrációja

A: nem szegmentálja a sejtet a modell semmilyen formában; B: a telofázisban lévő osztódó sejtnek csak az egyik leánysejtjét szegmentálja a modell; C: bár a modell helyesen megtalálta az osztódó sejtet, a szegmentáció csak a kromatin részt foglalja magába, azonban a képfeldolgozás további lépései érdekében előnyösebb, ha a sejt minél nagyobb részét szegmentáljuk.

Ennek érdekében következő lépésként a kiválasztott modell továbbfejlesztését tűztük ki célul. Ez és az ehhez hasonló általános sejtmag-, vagy sejttest szegmentáló modellek tanító adathalmazainak képein túlnyomó többségben normál, nem osztódó sejtek szerepelnek, sőt, a mitotikus sejtek a szegmentáció céljától függően akár úgy is jelenhetnek meg, mint nem kívánatos osztály. Ezért kiegészítettük az eredeti modell tanítóhalmazát az általunk használt képek javított szegmentációjával.

A képek manuális szegmentálásához szintén egy, a kutatócsoportunk által fejlesztett program, az AnnotatorJ (23) volt segítségünkre. Az ImageJ nevű Java alapú platformra íródott kiegészítő segítségével új kontúrokat rajzolhattunk ott, ahol nem talált sejteket a modell, a hibás kontúrokat pedig ki tudtuk javítani. Így képesek voltunk egy mitotikus sejteket nagyobb arányban tartalmazó adathalmaz létrehozására. Ezen az új adathalmazon újra tanítottuk a modellünket, azonban ez sem hozta el a várt sikert. A modellt általa még nem látott képeken tesztelve nem találtunk jelentős kvalitatív változást: a mitotikus sejtek szegmentációját nem sikerült jelentősen javítani, az optimalizált modell hasonló típusú hibákat követett el, mint az eredeti.

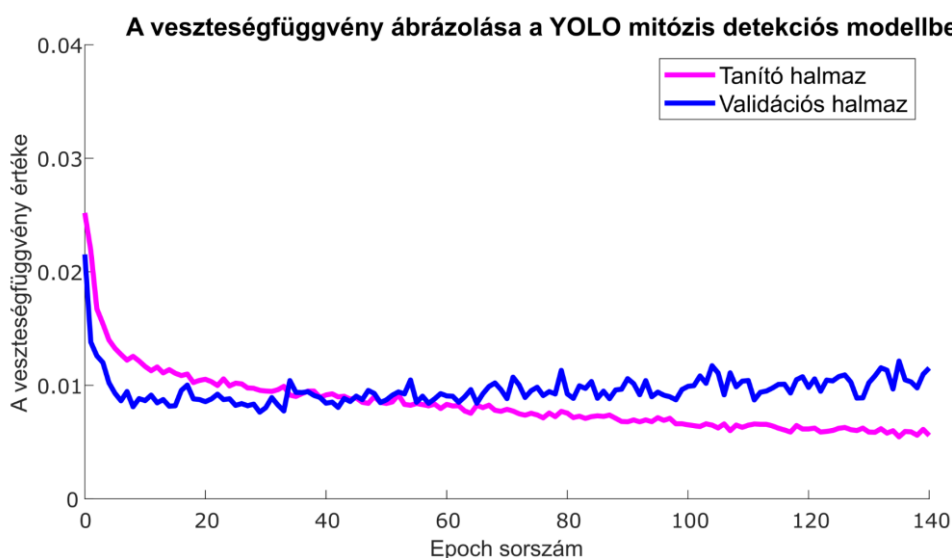
4.3 A mitotikus sejtek azonosítása a YOLO detekciós algoritmus segítségével

A következőkben tehát megpróbáltunk egy alternatív lehetőséget találni arra, hogy azonosítani tudjuk a mitotikus sejteket a hematoxilin-eozin-festett szöveti metszeteken.

A YOLO algoritmusra esett a választásunk, amely egy olyan objektum detekciós keretrendszer, amely a detekciók befoglaló tégalapjait és a prediktált osztály (ebben az esetben, hogy mitotikus sejtet talált-e) valószínűségét adja vissza.

A modellt szeretnénk volna egy olyan adathalmazon tanítani, amelyen rendelkezésre állnak patológusok által validált mitotikus sejt annotációk. Ezek közül a MIDOG hematoxin-eozin-festett emlő tumor képgyűjteményt választottuk.

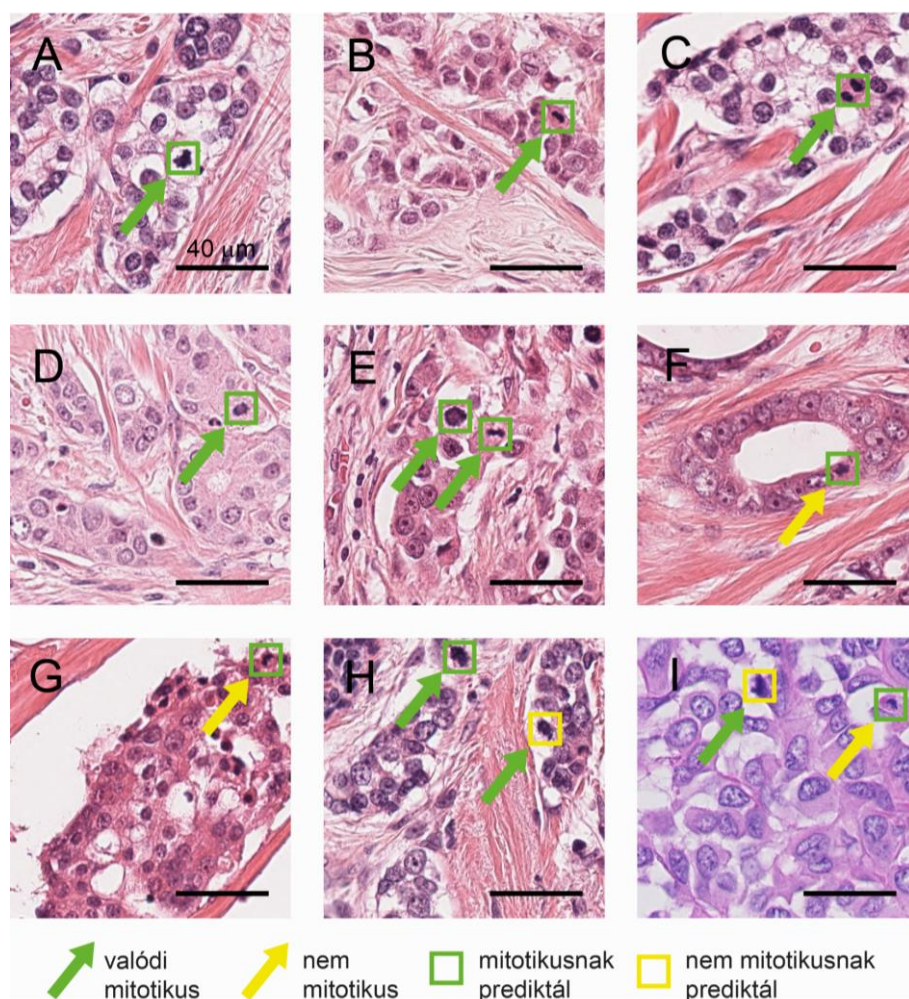
Az első feladatunk az volt, hogy a teljes metszetet tartalmazó képeken detektáljuk a mitotikus képeket. Ehhez a képeket tanító, validációs és teszt halmazokra osztottuk (ld. Anyagok és módszerek Mitotikus sejt detekció és klasszifikáció a YOLO detekciós algoritmussal fejezet), majd kisebb képekre daraboltuk az annotációban szereplő környezetét, ezt használtuk a módszer bemeneteként. Öt különböző architektúrát teszteltünk, a teszt és a validációs halmaz segítségével optimalizáltuk a modelleket, majd kiválasztottuk közülük a legjobbat a Jaccard-index hasonlósági metrika alapján. A tanító halmazon mért veszteségfüggvény folyamatosan csökkent a tanítás során, és konvergenciát mutatott (8. ábra), ami sikeres tanítást jelez. A validációs halmazon mért veszteségfüggvény is konvergált, ám a nagyobb sorszámoknál az enyhe növekedés túlillesztésre utal.



8. ábra: A tanító és a validációs halmaz veszteségfüggvényeinek ábrázolása a YOLO mitózis detekciós modellben

Magenta színnel jelöltük a tanító halmazon, kék színnel a validációs halmazon mért veszteségfüggvény értéket

A modell 0.5 IoU (metszet/unió) küszöbértéken (Jaccard-index) 0.927 AP50 detekciós eredményt ad a teszthalmazon, amely meglehetősen jó alkalmazhatóságot jelent mitózis detekciós feladatra (9. ábra). A 0.608 mAP érték azt mutatja, hogy a modell tovább optimalizálható, ugyanis nagyobb metszet per unió küszöb értékeken romlik a precízió. Ennek oka lehet, hogy a felhasznált MIDOG annotációk fix 50x50 befoglaló téglalapok, de a detekció lehet más méretű is, esetekben előfordulhat, hogy a detekció jobb, mint az annotáció.

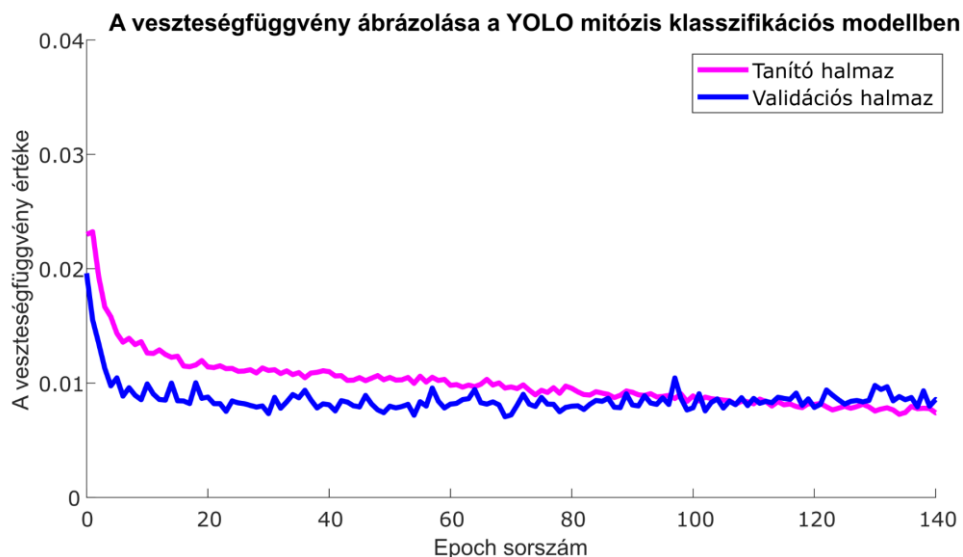


9. ábra: YOLO mitotikus sejt detekció példák a MIDOG adathalmaz képein

Nyíllal jelöltük az eredeti annotációt, négyzetekkel a detektált befoglaló téglalapokat. A zöld szín mitotikus sejtre, a sárga nem mitotikus sejtre utal. A-E) helyes predikciók F-G) hamis pozitív predikciók H) helyes predikció (fent), hamis negatív predikció (lent) I) hamis negatív (balra) és hamis pozitív (jobbra) predikció

4.4 YOLO mitotikus fázis klasszifikáció

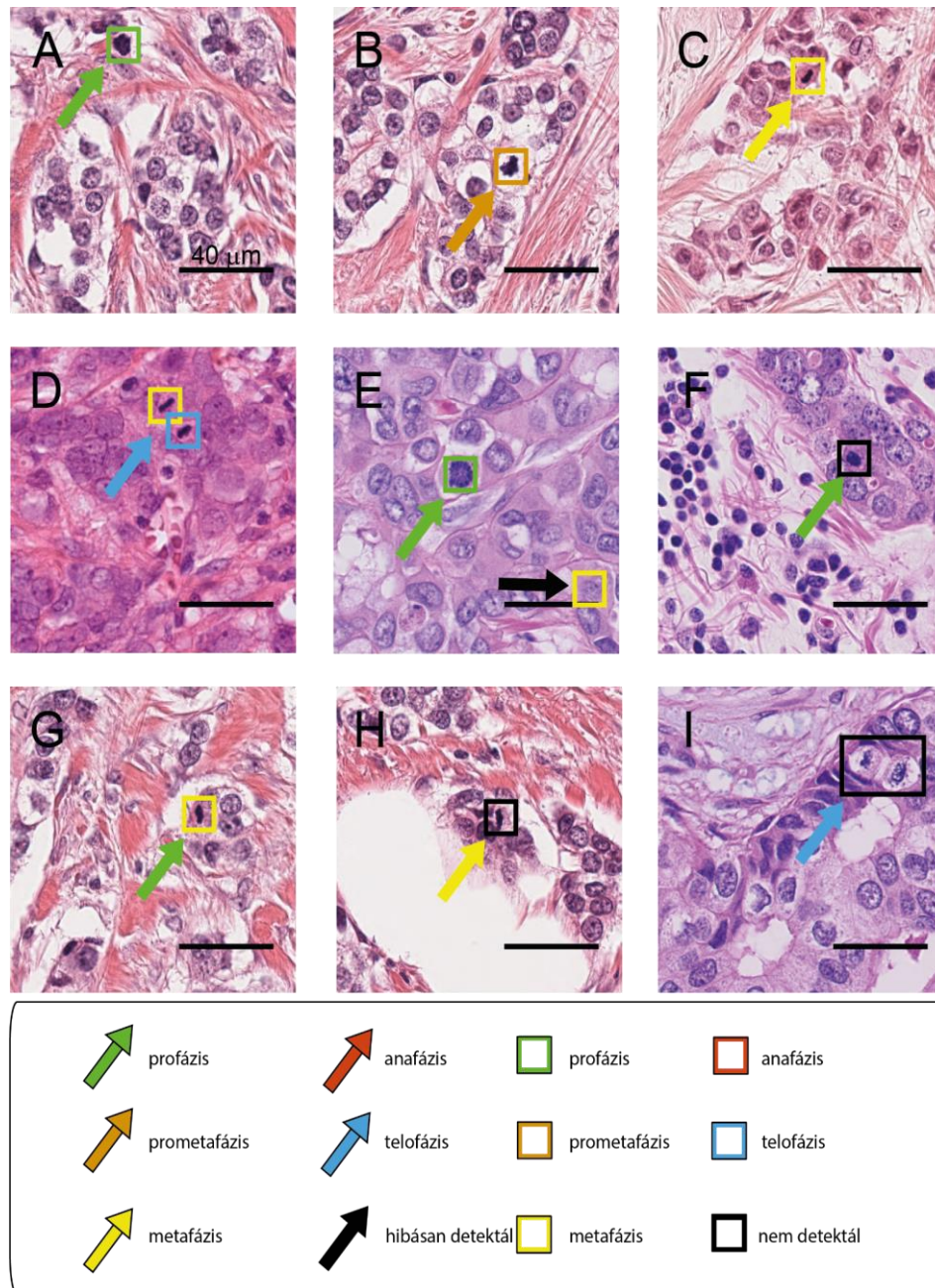
A YOLO módszer tehát hatékonyan tudja detektálni a MIDOG adatbázis szöveti metszetein a mitotikus sejteket. Következő lépésként arra voltunk kíváncsiak, hogy a modell milyen eredménnyel alkalmazható a mitózis hagyományos fázisainak klasszifikációjára. Ennek bemenetül egy olyan tanító halmazt készítettünk az annotációs stratégiánk (4. ábra) segítségével, amelyben a MIDOG adatbázisban annotált sejteket a mitózis hagyományos fázisaira osztottunk fel, valamint annotáltunk interfázisos sejteket is. Ezen kívül kialakítottunk egy olyan osztályt, amibe az atípusos mitotikus sejteket helyeztük, amelyeket nem használtunk fel a modellépítéshez. A veszteségfüggvények értékei a klasszifikációs YOLO modell esetében is hasonló konvergenciát mutattak (10. ábra).



10. ábra: A tanító és a validációs halmaz veszteségfüggvényeinek ábrázolása a YOLO klasszifikációs modellben

Magenta színnel jelöltük a tanító halmazon, kék színnel a validációs halmazon mért veszteségfüggvény értéket

Ez a modell 0.5 IoU (metszet/unió) küszöbértéken (Jaccard-index) 0.235 mAP (mean average precision) értéket ad az összes osztályt együtt tekintve. Ez nem tekinthető különösebben jó klasszifikációs modellnek, hiszen bár az esetek egy részében helyes predikciót ad a mitotikus sejtek osztályára, több esetben elvéti az osztály, vagy pedig helytelenül detektálja a sejteket (11. ábra), így bár ígéretes a módszer, a konkrét modell még optimalizálásra szorul (bővebben ld. Diskusszió A YOLO detekciós modell optimalizálása fejezet).



11. ábra: YOLO mitotikus sejt klasszifikáció példák a MIDOG adathalmaz képein
 Nyíllal jelöltük az eredeti annotációt, négyzetekkel a detektált befoglaló téglalapokat. A zöld szín profázisos, a narancssárga prometafázisos, a citromsárga metafázisos, a piros anafázisos, a kék telofázisos mitotikus sejtre, a fekete pedig nem mitotikus sejtekre, a befoglaló téglalapok esetében pedig nem detektált sejtre utal. A-C) helyes predikciók D) hibás befoglaló téglalap E) egy helyes (fent) és egy helytelen (lent) predikció F-I) hibás predikciók

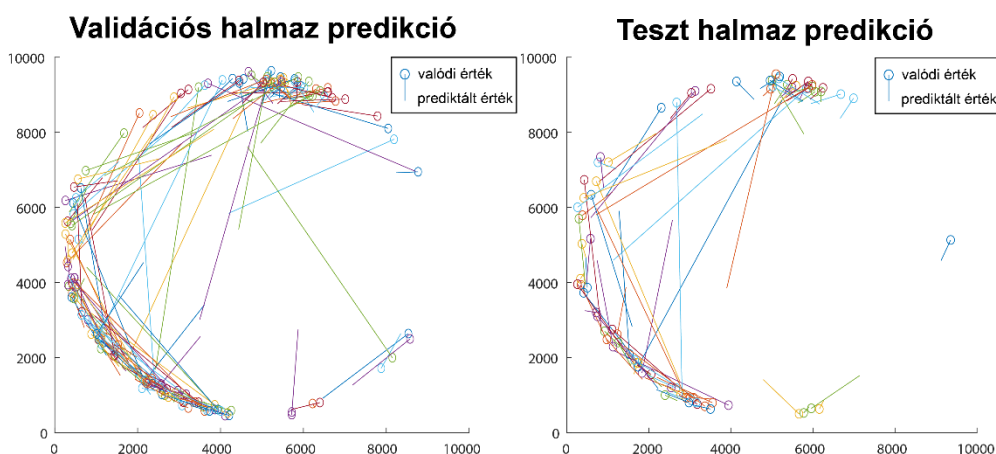
4.5 Mitotikus sejtek fázisának folytonos skálán való predikciója regressziós sík modell segítségével

A mitotikus sejt detektáló és fázisspecifikus lézer mikrodisszekciós munkafolyamatunk következő lépése az volt, hogy a klasszifikáció diszkrét osztályai helyett egy folytonos skálán, regresszióval állapítsuk meg, hogy a vizsgálni sejtek hol tartanak a mitózis folyamatában,

hiszen utóbbi egy folytonosan változó folyamat, így vizsgálata is pontosabb eredményt ígér ilyen skálán. A rendelkezésünkre álló MIDOG adathalmaz teljes tárgylemezt ábrázoló képeiből ennek érdekében csak mitotikus sejteket tartalmazó képkivágásokat készítettünk az eredeti annotáció alapján, és a korábban meghatározott stratégiánk segítségével elkészítettünk egy új, regressziós sík alapú annotációt az Advanced Cell Classifier program segítségével (4. ábra), melynek során a sejteket tartalmazó képkivágásokat elhelyeztük a mitózis folyamatát ábrázoló 4500 pixel sugarú kör lap kerületén.

Ez alapján elkészítettük a regressziós sík modellt (részletesen ld. Anyagok és módszerek fejezet), amely a validációs halmazon 2685 RMSE (root mean square error, négyzetes hiba) értéket adott (12/A. ábra). Ez azt jelenti, hogy a sejtek regressziós síkon prediktált koordinátái átlagosan ennyi pixel távolságra vannak az eredeti koordinátáktól. Ugyanez az érték a teszt halmazon 2589 RMSE (12/B. ábra). A regressziós sík méretét tekintve ezek viszonylag nagy értéknek számítanak, ám az ábrákat kiértékelve láthatjuk, hogy a sejtek egy részét helyesen az eredeti annotáció kis környezetébe, azaz a mitózis hagyományos fázisai szerinti felosztás alapján azonos fázisba prediktálja a modell, míg egy másik részét nagyobb távolságokra, akár az eredeti fázison kívülre is prediktálhatja. Az ábrán továbbá megfigyelhető az adathalmaz eloszlásának kiegyensúlyozatlansága (ld. még 5. ábra), különösképpen kiemelve a profázist és a metafázist, amely nagymértékben felülreprezentált a többi osztályhoz képest. Ez abból fakadhat, hogy a mitotikus folyamatok során ebben a két fázisban tölt el a legtöbb időt a sejt, az ana- és telofázis pedig egy viszonylag rövid folyamat.

MIDOG regressziós modell

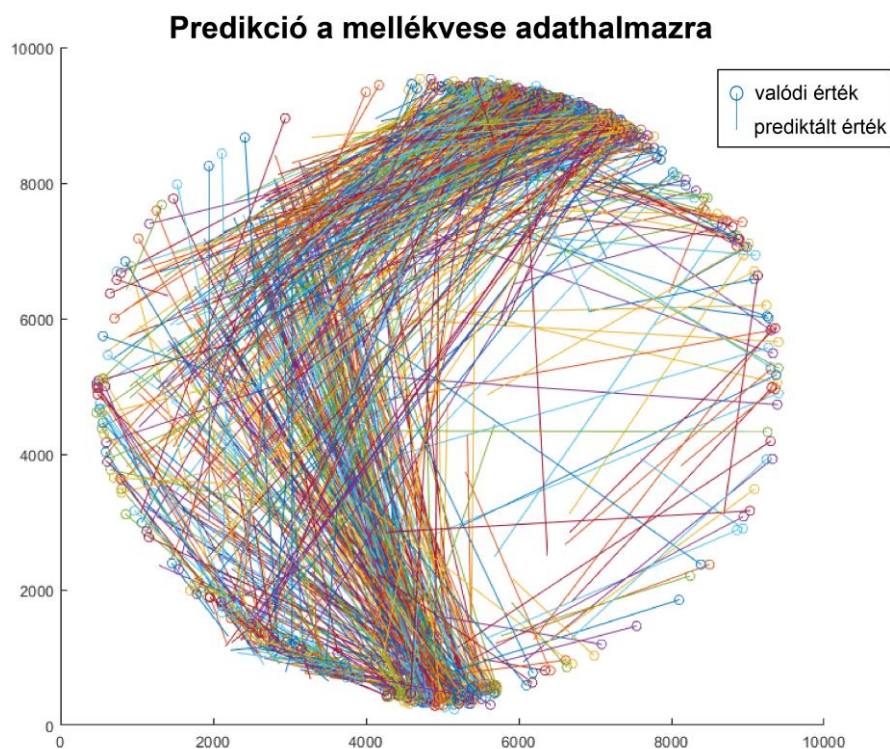


12. ábra: A MIDOG adathalmazon alapuló regressziós modell predikciója a MIDOG validációs és teszt halmazon

A 10.000x10.000 pixeles regressziós síkon a mitózis folyamatát kör alakban ábrázolva várjuk a predikciókat. A színes vonalak egy-egy sejt valós (körrel jelölve) és prediktált (szabdon végződő vonal) koordinátájának eltérését jelzik.

Ezután azt szeretnénk volna tesztelni, hogy a MIDOG adatok alapján készített modell mennyire képes másfajta szövetből származó mitotikus sejtek fázisának meghatározására, így a mellékvese adathalmazunkból elkülönítettük a mitotikus sejteket, és regressziós sík alapú annotációt készítettünk belőle (5. ábra), majd prediktáltuk a sejtek helyét a regressziós síkon a MIDOG RS modell segítségével.

Az eredményt 4386-os RMSE értékkel jellemezhetjük, ami meglehetősen magas. A predikciókat mutató ábra (13. ábra) is azt jelzi, hogy sok olyan sejt van, amit a valódi mitotikus fázisán kívülre prediktál a modell. Ez alapján arra a következtetésre jutottunk, hogy a modell változtatások nélkül nem használható általánosan más szövettípusokban. Ez következhet a képek felbontásából fakadó különbségekből, az eltérő szöveti felépítésből, valamint a különböző adathalmazok más elemszámából, és azok kiegyensúlyozatlanságából (5. ábra) egyaránt. Ebből fakadóan arra a következtetésre jutottunk, hogy érdemes továbbfejleszteni a modellt akár más adathalmazokkal való bővítéssel vagy stílus transzferrel (bővebben ld. Diskusszió Regressziós sík analízis fejlesztése és optimalizálása fejezet).



13. ábra: Az elkészített modell alkalmazása a mellékvese adathalmazon

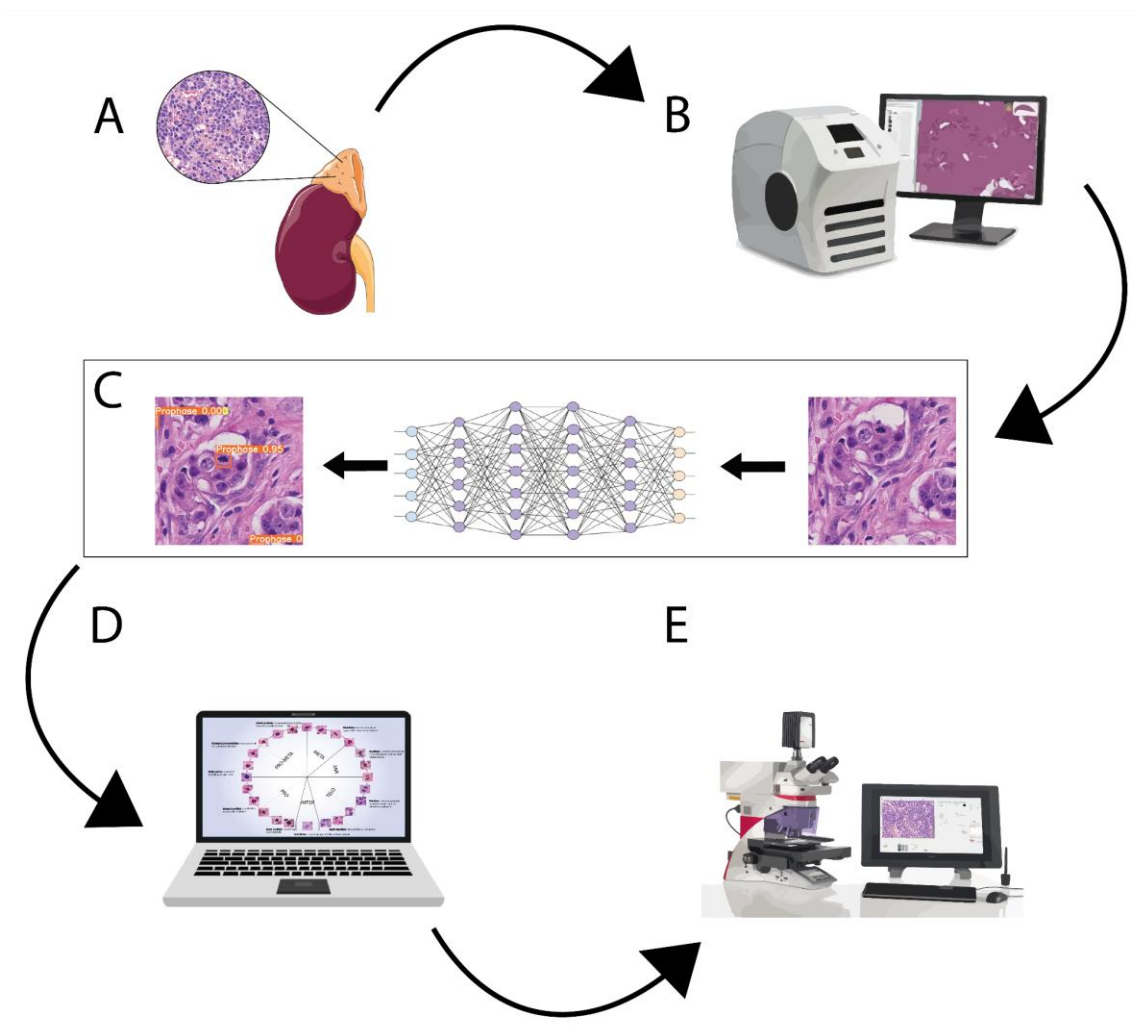
4.6 Összetett munkafolyamat létrehozása mitotikus sejtek fázis-specifikus lézer-mikrodisszekciója érdekében

Az előző kísérleteink alapján tehát képesek vagyunk hematoxilin-eozin festett metszetek teljes tárgylemezt ábrázoló képein detektálni a mitotikus sejteket YOLO módszer segítségével, majd azokat diszkrét, illetve folytonos osztályokba sorolni a YOLO és a regressziós sík módszerrel. A további megoldandó feladat tehát az a mitotikus sejtek fázis-specifikus gyűjtését célzó munkafolyamatban az volt, hogy képesek legyünk fizikailag is kivágni a választott sejteket egy adott mintából lézer-mikrodisszekció segítségével.

Ehhez a rutin patológiai protokoll következő lépéseit volt szükséges megváltoztatni (24):

- A mintákat hagyományos üveg tárgylemez helyett PEN (polietilén-naftalát) membránt tartalmazó üveg tárgylemezekre szükséges elhelyezni a mikrotómos metszés után, mivel a lézer- mikrodisszekció ezen működik optimálisan.
- A minták lefedése frissen kell, hogy történjen, és a mikroszkópos képalkotás után rögtön le kell oldani a fedőlemezt xilol segítségével, hiszen az a kivágást akadályozná, és minél tovább van rajta, annál jobban megszilárdul a fedőanyag.
- A mintákat szárazon, hűtőben szükséges tárolni.

Ezek után a változtatások után teszteltük a sejt kivágás hatékonyságát szöveti mintákon, de azt tapasztaltuk, hogy a festés során használt oldószerek valamelyike bejuthat a membrán és a tárgylemez közé, ami megakadályozza a lézer munkáját. Ezt úgy küszöböltük ki, hogy pengével elválasztottuk a membránt a tárgylemeztől, és kiszárítottuk azt. Egy másik probléma az volt, hogy vélhetően a fedőanyag maradványa olyan réteget képzett a szövet felületén, amely megakadályozta a kivágást. Ennek megoldására kétórás xilolos áztatást majd szárítást alkalmaztunk. Ezekkel a módosításokkal elértük, hogy a képfeldolgozási lépések után képesek legyünk a lézerrel történő egysejt-kivágásra formalin-fixált hematoxilin-eozin-festett szövetmintákon. Így összességében a világon először megalkottunk egy olyan munkafolyamatot (14. ábra), amellyel sikeresen tudunk mitotikus sejteket detektálni és a mitózis folyamatának morfológiai változásai alapján automatikusan, tetszőleges felbontással csoportokat elkülöníteni, majd lézer-mikrodisszekció segítségével kivágni a szövetből. Ez képessé tesz minket további kísérletek elvégzésére a kivágott mintákon, úgy, mint a mitózis fázisaiban a sejtek proteomjának és transzkriptomjának a meghatározása, melynek segítségével további ismeretre tudunk szert tenni a tumoros sejtek osztódásával kapcsolatban.



**14. ábra: Mitotikus fázis alapú egysejt-kivágás munkafolyamat vázlatos bemutatása
hematoxilin-eozin festett formalin-fixált paraffinba ágyazott metszeteken**

A munkafolyamat a következő lépések alkalmazásával valósul meg: A) Formalin-fixált, paraffinba ágyazott hematoxilin-eozin festett szöveti metszet előkészítése B) Mikroszkópos képalkotás a teljes tárgylemezről C) Mitotikus sejtek detekciója YOLO módszer segítségével D) Regressziós sík alapú mitotikus fázis predikció E) A kiválasztott sejtcsoportok elkülönítése további vizsgálatok céljából lézer-mikrodisszekcióval

5. Diszkusszió

Kutatócsoportunkban azt a kérdést szeretnénk megválaszolni hosszú távon, hogy a mitózis alatt milyen fehérjéknek és mRNSeknek a szintje változik szignifikánsan, és ezeknek a változásoknak milyen a dinamikája. Ezzel szeretnénk közelebb kerülni a sejtosztódás, ebből kiindulva pedig a tumoros sejtek osztódásának mélyebb megértéséhez. Csoportunk korábbi publikációi között megtalálható fluoreszcens fehérjékkel módosított sejtvonalak mikroszkópos képeinek elemzése, és a képek alapján a kiválasztott tulajdonságokkal rendelkező sejtek lézer mikrodisszekciós elkülönítése további vizsgálatok céljából (14). A sejtvonalak vizsgálatának egyik nagy előnye, hogy rutinszerűen vizsgálhatóak bennük dinamikus változások, például a sejtosztódás, hiszen akár képalkotás közben is sokáig fenntarthatók megfelelő körülmények között. Ha viszont fixált szövetet szeretnénk vizsgálni, akkor ez a lehetőség általában nem áll fenn, más megoldást kell keresnünk ahhoz, hogy csupán egyetlen kép alapján is meg tudjuk mondani, hogy az adott sejt éppen milyen mitotikus fázisban tart. Habár a sejtciklust vagy a mitózist lehetséges immunhisztokémiai módszerekkel viszonylag specifikusan jelölni (pl. Ki67 vagy PHH3 jelölés) (1, 2), a mitózis egyes fázisainak azonosításhoz jelenleg nem áll rendelkezésre immunfestési lehetőség. Így a következő járhatónak tűnő út, hogy a rutin patológiai festésként alkalmazott hematoxin-eozin (HE) jelölést alkalmazva a mitózis morfológiai jegyeire alapozva képfeldolgozási módszerekkel állapítsuk meg, hogy a vizsgált sejtek hol tartanak a mitózis folyamatában. Erre van precedens az orvosi gyakorlatban is (8), bár még nincs olyan automatizált technika, amely általánosan kiszorította volna a patológusok manuális vizsgálatát. Első lépésként tehát arra vállalkoztunk, hogy a rutin HE festett metszeteken mélytanulási módszerekkel szegmentáljuk a sejteket, hogy aztán mély- vagy gépi tanulási modellek építésével megtaláljuk a mitotikus sejteket. Kísérleteinkhez egy ismert nagy osztódási potenciállal rendelkező tumor típusnak, a Hodgkin-límfómának egy mellékvese áttét szövetét választottuk, mivel azt vártuk, hogy ebben a szövetben viszonylag sok osztódó sejtet találunk jól azonosítható szöveti környezetben. Bár találtunk egy normál sejtekre jól működő modellt a BIAS programban rendelkezésre állók közül, amit megpróbáltunk optimalizálni, sajnos nem értük el a kívánt eredményt a mitotikus sejtek tekintetében. Bár a modell továbbfejlesztése lehetséges lenne további, nagy mennyiségű mitotikus sejtet tartalmazó mintával, mégsem emellett az út mellett döntöttünk, mivel nem találtunk nagyszámú, ellenőrzött annotációval ellátott mitotikus sejtet tartalmazó, nyíltan hozzáférhető adathalmazt az adott szövettípusban, egy ilyen elkészítése a jelenlegi eszközökkel pedig hosszadalmas folyamat lenne.

Ezért a szegmentáció alapú mitózis analízis helyett a detekció alapú módszerekre fordítottuk a figyelmünket, melynek előnye, hogy jellegükből adódóan nem szükséges a kép minden objektumát azonosítani a tanító halmazban, hanem elég akár egy kívánt detektálandó osztályra fókuszálni. Ezek közül a YOLO v5 módszerre esett a választásunk, mivel ez a típusú detektor képes az objektumok nagyfokú általánosítására, valamint a mitotikus sejtek detektálása után az egyes osztályok befoglaló téglalapját kapjuk vissza a hozzájuk tartozó valószínűségi értékekkel.

5.1 A YOLO detekciós modell optimalizálása a MIDOG adathalmazon

Ennek a feladatnak a megoldásához egy olyan adathalmazt kerestünk, ami rendelkezik hiteles mitotikus sejt annotációval teljes tárgylemezről készült tumoros szöveteken. Az irodalomkutatás során azt tapasztaltunk, hogy kevés ilyen adathalmaz áll rendelkezésre (15, 25, 26), és azok is főleg emlő tumorból származnak, így ezek közül a legfrissebbet és a legnagyobb elemszámút, a MIDOG21 képgyűjteményt választottuk. Ennek kapcsán alapos irodalmi utánajárással sikerült meghatároznunk egy általános annotációs stratégiát, amely a legtöbb szövettípusra alkalmazható, a jövőben bármikor felhasználható. Másodszor érdemes megemlíteni a MIDOG adathalmaz felhasználásával készített detekciós algoritmust (9. ábra), amely meglepően jó eredményeket hozott a mitotikus sejtek detekciójára. Azonban a jövőbeni alkalmazhatósága érdekében érdemes törekedni a különböző halmazok (tanító, validációs, teszt) fázisok szerinti kiegyensúlyozására. A jelenlegi eredményeinknél megfigyelhető volt, hogy bár konvergált a validációs halmazon mért veszteségfüggvény (8. ábra), láthattunk egy enyhe növekedést, amely abból fakadhat, hogy több tanító képpel dolgoztunk a validációs adathalmazhoz képest.

Mivel a YOLO módszer nem csak egy osztály, hanem akár több detekciójára is alkalmazható, tulajdonképpen klasszifikációs probléma is megoldható vele, így további kísérleteinkben teszteltük az alkalmazhatóságát a sejtek mitotikus fázisai szerinti klasszifikációjára (11. ábra). Ezzel kapcsolatban azt tapasztaltuk, hogy a modell közepesnek mondható eredménnyel prediktálja a teszthalmaz képeinek fázisait, így érdemes figyelmet fordítani az osztályozás pontosabbá tételére, külön odafigyelve arra, hogy minden mitotikus osztályból elegendő elemszámú kép kerüljön be mind a három halmazba (tanító, validáló, teszt), hiszen az eddigi algoritmus randomizált módon válogatta be a képeket. Ezzel elérhető, hogy a tanítási és a validációs görbe megfelelően kövessék egymást. Valamint érdemes a jövőben a határoló téglalapok méretein is változtatni, flexibilissé tenni őket, hiszen a MIDOG adathalmaz fix határoló téglalapokkal dolgozik, amely rontotta az eredményeket azzal, hogy a pl. telofázisos sejteket a legtöbb esetben két különálló osztályban elhelyezkedő sejtnak detektálta.

5.2 Regressziós sík analízis fejlesztése és optimalizálása

A mitózis folyamatának folytonos skálája miatt szerettünk volna túllépni a klasszifikáció diszkrét osztályain. A csoportunk által fejlesztett regressziós síkkal képesek voltunk a mitózist egy finom átmenetektől álló, folytonos skálán elemezni, és elkészíteni az első regressziós sík alapú szöveti mitotikus annotációt (5. ábra), valamint a MIDOG adathalmazra támaszkodó regressziós sík modellt (12. ábra). Ez a jövőben optimalizálható a jobb RMSE érték elérése érdekében elsősorban a tanító adat növelésével, amelyhez további validált mellrák adatok állnak rendelkezésre akár a régebbi TUPAC16 (25) és MITOS2012 (ICPR) (26) kihívások képeit tartalmazó adathalmaz formájában, valamint a közeljövőben megrendezésre kerül a MIDOG 2022 kihívás, amely adatbázisa az elődjéhez hasonlóan nyílt hozzáférésű lesz, ezzel a regressziós síkon lévő adatok eloszlásának kiegyensúlyozatlansága is javítható lenne.

5.3 Hisztológiai minták proteomikai és transzkriptomikai analízisének előkészítése

A fent említett mélytanuló algoritmusok kidolgozása után lehetőségünk nyílt az egysejt szintű lézer-mikrodisszekciós kivágásra, amely jelentősen hozzájárul a jelenleg is zajló egysejt szintű analízisek fejlesztéséhez, mind proteomikai (27), mind transzkriptomikai (28) kísérletek révén, ehhez viszont optimalizálni kellett a szövetelőkészítési folyamat néhány lépését, amit sikeresen megtettünk, így kidolgoztunk egy komplett munkafolyamatot a szövetelőkészítéstől a mitotikus sejtek detekcióján és fázisának tetszőleges felbontású azonosításán keresztül a kivágásukig (14. ábra). Ezzel közelebb juthatunk a mitózis, valamint a tumoros folyamatok jobb megértéséhez, a proteomikai és transzkriptomikai változások feltérképezéséhez, a hibás sejt szintű folyamatok mélyebb elemzéséhez. Ezáltal akár új rákterápiás célpontok fedezhetők fel, az analízis áteresztőképességének növelésével pedig lehetőség nyílna a különböző tumor típusok, vagy akár egyedi tumorok vizsgálatára és a személyre szabottabb gyógyítás fejlesztésére.

6. Köszönetnyilvánítás

Elsősorban szeretném megköszönni **Dr. Horváth Péternek**, hogy bizalommal fordult felém, és lehetőséget, valamint helyet biztosított a TDK dolgozatom elkészítéséhez. Fontos megemlítenem Dr. Sükösd Farkas patológus professzort, aki rendelkezésünkre bocsájtotta a mellékvese tumor mintákat a kísérletek elvégzéséhez.

Továbbá köszönettel tartozom **Csapóné Dr. Miczán Viviennek**, hogy türelemmel és odaadással segítette munkámat, felkészített a különböző szakmai előadásokra és hasznos tanácsokkal látott el, amelyek a jövőben a hasznomra válhatnak.

Egyúttal szeretném megköszönni az új szegmentációs modell elkészítését Dr. Hollandi Rékának, a felhasznált detekciós algoritmus elkészítését Grexa Istvánnak, a regressziós sík modelljének elkészítését Dr. Koós Krisztiánnak. Ezen felül a BIOMAG csoport minden tagjának, külön kiemelve Beleon Attilát, Dr. Mígh Edét és Kaptás Flórát, a Deep Visual Proteomics csapatot, akik hozzájárultak a dolgozatom létrejöttéhez.

7. Tételes szerzői hozzájárulás

Alulírott Iván Zsanett Zsófia nyilatkozom, hogy a 36. OTDK-ra benyújtott **Mitotikus sejtek azonosítása humán daganatos szövetekben mélytanuló algoritmusok használatával** című dolgozatomban leírtak esetében az alábbi munkafolyamatok elvégzésében vettem részt:

Rész/feladat:	Felkészítés után önállóan végeztem a dolgozathoz szükséges mennyiségben (%-ban megadva)	Folyamatos segítséggel végeztem a dolgozathoz szükséges mennyiségben (%-ban megadva)
Képfeldolgozási munkafolyamat optimalizálása	100	0
Szegmentációs modell továbbfejlesztése ¹	50	50
Tanító adat elkészítése	100	0
Detekciós algoritmus alapú és regressziós sík alapú modellek elkészítése	0	100
Kiválasztott mellékvese minta előkészítése, festése és optimalizálása a vágáshoz ²	20	80
Ábrakészítés	90	10

Magyarázat:

¹ A kiválasztott modell továbbfejlesztéséhez szükséges manuális szegmentációt én végeztem, az új adathalmazon történő újratanítást Dr. Hollandi Réka.

² Az optimalizálásához szükséges xilolos áztatást és azt követően a minta kiszárítását én végeztem, a minta előkészítése és a festési lépések a Szegedi Tudományegyetem Szent-Györgyi Albert Klinikai Központ Patológiai Intézetében alkalmazott protokoll szerint valósultak meg.

Kelt: 2022. 05.30.

Pályamunkát benyújtó hallgató (Iván Zsanett Zsófia)
aláírása

A nyilatkozatban foglaltak helyességét megerősítem:

Témavezető (Dr. Horváth Péter, ELKH SZBK)
aláírása

Irodalomjegyzék

1. A. Ibrahim, A. Lashen, M. Toss, *et al.*, *J. Clin. Pathol.*, in press, doi:10.1136/jclinpath-2021-207742.
2. L. H. Lee, H. Yang, G. Bigras, Current breast cancer proliferative markers correlate variably based on decoupled duration of cell cycle phases. *Sci. Rep.* **4**, 1–8 (2014).
3. T. A. Donovan, F. M. Moore, C. A. Bertram, *et al.*, Mitotic Figures—Normal, Atypical, and Imposters: A Guide to Identification. *Vet. Pathol.* **58**, 243–257 (2021).
4. T. J. O’Leary, M. W. Steffes, Can you count on the mitotic index? *Hum. Pathol.* **27**, 147–151 (1996).
5. D. Mahdessian, A. J. Cesnik, C. Gnann, *et al.*, Spatiotemporal dissection of the cell cycle with single-cell proteogenomics. *Nature.* **590**, 649–654 (2021).
6. D. Gisselsson, Classification of chromosome segregation errors in cancer. *Chromosoma.* **117**, 511–519 (2008).
7. R. Nateghi, H. Danyali, M. S. Helfroush, A deep learning approach for mitosis detection: Application in tumor proliferation prediction from whole slide images. *Artif. Intell. Med.* **114**, 102048 (2021).
8. J. van der Laak, G. Litjens, F. Ciompi, Deep learning in histopathology: the path to the clinic. *Nat. Med.* **27**, 775–784 (2021).
9. X. Pan, Y. Lu, R. Lan, *et al.*, Mitosis detection techniques in H&E stained breast cancer pathological images: A comprehensive review. *Comput. Electr. Eng.* **91**, 107038 (2021).
10. D. C. Cireşan, A. Giusti, L. M. Gambardella, *et al.*, Mitosis detection in breast cancer histology images with deep neural networks. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. **8150 LNCS**, 411–418 (2013).
11. B. Wu, T. Kausar, Q. Xiao, *et al.*, FF-CNN: An efficient deep neural network for mitosis detection in breast cancer histological images. *Commun. Comput. Inf. Sci.* **723**, 249–260 (2017).
12. C. Li, X. Wang, W. Liu, *et al.*, DeepMitosis: Mitosis detection via deep detection, verification and segmentation networks. *Med. Image Anal.* **45**, 121–133 (2018).
13. H. Wang, A. Cruz-Roa, A. Basavanahally, *et al.*, Cascaded ensemble of convolutional neural networks and handcrafted features for mitosis detection. *Med. Imaging 2014 Digit. Pathol.* **9041**, 90410B (2014).

14. A. Mund, F. Coscia, R. Hollandi, *et al.*, *bioRxiv*, in press (available at <https://www.biorxiv.org/content/10.1101/2021.01.25.427969v1%0Ahttps://www.biorxiv.org/content/10.1101/2021.01.25.427969v1.abstract>).
15. MIDOG Challenge 2021 – MItosis Domain Generalization Challenge (2021).
16. J. Redmon, S. Divvala, R. Girshick, *et al.*, You only look once: Unified, real-time object detection. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* **2016-Decem**, 779–788 (2016).
17. A. Szkalisity, F. Piccinini, A. Beleon, *et al.*, Regression plane concept for analysing continuous cellular processes with machine learning. *Nat. Commun.* **12** (2021), doi:10.1038/s41467-021-22866-x.
18. F. Piccinini, T. Balassa, A. Szkalisity, *et al.*, Advanced Cell Classifier: User-Friendly Machine-Learning-Based Software for Discovering Phenotypes in High-Content Imaging Data. *Cell Syst.* **4**, 651-655.e5 (2017).
19. C. A. Bertram, M. Aubreville, C. Marzahl, *et al.*, A large-scale dataset for mitotic figure assessment on whole slide images of canine cutaneous mast cell tumor. *Sci. Data.* **6**, 1–9 (2019).
20. M. K. Gandhi, J. T. Tellam, R. Khanna, Epstein-Barr virus-associated Hodgkin’s lymphoma. *Br. J. Haematol.* **125** (2004), pp. 267–281.
21. C. Szegedy, V. Vanhoucke, S. Ioffe, *et al.*, Rethinking the Inception Architecture for Computer Vision. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* **2016-Decem**, 2818–2826 (2016).
22. R. Hollandi, A. Szkalisity, T. Toth, *et al.*, nucleAIzer: A Parameter-free Deep Learning Framework for Nucleus Segmentation Using Image Style Transfer. *Cell Syst.* **10**, 453-458.e6 (2020).
23. R. Hollandi, Á. Diósdí, G. Hollandi, *et al.*, Annotator J: An image J plugin to ease hand annotation of cellular compartments. *Mol. Biol. Cell.* **31**, 2179–2186 (2020).
24. J. W. Foley, C. Zhu, P. Jolivet, *et al.*, Gene expression profiling of single cells from archival tissue with laser-capture microdissection and Smart-3SEQ, 1816–1825 (2019).
25. M. Veta, Y. J. Heng, N. Stathonikos, *et al.*, Predicting breast tumor proliferation from whole-slide images: The TUPAC16 challenge. *Med. Image Anal.* **54**, 111–121 (2019).
26. C. Li, C.-Q. Cao, Y.-F. Gao, Mitosis Detection in Breast Cancer Using Histological Images, 687–694 (2017).
27. F. Coscia, S. Doll, J. M. Bech, *et al.*, A streamlined mass spectrometry–based proteomics workflow for large-scale FFPE tissue analysis. *J. Pathol.* **251**, 100–112

(2020).

28. I. Wimmer, A. R. Tröschner, F. Brunner, *et al.*, Systematic evaluation of RNA quality, microarray data reliability and pathway analysis in fresh, fresh frozen and formalin-fixed paraffin-embedded tissue samples. *Sci. Rep.* **8**, 1–17 (2018).