
TEZĂ DE DOCTORAT - REZUMAT

Carcinomul bazocelular: studiu morfologic, imagistic și asistat de inteligența artificială

Doctorand **Raluca Maria Bungărdean**

Conducător de doctorat Prof.dr. **Maria Crișan**



UMF
UNIVERSITATEA DE
MEDICINĂ ȘI FARMACIE
IULIU HAȚIEGANU
CLUJ-NAPOCA

CUPRINS

| | |
|---|----|
| INTRODUCERE | 15 |
| STADIUL ACTUAL AL CUNOAȘTERII | |
| 1. Carcinomul bazocelular generalități | 19 |
| 1.1. Definiție și incidență | 19 |
| 1.2. Factori de risc | 19 |
| 1.3. Diagnostic clinic și paraclinic | 21 |
| 1.4. Tratament | 22 |
| 2. Diagnosticul histopatologic al CBC | 25 |
| 2.1. Aspect caracteristic și clasificare pe subtipuri | 25 |
| 2.2. Imunohistochimie și diagnostic diferențial | 28 |
| 2.2.1. Imunohistochimie | 28 |
| 2.2.2. Diagnostic diferențial | 30 |
| 3. Ultrasonografia de înaltă frecvență | 31 |
| 3.1. Istoric și definiție | 31 |
| 3.2. HFUS în diagnosticul CBC | 31 |
| 4. Inteligența artificială în diagnosticul anatomopatologic al CBC | 35 |
| CONTRIBUȚIA PERSONALĂ | |
| 1. Ipoteza de lucru/obiective | 41 |
| 2. Metodologie generală | 43 |
| 2.1. Colectarea datelor | 43 |
| 2.2. Avizul comisiilor de etică | 43 |
| 2.3. Prelucrarea histopatologică | 44 |
| 2.4. Realizarea lamelor digitale prin scanare | 45 |
| 3. Studiul 1 – Aspecte morfologice ale vascularizației carcinomului bazocelular | 47 |
| 3.1. Introducere | 47 |
| 3.2. Ipoteza de lucru | 48 |
| 3.3. Proveniența și morfologia CBC | 48 |
| 3.4. Caracteristicile morfologice ale vascularizației CBC | 49 |
| 3.5. Ultrasonografia și histologia | 52 |
| 3.6. Utilitatea HFUS în dermatologie | 52 |
| 3.7. HFUS a CBC: morfologie și vascularizație | 53 |
| 3.8. Importanța Doppler-ului în diagnosticul CBC | 55 |
| 3.9. Avantajele și dezavantajele HFUS în evaluarea CBC | 56 |
| 3.10. Alte abordări în caracterizarea vascularizației CBC | 57 |
| 3.11. Concluzii | 57 |
| 4. Studiul 2 – Ecografia de înaltă frecvență: o investigație neinvazivă esențială pentru evaluarea preterapeutică a carcinomului bazocelular | 59 |
| 4.1. Introducere | 59 |
| 4.2. Ipoteza de lucru | 60 |
| 4.3. Material și metodă | 61 |
| 4.4. Rezultate | 62 |

| | |
|---|------------|
| 4.5. Discuții | 67 |
| 4.6. Concluzii | 68 |
| 5. Studiul 3 – Deep learning cu transfer learning în anatomie patologică. | 71 |
| Studiu de caz: clasificarea carcinomului bazocelular | 71 |
| 5.1. Introducere | 71 |
| 5.2. Ipoteza de lucru | 75 |
| 5.3. Material și metodă | 76 |
| 5.3.1. Material | 76 |
| 5.3.2. Metodă | 78 |
| 5.3.3. Transferul, formarea și evaluarea rețelelor | 79 |
| 5.3.4. Aplicația propriu-zisă | 80 |
| 5.3.5. Evaluarea clinică | 80 |
| 5.3.6. Evaluarea statistică | 80 |
| 5.4. Rezultate | 81 |
| 5.5. Discuții | 86 |
| 5.6. Concluzii | 91 |
| 6. Studiul 4 – Subtipurile de carcinom bazocelular nodular și micronodular sunt tumori diferite în funcție de arhitectura lor morfologică și de interacțiunea lor cu stroma înconjurătoare | 93 |
| 6.1. Introducere | 93 |
| 6.1.1. Stadiul actual al cunoașterii | 93 |
| 6.1.2. Insulele tumorale | 95 |
| 6.1.3. Retracția stromală | 96 |
| 6.1.4. Stroma tumorală | 97 |
| 6.2. Ipoteza de lucru | 98 |
| 6.3. Material și metodă | 98 |
| 6.3.1. Material | 98 |
| 6.3.2. Metode | 100 |
| 6.3.2.1. Pregătirea setului de date | 100 |
| 6.3.2.2. Analiza morfometrică | 103 |
| 6.3.2.3. Analiza segmentării semantice | 103 |
| 6.3.2.4. Segmentarea automată a WSI | 105 |
| 6.4. Rezultate | 105 |
| 6.5. Discuții | 108 |
| 6.5.1. Observații generale | 108 |
| 6.5.2. Insulele tumorale (inclusiv palisadarea periferică) | 109 |
| 6.5.3. Retracția stromală | 111 |
| 6.5.4. Stroma tumorală | 112 |
| 6.5.5. Cercetări de viitor | 113 |
| 6.5.6. Scurt rezumat al analizei rezultatelor | 113 |
| 6.6. Concluzii | 114 |
| 7. Concluzii generale | 115 |
| 8. Originalitatea și contribuțiile inovative ale tezei | 117 |
| REFERINȚE | 119 |

Cuvinte cheie: carcinom bazocelular, ultrasonografie de înaltă frecvență, morfologie, histopatologie, patologie digitală, machine learning, deep learning, transfer learning, rețele neuronale convoluționale.

INTRODUCERE

Cancerul cutanat reprezintă o problemă complexă și de actualitate de sănătate publică. Dintre toate cancerurile cutanate, cel mai răspândit la nivel global este carcinomul bazocelular (CBC). Diagnosticul clinic dermatologic al acestui cancer poate fi îmbunătățit utilizând noi tehnici imagistice precum ultrasonografia cutanată de înaltă frecvență. Informațiile oferite de această investigație au impact asupra tratamentului. Actualmente, standardul de aur în diagnosticul CBC este examenul histopatologic. O dată cu apariția patologiei digitale, imaginile histopatologice pot fi analizate utilizând metode computaționale.

Cercetarea doctorală explorează morfologia CBC și rafinarea diagnosticului acestui cancer prin convergența patologiei digitale cu inteligența artificială (IA). Aceasta valorifică tehnici avansate de *deep learning* (DL) și de *transfer learning* (TL) pentru a analiza imaginile histologice, oferind o abordare inovatoare atât pentru diagnosticul anatomopatologic, cât și pentru cercetare.

STADIUL ACTUAL AL CUNOAȘTERII

Carcinomul bazocelular are o incidență în creștere la nivel mondial, fiind diagnosticat la persoane din ce în ce mai tinere. Cu toate că metastazarea acestei malignități este redusă, apariția ei preponderent la nivelul capului și gâtului, caracterul distructiv local și faptul că este cea mai frecventă malignitate a tegumentului, reprezintă provocări importante în patologia oncologică cutanată.

Celula de origine a CBC rezidă în celulele bazale ale foliculului pilos. Apariția CBC proponderent pe zonele fotoexpuse ale pielii denotă legătura stransă dintre această malignitate și radiațiile ultraviolete (UV), în special cele de tip B. Pe lângă examenul clinic și dermatoscopic, diagnosticul clinic al CBC poate fi îmbunătățit cu diverse tehnici de imagistică medicală, precum ultrasonografia de înaltă frecvență, microscopia confocală sau tomografia în coerență optică. Cu toate că examenul histopatologic își menține statutul de standard de aur în diagnosticul acestui cancer, progresele în tehnicile de imagistică prezintă corelații satisfăcătoare din punct de vedere statistic cu histopatologia CBC, îmbunătățind diagnosticul clinic și extinzând metodele de tratament dincolo de intervențiile chirurgicale.

Din punct de vedere clinic, având la baza riscul de recurență și prognosticul, CBC poate fi clasificat în două mari categorii: ușor de tratat sau cu risc redus și dificil de tratat sau cu risc ridicat. Această clasificare facilitează planificarea tratamentului. Ecografia cutanată oferă informații privind dimensiunea, aspectul, invazia și subtipul histologic de CBC, prezentând o bună corelație cu aspectul microscopic al acestei malignități. Modurile de sonografie în scară gri, Doppler, elastografie și cu substanță de contrast contribuie la evaluarea completă a CBC. Doppler evidențiază în mod eficient vascularizația, în timp ce elastografia și ultrasonografia cu substanțe de contrast ajută la clasificarea leziunilor ca fiind maligne.

Din punct de vedere histopatologic, CBC este reprezentat de insule de celule bazale cu palisadare periferică înconjurată de o stromă fibro-mixoidă, cu retracții stromale peritumorale. Clasificat histologic în 10 subtipuri, CBC este grupat în funcție de riscul de recidivă în categorii cu risc ridicat și cu risc scăzut, similare grupelor clinice. În cadrul unei tumori pot coexista mai multe subtipuri.

Integrarea IA în anatomia patologică este facilitată de patologia digitală care presupune obținerea de imagini digitale de înaltă rezoluție ale țesuturilor prin scanarea lamelor histopatologice. Aceste imagini pot fi ulterior interpretate prin intermediul unor programe de analiză a imaginilor. IA valorifică algoritmi de recunoaștere a tiparelor, simulând gândirea umană. În domeniul învățării automate cu scopul diagnosticului CBC, se pot utiliza două abordări distincte: învățarea supravegheată care presupune antrenarea algoritmilor cu seturi de date etichetate sau învățarea nesupravegheată care identifică tipare fără date etichetate. DL este o tehnică de învățare supravegheată care folosește rețelele neuronale (RN) artificiale pentru a analiza imagini prin identificarea de tipare. Utilizarea de straturi neuronale multiple determină o învățare complexă capabilă să identifice caracteristici detaliate. În analiza imaginilor histologice, cele mai comune tehnici de IA sunt RN convoluționale, DL și TL. RN convoluționale sunt eficiente în sarcinile de analiză a imaginilor. Alegerea metodei de IA în evaluarea imaginilor histologice variază în funcție de complexitatea sarcinii, datele disponibile, resursele de calcul și expertiza echipei de dezvoltare. Învățarea prin transfer, care utilizează modelele preinstruite pentru sarcini specifice și modelele hibride care combină algoritmi tradiționali cu DL, sunt abordări comune. Aplicațiile IA în diagnosticarea CBC implică DL și RN convoluționale pentru a identifica modelele specifice tumorii, comparând imaginile cu o bază mare de date pentru o acuratețe sporită. Necesitatea unor seturi de date mari și diverse precum și a unor imagini de înaltă calitate reprezintă principalele provocări ale utilizării IA în anatomia patologică. Instrumentele bazate pe IA în diagnosticul histologic sunt adjuvanți ai medicilor oferind sprijin în diagnostic sau în învățare.

CONTRIBUȚIA PERSONALĂ

Obiectivul principal al cercetării doctorale a constat într-o caracterizare complexă și modernă a celei mai răspândite malignități cutanate, carcinomul bazocelular. Structura cercetării cuprinde o revizuire a literaturii și trei studii de cercetare aplicată, fiecare abordând diverse aspecte ale morfologiei CBC, după cum urmează:

1. **Vascularizația CBC:** este abordată printr-o recenzie descriptivă în primul studiu al tezei de doctorat, care analizează rolul acesteia în dezvoltarea, creșterea și extinderea CBC, explicând comportamente specifice ale acestei tumori cum sunt potențialul redus de metastazare și creșterea lentă.
2. **Corelația dintre ultrasonografie și histologie:** este apreciată printr-un studiu retrospectiv (studiul 2) care evaluează profunzimea invaziei CBC determinată ecografic și histologic. Acest studiu demonstrează o corelație puternică între cele două măsurători și explică importanța acesteia în ceea ce privește abordul terapeutic al acestui cancer.
3. **Clasificarea histologică asistată de IA:** Studiul 3 a implicat dezvoltarea, implementarea și validarea unei aplicații de inteligență artificială conceput pentru a clasifica imaginile microscopice ale CBC în diferite subtipuri și variante. Această aplicație utilizează cele mai recente tehnici de clasificare a imaginilor (DL cu TL al RN convoluționale) și îmbunătățește consensul în clasificarea CBC.
4. **Explorarea nuanțelor morfologice a CBC utilizând IA:** Studiul 4 s-a axat pe utilizarea IA cu scopul de a dezvălui variații morfologice subtile în cadrul a două subtipuri de CBC care fac parte din clase diferite de risc: nodular și micronodular. Această abordare aduce informații noi referitoare la asemănări (palisadarea periferică) și diferențe (retracția stromală, celulele insulelor tumorale) ale morfologiei acestor subtipuri, relevând aspecte nedetectabile prin examinarea uzuală histopatologică.

Prin integrarea acestor elemente, s-a realizat o investigație cuprinzătoare a morfologiei CBC utilizând tehnici moderne.

Studiu 1. Aspecte morfologice ale vascularizației carcinomului bazocelular

Primul studiu al cercetării doctorale este o recenzie (*review*) descriptivă a literaturii de specialitate în ceea ce privește morfologia vascularizației CBC și implicațiile acesteia în invazivitate, creștere și metastazare.

Scopul acestui studiu este explicarea unor caracteristici comportamentale (rata scăzută de metastazare, creșterea lentă și agresivitatea locală) ale CBC prin modelul vascularizației. De asemenea se realizează o analiză a utilității practice a cunoașterii distribuției vaselor de sânge în ceea ce privește măsurători imagistice, diferențierea preterapeutică a subtipurilor cu risc redus de cele cu risc ridicat și despre impactul asupra abordului chirurgical.

Examinarea vascularizației tumorii se structurează în jurul a două linii principale de investigație: metode clinice de investigație (arhitectura, dispoziția și lungimea vaselor cuantificată prin ecografie Doppler, dermatoscopie și capilaroscopie) și metode paraclinice histologice de investigație (densitatea microvaselor și proliferarea celulelor endoteliale studiate prin colorații imunohistochimice).

Agresivitatea locală a CBC este atribuită vascularizației sale intense, caracterizată de o arhitectură haotică, tortuoasă și arborizantă observată cu precădere la granița între tumoră și stromă, la frontul de invazie al insulelor de carcinom. La periferia tumorii, densitatea microvaselor este ridicată, iar proliferarea celulelor endoteliale este maximă. Densitatea microvaselor variază în funcție de subtipul de CBC, fiind mai abundentă în subtipurile cu risc ridicat. Această constatare este susținută prin studii ale densității microvaselor și prin investigații precum capilaroscopia și ecografia Doppler. O angiogeneză redusă la periferia tumorii, explică cazurile de CBC de risc redus, cu creștere indolentă. O densitate vasculară redusă sau absentă în mijlocul tumorii poate cauza ulcerările centrale frecvent observate macroscopic în acest cancer.

Rata scăzută de metastazare a CBC este legată de regresia sau absența structurilor vasculare în cadrul insulelor de celule bazale, ipoteză susținută de studiile imunohistochimice și de studii comparative cu alte tumori cutanate precum carcinomul scuamos sau melanomul, care prezintă o rată de metastazare mai crescută.

Ecografia Doppler și cu agenți de contrast evidențiază un flux sangvin proeminent în porțiunea inferioară a CBC, cu unul sau doi pediculi vasculari. Evaluarea Doppler a vascularizației CBC poate ghida procedurile terapeutice prin îmbunătățirea măsurătorilor atunci când ea este dificil de realizat în scala gri sau ajută la estimarea pierderilor de sânge în timpul intervenției. De asemenea poate fi de folos în încadrarea clinică a unui CBC în grupe ușor sau dificil de tratat, iar astfel are importanță în planificarea tratamentului optim. Dermatoscopia evidențiază modele caracteristice ale vaselor de sânge superficiale și este utilă în diagnosticul clinic al CBC. Microscopia confocală evidențiază vase superficiale, orizontale, neregulate, care unori se distribuie paralel.

Analiza vascularizației CBC oferă informații despre evoluția și progresia acestui cancer și poate îmbunătăți strategia terapeutică.

Studiu 2. Ecografia de înaltă frecvență: o investigație neinvazivă esențială pentru evaluarea preterapeutică a carcinomului bazocelular

Ultrasonografia de înaltă frecvență este o investigație utilă în evaluarea leziunilor cutanate fiind de folos în diagnosticul clinic, în evaluarea tratamentului și în identificarea

recidivelor. HFUS pe scară gri oferă date semnificative privind morfologia CBC (structură, contur, dimensiune, relația cu structurile adiacente). Examinarea pe scară gri se poate completa cu sonografie multimodală (Doppler, elastografie). O măsurătoare importantă în planificarea tratamentului CBC este indexul sonografic al profunzimii tumorii. Terapiile nechirurgicale pot fi luate în considerare în cazul CBC cu risc redus, iar CBC cu o adâncime mai mică de 2 mm pot fi tratate folosind terapia fotodinamică. Având în vedere predilecția acestei tumori în regiunea capului și a gâtului un tratament optim ia în considerare atât o cura oncologică cât și un rezultat estetic favorabil cu o pierdere minimă de țesut sănătos.

Scopul acestui studiu constă în evaluarea corelației între sonografie și histopatologie în ceea ce privește adâncimea CBC.

Studiul este retrospectiv și include 90 de cazuri de CBC diagnosticat la Spitalul Clinic Județean de Urgență Cluj. Toate examinările sonografice au fost realizate cu o sondă de înaltă rezoluție, de 18 MHz. Din buletinul ecografic au fost extrase informații despre aspectul, dimensiunile, vascularizația (Doppler) și rigiditatea (elastografie) tumorii, cel mai relevant pentru studiu fiind indicii de adâncime măsurat în milimetri. Lamele histopatologice au fost reevaluate și, utilizând programul Leica LAS EZ, s-au măsurat în milimetri indicii de adâncime histologică. Studiul a cuprins diferite subtipuri și variante de CBC iar pentru analiza statistică s-a utilizat coeficientul de corelație Pearson și testul Student t.

Dintre cele 90 de cazuri incluse în studiu, majoritatea au fost subtipuri cu risc redus iar subtipul nodular cu variantele sale a fost cel mai frecvent subtip histologic. Majoritatea leziunilor au avut doi pediculi vasculari. Toate tumorile au avut aport sangvin crescut în porțiunea inferioară și rigiditate crescută. Rezultatele statistice arată o corelație puternică (Pearson's $r=0,92$) între măsurătorile ultrasonografice și histologice cuantificate prin indicii de adâncime, demonstrând o relație predictivă între cele două abordări de investigare.

Importanța cercetării constă în implicațiile clinice directe pentru managementul pacienților prin integrarea ecografiei în diagnosticul clinic dermatologic, datorită înaltei corelații cu histopatologia în ceea ce privește cel mai important indice în stadializarea CBC, adâncimea invaziei.

Studiu 3. Deep learning cu transfer learning în anatomie patologică. Studiu de caz: clasificarea carcinomului bazocelular

Achiziția de imagini de tip secțiune integrală (*whole slide*) prin scanarea lamelor histopatologice a permis studierea metodelor automate de analiză de imagini. Utilitatea acestor aplicații constă în interpretarea caracteristicilor morfologice și a tiparelor arhitecturale ale diferitelor tumori, prin recunoașterea, analiza și interpretarea de modele. Din puncte de vedere histopatologic, CBC prezintă multiple subtipuri și variante. Vastitatea de modele vizuale din preparatele histopatologice sunt o provocare pentru programele de analiză a imaginilor. Viteza de analiză a modelelor automate bazate pe DL sunt mai rapide decât modelele manuale. Aceste programe computaționale contribuie cu date cantitative, ajută la recunoașterea modelelor vizuale și asigură coerența rezultatelor.

Scopul acestui studiu este de a concepe și implementa o aplicație de clasificare fiabilă, bazată pe DL cu TL care să poată recunoaște în mod eficient variante și subtipuri ale CBC.

Setul de date cuprinde 102 cazuri consecutive de CBC diagnosticate în secția de Anatomie patologică a Spitalului Clinic Municipal Cluj-Napoca. Lamele histopatologice au fost scanate și transformate în imagini digitale. Din aceste imagini s-au extras 2249 de parcele care au fost

ulterior evaluate și încadrate de către trei medici anatomopatologi în diversele subtipuri și variante histologice ale CBC. Ulterior, trei RN de DL, fiecare cu arhitectură diferită, au fost modificate folosind conceptul de TL pentru clasificarea CBC. Cele mai performante modele au format o aplicație autonomă capabilă să clasifice imaginile în timp real. O parte din cele 2249 de imagini au fost folosite pentru antrenare iar restul pentru validare. Pentru a obține o putere statistică adecvată, cele trei modele au fost rulate independent de 100 de ori cu validare încrucișată de 10 ori. Acordul dintre RN și medicii anatomopatologi a fost evaluat utilizând coeficientul kappa al lui Cohen deoarece această analiză ia în considerare și posibilitatea ca acordul să apară din întâmplare. Pentru a testa și mai mult programul conceput, am utilizat conceptul de *explainable AI* (XAI) care își are originea în procesul decizional al minții umane, bazat pe argumente.

Aplicația concepută a avut o acuratețe ridicată (medie acuratețe/sensibilitate/specificitate/aria de sub curbă = 82%/72%/98%/0,99) și un acord bun ($\kappa=0,6862$) cu anatomopatologii în ceea ce privește clasificarea în variante și subtipuri a CBC. Aplicația a avut o performanță de acord mai mare a rețelelor neuronale decât au avut anatomopatologii între ei (media acord rețelele = 0,7533, media acord anatomopatologi = 0,7288). Cu toate că studiul a demonstrat o încadrare corectă pentru majoritatea imaginilor analizate, ocazional au fost observate etichetări eronate. Fiecare confuzie a fost detaliată și dezbătută în capitolul Discuții.

Rezultate obținute sunt promițătoare și ar crește acuratețea și eficiența clasificării acestei tumori, scăzând variabilitatea inter- și intra-observator.

Studiu 4. Subtipurile de carcinom bazocelular nodular și micronodular sunt tumori diferite în funcție de arhitectura lor morfologică și de interacțiunea lor cu stroma înconjurătoare

Cea mai recentă clasificare a tumorilor cutanate realizată de către Organizația Mondială a Sănătății recunoaște 10 subtipuri de CBC, însă aceste subtipuri au variante multiple. Cu toate că există un consens ridicat între anatomopatologi în încadrarea CBC în grupe de risc, același lucru nu este valabil în ceea ce privește încadrarea în subtipuri și variante. Subtipul nodular de CBC este cel mai frecvent și este încadrat în grupa de risc redus. Subtipul micronodular are o arhitectură similară însă face parte din grupa de risc ridicat. Diverse cercetări din literatura de specialitate au observat anumite diferențe între cele două subtipuri în ceea ce privește prezența beta-cateninei, expresia bcl-2, prezența retracției stromale, prezența lamininei-5, tipul de metaloproteinaze prezente peritumoral, tipul de stromă sau cantitatea și tipul de celule inflamatorii. Metodele de analiză a imaginilor care au la bază IA, în mod particular DL, sunt promițătoare în înțelegerea și reconstrucția imaginilor, venind în ajutorul anatomopatologilor prin îmbunătățirea acurateții și eficienței diagnosticului.

Studiul are ca scop evaluarea morfologică a unui subtip de risc redus (nodular) și a unuia cu risc ridicat (micronodular) pentru a determina dacă o analiză folosind IA poate aduce informații noi în ceea ce privește asemănările și deosebirile morfologice dintre aceste subtipuri.

Pentru realizarea sarcinii propuse au fost segmentate 417 imagini preluate din imagini digitale întregi (*whole slide image*) obținute prin scanarea preparatelor histopatologice. Segmentarea a avut la bază patru componente morfologice distincte ale CBC: celulele tumorale excluzând cele din palisadarea periferică, celulele tumorale din palisadarea periferică, retracția stromală și stroma. Analiza statistică a avut la bază două abordări: o analiză morfometrică tradițională și o analiză prin segmentare semantică utilizând DL. Abordarea morfometrică tradițională a obținut rezultate descriptive, iar cea prin segmentare utilizând deep-learning a

obținut rezultate inferențiale. Analiza a fost completată cu segmentare automată a unor imagini digitale întregi (*whole slide*).

Din cele 417 imagini, 216 au fost din subtipul nodular iar 201 din cel micronodular. Studiul a demonstrat similitudini între cele două subtipuri în special în ceea ce privește zona de palisadare periferică. Au fost observate diferențe calitative în ceea ce privește retractoria stromală și diferențe morfologice în ceea ce privește celulele tumorale din interiorul insulelor de carcinom: celulele subtipului nodular sunt mai alungite, au orientarea în aceeași direcție și prezintă matrice intercelulară mai abundentă, iar cele din subtipul micronodular sunt mai rotunjite, sunt dezorganizate și au stromă intercelulară în cantitate mai redusă decât subtipul nodular.

În concluzie, analiza imaginilor histopatologice asociind DL și abordarea morfometrică convențională au caracterizat suplimentar morfologia subtipurilor nodular și micronodular de CBC.

Concluzii, originalitatea și contribuțiile inovative ale tezei

Cercetarea doctorală s-a axat pe caracterizarea comprehensivă a morfologiei celei mai frecvente tumori cutanate maligne, CBC. Principalele concluzii ale cercetării doctorale cuprind:

1. Analiza vascularizației: Vascularizația particulară a CBC explică agresivitatea sa locală și rata scăzută de metastazare.
2. Corelația dintre ecografie și histologie: Este prezentă o corelație puternică (Pearson $r=0,92$) între adâncimea CBC măsurată ultrasonografic și histopatologic.
3. Aplicarea IA pentru clasificarea CBC: Am dezvoltat o aplicație bazată pe DL cu TL pentru a clasifica 10 subtipuri de CBC în timp real. Am obținut o precizie ridicată (ACC/SN/SP/AUC= 82%/72%/98%/0,99) în diagnosticul histopatologic.
4. Aplicarea IA pentru studiu morfologic al CBC: Prin asocierea unei abordări morfometrice standard cu caracteristici de textură cu segmentarea semantică cu analiză prin *machine learning* am caracterizat un subtip de risc redus și unul de risc înalt de CBC observând noi diferențe morfologice în ceea ce privește celulele tumorale și retractoria stromală.

Originalitate și contribuții inovative:

- Am conceput, validat și implementat o aplicație autonomă de clasificarea a subtipurilor și variantelor de CBC, bazată tehnici de IA de ultimă generație, valorificând inclusiv conceptul de IA explicabilă.
- Am oferit în mod gratuit aplicația rezultată în urma studiului, oferind un instrument de primă sau a doua opinie medicilor anatomopatologi, dar și de învățare pentru studenți, rezidenți sau specialiști.
- Utilizând IA, am furnizat informații noi privind morfologia subtipurilor nodular și micronodular de CBC, evidențiind asemănări și diferențe care nu pot fi detectate prin examinarea histopatologică clasică.

Prin integrarea acestor elemente, cercetarea doctorală oferă o investigație cuprinzătoare și modernă a morfologiei CBC, utilizând tehnologii de ultimă generație.

PHD THESIS - RESUME

Basal cell carcinoma: a morphological, imaging and artificial intelligence- assisted study

PhD Student **Raluca Maria Bungărdean**

Scientific supervisor Prof.dr. **Maria Crișan**



TABLE OF CONTENTS

| | |
|--|----|
| INTRODUCTION | 15 |
| STATE OF THE ART | |
| 1. Basal cell carcinoma – general aspects | 19 |
| 1.1. Definition and epidemiology | 19 |
| 1.2. Risk factors | 19 |
| 1.3. Diagnosis | 21 |
| 1.4. Treatment | 22 |
| 2. The histopathology of BCC | 25 |
| 2.1. Microscopy and classification | 25 |
| 2.2. Immunohistochemistry and differential diagnosis | 28 |
| 2.2.1. Immunohistochemical analysis | 28 |
| 2.2.2. Differential diagnosis | 30 |
| 3. High frequency ultrasound | 31 |
| 3.1. History and definition | 31 |
| 3.2. HFUS – BCC aspects | 31 |
| 4. Artificial intelligence in the pathological diagnosis of BCC | 35 |
| PERSONAL CONTRIBUTION | |
| 1. Working hypothesis | 41 |
| 2. General Methodology | 43 |
| 2.1. Data collection | 43 |
| 2.2. Ethics Committees | 43 |
| 2.3. Histopathological processing | 44 |
| 2.4. Digitisation of slides | 45 |
| 3. Study 1 – Morphological aspects of basal cell carcinoma vascularization | 47 |
| 3.1. Introduction | 47 |
| 3.2. Working hypothesis | 48 |
| 3.3. BCC - origin and morphology | 48 |
| 3.4. Morphological characteristics of the CBC vascularization | 49 |
| 3.5. Ultrasonography and histology | 52 |
| 3.6. Usefulness of HFUS in dermatology | 52 |
| 3.7. HFUS of BCC: morphology and vascularization | 53 |
| 3.8. Importance of Doppler in the diagnosis of BCC | 55 |
| 3.9. Advantages and disadvantages of HFUS in the assessment of BCC | 56 |
| 3.10. Other approaches to characterise the vascularization of BCC | 57 |
| 3.11. Conclusions | 57 |
| 4. Study 2 – High-frequency ultrasound: an essential non-invasive tool for the pre-therapeutic assessment of basal cell carcinoma | 59 |
| 4.1. Introduction | 59 |
| 4.2. Working hypothesis | 60 |
| 4.3. Materials and methods | 61 |
| 4.4. Results | 62 |

| | |
|--|------------|
| 4.5. Discussions | 67 |
| 4.6. Conclusions | 68 |
| 5. Study 3 – Deep learning with transfer learning in pathological anatomy. Case study: classification of basal cell carcinoma | 71 |
| 5.1. Introduction | 71 |
| 5.2. Working hypothesis | 75 |
| 5.3. Materials and methods | 76 |
| 5.3.1. Material | 76 |
| 5.3.2. Method | 78 |
| 5.3.3. Networks transfer, training, and evaluation | 79 |
| 5.3.4. Standalone application | 80 |
| 5.3.5. Clinical evaluation | 80 |
| 5.3.6. Statistical assessment | 80 |
| 5.4. Results | 81 |
| 5.5. Discussions | 86 |
| 5.6. Conclusions | 91 |
| 6. Study 4 – Nodular and micronodular basal cell carcinoma subtypes are different tumours based on their morphological architecture and their interaction with the surrounding stroma | 93 |
| 6.1. Introduction | 93 |
| 6.1.1. State of the art | 93 |
| 6.1.2. Tumor Island | 95 |
| 6.1.3. Peritumoral Cleft | 96 |
| 6.1.4. Tumor Stroma | 97 |
| 6.2. Working hypothesis | 98 |
| 6.3. Materials and methods | 98 |
| 6.3.1. Material | 98 |
| 6.3.2. Method | 100 |
| 6.3.2.1. Dataset Preparation | 100 |
| 6.3.2.2. Morphometric Analysis | 103 |
| 6.3.2.3. Semantic Segmentation Analysis | 103 |
| 6.3.2.4. WSI Automatic Segmentation | 105 |
| 6.4. Results | 105 |
| 6.5. Discussions | 108 |
| 6.5.1. General Remarks | 108 |
| 6.5.2. Tumor Islands (Including Peripheral Palisade) | 109 |
| 6.5.3. Peritumoral Cleft | 111 |
| 6.5.4. Tumor Stroma | 112 |
| 6.5.5. Future Research | 113 |
| 6.5.6. Brief Summary | 113 |
| 6.6. Conclusions | 114 |
| 7. General conclusions | 115 |
| 8. Originality and innovative contributions of the thesis | 117 |
| REFERENCES | 119 |

Keywords: basal cell carcinoma, high frequency ultrasound, morphology, histopathology, digital pathology, machine learning, deep learning, transfer learning, convolutional neural networks.

INTRODUCTION

Skin cancer is a complex and current public health problem. Of all skin cancers, the most widespread globally is basal cell carcinoma (BCC). Clinical dermatological diagnosis of this malignancy can be improved using new imaging techniques such as high-frequency ultrasound. The information provided by this investigation has an influence on treatment choice. Currently, the gold standard in BCC diagnosis is histopathological examination. With the advent of digital pathology, histopathological images can be analysed using computational methods.

This PhD research explores the morphology of BCC and the diagnostic refinement of this cancer by converging digital pathology with artificial intelligence (AI). It leverages advanced deep learning (DL) and transfer learning (TL) techniques to analyse histological images, providing an innovative approach for both pathological diagnosis and research.

STATE OF THE ART

BCC is growing in incidence worldwide and is being diagnosed in younger and younger patients. Although the metastatic rate of this malignancy is low, its prevailing occurrence in the head and neck region, its local destructiveness and the fact that it is the most common malignancy of the skin, represent major challenges in cutaneous oncological pathology.

The cell of origin of BCC resides in the basal cells of the hair follicle. The occurrence of BCC predominantly on photo-exposed areas of the skin indicates the close link between this malignancy and ultraviolet (UV) radiation, especially type B radiation. In addition to clinical and dermoscopic examination, clinical diagnosis of BCC can be improved with various medical imaging techniques, such as high-frequency ultrasound, confocal microscopy or optical coherence tomography. Although histopathological examination maintains its status as the gold standard in the diagnosis of this cancer, advances in imaging techniques show statistically satisfactory correlations with BCC histopathology, improving clinical diagnosis and extending treatment methods beyond surgery.

Based on recurrence risk and prognosis BCC can be clinically classified into two broad categories (easy to treat or low risk and difficult to treat or high risk). This classification facilitates treatment planning. Skin ultrasound provides information on the size, appearance, invasiveness and subtype of BCC, showing a good correlation with the microscopic appearance of this malignancy. Grayscale, Doppler, elastography and contrast-enhanced sonography modes contribute to the complete evaluation of BCC. Doppler effectively highlights the vasculature, while elastography and contrast-enhanced ultrasonography help classify lesions as malignant.

Histopathologically, BCC is represented by islands basal cells with peripheral palisades surrounded by a fibromyxoid stroma with peritumoral stromal retractions. Histologically classified into 10 subtypes, BCC is grouped according to risk of recurrence into high-risk and low-risk categories, similar to clinical groups. Several subtypes may coexist within a tumour.

Integration of AI in the field of pathology is facilitated by digital pathology which involves obtaining high-resolution digital images of tissues by scanning histopathological slides. These images can then be interpreted using image analysis software. AI leverages pattern recognition algorithms, simulating human thought. In the field of machine learning for the purpose of BCC diagnosis, two distinct approaches can be used: supervised and unsupervised learning. Supervised learning involves training algorithms with labelled datasets. Unsupervised learning

implies pattern identification without labelled data. DL is a supervised learning technique that uses artificial neural networks (NN) to analyse images by identifying patterns. The use of multiple neural layers results in complex learning capabilities of identifying detailed features. In histological image analysis, the most common AI techniques are convolutional NNs, DL and TL. Convolutional NNs are effective in image analysis tasks. The choice of AI method in histological image evaluation varies depending on the complexity of the task, available data, computational resources and expertise of the development team. Some of the most common approaches are TL, which uses pre-trained models for specific tasks, and hybrid models that combine traditional algorithms with DL. AI applications in BCC diagnostics involve DL and convolutional NN to identify tumor-specific patterns, comparing images against a large database for increased accuracy. The need for large and diverse datasets as well as high-quality images are the main challenges of using AI in pathology. AI-based tools in histological diagnosis are aids for doctors providing diagnostic or learning support.

PERSONAL CONTRIBUTION

The main objective of the PhD research was a complex and modern characterization of the most common skin malignancy, BCC. The structure of the research comprises a literature review and three applied research studies, each addressing different aspects of BCC morphology, as follows:

1. **BCC vascularization:** is addressed by a descriptive review in the first study of the PhD thesis, which analyses its role in the development, growth and spread of BCC, explaining specific behaviours of this tumour such as low metastatic potential and slow growth.

2. **The correlation between ultrasonography and histology:** is assessed by a retrospective study (study 2) that evaluates the depth of invasion of BCC determined sonographically and histologically. This study demonstrates a strong correlation between the two measurements and explains the importance of this correlation in terms of the therapeutic approach to this cancer.

3. **AI-assisted histological classification:** Study 3 involved the development, implementation and validation of artificial intelligence software designed to classify microscopic images of BCC into different subtypes and variants. This application uses the latest image classification techniques (DL with TL of convolutional NN) and improves consensus in BCC classification.

4. **Exploring morphological nuances of BCC using AI:** Study 4 focuses on using AI to reveal subtle morphological variations within two subtypes of BCC belonging to different risk classes: nodular and micronodular. This approach brings new insights into similarities (peripheral palliation) and differences (stromal shrinkage, tumor islet cells) in the morphology of these subtypes, revealing aspects not detectable by the usual histopathological examination.

By integrating these elements, a comprehensive investigation of BCC morphology was achieved using modern techniques.

Study 1. Morphological aspects of basal cell carcinoma vascularization

The first study of the PhD research is a descriptive review of the literature on the morphology of BCC vasculature and its implications for invasiveness, growth and metastasis.

The aim of this study is to explain some behavioural characteristics (low metastatic rate, slow growth and local aggressiveness) of BCC through the pattern of vascularization. An analysis of the practical utility of knowing the distribution of blood vessels in terms of imaging

measurements, pretherapeutic differentiation of low-risk from high-risk subtypes and the impact on surgical approach is also performed.

The examination of tumor vascularization is structured around two main lines of investigation: clinical methods of investigation (architecture, arrangement and length of vessels quantified by Doppler ultrasound, dermoscopy and capillaroscopy) and paraclinical methods of investigation (microvessel density and endothelial cell proliferation studied by immunohistochemical staining).

The local aggressiveness of BCC is attributed to its intense vascularization, characterized by a chaotic, tortuous and arboreal architecture seen predominantly at the border between tumor and stroma, right at the invasion front of carcinoma islands. At the periphery of the tumour, microvessel density is high and endothelial cell proliferation is maximal. Microvessel density varies with BCC subtype, being more abundant in high-risk subtypes. This finding is supported by microvessel density studies and investigations such as capillaroscopy and Doppler ultrasound. Reduced angiogenesis at the periphery of the tumour explains low-risk BCC cases with indolent growth. A reduced or absent vascular density in the middle of the tumour may cause the central ulcerations commonly seen macroscopically in this cancer.

The low metastatic rate of BCC is related to regression or absence of vascular structures within basal cell islands, a hypothesis supported by immunohistochemical studies and comparative studies with other skin tumours such as squamous cell carcinoma or melanoma, which have a higher metastatic rate.

Doppler and contrast-enhanced ultrasound reveals prominent blood flow in the lower portion of the BCC with one or two vascular pedicles. Doppler assessment of the CBC vasculature can guide therapeutic procedures by improving measurements when it is difficult to perform in the grey scale or help estimate blood loss during surgery. It can also be useful in the clinical classification of a BCC into easy or difficult to treat groups, and thus has importance in optimal treatment planning. Dermoscopy reveals characteristic patterns of superficial blood vessels and is useful in the clinical diagnosis of BCC. Confocal microscopy reveals superficial, horizontal, irregular vessels that sometimes run parallel.

Analysis of the BCC vasculature provides information about the course and progression of this cancer and may improve therapeutic strategy.

Study 2. High-frequency ultrasound: an essential non-invasive tool for pre-therapeutic evaluation of basal cell carcinoma

High-frequency ultrasonography is a useful investigation in the evaluation of skin lesions and is instrumental in clinical diagnosis, treatment evaluation and identification of recurrences. Grey-scale HFUS provides significant data on BCC morphology (structure, contour, size, relationship to adjacent structures). Grey scale examination can be complemented with multimodal sonography (Doppler, elastography). An important measure in BCC treatment planning is the sonographic index of tumour depth. Nonsurgical therapies may be considered for low-risk BCC, and BCC with a depth of less than 2 mm can be treated using photodynamic therapy. Given the predilection of this tumour in the head and neck region an optimal treatment considers both an oncological cure and a favourable aesthetic outcome with minimal loss of healthy tissue.

The aim of this study is to evaluate the correlation between sonography and histopathology in terms of BCC depth.

The study includes 90 cases of BCC diagnosed at the Cluj County Emergency Hospital. All sonographic examinations were performed with a high-resolution 18 MHz probe. Information about the appearance, size, vascularity (Doppler) and stiffness (elastography) of the tumour was extracted from the sonographic report, the most relevant for the study being the depth index measured in millimetres. Histopathological slides were re-evaluated and using Leica LAS EZ software, histological depth indices were measured in millimetres. The study included different subtypes and variants of BCC and Pearson correlation coefficient and Student t-test were used for statistical analysis.

Of the 90 cases included in the study the majority were low-risk subtypes and the nodular subtype with its variants was the most common histological subtype. The majority of lesions had two vascular pedicles. All tumors had increased blood supply in the lower portion and increased stiffness. Statistical results show a strong correlation (Pearson's $r=0.92$) between ultrasonographic and histological measurements quantified by depth index, demonstrating a predictive relationship between the two investigative approaches.

The relevance of the research lies in the direct clinical implications for patient management by integrating ultrasonography into clinical dermatological diagnosis, due to the high correlation with histopathology regarding the most important index in staging BCC, depth of invasion.

Study 3. Deep learning with transfer learning in pathological anatomy. Case study: classification of basal cell carcinoma

Acquisition of whole slide images by scanning histopathological slides allowed the study of automated image analysis methods. The utility of these applications lies in the interpretation of morphological features and architectural patterns of different tumours through pattern recognition, analysis and interpretation. Histopathologically, BCC shows multiple subtypes and variants. The vastness of visual patterns in histopathological preparations are a challenge for image analysis programs. The analysis speed of DL-based automated patterns are faster than manual patterns. These softwares contribute quantitative data, help to recognize visual patterns and ensure consistency of results.

The aim of this study is to design and implement a reliable DL-based classification application with TL that can efficiently recognize variants and subtypes of CBC.

The dataset comprises 102 consecutive cases of BCC diagnosed in the Pathology Department of the Municipal Clinical Hospital Cluj-Napoca. Histopathological slides were scanned and transformed into digital images. From these images, 2249 patches were extracted and subsequently evaluated and classified by three pathologists into the various histological subtypes and variants of BCC. Subsequently, three DL NNs, each with different architecture, were modified using the TL concept for BCC classification. The best performing models formed a stand-alone application capable of classifying images in real time. Part of the 2249 images were used for training and the rest for validation. To obtain adequate statistical power, the three models were run independently 100 times with 10-fold cross-validation. Agreement between NNs and pathologists was assessed using Cohen's kappa coefficient because this analysis also takes into account the possibility that agreement may occur by chance. To further test the designed program, we used the concept of explainable AI (XAI) which originates from the human mind's argument-based decision-making process.

The designed application had a high accuracy (mean accuracy/sensitivity/specificity/area under the curve = 82%/72%/98%/0.99) and a good agreement (kappa=0.6862) with pathologists in terms of variant and subtype classification of

BCC. The application had a higher agreement performance of NN than pathologists had with each other (mean network agreement = 0.7533, mean pathologist agreement = 0.7288). Although the study demonstrated correct classification for the majority of images analysed, occasionally mislabelling was observed. Each mislabelling has been detailed and discussed in the Discussion chapter.

The results obtained are promising and would increase the classification accuracy of this tumor, decreasing inter- and intra-observer variability.

Study 4. Nodular and micronodular basal cell carcinoma subtypes are different tumours based on their morphological architecture and their interaction with the surrounding stroma

The most recent classification of skin tumours by the World Health Organization recognises 10 subtypes of BCC, but these subtypes have multiple variants. Although there is a high consensus among pathologists in classifying BCC into risk groups, the same cannot be said for classification into subtypes and variants. The nodular subtype of BCC is the most common and is classified as low risk. The micronodular subtype has a similar architecture but belongs to the high risk group. Various studies in the literature have observed some differences between the two subtypes in terms of the presence of beta-catenin, bcl-2 expression, presence of stromal shrinkage, presence of laminin-5, type of metalloproteinases present peritumorally, type of stroma or amount and type of inflammatory cells. Image analysis methods based on AI, in particular DL, are promising in understanding and reconstructing images, helping pathologists by improving diagnostic accuracy and efficiency.

This study aims to evaluate the morphology of a low-risk (nodular) and a high-risk (micronodular) subtype to determine whether an analysis using AI can provide new information regarding morphological similarities and differences between these subtypes.

For the proposed task, 417 images were segmented from whole slide digital images obtained by scanning histopathological glass slides. Segmentation was based on four distinct morphological components of BCC: tumor cells excluding peripheral palisading, tumor cells in peripheral palisading, stromal retraction and stroma. The statistical analysis was based on two approaches: a traditional morphometric analysis and a semantic segmentation analysis using DL. The traditional morphometric approach obtained descriptive results and the segmentation approach using deep-learning obtained inferential results. The analysis was complemented with automatic segmentation of whole slide digital images.

Of the 417 images, 216 were of the nodular subtype and 201 of the micronodular subtype. The study showed similarities between the two subtypes especially in the peripheral palisading area. Qualitative differences in stromal shrinkage and morphological differences in tumour cells within carcinoma islands were observed: cells of the nodular subtype are more elongated, have orientation in the same direction and show more abundant intercellular matrix, while those of the micronodular subtype are more rounded, disorganised and have less intercellular stroma than the nodular subtype.

In conclusion, the analysis of histopathological images associating DL and conventional morphometric approach provided new morphological information about the morphology of nodular and micronodular subtypes of BCC.

Conclusions, originality and innovative contributions of the thesis

This PhD research focuses on the comprehensive characterization of the morphology of the most common malignant skin tumor, BCC. The main findings of the PhD research include:

1. Vascularization analysis: The particular vascularization of BCC explains its local aggressiveness and low metastasis rate.
2. Correlation between ultrasonography and histology: There is a strong correlation (Pearson $r=0.92$) between BCC depth measured on ultrasound and histological examination.
3. AI application for BCC classification: We developed a DL-based application with TL to classify 10 BCC subtypes in real time. We achieved high accuracy (ACC/SN/SP/AUC = 82%/72%/98%/0.99) in histopathological diagnosis.
4. Application of AI for morphological study of BCC: By combining a standard morphometric approach with texture features with semantic segmentation with machine learning analysis we characterized a low-risk and a high-risk subtype of BCC by observing novel morphological differences in tumor cells and stromal shrinkage.

Originality and innovative contributions:

- We have designed, validated and implemented a stand-alone application for classifying subtypes and variants of BCC based on state-of-the-art AI techniques, leveraging including the concept of explainable AI.
- We offered the resulting application free of charge, providing a first or second opinion tool for pathologists, as well as a learning tool for students, residents or specialists.
- Using AI, we provided new information on the morphology of nodular and micronodular subtypes of BCC, highlighting similarities and differences that cannot be detected by classical histopathological examination.

By integrating these elements, the PhD research provides a comprehensive and modern investigation of BCC morphology using state-of-the-art technologies.