

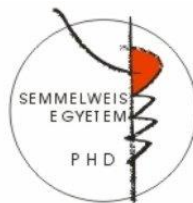
# Infravörös spektroszkópia alkalmazása sejtvonalak és tumorszövetek elkülönítésére

Doktori tézis

**Kontsek Endre**

Semmelweis Egyetem

Patológiai tudományok Doktori Iskola



## **Témavezető:**

Dr. Kiss András D.Sc., egyetemi tanár

## **Hivatalos bírálók:**

Dr. Pálfi Emese Ph.D., egyetemi adjunktus

Dr. Tornai Dávid Ph.D., tudományos munkatárs

## **Szigorlati bizottság elnöke:**

Dr. Sági Zoltán D.Sc., egyetemi tanár

## **Szigorlati bizottság tagja:**

Dr. Juriga Dávid Ph.D., egyetemi adjunktus

Budapest

2023

## 1. Bevezetés

Disszertációmban a jelölésmentes infravörös spektroszkópia patológiai mintákon történő alkalmazhatóságát vizsgáltam meg. A látható fény tartományán túl detektálható spektrumok kiszélesítik a morfológiai vizsgálatok horizontját. Legyen ez a kérdés klasszikus műtétechnika, úgy mint a rezekciós, sebészi szél meghatározása vagy egy, a klasszikus patológiai preanalitika során bevezethető előszűrés későbbi molekuláris vizsgálatokhoz.

A rezgési spektroszkópia a vegyületek rezgéseinek energiaváltozásaival foglalkozik. Egy molekulában az atomok folytonos mozgásban vannak. A rotáció és a transláció a legkisebb energiát igénylő mozgások, melyek során a kötésben résztvevő atomok távolsága és a kötésszögek nem változnak meg érdemleges léptékben.

A rezgési spektroszkópia két elkülönült doménje a Raman, illetve az infravörös (*infrared* - IR) spektroszkópia. Az infravörös és a Raman-képződés egyaránt a kémiai analízis technikája. A mintával való kölcsönhatás során a fényszórás típusa különbözteti meg a kettőt.

A gyakorlatban három jelentős infravörös tartományt szokás megkülönböztetni:

- közeli IR (*near infrared*, NIR): 0,78-2,5  $\mu\text{m}$ ,
- közép vagy más néven analitikai IR tartomány (*mid infrared*, MIR, de leggyakrabban IR-ként jelölik): 2,5-50  $\mu\text{m}$ ,
- távoli IR tartomány (*far infrared*, FIR): 50-1000  $\mu\text{m}$ .

A közép-infravörös tartomány ( $4000\text{-}650\text{ cm}^{-1}$ ) magában foglalja az úgynevezett ujjlenyomat tartományt ( $1800\text{-}650\text{ cm}^{-1}$ ), ahol a lipidekre, fehérjékre, amid I/II-re és nukleinsavakra jellemző csúcsok találhatóak (30). Ezt a tartományt azért nevezik így, mert az ebben a tartományban megjelenő abszorpciós sávok gyakran egyediek, tehát egy adott vegyületre karakterisztikusak, ahogyan az embereknél az ujjlenyomat. Az ujjlenyomat tartomány különösen hasznos az ismeretlen vegyületek azonosításában és az anyagkeverékek jellemzésében.

Az ujjlenyomat-tartomány jelentősége az infravörös spektroszkópiában abban rejlik, hogy nagyszámú abszorpciós sávot tartalmaz, amelyek a molekula alapvető rezgési módusainak köszönhetőek. Ezek a rezgési módusok érzékenyek a molekula kémiai szerkezetére, és részletes információt szolgáltathatnak a mintában jelen lévő funkciók csoportokról. Például a szén-hidrogén kötésre jellemző nyújtási sáv, az ujjlenyomat tartományban jó indikátor a szénhidrogének jelenlétére, míg a szén-oxigén kettős kötésre jellemző nyújtási sáv jó indikátor a karbonil funkciók csoportok jelenlétére.

Az ujjlenyomat tartomány az összetett keverékek elemzésében is hasznos. Például a biofolyadékok elemzése során az ujjlenyomat régió információt szolgáltat a különböző biomolekulák: fehérjék, lipidek és szénhidrátok jelenlétéről. Hasonlóképpen, a környezeti minták elemzése során az ujjlenyomat régió információt adhat a különböző szennyező anyagok és egyéb szennyeződések jelenlétéről.

Számos adatelemzési/matematikai technika használható az ujjlenyomat régió elemzésére, többek között a Fourier-

transzformációs infravörös (FTIR) spektroszkópia, a csillapított teljes reflexiós spektroszkópia és a diffúz reflexiós spektroszkópia. Mindegyik technikának megvannak a maga előnyei és korlátai, és az adott technika kiválasztása a specifikus alkalmazástól és a minta típusától is függ.

A spektrumokat gyakran érdemes matematikailag előkezelni. A vizualizáláshoz elterjedt művelet a spektrum átlagát kiszámolni és ezt hőterképként bemutatni, jellegzetesen ilyenek az abszorbciantérképek, illetve a hőkamerák képei. Az esetleges műtermékek, zajok szűrésének széles irodalma van.

A nyers adatok elkészítése még csak a munka egyik fele, az adatelemzés hasonló fontossággal bíró művelet. Az alapszintű statisztikai műveleteken túl széles eszköztár fejlődött ki a spektrális adatok előkezelésére és feldolgozására.

A legkézenfekvőbb csoportosítás, hogy az adott modell igényel-e felügyelt tanítást vagy az elkészült csoportokat utólag feleltetjük-e meg a minták csoportjainak.

## 2. Célkitűzések

Munkacsoportunk célként tűzte ki az infravörös spektroszkópia patológiai alkalmazásának vizsgálatát. A kutatási erőfeszítések hosszútávú célja az intraoperatív diagnosztikába való bevezetés, azonban jelen doktori munka keretében legfőképp a megalapozó vizsgálatok kérdéseit helyeztük górcső alá. A bevezetésben áttekintett rezgési spektroszkópiai módszerek és a sokváltozós adatelemzési módszerek mellett a jelenleg fejlesztés alatt álló intraoperatív alkalmazásokra is áttekintést kívántunk biztosítani.

Az alábbi specifikus célkitűzésekkel indultunk neki a munkának:

- Jelölésmentes infravörös spektroszkópia alkalmazhatóságának vizsgálata normál és transzformált sejtek vonalak elkülönítésére
- Sejtes rendszerek ideális mérési paramétereinek meghatározása és az elérhető adatfeldolgozási megoldások optimalálása, továbbfejlesztése
- FFPE mintákon való alkalmazhatóság vizsgálata, tekintettel a mintaelőkészítés lehetőségeire
- Tüdőrákok szövettani altípusainak meghatározása biopsziás és sebészi rezekciók patológiai FFPE (formalin fixált, paraffinba ágyazott) szövetmintáiról készített jelölésmentes közép-infravörös karakterisztikájuk alapján
- Minta- és pixelszintű adatkiértékelési megközelítések összevetése a jobb szeparáció eléréséhez

### 3. Módszerek

Disszertációm a két megjelent publikációmra épül, melyben először sejtvonalakat vizsgáltam, majd különböző szövettani mintázatú FFPE tüdőrák mintákat. A műtéti helyszíni intraoperatív döntéstámogatás előfutáraként egy kevésbé komplex (sejtvonalak) és egy komplexebb (FFPE tüdő szubtípusok) környezetben kívántuk bemutatni a patológiai kérdésekre történő válaszadási képességet. Az előkísérletekhez HT-1080 (ATCC, Manassas, Virginia, USA) fibroszarkóma sejtvonalat használtunk.

Az A-375 melanoma és a HT-29 humán vastagbélrák sejteket az ATCC-től (Manassas, Virginia, USA) származnak. A CCD986SK fibroblasztokat nem neoplasztikus sejtvonalként választottuk ki az összehasonlításához (szintén az ATCC-től). A sejtek fenntartásához DMEM (Lonza Group Ltd, Basel, Svájc) mellé 10%-os FBS (Euroclone Ltd, Pero, Olaszország) puffert adtunk. A sejteket 37 °C-on és 5% CO<sub>2</sub>-on inkubáltuk. A sejteket 95%-os etanol (EtOH) segítségével fixáltuk, majd PBS-vel mostuk és hagytuk megszáradni.

Összesen 30 FFPE tüdőrák eset blokkjaiból választottunk ki a Semmelweis Egyetem II. Sz. Patológiai Intézet (később Patológiai, Igazságügyi és Biztosítás Orvostani Intézet) és a Pécsi Tudományegyetem Patológiai Intézet archívumaiból, minden altípusból 10-10 darabot. A paraffinba ágyazott blokkokból 2 µm-es metszeteket készítettünk, és kétszer 10 percig xilolban deparaffináltuk. A humán FFPE minták felhasználását a Tudományos és

Kutatás-ethikai Bizottság (61303-2/2018/EKU) engedélyezte. A rutin munkafolyamatban a praiffinba ágyazott blokk elkészítése a patológiai indítást követően 1 munkanapot vesz igénybe. Az alábbi három szövettani altípust választottuk ki: laphámrák (SQ), adenokarcinóma (LUAD) és kissejtes karcinóma (SCLC).

A spektrumok gyűjtésére Fourier-transzformációs közép-infravörös képalkotást használtunk transzfleksziós optikai elrendezéssel. A Spotlight 400 mikroszkópot (Perkin Elmer Inc., Waltham, Massachusetts, USA) a Spectrum 400 spektrofotométerhez csatlakoztatva készítettük a képeket. A MCT detektor  $4000\text{-}648\text{ cm}^{-1}$  hullámhossz-tartományban.

A sejtes minták esetén a teljes és az ujjlenyomat régiót is elemeztük, míg a tüdőtumороk esetén csak az ujjlenyomat régióra koncentráltunk.

## 4. Eredmények

A mérési időket is figyelembe véve az aprólékosabb pixelméret (6,25  $\mu\text{m}$ ) mellett döntöttünk, mert így lesz a legrészletgazdagabb a kép, illetve, így nagyobb számú spektrum nyerhető ki.

A klasszikus fixálások közül a paraformaldehid és az etanol került kiválasztásra. A HT-29 és az A-375 sejtvonalakról készült 2-2 metszetet mindkét módszerrel fixáltuk. A fixálásokat egy főkomponens elemzéssel vizuálisan is szeparáltuk, úgy, hogy a spektrumok értékeit kiszíneztük a diagrammon.

Új megközelítésünkben tumoros sejtvonalak MIR spektrumait használtuk fel, melyeken előzetesen légköri korrekciót és zajcsökkentést eszközöltünk. Mivel a sejtek kitapadása az alumínium gőzölt felületre nem egy szorosan kontrollált folyamat, ezért egyaránt előfordulnak sejtmentes és igen sejtűs területek is.

Hat csoportra osztottuk a spektrumokat a teljes tartományra (4000-648  $\text{cm}^{-1}$ ) számított átlag transzmittancia (T%) értékük szerint. Első lépésben a nagy számú spektrumok előszűrése szükséges, hogy eltávolítsuk a túl magas átlagos T%-ot mutató, sejtmentes területekről származó spektrumokat, majd ezt követte a vastag sejtrétegeket (azaz átfedő sejteket) reprezentáló pixelek spektrumainak eltávolítása, melyek túl alacsony átlagos T%-ot eredményeztek.

Összesen 12 főkomponens elemzést készítettünk el. A daganatos és normál sejtek infravörös spektrumainak 5 %-



os tartományait főkomponens elemzéssel értékelve az 50-55 T%-os tartomány (a hat szisztematikusan vizsgált 5 T%-os spektrumtartományból) eredményezte a legjobb elválasztást. Az 1800-648  $\text{cm}^{-1}$  tartomány jobban teljesített, mint a teljes spektrumtartomány.

Az ujjlenyomat régió elnyelési sávjai kifejezettebb módon reprezentálták a minták biológiai komplexitását, és jobb elválasztást tettek lehetővé. A téves osztályozások számát a tényleges osztályokkal együtt kell értelmezni. Emiatt az lineáris diszkriminancia analízis (*linear discriminant analysis* - LDA) a kiválasztott felügyelt módszer a sejtek osztályozására.

A tumoros sejtek esetén nagyobb keveredés volt, így alacsonyabb pontosságot ért el a modell, számszerűen 93,92%-t. Az ujjlenyomat régió elemzése mindig nagyobb pontosságot eredményezett, mint a teljes tartomány esetében. A felügyelt LDA-modellek kielégítő eredményeket mutattak az összes adatkészleten, mind a pontosság, mind a Cohen Kappa értékek tekintetében.

Összesen 5 analízist végeztünk három LDA (lineáris, kvadratikus és Mahalanobis-távolságot használó) és két támasztó vektor gép (*support vector machine* - SVM) (lineáris nu-SVC és lineáris C-SVC beállítással) modellel. Az említett adatbázis másik részét ezeken a modelleken futtattuk le, és prediktáltuk a szövettani altípust. Minden spektrumot külön-külön osztályoztunk a fent említett öt matematikai modellel. Az eredményeket esetenként tévesztési mátrixokba rendeztük. Az egyes mintákból (mintánként 3840 spektrum) a helyesen osztályozott

spektrumok arányát tekintettük a teljes eset osztályozás döntő tényezőjének.

Összességében és összhangban a betegalapú megközelítéssel, a lineáris C-SVC SVM modell bizonyult ismét a legjobbnak, 81,645 % és 88,885 % közötti szenzitivitással és 90,484 % és 94,784 % közötti specificitással a szövettani altípusok tekintetében. A kvadratikus LDA modell nagyobb szenzitivitást ért el az SQ előrejelzések meghatározása során, mint a lineáris C-SVC SVM, de alacsonyabb specificitással. Hasonlóképpen, a LUAD minták spektrumai esetében a kvadratikus LDA 99%-os specificitást adott, de szenzitivitásban elmaradt az SVM-től. A kvadratikus LDA 95,107%-os pozitív prediktív értékkel rendelkezik LUAD esetében, ami jobb, mint a C-SVC SVM 82,885%-os teljesítménye.

## 5. Következtetések

### Az alábbi új megállapítások tesszük:

1. Az infravörös spektroszkópia transzflektanciás mérési elrendezését különböző szervek, eltérő entitású daganatos sejtvonalai diagnosztikus vizsgálatára bevezetve, kifejlesztettük az átlag transzmittancia alapú szűrést. Ezáltal a zaj és háttér spektrumot sikerült minimalizálni.
2. A daganatos sejtvonalakon az átlagtranszmittancia sávokat 5 T%-os sávokon kiértékelve az optimális tartománynak az 55-60 T%-ot tartjuk.
3. Az eltérő entitású daganatos sejtvonalakon a teljes és ujjlenyomat régió összevetése, valamint a felügyelt és nem felügyelt sokváltozós adatelemzési módszerünk egyaránt kimutatta, hogy az ujjlenyomat régió a daganatos sejtvonalak jobb elkülönítését eredményezi.
4. Infravörös spektroszkópiai elemzéssel a tüdőrák különböző szövettani altípusainak szétválasztását sikerült deparaffinált mintákon megvalósítani, mellyel a jövőben több biopsziás szövetszövetminta maradhat molekuláris tipizálásra.
5. FFPE mintákon a tüdőrák különböző szövettani altípusainak elkülönítésére jelölésmentes módszert fejlesztettünk ki a kiértékelhető spektrumok alapján.

## 6. Saját publikációk jegyzéke

### Disszertációhoz kapcsolódó közlemények

**E. Kontsek**, A. Pesti, M. Bjornstedt, T. Uveges, E. Szabo, T. Garay, P. Gordon, S. Gergely, A. Kiss, Mid-Infrared Imaging Is Able to Characterize and Separate Cancer Cell Lines. *Pathol Oncol Res*, 2020. 26(4): p. 2401-2407. (2020) <https://doi.org/10.3389/pore.2022.1610439>

**IF=3,201**

**E. Kontsek**, A. Pesti, J. Slezsak, P. Gordon, T. Tornoczki, G. Smuk, S. Gergely, A. Kiss, Mid-Infrared Imaging Characterization to Differentiate Lung Cancer Subtypes. *Pathol Oncol Res*, 2022. 28: p. 1610439. <https://doi.org/10.1007/s12253-020-00825-z>

**IF=2,874**

### Disszertációhoz közvetlenül nem kapcsolódó közlemények

Balogh, Judit ; Szócska, Miklós ; Palicz, Tamás ; **Kontsek, Endre** ; Pollner, Péter ; Varga, Gergely ; Ugrin, Irina ; Davidovics, Krisztina ; Joó, Tamás A mesterséges intelligencia alapú megoldások fejlesztése és bevezetése az egészségügyben – kézműves manufaktúrától a gyártósról? *IME: INTERDISZCIPLINÁRIS MAGYAR EGÉSZSÉGÜGY / INFORMATIKA ÉS MENEDZSMENT AZ EGÉSZSÉGÜGYBEN 21 : 2 pp. 56-63. , 8 p. (2022)* <https://doi.org/10.53020%2FIME-2022-206>

Pataki, Bálint Ármin ; Olar, Alex ; Ribli, Dezső ; Pesti, Adrián ; **Kontsek, Endre** ; Gyöngyösi, Benedek ; Bilecz, Ágnes ; Kovács, Tekla ; Kovács, Kristóf Attila ; Kramer,

Zsófia ; Kiss, András ; Szócska, Miklós ; Pollner, Péter ; Csabai, István HunCRC: annotated pathological slides to enhance deep learning applications in colorectal cancer screening SCIENTIFIC DATA 9 : 1 Paper: 370 , 7 p. (2022) <https://doi.org/10.1038/s41597-022-01450-y> **IF: 8,501**

Elek, Zsuzsanna ; Rónai, Zsolt ; Keszler, Gergely ; Harsányi, László ; **Kontsek, Endre** ; Herold, Zoltán ; Herold, Magdolna ; Somogyi, Anikó ; Bánlaki, Zsófia Correlation between Expression Profiles of Key Signaling Genes in Colorectal Cancer Samples from Type 2 Diabetic and Non-Diabetic Patients LIFE-BASEL 10 : 9 Paper: 216 (2020) <https://doi.org/10.3390/life10090216> **IF: 3,817**

Lendvai, G. ; Szekerczés, T. ; Illyés, I. ; Dóra, R. ; **Kontsek, E.** ; Gógl, A. ; Kiss, A. ; Werling, K. ; Kovalszky, I. ; Schaff, Z. ; Borka, K. Cholangiocarcinoma: Classification, Histopathology and Molecular Carcinogenesis PATHOLOGY AND ONCOLOGY RESEARCH 26 : 1 pp. 3-15. , 13 p. (2020) <https://doi.org/10.1007/s12253-018-0491-8> **IF: 3,201**

Lendvai, Gábor ; Szekerczés, Tímea ; **Kontsek, Endre** ; Selvam, Arun ; Szakos, Attila ; Schaff, Zsuzsa ; Björnstedt, Mikael ; Kiss, András The Effect of Methylselenocysteine and Sodium Selenite Treatment on microRNA Expression in Liver Cancer Cell Lines PATHOLOGY AND ONCOLOGY RESEARCH 26 : 4 pp. 2669-2681. , 13 p. (2020) <https://doi.org/10.1007/s12253-020-00870-8> **IF: 3,201**

Lendvai, Gábor ; Szekerczés, Tímea ; Gyöngyösi, Benedek ; Schlachter, Krisztina ; **Kontsek, Endre** ; Pesti,

Adrián ; Patonai, Attila ; Werling, Klára ; Kovalszky, Ilona ; Schaff, Zsuzsa ; Kiss, András MicroRNA Expression in Focal Nodular Hyperplasia in Comparison with Cirrhosis and Hepatocellular Carcinoma. PATHOLOGY AND ONCOLOGY RESEARCH 25 : 3 pp. 1103-1109. , 7 p. (2019) <https://doi.org/10.1007/s12253-018-0528-z> **IF: 2,826**