

TDK-dolgozat

Tumay Ádám

2024.



Budapesti Műszaki és Gazdaságtudományi Egyetem
Villamosmérnöki és Informatikai Kar
Méréstechnika és Információs Rendszerek Tanszék

Szív szegmentációja modell alapú neurális megközelítéssel

Készítette:

Tumay Ádám
VI. évfolyam

Konzulens:

Hadházi Dániel,
tudományos segédmunkatárs

2024.

Tartalomjegyzék

Kivonat	i
Abstract	ii
1. Bevezetés	1
1.1. Kapcsolódó Irodalom	2
2. Adathalmazok	4
2.1. JSRT	5
2.2. Privát adathalmaz	6
2.3. VinDr-CXR	7
3. Javasolt szegmentációs módszer	8
3.1. Áttekintés	8
3.2. Active Shape Modellek	9
3.2.1. Áttekintés	9
3.2.2. Alakzat rekonstrukciója részleges információ alapján	10
3.2.3. Kontrollpontok kijelölése	11
3.2.4. Alakmodell felépítése	12
3.3. Modell inicializáció	13
3.3.1. Sarokpontok becslése	13
3.3.2. Alakzat rekonstrukciója sarokpontokból	14
3.4. Szegmentáló modell finomító lépése	15
3.4.1. Mély neurális hálózat alapú finomítás	16
3.4.2. Optimalizáció	16
3.5. Modell tanítása	19
4. Kiértékelés	21
4.1. Alkalmazott kiértékelési módszerek	21
4.2. Abláció, egyes paraméterek hatása a modell teljesítményére	22
4.2.1. Optimális hiperparaméterek megtalálása	22
4.2.2. Likelihood regularizációs tag hatása	23
4.2.3. Jólformáltsági kényszerek hatása	24
4.2.4. Inicializációs maximum likelihood becslés hatása	26
4.2.5. Tüdőszegmentáció felhasználása az előfeldolgozás során	26
4.2.6. Elemenkénti likelihood limit hatása	27
4.2.7. Az ADMM optimalizáció büntetősúlyai	27
4.2.8. Konfidenciát becselő eljárás hatása	29

4.2.9. Neurális háló által prediktált görbe újraparaméterezésének ha- tása	29
4.2.10. Tisztán neurális baseline hálózat	30
4.3. Eredmények	30
5. Összefoglalás	32
Irodalomjegyzék	35

Kivonat

Napjainkban a szív- és érrendszeri megbetegedések (CVD-k) a vezető halálozási okok világszerte. Az ezekhez kapcsolódó egyes elváltozások (például kardiomegália, mellhártyagyulladás, tüdőödéma) kimutathatók a Posterior Anterior (PA) mellkas röntgenfelvételeken (mely az egyik legolcsóbb és leggyakrabban alkalmazott radiológiai képalkotó eljárás). Egy, a szív szegmentálására szolgáló robusztus módszer kifejlesztésével lehetővé válik az ilyen elváltozások automatikus szűrése. Ebben a dolgozatban egy új, modell alapú mélytanulási megközelítést mutatunk be Active Shape Modellek (ASM), mély konvolúciós neurális hálózatok (CNN) és szakértői képfeldolgozás kombinálásával, ahol az egyes komponensek összehangolását ADMM (Alternating Direction Method of Multipliers) alapú optimalizációval végezzük. A javasolt módszert különböző adatkészleteken értékeljük ki egy tisztán neurális hálózat alapú megközelítéssel szemben, mely architektúrájában hasonló a javasolt megoldáshoz. Az eredmények azt mutatják, hogy a javasolt szegmentációs eljárás robusztusabb, jobb általánosítást biztosít az adathalmazok között, különösen az olyan esetekben, amikor a szegmentáció kihívást jelent a célterületen található elváltozások miatt. Az összehasonlítást a JSRT-, valamint egy olyan privát- adathalmazon végezzük, amely felvételeit itthon, napi klinikai rutin során rögzítették. Kvantitatívan a javasolt módszer a hazai, orvosi gyakorlatban előforduló példákat tartalmazó adathalmazon **0.895 ± 0.036** Dice score-t és **$16.46 \text{ mm} \pm 5.10 \text{ mm}$** Hausdorff távolságot ért el, a tisztán neurális benchmark model **0.869 ± 0.075** Dice score és **$24.66 \text{ mm} \pm 11.09 \text{ mm}$** Hausdorff távolság eredményeivel szemben.

Abstract

Nowadays, Cardiovascular Diseases (CVDs) are the most common cause of deaths worldwide. Some lesions connected to CVDs (such as Cardiomegaly, pleural effusion, pulmonary edema) can be detected on Posterior Anterior (PA) chest X-ray projection images (one of the cheapest and most common modality in radiology). By developing robust methods for the segmentation of the heart as a computer aided diagnostic system (CADx) the automatic screening for such lesions becomes possible. In this paper, we present a novel model-based deep learning approach by combining Active Shape Models (ASM), Deep Convolutional Neural Networks (CNN) and image processing expert knowledge through numerical optimization solved by Alternating Direction Method of Multipliers (ADMM) based optimization scheme. The proposed method is evaluated on various datasets against a purely neural network-based approach that is architecturally similar to the proposed solution. Results indicate that our model is more robust, provides a better generalization across datasets especially when segmentation is challenging due to lesions obscuring the target area. Comparison is done on the JSRT dataset as well as on a private dataset containing samples acquired in domestic daily medical practice. Quantitatively on the dataset coming from daily medical practice, the proposed method achieved a Dice Score of **0.895 ± 0.036** and a Hausdorff distance of **$16.46 \text{ mm} \pm 5.10 \text{ mm}$** , while the purely neural benchmark model managed to only achieve a Dice score of **0.869 ± 0.075** and a Hausdorff distance of **$24.66 \text{ mm} \pm 11.09 \text{ mm}$** .

1. fejezet

Bevezetés

Napjainkban a szív- és érrendszeri megbetegedések (CVD-k) világszerte a leggyakoribb halálokok. A WHO 2019-es becslése alapján 17,9 millió ember halt meg CVD-ben (mely az összes halálozás 32%-a) [1]. Az elmúlt évtizedekben az ezen megbetegedésekhez köthető halálozási arány sikeresen csökkent a magasabb jövedelmű országokban (100 000 főre vetítve 271,8-ról 128,5-re), míg az alacsonyabb és közepes jövedelmű régiókban továbbra is gyakori halálozási okok maradtak (368,2-ről 316,9-re változott 1990 és 2017 között) [2]. A Covid-19 világjárvány globálisan megtörte a javuló tendenciát (például az Egyesült Államokban a 100 000 főre jutó, életkorral korrigált halálozási arány 456,7 volt 2010-ben, 416 volt 2019-ben és 454,5 volt 2022-ben) [3].

A CVD-k korai felismerése hatékonyan csökkentheti az életkorral korrigált halálozási arányt, valamint a kapcsolódó orvosi kezelések számát és költségeit [4]. A CVD-k diagnózisát nehezíti, hogy korai stádiumban jellemzően tünetmentesek, az időbeni felismeréshez szűrővizsgálatokra van szükség, tünetek általában már csak az életet veszélyeztető állapotokban észlelhetők [5]. A CVD-k okozta halálozás csökkenése valószínűleg a lakossági szűrővizsgálatok eredménye, mivel a magas jövedelmű országokban a megbetegedések gyakorisága csak kis mértékben csökkent [6]) a kapcsolódó halálozási gyakoriság intenzív visszaesése ellenére. Az alacsonyabb jövedelmű országok polgárai nehezebben jutnak hozzá ezekhez a vizsgálatokhoz, viszont rendszerezések a fertőzések okozta súlyos légúti megbetegedések (pl. tuberkulózis, vírusos tüdőgyulladás), valamint a Covid-19 fertőzések továbbra is gyakoriak világszerte, és az ilyen betegségek orvosi kezelése során gyakran alkalmazzák a mellkasröntgen modalitást. Annak ellenére, hogy ez nem a CVD-szűrés elsődleges módszere, a különböző szívbetegségek egyes tünetei (pl. mellhártyagyulladás, Kerley B-vonalak, a szív- és mellkasi arány (CTR) megnagyobbodása) az esetek többségében azonosíthatók a felvételeken [7]. E tünetek némelyikének automatizált felismeréséhez szükséges a szív pontos szegmentálása, amely nem mindig egyértelmű feladat: A szív árnyékának határa nem jól meghatározott (pl. a cardio-caudalis irány mentén), és a kontúrok sem mindig élesek a mellkasröntgen felvételek perspektivikus geometriája miatt még egészséges páciensek esetén sem. Súlyos betegségek fennállása esetén (pl. mellhártyagyulladás, tüdőgyulladás, megemelkedett rekeszizom, tüdőödéma) a szív körvonala részben ki lehet takarva, mely megnehezíti a kézi szegmentálást. Következésképp egy olyan számítógépes diagnosztikai (CADx) rendszer, amely képes a szívárnyékot kellő pontossággal és robusztussággal szegmentálni, hatékony eszköz lehet a CVD-k mellkas röntgenfelvétel alapján történő diagnosztizálásában.

1.1. Kapcsolódó Irodalom

A szív szegmentálása mellkas röntgenfelvételeken kevésbé kutatott probléma összehasonlítva a modalitáson alkalmazott más elváltozás-detektálási feladatokkal, vagy a szív szegmentációjával más modalitásokon. Az eddig publikált megközelítések két fő kategóriába sorolhatók: a szakértői képfeldolgozáson és tisztán gépi tanuláson alapuló módszerekébe. Az első kategóriában többségében statisztikai modell alapú megközelítéseket használnak (a szív kontúrjának nehéz felismerhetősége miatt). Dallal et al. [8] és Candemir et al. [9] egy multi-atlaszon alapuló regisztrációs módszert javasolt. Dallal et al. [8] csak a regisztrált tüdőmaszkokat használja, míg Candemir et al. [9] a regisztrált szívmaszkokat is. Ginneken et al. [10] Active Shape Model (ASM) és Active Appearance Model (AAM) alapú szegmentálási módszert javasolt a mellkas röntgenfelvételek anatómiai struktúráinak lokalizálására. Ez a módszer egy explicit Gauss eloszlású változókkal paraméterezett alakmodellt tanul egy mintahalmazból, mely konstrukció értelmezhető és átlátható viselkedést biztosít.

Napjainkban egyre népszerűbbek a második kategóriába tartozó, tisztán gépi tanuláson alapuló megoldások. Ennek oka, hogy ezen megközelítés nem igényel szakértői tudást, a predikciók minősége általában hatékonyan javítható a tanításra használt minták számának növelésével. Novikov et al. [11] konvolúciós neurális hálózat (CNN) architektúrákat javasolt a tüdő, a kulcscsont, valamint a szív szegmentálására mellkas röntgenfelvételeken, és kimutatta, hogy az általános CNN-megoldások megfelelő módosításai javíthatják a szegmentáció pontosságát. Maga [12] egy Attention UNet alapú architektúrát javasolt a szív és a tüdő szegmentálására, és sikeresen csökkentette a tanítás mintaigényét. Frid-Adar et al. [13] azt vizsgálta, hogy a transzfer tanítás hogyan javíthatja a pontosságot. Goméz et al. [14] pedig módosításokat javasolt a CNN-architektúrákon annak érdekében, hogy a hálók nagyobb felbontású bemeneti felvételeken is hatékonyan működjenek. Kholiavchenko et al. [15] azt kutatta, hogyan javíthatja a CNN-alapú szegmentálás pontosságát a ground truth szegmentációs maszkok kontúrjainak közvetlen felhasználása (ahelyett, hogy csak a szegmentációs maszkokat használnák).

Általánosan, a CNN-alapú megoldások nagyobb pontosságot értek el a JSRT adathalmazon [16], amely jó minőségű analóg röntgenfelvételekből áll, valamint a szív kontúrjának detekcióját jelentősen megnehezítő elváltozásokkal rendelkező mintákat nem tartalmaz. Ezen módszerek robusztussága azonban nem megfelelő a szív szegmentálásához a gyakorlati klinikai rutin során készített mellkas röntgenfelvételeken. Ebben a dolgozatban egy modell alapú neurális szegmentálási módszert javasolunk, amely egyesíti a szakértői képfeldolgozás és a neurális paradigma előnyeit. Az ezzel kapcsolatos kutatásunk korábbi eredményeit a [17] cikkben¹ publikáltuk.

A dolgozat felépítése a következő: Először az adathalmazokat mutatjuk be a 2. fejezetben, melyet az általunk javasolt szegmentációs módszer részletezése követ (3. fejezet). Ennek folyamán ismertetjük a használt alakmodellt és annak felépítését a rendelkezésre álló adatokból (3.2. fejezet), majd a modell inicializációjának módját a 3.3. fejezetben. A szegmentáló modell finomító lépését és az ehhez tartozó optimalizációs eljárást a 3.4. fejezet mutatja be, végül ismertetjük az ezekhez használt neurális hálózatok tanításáról is (3.5. fejezet). A kiértékelés folyamán először az

¹Az ebben a tanulmányban bemutatott kutatás a Gamax Laboratory Solutions MAIA - Medical AI Assistant kutatási, fejlesztési projektjének részeként készült, amelyet a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal 2020-1.1.2-PIACI-KFI-2021-00291 számú pályázata finanszírozott.

alkalmazott metrikákat taglaljuk (4.1. fejezet), melyet a hiperparaméterek optimalizációja, és az eljárás komponenseinek hatásvizsgálata követ (4.2. fejezet). Végezetül összehasonlítjuk az eljárás eredményeit egy tisztán neurális megközelítésével (4.3. fejezet), majd a dolgozatot a 5. fejezetben foglaljuk össze.

2. fejezet

Adathalmazok

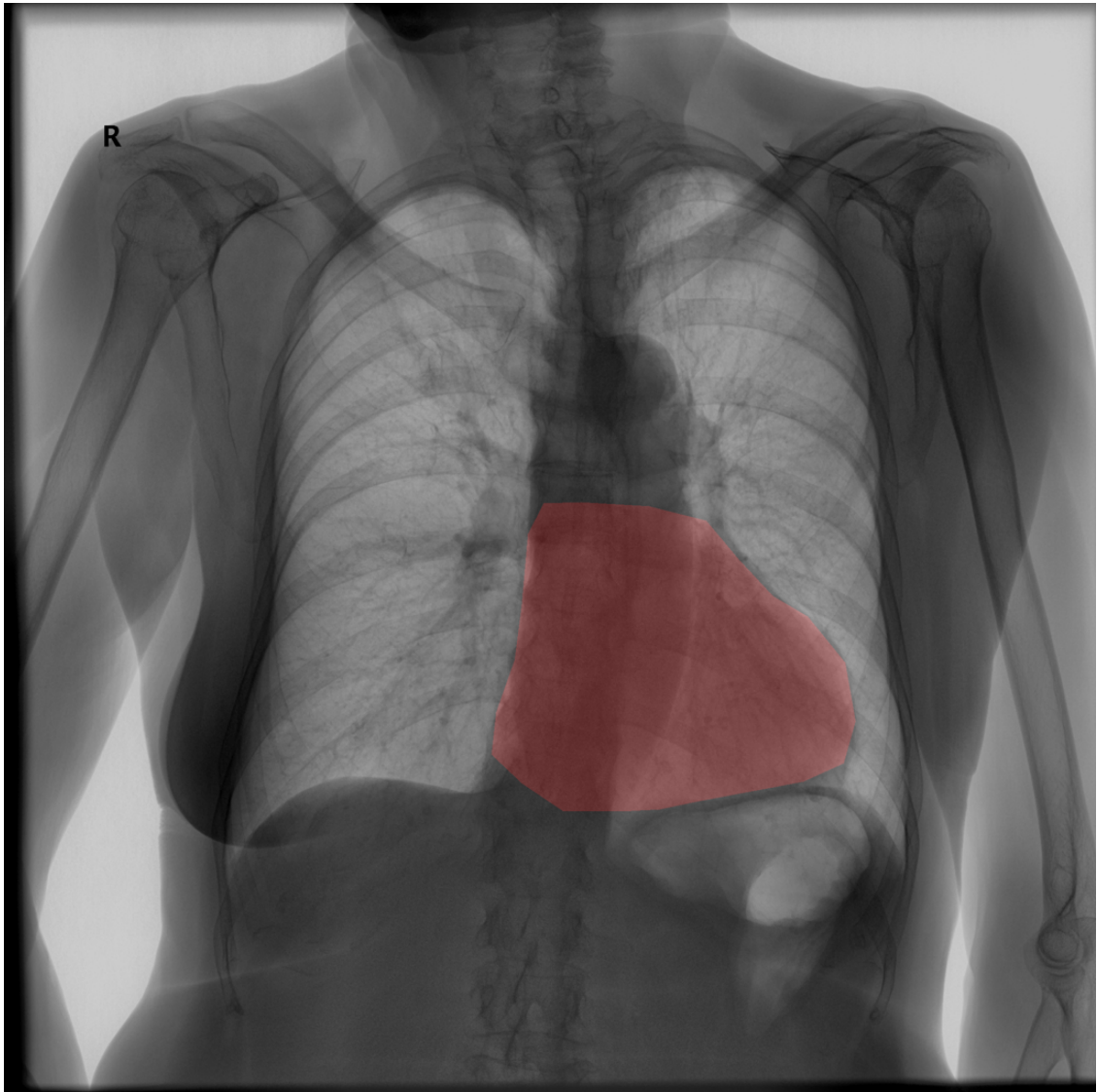
Bár a szív szegmentációját végző eljárás bemenetét képező felvételek csupán a mellkas vetületeit tartalmazzák, az általánosítást gyakran nehezíti, hogy a páciensek beállítása sok esetben nem megfelelő, ráadásul az egyes adathalmazok tipikusan egy-egy specifikus földrajzi terület/kórház jellemző vonásait is magukon hordozzák. Ilyen vonások például az egyes szervek méretei, elhízás, kóros elváltozások típusa és relatív gyakorisága. Ehhez az alapvetően nagy varianciához továbbá hozzájárulnak a modalitás jellegzetességei, a felvételkedészítés pontos módja (röntgensugár energiaspektruma, effektív sugárkeménység, röntgencső árama, analóg/digitális detektor, anti-scatter grid esetleges alkalmazása), mely nagy hatással van a különböző képek mind abszolút, mind relatív intenzitására, kontúrjaik elmosódottságára, felvételek zajának energiájára. További nehezítő tényező, hogy jellemzően minél jobb minőségű, részletesebb annotációval rendelkezik egy adathalmaz, mérete annál kisebb, a felvételek jellegét tekintve pedig kevésbé változatos, így az ezekre tanított modellek hajlamosabbak a túlilleszkedésre. Olyan szegmentáló eljárás készítésére van tehát szükség, amely képes kevés homogén felvétel alapján jól általánosítani azok valódi háttéreloszlására. Ennek bemutatása előtt azonban először tekintsük át az adathalmazok jellemzőit, melyek felvételeire példák a 2.1, 2.2, 2.3 ábrákon láthatók.



2.1. ábra. A JSRT adathalmaz egy felvétele, a ground truth szegmentációs maszk a vörös színcsatornán látható.

2.1. JSRT

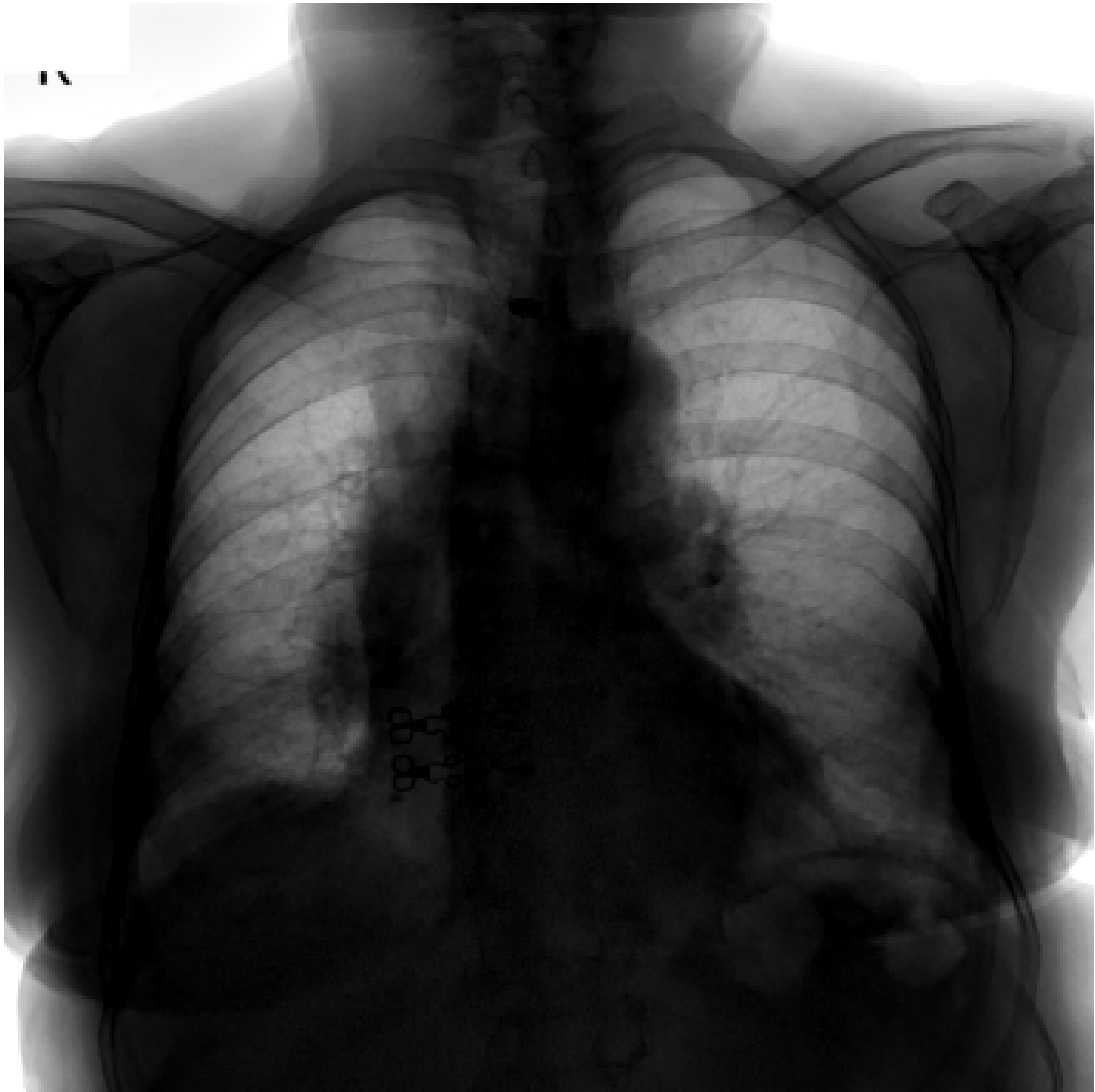
A Japán Radiológiai Társaság által készített [16] 247 db analóg PA mellkasröntgen felvételt tartalmazó adathalmaz. A képekhez tüdő- és szívszegmentációk is tartoznak melyek lehetővé teszik a felügyelt tanulást, és kvantitatív kiértékelést. A tanítást és a jó általánosítóképesség elérését nehezíti a minták alacsony száma, homogenitása. Továbbá adathalmaz felvételei nem reprezentatívak a mai, általános hazai betegpopulációra nézve az azokról napi klinikai rutin során készült felvételek jellemzőinek tekintetében.



2.2. ábra. A privát adathalmaz egy felvétele. A szegmentálási feladat nehezebb a JSRT adathalmazból származó mintához képest (2.1). A szegmentált szívárnyék a vörös színcsatornán látható.

2.2. Privát adathalmaz

Hazai, napi klinikai gyakorlatból származó 22 db digitális röntgenfelvétel. Az adathalmazhoz szakértők által annotált szívárnyék maszkok tartoznak. Kevésbé homogén, mint a JSRT, azonban a minták száma túlságosan alacsony ahhoz, hogy tanítani lehessen rajta. Elsősorban a betanított modellek minősítésére, összehasonlítására használjuk.



2.3. ábra. Példa felvétel a VinDr-CXR adathalmazra, látható, ahogy a szegmentálási feladat még nehezebb, valamint, hogy ezen adathalmazhoz már nem tartozik annotáció.

2.3. VinDr-CXR

Vietnámi kórházakban rögzített [18] 15.000 db PA felvételtől álló adathalmaz, mely kellően változatos a különböző kóros eseteket illetően, azonban a szívárnyék szegmentációjához maszkok nem tartoznak, a felvételeken felügyelt módon modell nem tanítható. Az adathalmazt főként kvalitatív kiértékelésre használjuk, annak változatossága miatt.

3. fejezet

Javasolt szegmentációs módszer

Mivel a napjainkban elterjedt tisztán neurális, a feladatot pixel szintű osztályozóként értelmező szegmentációs modellek nem biztosítanak kellő robusztusságot, így a fejezet folyamán egy új, hibrid szegmentáló eljárást mutatunk be, amely kombinálja a neurális hálózatok kifejező erejét a statisztikai, szakértői képfeldolgozási modellek robusztusságával.

3.1. Áttekintés

A javasolt módszer a következő lépésekből áll:

1. A bemeneti DICOM felvétel előfeldolgozása során a normalizáció min-max módszerrel történik, ahol a minimum és maximum értékeket az intenzitások 3. és 97. percentiliseiből vesszük, továbbá egy Gauss-kernel segítségével elmoszuk a felvételt a nagyfrekvenciás komponensek kiszűrése és így a robusztus működés biztosítása érdekében. A felvételen szegmentáljuk a tüdőt egy szakértői képfeldolgozáson alapuló eljárással, majd ezen szegmentált terület segítségével vágjuk a felvételt, hogy az csak a feladat szempontjából releváns képleteket tartalmazzon.
2. Neurális hálózat segítségével lokalizáljuk a szív sarokpontjait, a módszer részletes leírása a 3.3 fejezetben található.
3. A sarokpontok alapján a többi kontrollpontot is inicializáljuk, a 3.2.2 fejezetben leírtak szerint.
4. A szegmentáció során az alakzatot iteratíván finomítjuk 3.4 szerint. Egy neurális hálózat becsüli meg, hogy hogyan kellene a kontrollpontokat mozgatni a jobb illeszkedéshez, majd egy optimalizációs algoritmus kiszámítja a tényleges mozgásokat, figyelembe véve az alakmodellből és egyéb szakértői forrásokból származó jólformáltsági kényszereket is.
5. A szegmentált körvonal utófeldolgozása Catmull-Rom spline [19] illesztésével a kontrollpontokra, melynek célja, hogy simább és jobban illeszkedő szegmentációt kapjunk, valamint bináris maszk számítása a kontúrból. A kontrollpontokat, és a szegmentált maszkot visszavetítjük a vágás előtti felvételre.

3.2. Active Shape Modellek

3.2.1. Áttekintés

Az Active Shape Modellek [20] a maszkokból származtatott kontrollpontokon a következő módon működnek: Legyen x_i az i -edik mintához tartozó kontrollpontok ki-lapított vektora, és $\bar{x}$ ezen vektorok átlaga a minták felett. Egy adott minta átlagtól való eltérését v_i jelöli. Minden kontrollponthoz súlyokat is rendelünk, melyeket w je-löl. Az eltérések kovariancia mátrixa a (3.1) egyenlet által adódik, ahol $W = \text{diag}(w)$:

$$S = \frac{1}{N} \sum_{i=1}^N v_i^T W^2 v_i \quad (3.1)$$

A kovariancia mátrix főkomponens-analízisével megkapjuk a sajátvektorok P mátrixát és a sajátértékek λ vektorát, amelyekre: ($\Lambda = \text{diag}(\lambda)$)

$$SP = P\Lambda; P^T P = I \quad (3.2)$$

Ez lehetővé teszi egy adott, kontúrponatok által meghatározott alakzat likeli-hoodjának ($\mathcal{L}$) kiszámítását. Ehhez először az alakzatot látens paraméterek terébe vetítjük, majd felhasználjuk azt a tényt, hogy a látens paraméterek normális elosz-lásúak, 0 átlaggal és λ varianciával.

$$\begin{aligned} b &= P^T W(x - \bar{x}) \\ \mathcal{L} &= c \cdot \exp(-(1/2)b^T \Lambda^{-1} b) \end{aligned} \quad (3.3)$$

A fenti egyenletben b a látens paraméterek vektora, c pedig a normalizációért felelős konstans. A likelihood függvény lehetővé teszi a szegmentáció minőségének értékelését az alakmodell alapján, valamint a látens paraméterek általi regularizá-ciót. Ennek módját részletesen a 3.4.2 fejezetben magyarázzuk el. Adott b látens paraméterek alapján a hozzá tartozó alakzat a következőképpen rekonstruálható:

$$x = W^{-1} P b + \bar{x} \quad (3.4)$$

Megjegyezzük, hogy a numerikus stabilitás és az alakzatok regularizációjának érdekében a magasabb rendű látens paramétereket és a hozzájuk tartozó sajátvekt-o-rokat és sajátértékeket gyakran kihagyják b -ból, P -ből és Λ -ból. Numerikus stabilitá-si problémák léphetnek fel, amennyiben, az S kovarianciamátrix rosszul kondicionált ($\max_i(\lambda_i)/\min_j(\lambda_j)$ magas), ekkor a kis sajátértékekhez tartozó látens paraméterek-re gyakran instabil értékek jönnek ki a látens paramétertérbe történő transzformáció során. Regularizáció során a látens térbeli alakzat likelihoodját szeretnénk maxima-lizálni egyes látens paraméterek elhagyásával (gyakorlatilag egyenértékű azok 0-nak vételével), miközben szempont, hogy a regularizált alakzat eukleidészi értelemben minél közelebb legyen az eredetihez. Ilyen esetekben is az optimális megoldás a kis sajátértékekhez tartozó látens paraméterek elhagyása. Precízen, mindkét esetben a regularizáció, illetve numerikus stabilitás biztosítása azon látens paraméterek elha-gyásával történik, amelyek a kovarianciamátrixot csak kis mértékben magyarázzák (explained variance). Ezen esetekben a paraméterek elhagyása a következőképp tör-ténik t vágási küszöb szerint:

$$\text{explained_variance}(\lambda_i) = \lambda_i / \sum_j \lambda_j \quad (3.5)$$

A (3.5) egyenletben látható módon megállapítjuk az egyes λ_i sajátértékekhez tartozó magyarázott varianciát, majd jelölje λ' azon legnagyobb sajátértékek halmazát, melyre:

$$\lambda' : \min |\lambda'| \text{ és } \sum_{\lambda_i \in \lambda'} \text{explained_variance}(\lambda_i) > t \quad (3.6)$$

Azaz a legkisebb olyan halmaz, amely által magyarázott összes variancia meghalad egy t küszöbértéket. Ekkor jelölje P' P mátrix azon oszlopait, melyek a λ' sajátértékekhez tartozó sajátvektorok, valamint $\Lambda' = \text{diag}(\lambda')$. A látens paramétertérbe történő transzformáció, valamint a rekonstrukció a következőképp fest:

$$\begin{aligned} b' &= P'^T W(x - \bar{x}) \\ x' &= W^{-1} P' b' + \bar{x} \end{aligned} \quad (3.7)$$

Mint látható, látens paraméterek elhagyása esetén is úgy használható a modell, mint alapesetben. Figyeljük meg azonban, hogy ekkor a rekonstrukció folyamán nem x -et, hanem csak egy közelítését x' -t kapjuk vissza. Mivel gyakran szükséges a látens paramétertér vágása már a numerikus stabilitás biztosítása miatt is (S általában igen rosszul kondicionált, többek közt ez az egyik fő motiváció a főkomponens analízis tömörítésre való alkalmazásakor), a továbbiakban az egyszerűség kedvéért b' -t, P' -t, λ' -t is azok vágatlan változataival (b -vel, P -vel, λ -val) jelöljük.

3.2.2. Alakzat rekonstrukciója részleges információ alapján

A (3.3)-as likelihood kifejezés lehetővé teszi egy alakzat rekonstrukcióját, még abban az esetben is, ha a kontúrponatok többsége hiányzik, maximum likelihood becslés felhasználásával. Ez hasznos többek között a neurális hálózatok által becsült sarokpontokból történő kontúrponatok számításához, valamint a finomításra használt hálózat predikciói megbízhatóságának meghatározásához. Legyen x_A a kontúrponatok ismert része, és x_B a becsülendő kontrollponatok. Hasonlóképpen, meg lehet határozni v_A és v_B eltérésvektorokat is. Továbbá, P^T fel van osztva P_A és P_B részekre úgy, hogy $P^T = P_A | P_B$, valamint hasonlóan járunk el a W súlymátrixszal is: $W = W_A | W_B$, és $b = P_A W_A v_A + P_B W_B v_B$, ahol a $|$ operátor az összefűzést jelenti. b -t behelyettesítve a (3.3) valószínűségi kifejezésbe, a maximum (log) likelihood becslés segítségével levezethetjük v_B -t:

$$\begin{aligned} \nabla_{v_B} \log \mathcal{L} &= 0 \\ \nabla_{v_B} \log(c \cdot \exp(-(1/2)b^T \Lambda^{-1} b)) &= 0 \quad \text{és } b = P^T W v = P_A W_A v_A + P_B W_B v_B \\ \nabla_{v_B} \log(c) + (-(1/2)(P_A W_A v_A + P_B W_B v_B)^T \Lambda^{-1} (P_A W_A v_A + P_B W_B v_B)) &= 0 \\ W_B^T P_B^T \Lambda^{-1} P_B W_B v_B + W_B^T P_B^T \Lambda^{-1} P_A W_A v_A &= 0 \end{aligned} \quad (3.8)$$

$$v_B = -(W_B^T P_B^T \Lambda^{-1} P_B W_B)^{-1} W_B^T P_B^T \Lambda^{-1} P_A W_A v_A \quad (3.9)$$

Azonban, ha az ismeretlen rész nagyobb dimenziójú, mint a látens paraméterek dimenziója (mint például a sarokpontokból történő alakzat becslése esetén), akkor az inverz mátrix nem létezhet alul determináltság miatt. Ennek kiküszöbölésére Tikhonov regularizációt alkalmazunk a likelihood függvényre, ahol az s hiperparaméter a regularizáció erősségeért felelős. Szemléletesen, ekkor az ismeretlen eltérésvektorokra (v_B -re) $1/s$ varianciájú normális prior eloszlást feltételezünk. Könnyen belátható, hogy ez a feltételezés a rekonstruált kontrollpontokat az átlagmodell megfelelő pontjaihoz viszi közelebb. Ekkor a módosított likelihood függvény, illetve az ebből levezett rekonstruált alakzat a következő:

$$\mathcal{L} = c \cdot \exp(-(1/2)(P_A^T W_A v_A + P_B^T W_B v_B)^T \Lambda^{-1} (P_A^T W_A v_A + P_B^T W_B v_B) + v_B^T (s \cdot I) v_B) \quad (3.10)$$

$$v_B = -(W_B^T P_B^T \Lambda^{-1} P_B W_B + sI)^{-1} W_B^T P_B^T \Lambda^{-1} P_A W_A v_A \quad (3.11)$$

3.2.3. Kontrollpontok kijelölése

Az ASM modellek landmark pontokon működnek, azonban a rendelkezésre álló adathalmazok csak pixel szintű szegmentációs maszkokat tartalmaznak. Ahhoz, hogy az alakmodell megfelelő robusztussággal működjön, a kontúrpontokat konzisztensen kell hozzárendelni a maszkokhoz, ami a legtöbb esetben az anatómiai szempontból fontos és jól azonosítható területek kiválasztásával történik a referencia maszkokon. Esetünkben egy ilyen módszer alkalmazása nehezen lenne kivitelezhető, mivel ezen pontok nagy részét még szakértők számára is nehéz pontosan beazonosítani a felvételeken. Ehelyett kezdetben a szív sarokpontjait lokalizáljuk, majd a kontrollpontokat egyenletes interpolációval számítjuk ki a körvonal mentén. Az ASM modell felépítésekor, a sarokpontok lokalizálása során olyan négyszögeket keresünk a szív szegmentációján belül, amelyek területe lehető legnagyobb. Jelöljük a sarokpontok koordinátáit leíró vektort x -el, továbbá a szív kontúrpontjainak halmazát K -val:

$$x^* = \arg \max_{x_0, x_1, x_2, x_3 \in K} \{T(x) \mid \forall k \in K : \exists 0 \leq i, j \leq 3 : ((k - x_i)^\perp \cdot (x_{(i+1 \bmod 4)} - x_i)) \\ ((k - x_j)^\perp \cdot (x_{(j+1 \bmod 4)} - x_j)) < 0\} \quad (3.12)$$

ahol T jelöli x sarokpontok által meghatározott négyszög területét, x_i az i -edik sarokpontot, $y^\perp$ y vektor merőlegesét. A keresést egy RANSAC [21] módszeren alapuló algoritmus valósítja meg. A szegmentáció folyamán pedig ezen sarokpontok pozícióját egy CNN prediktálja.

A sarokpontok kijelölése után jelölje $ind(K_a)$ az a sarokponthoz tartozó kontúrpont indexét. Ekkor két sarokpont közötti távolság a kontúr mentén:

$$D(a, b) = \sum_{i=ind(K_a)}^{ind(K_b)-1} \|K_{i+1} - K_i\|_2 \quad (3.13)$$

Ekkor, ha két sarokpont között n további kontúrpointot szeretnénk felvenni, azt a következőképp tehetjük meg:

$$K_{n_i} : \arg \min_{n_i} \left\{ n_i \mid \sum_{j=ind(K_a)}^{n_i-1} \|K_{i+1} - K_i\|_2 \geq D(a, b)/(n+1) \cdot i \right\} \quad (3.14)$$

Ahol n_i az i -edik interpolált kontúrpoint indexét jelöli. Az eljárást megismételve minden egymást követő sarokpont párra megkapjuk a kontrollpontokat.

A kontrollpontok kijelölésére egy másik vizsgált lehetőség volt, hogy azokat egy kezdeti inkonzisztens állapotból, konvex optimalizáció segítségével iteratíván igazítjuk egymáshoz a maszkokon. Formálisan:

$$\begin{aligned} \arg \min_t \sum_{i,j} \|x^i - x^j\|_2^2 - \alpha \sum_i \sum_l \|x_{l+1}^i - x_l^i\|_2 \\ s.t. : x^i = s^i(t^i) \\ x^j = s^j(t^j) \end{aligned} \quad (3.15)$$

Ahol x^k a k -edik maszk kontúrpointjainak koordinátáit jelöli, x_m^k pedig az k -edik maszk m -edik kontúrpointjának koordinátáját, $s^k : \mathbb{R}^N \rightarrow \mathbb{R}^{2N}$ a k -edik maszk kontúrját leíró Catmull-Rom spline, amely biztosítja, hogy a az egyes kontúrpointok valóban a kontúron legyenek (ahol N a kontúrpointok számát jelöli), t^k vektor pedig ennek egy paraméterezése. Szemléletesen az optimalizációs kifejezés első tagja biztosítja, hogy az egyes kontúrpointok közel, remélhetőleg közel azonos anatómiai szemantikájú pontokon legyenek a maszkok között, míg a második kifejezés azért felel, hogy az egyes maszkokhoz rendelt kontúrpoint lehelyezések minél pontosabban rekonstruálják az adott maszkot. A két párhuzamosan optimalizálandó kifejezés között egy α súly paraméter teremt kapcsolatot. t^k -n keresztül lehetőség nyílik a kifejezés optimalizációjára, pl.: gradient descent-en keresztül, azonban azt tapasztaltuk, hogy mivel összességében a probléma nem konvex, és az sem mondható el, hogy a lokális minimumok jól közelítenék a valódi optimumot, így az optimalizáció során az eljárás bár konvergál, jellemzően hamar "beragad" egy, az optimumtól távol található lokális minimumban. A nem megfelelő konvergencia tulajdonságok miatt a gyakorlatban az előző, sarokpontok szakértői becslésén alapuló megközelítést alkalmazzuk a kontrollpontok kijelölésére.

3.2.4. Alakmodell felépítése

Először a rendelkezésre álló maszkok felhasználásával létrehozuk a landmark pontokat a 3.2.3 szakaszban leírt módszerrel. Ezután a kontúrpointok vektorait iteratíván egymáshoz igazítjuk úgy, hogy először kiszámítjuk az átlagos maszkot, azaz az egyes maszkok kontúrpointjait jelölő vektorainak átlagát vesszük. Ezután egy affin hasonlósági transzformációval igazítjuk az egyes alakzatokat az előzőleg kiszámított átlaghoz. Ezt addig ismételjük, amíg konvergencia nem történik, miközben minden iteráció után a kontúrpointok vektorait újraskálázzuk a középpont körül, úgy, hogy a kontúrpointok középponttól való négyzetes euklideszi távolságainak összege az iterációk folyamán állandó maradjon. A lépés ahhoz szükséges, hogy az egymásra igazítási

lépés azonos eredményre vezessen az alakmodell felépítése és gyakorlati alkalmazása (azaz a modell új felvételeken történő szegmentálása) során. Az iteratív egymásra igazítást végző algoritmus ciklusmagja a (3.16) egyenletben látható formalizálva.

$$\begin{aligned}
\bar{x} &= 1/n \cdot \sum_i^n x_i \\
f &= \min_{f_{M,O}(x_i)=Mx_i^T+O} \|f_{M,O}(x_i) - \bar{x}\|_2^2 \\
x'_i &= f(x_i) \cdot \|x_i\|_2^2 / \|f(x_i)\|_2^2 \\
x_i &= x'_i
\end{aligned} \tag{3.16}$$

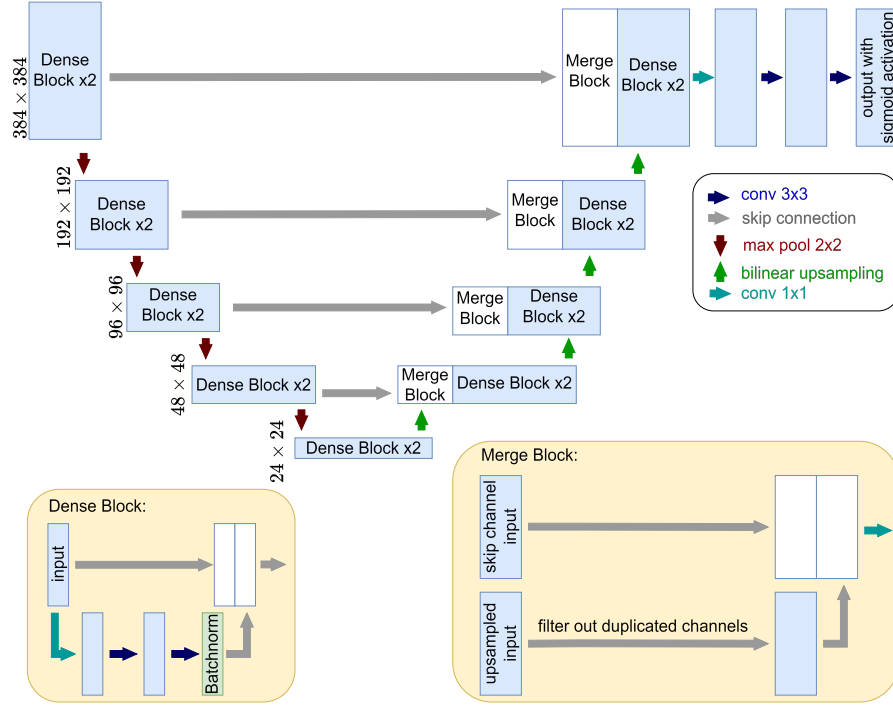
A landmark pontok azonosítása és egymásra igazítása után az alakmodellt az igazított pontokból (az utolsó iterációban kapott x_i -kből) lehet felépíteni a 3.2.1 szakaszban leírtak szerint. Továbbá súlyokat rendelünk a landmark pontokhoz a referencia szegmentációk bizonytalanságának modellezéséhez (W), a szív oldalsó pontjai nagyobb súlyokat kapnak, mert ezen régiók pontosabban lokalizálhatók, mint a körvonal felső és alsó részei, továbbá ezek pontos azonosítása a gyakorlati alkalmazás során is fontosabb feladat. Az ily módon felépített modellben összesen 20 kontrollpont került kijelölésre. Ennyi pont alkalmas a szívárnyék körvonal megfelelő minőségű rekonstrukciójára, valamint hatékonyan kezelhető a kontúrpointok lokalizációjának és a belőlük alkotott prediktált körvonalak optimalizációjának szemszögéből.

3.3. Modell inicializáció

A szegmentáló modell inicializációs lépése a sarokpontok neurális hálózat általi becsléséből, majd a 3.2.2 szakaszban leírt algoritmus használatából áll a többi kontúrpoint rekonstruálásához. Az inicializációt szemléletesen a 3.2 ábra mutatja be.

3.3.1. Sarokpontok becslése

A szív sarokpontjainak becslésére használt neurális hálózat egy DenseUNet által inspirált UNet-szerű architektúra, amely Dense blokkokból épül fel [22] [23]. Itt globális feature csatornák segítik a hatékony visszaterjesztést, és lehetővé teszik, hogy a Dense blokkok több korábbi réteg aktivációit használhassák fel, így jelentősen növelve a hálózat teljesítményét. Azonban, ellentétben a DenseNet-tel, ezt a hálózatot szegmentációra használjuk, nem pedig feature extrakcióra, így a hálózat dekódoló ágán található blokkok aktivációinak felbontása is növekszik a csatornák számával együtt. Összességében ez azt jelenti, hogy az aktivációs térképek mérete a hálózat minden szintjével köbösen növekszik, túlságosan erőforrás-igényessé téve az architektúrát. Ennek megoldására 1x1-es konvolúciós rétegeket használunk minden upsampling réteg előtt, hogy csökkentsük a globális feature csatornák számát. Ez néhány blokk kimenetének teljes kiszűréséhez vezethet, ám ezt a gyakorlatban nem találtuk problémának. Az architektúra a 3.1 ábrán látható. A sarokpontok becslése esetén a hálózat egyetlen bemeneti csatornával rendelkezik, amelyen a normalizált input felvétel található, és négy kimeneti csatornával, amelyek a sarokpontoknak felelnek meg.



3.1. ábra. Az alkalmazott DenseUNet architektúra.

A hálózat kimeneti csatornáit a kontúrponatok elhelyezkedésének sűrűségfüggvényeként értelmezzük, melynek a várható értékét (tömegközéppontját) vesszük következtetés során. Ennek motivációja, hogy robusztusan értelmezzük a kontúrponokat, esetleges, néhány pixelre kiterjedő hallucinációk ne vezethessék félre a predikciót. A hálózat betanításához a ground truth pontokat Gauss-kernellel moszuk el fix varianciával. A módszer másik előnye, hogy a következtetés során mért variancia megfigyelésével értékelni lehet a becslés bizonyosságát.

3.3.2. Alakzat rekonstrukciója sarokpontokból

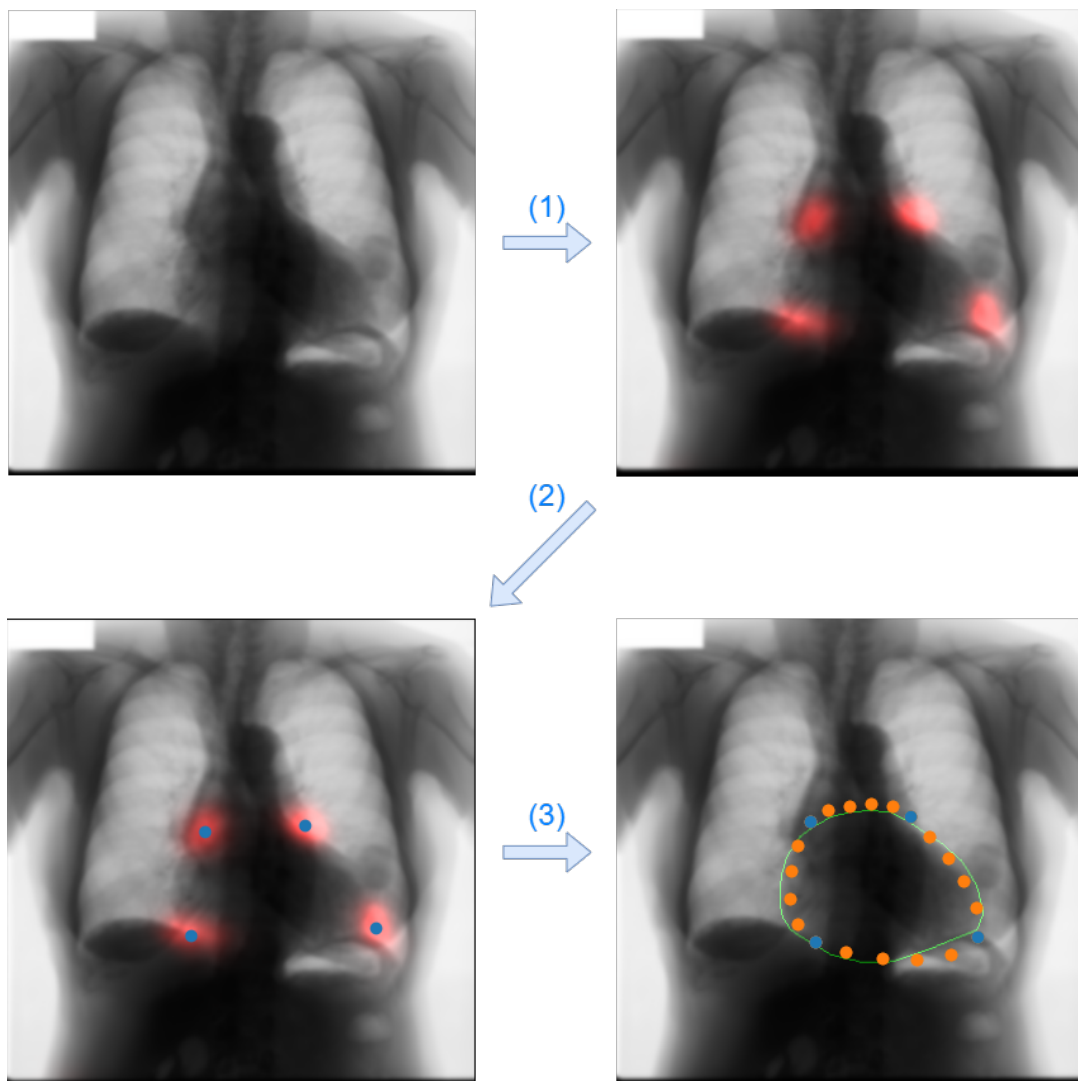
A szív sarokpontjainak becslése után meghatározzuk egy olyan affin transzformáció paramétereit, amelyek a legjobban igazítja a kapott pontokat (x_A) az átlagos modell sarokpontjaihoz ($\overline{x_A}$) a kettő közötti euklideszi távolság négyzetének szempontjából:

$$f = \min_{f_{M,O}(x_A) = Mx_A^T + O} \|f_{M,O}(x_A) - \overline{x_A}\|_2^2 \quad (3.17)$$

A transzformáció szabadságfokai itt is ugyanúgy korlátozva vannak, ahogy a 3.2.4 fejezetben is (M mátrix csak elforgatás és skálázás műveletek kompozíciójából áll, O vektor pedig eltolást valósít meg). A sarokpontokhoz tartozó különbségvektorokat úgy számítjuk ki, hogy az ily módon transzformált prediktált sarokpontokból kivonjuk az átlagos sarokpontokat. $v_A = f(x_A) - \overline{x_A}$ Eképpen megkapjuk a (3.11) egyenletbeli v_A paramétert, v_B -t pedig a korábban levezetett képlet alapján számíthatjuk:

$$v_B = -(W_B^T P_B^T \Lambda^{-1} P_B W_B + sI)^{-1} W_B^T P_B^T \Lambda^{-1} P_A W_A v_A \quad (3.18)$$

v_A és v_B összefűzésével $v = v_A|v_B$, valamint a modell átlagának hozzáadásával az alakzatot a modellezési térben rekonstruáljuk $x = v + \bar{x}$. Végezetül az predikciók és az alakmodell egymásra igazításhoz használt transzformáció inverzával vissza-transzformáljuk az alakzatot képtérbe ($\Xi = f^{-1}(x)$).



3.2. ábra. Az inicializáció menetét bemutató ábra. Először (1) a sarokpontokat becsüljük 3.3.1 fejezet szerint egy neurális hálózattal, majd ezeknek vesszük a tömegközéppontját (2), amiket az ábrákon kék pontokkal jelöljük. Végezetül a sarokpontokból rekonstruáljuk a többi kontúrponthelyezkedését, az (3.11) egyenlet segítségével (3), melyeket narancssárga pontok jelölnek, a ground truth kontúrt zölddel ábrázoljuk.

3.4. Szegmentáló modell finomító lépése

A 3.3 fejezetben meghatározott inicializáció után kapott alakmodell iteratív javítása két lépésből áll. Először egy neurális hálózat becsüli meg, hogy hogyan kellene

mozgatni a landmark pontokat egy jobban illeszkedő szegmentáció érdekében, majd a robusztusság, az alaki kényszerek, valamint további jólformáltsági követelmények biztosítása numerikus optimalizációval történik ADMM [24] alkalmazásával.

3.4.1. Mély neurális hálózat alapú finomítás

A kontúrponthoz finomítására használt neurális hálózat architektúrája azonos a sarokpontok becslésére használt hálózattal, leszámítva a bemeneti és kimeneti csatornák számát. Itt minden landmark ponthoz egy megfelelő bemeneti és kimeneti csatorna tartozik. Egy landmark ponthoz tartozó bemenet az eredeti kép, amely a pont által jelölt pozíció körül van maszkolva. A elvárt kimenet ismét egy sűrűségfüggvény, amely a pont új helyét adja meg, a végső becslés pedig a becsült eloszlás tömegközéppontja. Egy másik lehetőség a maszkolás helyett cropping alkalmazása a pont körül, azonban azt találtuk, hogy a maszkolás használata pontosabb predikcióhoz vezet, valamint segíti az adathalmazok közötti általánosítást. Továbbá azt is tapasztaltuk, hogy előnyös egyetlen hálózatot használni több csatornával az összes kontúrponthoz egyidejű becslésére, mintsem külön-külön tanítani az egyes pontokhoz tartozó javításokat végző hálókat, mivel így a hálózat képes a maszkolt bemenetekből rekonstruálni a kép nagy részét, és megtanulni a landmark pontokat lokalizáló közös és egyedi feature-öket. A kontúrponthoz egyidejű becslésekor nehézséget okozhat, hogy a pontbecslő neurális hálózat a kontúrponthoz eloszlására egy implicit alakmodellt tanul, amely nem fog jól illeszkedni az általunk felépített ASM modellhez. A problémát a becsült pontok lokális lineáris reinterpolációjával korrigáljuk az optimalizáció során.

3.4.2. Optimalizáció

A javítóvektorok kiszámítása után a hosszúságukat leskálázzuk, a robusztusabb működés érdekében. A predikciók megbízhatósági értékeit a következő módon számítjuk ki: (1) Minden landmark pontra kiszámítjuk az eloszlás sűrűségfüggvényének varianciáját, amit a finomításért felelős neurális hálózat ad (σ^2 vektor). (2) Figyelembe vesszük, hogy mennyire konzisztensek a predikciók a landmark pontokat az alakmodell látens paraméterein keresztül vizsgálva. Ehhez a 3.2.2 szakaszban leírt rekonstrukciós algoritmust használjuk, ahol minden landmark pont helyét a többi pont helyzetéből becsüljük újra a (3.11) egyenlet segítségével, és megmérjük ezek távolságát a neurális hálózat becslésének helyétől (τ vektor). Ezek alapján a megbízhatósági értékeket a landmark pontok és az újrabecsült pozíciók közötti távolság függvényében számítjuk. Az (1) és (2) megbízhatósági értékeket normalizáljuk, hogy valószínűségeket reprezentáljanak, majd az egyes valószínűségi értékeket fuzionáljuk. Az egyes eredmények normalizálásának és fúziójának menetét a (3.19) egyenlet mutatja be, ahol a vektorokra az osztás és szorzás műveletek elemenkénti operátorokként vannak értelmezve. A következő részben ezt a megbízhatósági vektort h -val jelöljük, továbbá $H = \text{diag}(h)$.

$$\begin{aligned}
\sigma_{norm}^2 &= \sigma^2 / \sum_i \sigma_i^2 \\
\tau_{norm} &= 1 - (\tau - \min(\tau)) / (\max(\tau) - \min(\tau)) \\
h_{joint} &= \sigma_{norm}^2 + \tau_{norm} \\
h &= h_{joint} / \sum_i h_{joint}^i \cdot \dim(h_{joint})
\end{aligned} \tag{3.19}$$

Az optimalizáció során feltételezzük, hogy az alakmodellünk optimális paraméterezése "közel" van a kontúrhoz (amit annak sokszögnek a kerülete definiál, amelynek csúcsai a neurális hálózat által megadott landmark pontok), valamint a likelihoodja elegendően nagy az alakmodell látens paramétereinek eloszlásában:

$$\begin{aligned}
\min . \quad & 0.5 \cdot d^T H d + 0.5 \cdot l \cdot b^T \Lambda b \\
s.t. \quad & d = W^{-1} P b + \bar{x} - \Phi(s) \\
& |b| \leq \kappa \sqrt{\lambda} \\
& s \in [-1, 1]^N \\
& \forall j \in [1, N/2] : s_{2 \cdot j - 1} = s_{2 \cdot j}
\end{aligned} \tag{3.20}$$

ahol κ és l a látens paraméterek regularizálásáért felel, κ konstans, l -t pedig az iterációk folyamán exponenciálisan ütemezzük. Mivel a javítópontok becslése egy neurális hálózat által, egy lépésben történik, előfordulhat, hogy ezen hálózat egy implicit alakmodellt tanul. Ennek korrigálása érdekében vezetjük be a Φ leképezést ($\Phi : \mathbb{R}^N \rightarrow \mathbb{R}^N$), amely interpolál a kontúr szomszédos kontrollpontjai között:

$$\Phi(s)_i = \begin{cases} s_i \cdot \Xi_i + (1 - s_i) \cdot \Xi_{\alpha(i+2, N)} & s_i \in [0, 1] \\ (-s_i) \cdot \Xi_i + (1 + s_i) \cdot \Xi_{\alpha(i-2, N)} & s_i \in [-1, 0] \end{cases} \tag{3.21}$$

ahol $s \in \mathbb{R}^N$ és $s_{2 \cdot j - 1} = s_{2 \cdot j}$ minden $j \in [1, N/2]$ esetén, Ξ_i a neurális hálózat predikciójának i -edik koordinátája és $\alpha(i, N) = i \bmod N$. Továbbá N a kontúrpontok számának kétszerese.

Egy másik feltételezésünk, hogy a rekonstruált alakzat sima, azaz alacsony görbülettel rendelkezik, és a kontrollpontok egyenletesen helyezkednek el a kontúr mentén:

$$\begin{aligned}
\min . \quad & 0.5 \cdot k_1 q^T D_1^T D_1 q + 0.5 \cdot k_2 q^T D_2^T D_2 q \\
s.t. \quad & q = W^{-1} P b + \bar{x}
\end{aligned} \tag{3.22}$$

ahol D_1 és D_2 az első- és másodrendű derivált operátorok mátrixai (3.23), k_1 és k_2 a regularizációs tagok súlyozására használt konstansok, melyeket lineárisan ütemezünk az iterációk során, q pedig a rekonstruált alakzatot reprezentálja.

$$D_1 = \begin{bmatrix} -1 & 0 & 1 & 0 & \dots & 0 \\ 0 & -1 & 0 & 1 & 0 & \dots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 1 & 0 & \dots & 0 & -1 & 0 \\ 0 & 1 & 0 & \dots & 0 & -1 \end{bmatrix} \quad (3.23)$$

$$D_2 = \begin{bmatrix} -2 & 0 & 1 & 0 & \dots & 0 & 1 & 0 \\ 0 & -2 & 0 & 1 & 0 & \dots & 0 & 1 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 1 & 0 & \dots & 0 & 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & \dots & 0 & 1 & 0 & -2 \end{bmatrix}$$

Ennek a több konvex részfeladatból álló problémának a hatékony megoldása érdekében az Alternating Direction Method of Multipliers (ADMM [24]) optimalizációs eljárást alkalmazzuk, amely ötvözi a dekomponálhatóságot és a gyors kezdeti konvergencia tulajdonságokat. Az ADMM hatékony alkalmazásához bevezetünk egy új látens paraméter változót b -hez (ezt jelölve a -val). Ennek segítségével a -t a likelihood függvény regularizációjára alkalmazzuk (3.20)-ben, b pedig egy konszenzus változó lesz, melyet az alakzat rekonstrukciójára és regularizációjára használunk mind (3.20)-ben, mind (3.22)-ben. Ez lehetővé teszi a részfeladatok optimumának analitikus kifejezését. Az így adódó teljes optimalizálási probléma és a hozzá tartozó skálázott augmentált Lagrange duális ($\mathcal{L}$) a következők:

$$\begin{aligned} \min. \quad & \frac{1}{2}(\|d\|_H^2 + l \cdot \|a\|_\Lambda^2 + k_1\|D_1q\|_2^2 + k_2\|D_2q\|_2^2) \\ \text{s.t.} \quad & d = W^{-1}Pb + \bar{x} - \Phi(s) \\ & a = b \\ & q = W^{-1}Pb + \bar{x} \\ & |b| \leq \kappa\sqrt{\lambda} \end{aligned} \quad (3.24)$$

$$\begin{aligned} \mathcal{L} = & \frac{1}{2}(\|d\|_H^2 + l \cdot \|a\|_\Lambda^2 + k_1\|D_1q\|_2^2 + k_2\|D_2q\|_2^2 \\ & + \rho_1\|d - W^{-1}Pb - \bar{x} + \Phi(s) + u_1\|_2^2 - \rho_1\|u_1\|_2^2 \\ & + \rho_2\|b - a + u_2\|_2^2 - \rho_2\|u_2\|_2^2 \\ & + \rho_3\|q - W^{-1}Pb - \bar{x} + u_3\|_2^2 - \rho_3\|u_3\|_2^2) \\ & + I_\infty(|b| > \kappa\sqrt{\lambda}) \end{aligned} \quad (3.25)$$

ahol u_1 , u_2 és u_3 a skálázott duális változók, ρ_1 , ρ_2 , ρ_3 a büntetési paraméterek, esetünkben 10^{-3} -ra rögzítve. Az $I_\infty(\chi)$ egy olyan indikátorfüggvényt jelöl, melynek értéke végtelen, ha χ igaz, különben 0.

A (3.25)-es skálázott augmentált Lagrangian függvény minimalizálását szekvenciálisan végezzük a változók szerint (egy adott változó frissítése során a többit aktuális értékeikkel helyettesítjük). Ekkor a változók frissítési sorrendje: $d, b, a, s, q, u_1, u_2, u_3$. Minden frissítési lépés optimuma analitikusan meghatározható,

és szükséges lebegőpontos számítási műveletek száma is minden lépésben alacsony marad (a legmagasabb költségű számítás az $N \times N$ mátrixok inverzének kiszámítása az optimalizáció során) [24]. A primál és duál változók frissítését az (3.26) egyenlet formalizálja.

$$\begin{aligned}
d &\leftarrow \rho_1(H + \rho_1 I)^{-1}(W^{-1}Pb + \bar{x} - \Phi(s) - u_1) \\
b &\leftarrow ((\rho_1 + \rho_3)P^T W^{-2}P + \rho_2 I)^{-1}(\rho_1 P^T W^{-1}(\Phi(s) + u_1 + d - \bar{x}) + \\
&\quad + \rho_2(b - u_2) + \rho_3 P^T W^{-1}(q + u_3 - \bar{x})) \\
a &\leftarrow \min(\kappa\sqrt{\lambda}, \max(-\kappa\sqrt{\lambda}, (\Lambda^{-1} \odot l + \rho_2 I)^{-1}\rho_2(a + u_2))) \\
s &\leftarrow \underset{s}{\operatorname{argmin}}(||d - W^{-1}Pb - \bar{x} + \Phi(s) + u_1||) \\
q &\leftarrow \rho_3(k_1 D_1^T D_1 + k_2 D_2^T D_2 + \rho_3 I)^{-1}(W^{-1}Pb + \bar{x} - u_3) \\
u_1 &\leftarrow u_1 + d - W^{-1}Pb - \bar{x} + \Phi(s) \\
u_2 &\leftarrow u_2 + b - a \\
u_3 &\leftarrow u_3 + q - W^{-1}Pb - \bar{x}
\end{aligned} \tag{3.26}$$

A duális változókat 0-ra inicializáljuk, míg az primál változókat a következő módon állítjuk be:

$$d \leftarrow 0; \quad b \leftarrow P^T W(\Xi - \bar{x}); \quad a \leftarrow 0; \quad s \leftarrow 0; \quad q \leftarrow \Xi \tag{3.27}$$

Az optimalizációt 500 iteráción keresztül futtatjuk, majd a végén a kimeneti alakzatot az a -ból primál változóból a következő módon rekonstruáljuk:

$$x \leftarrow W^{-1}Pa + \bar{x} \tag{3.28}$$

Egy felvétel szegmentálása során összesen 100 finomítási lépést hajtunk végre (beleértve mind a neurális hálózat predikcióit, mind az imént leírt optimalizációt). A regularizációs együtthatókat (l , k_1 , k_2) fokozatosan ütemezzük a finomítási lépések során.

3.5. Modell tanítása

Az alakmodell létrehozásához és a neurális hálózatok betanításához a JSRT adathalmazt [16] használtuk, amely 247 analóg PA mellkasröntgen felvételtől és manuálisan létrehozott szívárnyék maszkból áll. Fontos megjegyezni, hogy az adatkészlet nem tartalmazza olyan betegek felvételeit, akik bizonyos, a szegmentálást nehezítő elváltozásokkal rendelkeznek (pl. tüdőödéma, aszimmetrikus rekeszizom, pleurális folyadékgyülem, kardiomegália), továbbá megfelelő a páciensek pozicionálása is a képeken. Az adathalmazt 200 felvételtől álló tanító, és 47 felvételtől álló validációs halmazokra osztottuk. Az alakmodellt a tanító halmaz ground truth maszkjaiból hoztuk létre (lásd 3.2.4), összesen 20 landmark pontot kialakítva ($N = 40$ az (3.21)-ben). A sarokpontok becsléséhez használt neurális hálót tanítása során a ground truth értékeket az alakmodell sarokpontjai képezték. A landmark pontok javítására használt neurális hálózatot is hasonló módon tanítottuk, azonban a bemenet esetén véletlenszerű Gauss-zajt adtunk minden kontrollpont pozíciójához, hogy szimulál-

juk a következtetés során előforduló hibákat. Mindkét esetben azonos előfeldolgozási és augmentációs transzformációkat alkalmaztunk a bemeneten: az bemeneti felvétel min-max normalizálása a hisztogram 3. és 97. percentilisei alapján és Gauss elmosás az előfeldolgozáshoz, majd véletlenszerű affin transzformációk és az intenzitások átskálázása az augmentációhoz.

4. fejezet

Kiértékelés

A modelleket a JSRT adathalmaz 47 felvételt tartalmazó teszt részén, valamint egy privát adathalmazon értékeltük ki, amely 22 mindennapi klinikai rutin során rögzített felvételből állt. Az új adathalmaz használatának motivációja az volt, hogy ezek a minták biztosan más eloszlásból származnak, mivel digitális röntgen felvételek, továbbá magasabb varianciát mutatnak a képminőség, megfigyelhető elváltozások és páciensek beállítási pontatlanságainak tekintetében. Egy ilyen, kevésbé homogén mintahalmaz jobban képes felmérni az egyes megoldások általánosítási képességét. A szegmentációk minőségének értékelésére Sørensen–Dice metrikát alkalmaztunk, amely a modellek kimenetét pixel-szintű osztályozóként értelmezi. Továbbá Hausdorff-távolságot is használtunk a prediktált és a ground truth maszkok közötti különbség mérésére annak érdekében, hogy értékeljük a módszerek robusztusságát.

4.1. Alkalmazott kiértékelési módszerek

Amint a rendelkezésre álló adathalmazokból látható, szegmentált szív annotációval mindössze két adathalmaz van ellátva, ráadásul azok is csupán 247, illetve 22 darab mintából állnak. Ebből kifolyólag az általánosítóképesség és robusztusság aspektusában az eredményeket többnyire kvalitatívan értékeljük. A tanítás folyamán, illetve amikor lehetőség van kvantitatív kiértékelésre, az egyes megoldásokat a Sørensen–Dice metrika segítségével minősítjük, ahol a szegmentációs eljárások kimeneti maszkjait pixelszintű osztályozásként értelmezzük, valószínűségek esetén az értékeket 0.5-ös szinten küszöbölve (4.1).

$$DSC = \frac{2TP}{2TP + FN + FP} \quad (4.1)$$

A választás oka, hogy valóban jól jellemzi a szegmentáció minőségét, összhangban áll a helyes kimenetre vonatkoztatott elvárásainkkal, továbbá minősítésre való használata elterjedt a szakirodalomban, így lehetővé téve az eredmények „state-of-the-art”-al való összehasonlítását is. A szegmentációk minőségének egy másik mérőszáma a Hausdorff távolság volt, mely a ground truth és prediktált maszkok pontjai közötti legnagyobb euklideszi távolságot veszi figyelembe.

$$Hd(X, Y) = \max \left\{ \sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y) \right\} \quad (4.2)$$

A metrika használatának motivációja, hogy segítségével a szegmentáció pontosságán felül annak robusztusságát is mérhetővé teszi.

4.2. Abláció, egyes paraméterek hatása a modell teljesítményére

Az ismertetett szegmentáló eljárás meglehetősen sok szabad paramétert, kiegészítő algoritmust használ a modell teljesítményének növeléséhez. A fejezet folyamán először hiperparaméter optimalizációt hajtunk végre a modell paraméterein, majd az egyes komponenseket külön-külön vizsgáljuk. Minden egyes komponens hatását úgy becsüljük, hogy az adott komponens kikapcsolása mellett a megmaradt hiperparamétereket újra optimalizáljuk. Amennyiben az adott komponens paraméterei folytonos értékűek, akkor csak azok változtatásával is megvizsgáljuk a eredmény módosulását, érzékenységi vizsgálatot végrehajtva. Az ablációs tanulmány során a következő komponenseket vizsgáljuk (a jelöléseket a (3.24) egyenletben definiáltuk):

1. Az optimalizációs feladat l likelihoodot súlyozó regularizációs paramétere és az iterációk folyamán történő ütemezése.
2. Az optimalizációs feladat k_1 és k_2 kimeneti ponthalmaz jólformáltságaért felelős tagjait (folytonosság és görbület) súlyozó paraméterei és az iterációk folyamán való ütemezésük.
3. A modell inicializációjáért felelős sarokpont becslésből teljes körvonalat származtató maximum-likelihood alapú eljárás hatása, átlagmodell illesztés helyett.
4. Szakértői tudószegmentáló eljárás alkalmazása az előfeldolgozás során a felvételek célterületre való vágásához.
5. Az optimalizációs feladat látens paraméterekre vonatkoztatott elemenkénti κ felső korlát alapú regularizációjának hatása.
6. Az optimalizációs feladat augmented lagrangian felírásában ρ_1, ρ_2, ρ_3 paraméterek meghatározása az eljárás gyors konvergenciájához.
7. Az optimalizációs feladat rekonstrukciós tagjában H konfidenciát reprezentáló súlymátrix becslése.
8. Az optimalizációs feladatban a neurális hálózat által becsült ponthalmaz újraparaméterezésének hatása (Φ függvény).
9. Tisztán neurális baseline modell, amely csak a sarokpontokat becslő, és javító pontokat adó neurális hálózatokat használja a szegmentáció során.

4.2.1. Optimális hiperparaméterek megtalálása

A hiperparaméter optimalizáció folyamán az egyes paraméterek értékét véletlenszerűen választottuk egy megfelelő tartományból. A tartományok meghatározásához

érzékenységvizsgálatot alkalmaztunk a folytonos paraméterekre (melyek eredményei a következő alfejezetekben ismertetésre kerülnek). A kiindulási fix paraméterkonfigurációt tapasztalati úton határoztuk meg a modellel való kísérletezés során, mely alapján a következő beállításokkal kaptuk a legjobb eredményeket: az optimalizációs probléma l változóját exponenciálisan ütemeztük 10^{-1} -től 10^{-3} -ig, a k_1 és k_2 változókat lineárisan ütemeztük 0-tól 0.03, valamint 0.05-ig, κ elemenkénti likelihood korlátot szabályozó paramétert pedig 2-nek rögzítettük. Így a modell az egyes adathalmazokon a következő eredményeket produkálta:

	JSRT		Privát adathalmaz	
	Dice Score	Hausdorff	Dice Score	Hausdorff
Kezdeti eredmények	0.940 ± 0.029	$10.34 \text{ mm} \pm 5.39 \text{ mm}$	0.892 ± 0.036	$18.96 \text{ mm} \pm 6.51 \text{ mm}$

4.1. táblázat. A kiindulási paraméterkonfiguráció eredményei az egyes adathalmazokon.

Az érzékenységvizsgálat elvégzése után meghatároztuk az egyes paraméterek értékeire vonatkozó lehetséges intervallumokat, majd ezek alapján hajtottuk végre a hiperparaméter optimalizációt, mely eredményei a 4.2 táblázatban láthatók. A továbbiakban helytakarékoság érdekében (ahol ezt explicit nem definiáljuk felül), l mindig exponenciálisan van ütemezve, ahol a kezdeti érték mindig 10^{-1} , a célértéket pedig a táblázatban közöljük. Hasonlóan járunk el a k_1 , k_2 változókkal, melyeket lineárisan ütemezünk 0-tól a célértékig. Összességében elmondható, hogy a JSRT adathalmazon a további hiperparaméter optimalizációval nem sikerült számottevő javulást elérni, inkább az eredmények szórása csökkent, a privát adathalmazon viszont minden metrikában sikerült jelentősen meghaladni a kezdeti paraméterkonfigurációval elért eredményeket.

	JSRT		Privát adathalmaz	
	Dice Score	Hausdorff	Dice Score	Hausdorff
$l: 10^{-2.311}, \kappa: 2.980, k_1: 0.111, k_2: 0.284$	0.939 ± 0.025	$10.30 \text{ mm} \pm 4.82 \text{ mm}$	0.893 ± 0.042	$16.49 \text{ mm} \pm 6.36 \text{ mm}$
$l: 10^{-2.116}, \kappa: 2.712, k_1: 0.166, k_2: 0.127$	0.937 ± 0.024	$10.43 \text{ mm} \pm 4.70 \text{ mm}$	0.888 ± 0.047	$16.49 \text{ mm} \pm 5.69 \text{ mm}$
$l: 10^{-2.779}, \kappa: 2.203, k_1: 0.152, k_2: 0.014$	0.940 ± 0.027	$10.26 \text{ mm} \pm 5.06 \text{ mm}$	0.893 ± 0.036	$18.27 \text{ mm} \pm 6.58 \text{ mm}$
$l: 10^{-2.733}, \kappa: 2.131, k_1: 0.148, k_2: 0.094$	0.940 ± 0.027	$10.29 \text{ mm} \pm 5.05 \text{ mm}$	0.893 ± 0.036	$18.09 \text{ mm} \pm 6.51 \text{ mm}$
$l: 10^{-2.591}, \kappa: 2.316, k_1: 0.271, k_2: 0.107$	0.939 ± 0.028	$10.32 \text{ mm} \pm 5.02 \text{ mm}$	0.895 ± 0.036	$16.96 \text{ mm} \pm 6.71 \text{ mm}$
$l: 10^{-2.022}, \kappa: 1.621, k_1: 0.057, k_2: 0.268$	0.936 ± 0.025	$10.63 \text{ mm} \pm 4.82 \text{ mm}$	0.887 ± 0.048	$16.50 \text{ mm} \pm 5.37 \text{ mm}$
$l: 10^{-2.843}, \kappa: 2.525, k_1: 0.053, k_2: 0.285$	0.939 ± 0.028	$10.43 \text{ mm} \pm 5.24 \text{ mm}$	0.891 ± 0.036	$19.03 \text{ mm} \pm 6.83 \text{ mm}$
$l: 10^{-2.805}, \kappa: 1.539, k_1: 0.237, k_2: 0.270$	0.940 ± 0.027	$10.30 \text{ mm} \pm 5.01 \text{ mm}$	0.893 ± 0.039	$17.53 \text{ mm} \pm 6.20 \text{ mm}$
$l: 10^{-3.250}, \kappa: 2.943, k_1: 0.251, k_2: 0.229$	0.939 ± 0.028	$10.68 \text{ mm} \pm 5.02 \text{ mm}$	0.892 ± 0.038	$19.48 \text{ mm} \pm 7.69 \text{ mm}$
$l: 10^{-3.497}, \kappa: 2.550, k_1: 0.066, k_2: 0.105$	0.939 ± 0.029	$10.52 \text{ mm} \pm 5.40 \text{ mm}$	0.887 ± 0.041	$20.87 \text{ mm} \pm 7.93 \text{ mm}$
$l: 10^{-3.574}, \kappa: 1.692, k_1: 0.173, k_2: 0.138$	0.940 ± 0.028	$10.51 \text{ mm} \pm 5.21 \text{ mm}$	0.884 ± 0.058	$21.15 \text{ mm} \pm 10.05 \text{ mm}$
$l: 10^{-3.958}, \kappa: 1.301, k_1: 0.228, k_2: 0.094$	0.940 ± 0.028	$10.37 \text{ mm} \pm 5.15 \text{ mm}$	0.884 ± 0.060	$20.02 \text{ mm} \pm 10.12 \text{ mm}$
$l: 10^{-3.356}, \kappa: 1.623, k_1: 0.045, k_2: 0.118$	0.940 ± 0.029	$10.40 \text{ mm} \pm 5.40 \text{ mm}$	0.892 ± 0.037	$19.01 \text{ mm} \pm 6.41 \text{ mm}$
$l: 10^{-2.447}, \kappa: 3.035, k_1: 0.240, k_2: 0.227$	0.940 ± 0.025	$10.21 \text{ mm} \pm 4.82 \text{ mm}$	0.894 ± 0.039	$16.59 \text{ mm} \pm 6.50 \text{ mm}$
$l: 10^{-1.582}, \kappa: 1.066, k_1: 0.125, k_2: 0.199$	0.933 ± 0.027	$11.00 \text{ mm} \pm 4.86 \text{ mm}$	0.880 ± 0.049	$16.96 \text{ mm} \pm 4.65 \text{ mm}$
$l: 10^{-1.955}, \kappa: 2.955, k_1: 0.253, k_2: 0.015$	0.936 ± 0.025	$10.65 \text{ mm} \pm 4.67 \text{ mm}$	0.885 ± 0.047	$16.46 \text{ mm} \pm 5.10 \text{ mm}$

4.2. táblázat. A szegmentációs eljárás hiperparaméter optimalizációjának eredménye különböző paraméter konfigurációk mellett.

4.2.2. Likelihood regularizációs tag hatása

Az alakmodellel történő regularizáció használatának kulcskérdése a rekonstrukciós és likelihood veszteségek összeegyeztetése. Túlságosan gyenge likelihood regularizáció

esetén (l értéke kicsi) az alakmodell nincs hatással az eljárás kimenetére még out-of-domain alakzatok esetén sem, míg ellenkező esetben a látens paraméterek 0 közeli értéket kapnak, a kimeneti szegmentáció pedig az átlagmodellhez tart, csökkentve a szegmentáció minőségét. Empirikusan a javítópontokat becsülő neurális hálózat jellemzően a kezdeti iterációkban ad kevésbé pontos predikciókat, hajlamosabb nem megfelelő kontúrt követni, így ezen fázisban erősebb regularizáció szükséges, mint a szegmentáció végén. A paraméter értéke így nem konstans az iterációk során, hanem dinamikusan ütemezzük azt, a szegmentáció előrehaladásával gyengítve a regularizációt. A 4.3 táblázatban látható, hogy konstans értékű paraméter helyett, azt exponenciálisan ütemezve jelentős teljesítménynövekedést sikerült elérni. Amennyiben a regularizációs tagot kikapcsoljuk, és így végzünk hiperparaméter optimalizációt az eljárás többi tagjára, a modell teljesítménye jelentősen visszaesik az irreguláris eseteket is tartalmazó privát adathalmazon, ahogy az a 4.4 táblázatban is látható.

	JSRT		Privát adathalmaz	
	Dice Score	Hausdorff	Dice Score	Hausdorff
Konstans l , 10^{-4}	0.940 ± 0.030	$10.43 \text{ mm} \pm 5.63 \text{ mm}$	0.877 ± 0.061	$23.07 \text{ mm} \pm 10.83 \text{ mm}$
Konstans l , 10^{-3}	0.940 ± 0.028	$10.40 \text{ mm} \pm 5.31 \text{ mm}$	0.882 ± 0.060	$20.84 \text{ mm} \pm 10.06 \text{ mm}$
Konstans l , 10^{-2}	0.935 ± 0.025	$10.68 \text{ mm} \pm 4.81 \text{ mm}$	0.877 ± 0.068	$17.84 \text{ mm} \pm 8.22 \text{ mm}$
Konstans l , 10^{-1}	0.931 ± 0.027	$11.17 \text{ mm} \pm 4.98 \text{ mm}$	0.877 ± 0.050	$17.41 \text{ mm} \pm 4.57 \text{ mm}$
Exponenciális l , $10^{-1} - 10^{-4}$	0.939 ± 0.030	$10.64 \text{ mm} \pm 5.61 \text{ mm}$	0.887 ± 0.040	$21.10 \text{ mm} \pm 8.06 \text{ mm}$
Exponenciális l , $10^{-1} - 10^{-3}$	0.940 ± 0.029	$10.34 \text{ mm} \pm 5.39 \text{ mm}$	0.892 ± 0.036	$18.96 \text{ mm} \pm 6.51 \text{ mm}$
Exponenciális l , $10^{-1} - 10^{-2}$	0.936 ± 0.025	$10.67 \text{ mm} \pm 4.82 \text{ mm}$	0.887 ± 0.046	$16.49 \text{ mm} \pm 5.31 \text{ mm}$

4.3. táblázat. Az egyes likelihood regularizációs paraméterkonfigurációk eredményei. Jól látható, hogy a megfelelően beállított exponenciális ütemezés túlszárnyalja a konstans beállítások eredményeit, vagy csak minimálisan marad el azoktól.

	JSRT		Privát adathalmaz	
	Dice Score	Hausdorff	Dice Score	Hausdorff
$l: 0, \kappa: 3.172, k_1: 0.017, k_2: 0.088$	0.939 ± 0.031	$10.46 \text{ mm} \pm 5.54 \text{ mm}$	0.875 ± 0.064	$23.95 \text{ mm} \pm 11.51 \text{ mm}$
$l: 0, \kappa: 3.130, k_1: 0.017, k_2: 0.021$	0.940 ± 0.031	$10.37 \text{ mm} \pm 5.64 \text{ mm}$	0.875 ± 0.066	$24.21 \text{ mm} \pm 11.66 \text{ mm}$
$l: 0, \kappa: 3.403, k_1: 0.249, k_2: 0.112$	0.939 ± 0.028	$10.42 \text{ mm} \pm 5.15 \text{ mm}$	0.881 ± 0.058	$22.46 \text{ mm} \pm 11.71 \text{ mm}$
$l: 0, \kappa: 1.521, k_1: 0.012, k_2: 0.014$	0.941 ± 0.030	$10.31 \text{ mm} \pm 5.61 \text{ mm}$	0.881 ± 0.061	$21.78 \text{ mm} \pm 10.16 \text{ mm}$
$l: 0, \kappa: 1.375, k_1: 0.228, k_2: 0.114$	0.941 ± 0.027	$10.36 \text{ mm} \pm 5.04 \text{ mm}$	0.884 ± 0.061	$20.41 \text{ mm} \pm 10.41 \text{ mm}$
$l: 0, \kappa: 1.510, k_1: 0.029, k_2: 0.060$	0.941 ± 0.030	$10.37 \text{ mm} \pm 5.54 \text{ mm}$	0.881 ± 0.061	$21.55 \text{ mm} \pm 10.12 \text{ mm}$
$l: 0, \kappa: 3.473, k_1: 0.054, k_2: 0.011$	0.940 ± 0.031	$10.42 \text{ mm} \pm 5.61 \text{ mm}$	0.875 ± 0.064	$23.92 \text{ mm} \pm 11.61 \text{ mm}$
$l: 0, \kappa: 1.660, k_1: 0.126, k_2: 0.196$	0.939 ± 0.029	$10.43 \text{ mm} \pm 5.30 \text{ mm}$	0.882 ± 0.060	$21.36 \text{ mm} \pm 10.36 \text{ mm}$

4.4. táblázat. Hiperparaméter optimalizáció eredménye kikapcsolt likelihood regularizáció mellett. Jól látható, hogy a komponens nélkül a modell teljesítménye jelentősen csökken, különösen a kevésbé homogén privát adathalmazon. A táblázatban l -re a 0 érték konstans értéket jelöl.

4.2.3. Jólformáltsági kényszerek hatása

Az optimalizáció folyamán nem csak az alakmodell látens paramétereit vesszük figyelembe a szegmentáció jóságának megítélésakor, hanem szakértői kényszereket is (a kontrollpontok mennyire egyenletesen helyezkednek el a kontúr mentén, valamint, hogy mekkora görbületű ívek jellemzik a kontúrt). Ezek regularizációjára vezetjük

be a k_1 és k_2 paramétereket, melyek súlyozzák a veszteség tagokat. A likelihood alapú regularizációhoz hasonlóan, itt sem konstans értékűek a paraméterek az iterációk folyamán, hanem lineárisan ütemezzük őket, a feldolgozás során folyamatosan erősítve azokat. Figyeljük meg, hogy ezen regularizáció ütemezése a likelihood alapúéval ellentétesen hat, melynek oka, hogy a kezdeti iterációkban az erős likelihood regularizáció biztosítja a jólformáltsági kritériumokat, így ezt tovább büntetni felesleges, míg az utolsó iterációk folyamán alacsonyabb likelihoodot megengedve a kimenetnek, ezen büntetések segítségével tudjuk biztosítani a kimenet jólformáltságát. A különböző paraméter konfigurációk mellett mért eredmények a 4.5 táblázatban találhatók, jól látható, hogy a paraméterek ütemezett erősítése növeli az eljárás teljesítményét, azonban ha azok értéke megfelelő tartományból van kiválasztva, az eljárás nem különösebben érzékeny rájuk. Amennyiben a komponenst kihagyjuk az optimalizációs feladatból, és így végzünk hiperparaméter optimalizációt a többi paraméterre, úgy a 4.6 táblázatban mért eredményeket kapjuk. Jól látható, hogy az eredmények nem változtak szignifikánsan a büntető tagok elhagyásának hatására, amely arra utal, hogy a már jelenlévő többi regularizációs tag (likelihood, látens paraméterek konfidencia szerinti vágása) képes kompenzálni az ilyen típusú jólformáltsági kényszerek hiányát. Miután ezen regularizációs tagok az alakmodellre építkeznek, az eredmények ennek jóságát is bizonyítják. Ennek ellenére érdemes lehet párhuzamosan alkalmazni jólformáltsági regularizációt is a modellben, növelve annak robusztusságát a kiértékelési halmazban sem szereplő outlier esetekre.

	JSRT		Privát adathalmaz	
	Dice Score	Hausdorff	Dice Score	Hausdorff
Konstans $k_1 = 0.03, k_2 = 0.05$	0.940 ± 0.029	$10.36 \text{ mm} \pm 5.35 \text{ mm}$	0.883 ± 0.060	$20.68 \text{ mm} \pm 9.99 \text{ mm}$
Konstans $k_1 = 0.015, k_2 = 0.025$	0.940 ± 0.029	$10.34 \text{ mm} \pm 5.41 \text{ mm}$	0.883 ± 0.060	$20.80 \text{ mm} \pm 10.07 \text{ mm}$
Konstans $k_1 = 0.06, k_2 = 0.1$	0.940 ± 0.028	$10.38 \text{ mm} \pm 5.31 \text{ mm}$	0.892 ± 0.037	$18.93 \text{ mm} \pm 6.53 \text{ mm}$
Lineáris $k_1 = 0.03, k_2 = 0.05$	0.940 ± 0.029	$10.54 \text{ mm} \pm 5.12 \text{ mm}$	0.893 ± 0.036	$18.41 \text{ mm} \pm 6.60 \text{ mm}$
Lineáris $k_1 = 0.06, k_2 = 0.025$	0.940 ± 0.028	$10.33 \text{ mm} \pm 5.27 \text{ mm}$	0.892 ± 0.037	$18.90 \text{ mm} \pm 6.57 \text{ mm}$
Lineáris $k_1 = 0.015, k_2 = 0.1$	0.940 ± 0.029	$10.33 \text{ mm} \pm 5.40 \text{ mm}$	0.892 ± 0.037	$19.04 \text{ mm} \pm 6.51 \text{ mm}$
Lineáris $k_1 = 0.06, k_2 = 0.1$	0.940 ± 0.028	$10.40 \text{ mm} \pm 5.33 \text{ mm}$	0.892 ± 0.037	$18.92 \text{ mm} \pm 6.51 \text{ mm}$
Lineáris $k_1 = 0.12, k_2 = 0.2$	0.940 ± 0.028	$10.48 \text{ mm} \pm 5.16 \text{ mm}$	0.892 ± 0.037	$18.85 \text{ mm} \pm 6.56 \text{ mm}$
Lineáris $k_1 = 0.5, k_2 = 0.5$	0.938 ± 0.025	$10.76 \text{ mm} \pm 4.60 \text{ mm}$	0.882 ± 0.059	$19.53 \text{ mm} \pm 9.06 \text{ mm}$

4.5. táblázat. Módszer érzékenysége különböző k_1, k_2 paraméter konfigurációk mellett azokat az iterációk során konstansnak véve, illetve lineárisan ütemezve 0-tól a cél értékig.

	JSRT		Privát adathalmaz	
	Dice Score	Hausdorff	Dice Score	Hausdorff
$l: 10^{-2.593}, \kappa: 1.971, k_1: 0.000, k_2: 0.000$	0.940 ± 0.026	$10.14 \text{ mm} \pm 4.94 \text{ mm}$	0.895 ± 0.039	$16.54 \text{ mm} \pm 6.64 \text{ mm}$
$l: 10^{-1.708}, \kappa: 2.728, k_1: 0.000, k_2: 0.000$	0.933 ± 0.026	$11.00 \text{ mm} \pm 4.89 \text{ mm}$	0.880 ± 0.049	$17.05 \text{ mm} \pm 4.71 \text{ mm}$
$l: 10^{-2.079}, \kappa: 1.933, k_1: 0.000, k_2: 0.000$	0.936 ± 0.025	$10.68 \text{ mm} \pm 4.74 \text{ mm}$	0.886 ± 0.047	$16.47 \text{ mm} \pm 5.39 \text{ mm}$
$l: 10^{-3.889}, \kappa: 2.480, k_1: 0.000, k_2: 0.000$	0.941 ± 0.030	$10.21 \text{ mm} \pm 5.53 \text{ mm}$	0.886 ± 0.041	$21.32 \text{ mm} \pm 8.48 \text{ mm}$
$l: 10^{-1.868}, \kappa: 3.431, k_1: 0.000, k_2: 0.000$	0.934 ± 0.026	$10.88 \text{ mm} \pm 4.87 \text{ mm}$	0.882 ± 0.048	$16.75 \text{ mm} \pm 4.74 \text{ mm}$
$l: 10^{-3.369}, \kappa: 1.495, k_1: 0.000, k_2: 0.000$	0.942 ± 0.028	$10.09 \text{ mm} \pm 5.24 \text{ mm}$	0.892 ± 0.038	$18.34 \text{ mm} \pm 6.35 \text{ mm}$
$l: 10^{-3.083}, \kappa: 3.047, k_1: 0.000, k_2: 0.000$	0.941 ± 0.028	$10.16 \text{ mm} \pm 5.18 \text{ mm}$	0.891 ± 0.039	$19.24 \text{ mm} \pm 7.52 \text{ mm}$
$l: 10^{-2.801}, \kappa: 0.741, k_1: 0.000, k_2: 0.000$	0.940 ± 0.026	$10.12 \text{ mm} \pm 4.86 \text{ mm}$	0.886 ± 0.047	$16.47 \text{ mm} \pm 5.30 \text{ mm}$

4.6. táblázat. Hiperparaméter optimalizáció eredménye kikapcsolt k_1 , és k_2 tagok mellett.

4.2.4. Inicializációs maximum likelihood becslés hatása

Az ASM eljárás inicializációja folyamán kezdetben csak a sarokpontokat lokalizáljuk, majd ezekből származtatjuk a kezdeti alakzatot maximum likelihood becslés felhasználásával (3.3.2 fejezet alapján). Egy egyszerűbb megoldás, ha kezdetben csak az átlagmodellt illesztjük a sarokpontokhoz Prokrusztész analízis szerint. Amennyiben így inicializáljuk a modellt, úgy a 4.7 táblázatban található eredményeket kapjuk. Mint látható, Az eljárás teljesítménye nem romlik számottevően, melynek a oka, hogy a ML becslés alapját képező 4 sarokpont nem hordoz számottevő információt a teljes alakzatról, ezáltal a Tikhonov regularizáció által kerülünk közelebb az átlagmodellhez, valamint, hogy a kezdeti javítóiterációkban erős likelihood regularizációt használunk a látens paraméterekre. A fentiek ellenére, a matematikailag precízebb megközelítésen felül a módszer előnye, hogy robusztussági tartalékot képez a szegmentálási módszerben olyan bonyolult esetekre is, ahol az inicializáció még nehezebb feladat lehet (például ha a szívárnyék valamely sarokpontja és egy anatómiai elváltozás pl. daganat árnyéka összevetül. Ilyen ritka esetek nem szerepeltek a kiértékelési halmaz elemei között sem).

	JSRT		Privát adathalmaz	
	Dice Score	Hausdorff	Dice Score	Hausdorff
$l: 10^{-2.902}, \kappa: 1.776, k_1: 0.252, k_2: 0.062$	0.940 ± 0.027	$10.31 \text{ mm} \pm 5.01 \text{ mm}$	0.893 ± 0.038	$18.17 \text{ mm} \pm 6.44 \text{ mm}$
$l: 10^{-2.551}, \kappa: 0.796, k_1: 0.208, k_2: 0.232$	0.940 ± 0.025	$10.12 \text{ mm} \pm 4.84 \text{ mm}$	0.888 ± 0.046	$16.72 \text{ mm} \pm 4.97 \text{ mm}$
$l: 10^{-3.391}, \kappa: 1.183, k_1: 0.186, k_2: 0.156$	0.940 ± 0.028	$10.38 \text{ mm} \pm 5.13 \text{ mm}$	0.883 ± 0.062	$19.56 \text{ mm} \pm 9.83 \text{ mm}$
$l: 10^{-2.629}, \kappa: 2.552, k_1: 0.040, k_2: 0.171$	0.939 ± 0.027	$10.35 \text{ mm} \pm 5.22 \text{ mm}$	0.893 ± 0.035	$18.03 \text{ mm} \pm 6.66 \text{ mm}$
$l: 10^{-3.076}, \kappa: 1.302, k_1: 0.168, k_2: 0.083$	0.941 ± 0.028	$10.38 \text{ mm} \pm 5.12 \text{ mm}$	0.892 ± 0.040	$17.77 \text{ mm} \pm 5.94 \text{ mm}$
$l: 10^{-2.448}, \kappa: 2.254, k_1: 0.028, k_2: 0.196$	0.939 ± 0.026	$10.38 \text{ mm} \pm 5.02 \text{ mm}$	0.894 ± 0.036	$17.07 \text{ mm} \pm 6.62 \text{ mm}$
$l: 10^{-2.191}, \kappa: 1.148, k_1: 0.130, k_2: 0.118$	0.937 ± 0.025	$10.40 \text{ mm} \pm 4.74 \text{ mm}$	0.889 ± 0.047	$16.33 \text{ mm} \pm 5.72 \text{ mm}$
$l: 10^{-2.120}, \kappa: 2.831, k_1: 0.156, k_2: 0.201$	0.937 ± 0.024	$10.47 \text{ mm} \pm 4.72 \text{ mm}$	0.888 ± 0.047	$16.51 \text{ mm} \pm 5.66 \text{ mm}$

4.7. táblázat. Hiperparaméter optimalizáció eredménye az inicializáció folyamán az átlagmodellt illesztve a sarokpontokra maximum-likelihood becslés helyett.

4.2.5. Tüdőszegmentáció felhasználása az előfeldolgozás során

A bemenetek előfeldolgozása során egy szakértői rendszer segítségével azonosítjuk a tüdőárnyék körvonalait, melyek alapján a bemeneti felvételt vágjuk, hogy az lehetőleg csak a célterület körülötti részt tartalmazza. Az eljárás motivációja, hogy így olyan felvételeken is pontos szegmentációra nyílik lehetőség, amelyekben a páciens beállítása nem pontos, így előfordulhat, hogy egyéb zavaró testrészek is láthatóak, potenciálisan rontva a pontbecslő neurális hálózat pontosságát. Amennyiben a komponens kikapcsoljuk, és így végezzük el a hiperparaméter optimalizációt, úgy a 4.8 táblázatban található eredményeket kapjuk. Jól látható, hogy a privát adathalmazon jelentősen romlanak az eredmények az eredeti modellhez képest, mely felvételeken valóban gyakoriak az ilyen hibák, a JSRT adathalmaz kiértékelési partícióján azonban pont, hogy javulnak az eredmények. Ennek oka vélhetően, hogy a javító pontbecsléseket végző neurális hálózat nem vágott felvételeken volt tanítva, valamint a tüdőszegmentációra történő vágás változtathat a felvételek képarányán, melyre sem az alakmodell, sem az igazítást végző affin transzformáció nincs felkészítve. Az egyes adathalmazok ground truth szegmentációit vizsgálva azt találtuk, hogy a szí-

vek szélességének és magasságának aránya jelentősen eltér az adathalmazok között, azonban a szegmentált szívek arányai jellemzően képesek lekövetni ezen eltéréseket, továbbá a tüdőszegmentáció elvégzése és a célterületre történő vágás az adathalmazok képarányán csak marginálisan módosít. A fentiek ellenére a jövőben érdemes lehet az alakmodell illesztési lépését felkészíteni a különböző képarányú célterületek kezelésére nem izotropikus skálázás alkalmazásával. Összességében azonban így is elmondható, hogy az előfeldolgozási lépés már jelen formájában is jóval többet javít az eljárás kimenetén, mint amennyit esetleg ront.

	JSRT		Privát adathalmaz	
	Dice Score	Hausdorff	Dice Score	Hausdorff
$l: 10^{-2.647}$, $\kappa: 1.958$, $k_1: 0.131$, $k_2: 0.288$	0.944 ± 0.021	$10.16 \text{ mm} \pm 4.96 \text{ mm}$	0.872 ± 0.086	$19.39 \text{ mm} \pm 11.45 \text{ mm}$
$l: 10^{-3.926}$, $\kappa: 3.464$, $k_1: 0.270$, $k_2: 0.075$	0.945 ± 0.022	$10.34 \text{ mm} \pm 5.01 \text{ mm}$	0.872 ± 0.082	$21.19 \text{ mm} \pm 12.44 \text{ mm}$
$l: 10^{-1.678}$, $\kappa: 0.939$, $k_1: 0.147$, $k_2: 0.241$	0.933 ± 0.027	$11.29 \text{ mm} \pm 5.08 \text{ mm}$	0.854 ± 0.088	$20.46 \text{ mm} \pm 10.10 \text{ mm}$
$l: 10^{-2.716}$, $\kappa: 3.338$, $k_1: 0.084$, $k_2: 0.239$	0.945 ± 0.022	$10.12 \text{ mm} \pm 5.03 \text{ mm}$	0.873 ± 0.085	$19.75 \text{ mm} \pm 11.51 \text{ mm}$
$l: 10^{-3.299}$, $\kappa: 1.346$, $k_1: 0.133$, $k_2: 0.160$	0.945 ± 0.023	$10.30 \text{ mm} \pm 5.20 \text{ mm}$	0.869 ± 0.086	$20.36 \text{ mm} \pm 12.51 \text{ mm}$
$l: 10^{-1.884}$, $\kappa: 2.177$, $k_1: 0.126$, $k_2: 0.249$	0.934 ± 0.026	$11.13 \text{ mm} \pm 5.00 \text{ mm}$	0.857 ± 0.088	$20.00 \text{ mm} \pm 10.14 \text{ mm}$
$l: 10^{-3.296}$, $\kappa: 3.240$, $k_1: 0.129$, $k_2: 0.244$	0.945 ± 0.023	$10.24 \text{ mm} \pm 5.17 \text{ mm}$	0.872 ± 0.084	$21.13 \text{ mm} \pm 12.28 \text{ mm}$
$l: 10^{-2.131}$, $\kappa: 1.025$, $k_1: 0.260$, $k_2: 0.198$	0.938 ± 0.024	$10.76 \text{ mm} \pm 4.82 \text{ mm}$	0.861 ± 0.090	$19.59 \text{ mm} \pm 10.45 \text{ mm}$

4.8. táblázat. A tüdőszegmentáció kikapcsolása mellett elért egyes paraméterkonfigurációk eredményei. Látható, hogy a JSRT adathalmazon szignifikáns teljesítménynövekedést sikerült elérni az egyes metrikákban, azonban a privát adathalmazon tapasztalt visszaesés ennél jóval jelentősebb.

4.2.6. Elemenkénti likelihood limit hatása

Az active shape modellek regularizációjának egy hagyományos módja a látens paraméterek elemenkénti vágással történő regularizációja, hogy azok értéke egy adott konfidenciaintervallumból kerüljön ki (feltételezzük, hogy a látens paraméterek egymástól független valószínűségi változók). A 4.9 táblázatban láthatók a paraméterre vonatkozó érzékenységi vizsgálat eredményei. Amennyiben a regularizációt kihagyjuk az optimalizációból (a konfidenciaintervallumot a szórás százszorosára állítjuk), és a modell hiperparaméter optimalizációját így végezzük el, úgy a 4.10 táblázatban foglalt eredményeket nyerjük. Összességében elmondható, hogy a likelihood alapú regularizáció nagyrészt képes kiváltani ezen regularizációt. Az egyes kiértékelési konfigurációk között a látens paraméterek elemenkénti korlátozásával, illetve anélkül elérhető legjobb eredmény csak marginálisan különbözik egymástól, azonban az eljárás robusztussági tartalékának növelése érdekében érdemes a két regularizációt párhuzamosan használni.

4.2.7. Az ADMM optimalizáció büntetősúlyai

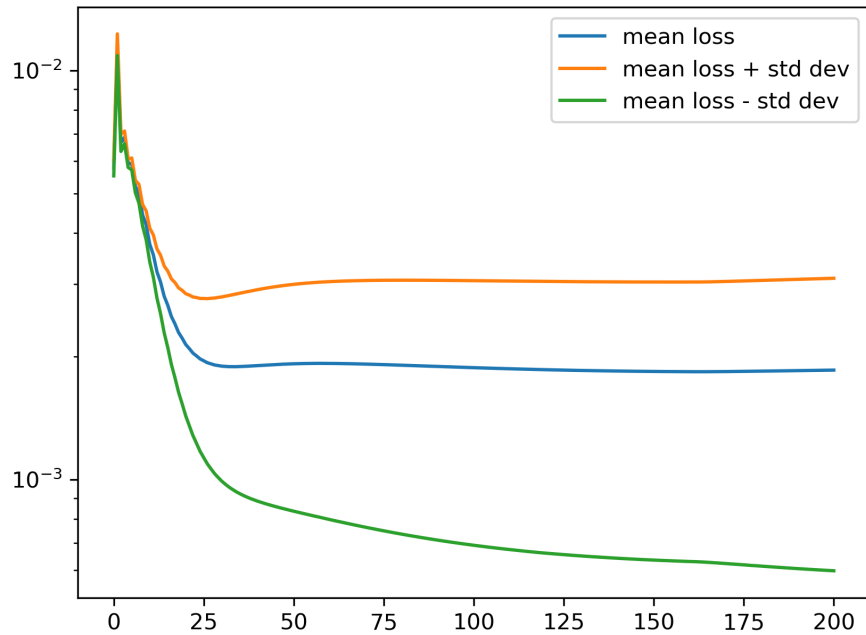
Az ADMM optimalizáció fontos paraméterei a ρ_1 , ρ_2 , ρ_3 augmentációs büntetősúlyok, melyek megfelelő megválasztása szükséges a gyors konvergencia elérése érdekében. Gyakori, hogy ezen súlyokat dinamikusan változtatják az iterációk folyamán, azonban azt találtuk, hogy az optimalizált veszteségfüggvény konvergenciája robusztus működést mutat konstans értékekkel is, melyeket empirikusan 10^{-3} -ra állítottunk be mindhárom paraméterre. A veszteségfüggvény értékének alakulása és konvergenciája az ADMM iterációk függvényében a 4.1 ábrán látható.

	JSRT		Privát adathalmaz	
	Dice Score	Hausdorff	Dice Score	Hausdorff
$\kappa: 0.5$	0.939 ± 0.027	$10.43 \text{ mm} \pm 5.07 \text{ mm}$	0.884 ± 0.047	$16.78 \text{ mm} \pm 5.06 \text{ mm}$
$\kappa: 1.5$	0.940 ± 0.028	$10.29 \text{ mm} \pm 5.38 \text{ mm}$	0.892 ± 0.038	$17.92 \text{ mm} \pm 6.08 \text{ mm}$
$\kappa: 2$	0.940 ± 0.029	$10.54 \text{ mm} \pm 5.12 \text{ mm}$	0.893 ± 0.036	$18.41 \text{ mm} \pm 6.60 \text{ mm}$
$\kappa: 2.5$	0.940 ± 0.029	$10.35 \text{ mm} \pm 5.39 \text{ mm}$	0.891 ± 0.038	$19.83 \text{ mm} \pm 7.48 \text{ mm}$
$\kappa: 3.5$	0.940 ± 0.029	$10.35 \text{ mm} \pm 5.39 \text{ mm}$	0.890 ± 0.038	$19.83 \text{ mm} \pm 7.48 \text{ mm}$

4.9. táblázat. Modell szenzitivitása a látens paraméterek szórásának κ szerinti regularizációjára.

	JSRT		Privát adathalmaz	
	Dice Score	Hausdorff	Dice Score	Hausdorff
$l: 10^{-2.239}, \kappa: 100.000, k_1: 0.068, k_2: 0.292$	0.938 ± 0.025	$10.41 \text{ mm} \pm 4.83 \text{ mm}$	0.892 ± 0.043	$16.43 \text{ mm} \pm 6.08 \text{ mm}$
$l: 10^{-2.095}, \kappa: 100.000, k_1: 0.039, k_2: 0.173$	0.937 ± 0.025	$10.57 \text{ mm} \pm 4.84 \text{ mm}$	0.888 ± 0.046	$16.44 \text{ mm} \pm 5.73 \text{ mm}$
$l: 10^{-3.140}, \kappa: 100.000, k_1: 0.283, k_2: 0.025$	0.940 ± 0.028	$10.47 \text{ mm} \pm 5.03 \text{ mm}$	0.883 ± 0.057	$20.82 \text{ mm} \pm 10.11 \text{ mm}$
$l: 10^{-2.362}, \kappa: 100.000, k_1: 0.285, k_2: 0.212$	0.940 ± 0.024	$10.16 \text{ mm} \pm 4.72 \text{ mm}$	0.892 ± 0.042	$16.48 \text{ mm} \pm 6.19 \text{ mm}$
$l: 10^{-3.393}, \kappa: 100.000, k_1: 0.270, k_2: 0.214$	0.937 ± 0.028	$10.95 \text{ mm} \pm 4.98 \text{ mm}$	0.881 ± 0.057	$21.71 \text{ mm} \pm 10.22 \text{ mm}$
$l: 10^{-3.227}, \kappa: 100.000, k_1: 0.206, k_2: 0.250$	0.939 ± 0.028	$10.66 \text{ mm} \pm 5.10 \text{ mm}$	0.881 ± 0.057	$21.51 \text{ mm} \pm 10.13 \text{ mm}$
$l: 10^{-2.606}, \kappa: 100.000, k_1: 0.228, k_2: 0.159$	0.941 ± 0.025	$10.15 \text{ mm} \pm 4.86 \text{ mm}$	0.895 ± 0.035	$17.24 \text{ mm} \pm 6.79 \text{ mm}$
$l: 10^{-2.385}, \kappa: 100.000, k_1: 0.072, k_2: 0.291$	0.939 ± 0.025	$10.32 \text{ mm} \pm 4.90 \text{ mm}$	0.894 ± 0.038	$16.63 \text{ mm} \pm 6.47 \text{ mm}$

4.10. táblázat. Hiperparaméter optimalizáció eredménye, a látens paraméterek elemenkénti korlátját kikapcsolva.



4.1. ábra. Az optimalizációs probléma veszteségének alakulása az ADMM iterációk függvényében. A méréseket több adathalmazból származó bemeneti felvételeken is elvégeztük. A mért veszteségek átlaga kék színnel látható, míg a szórásstartomány határait narancssárga, illetve zöld színnel jelöljük.

4.2.8. Konfidenciát becslő eljárás hatása

Az optimalizáció során a veszteségfüggvényben a javítópontokat adó neurális hálózat, és az alakmodell által rekonstruált alakzat pontjai közötti számolt négyzetes távolságot egy H diagonális mátrixszal súlyozzuk melynek elemei a neurális hálózat és az alakmodell konfidenciáját reprezentálják az egyes kontrollpontokra. Ezen konfidenciaérték számítása egy heurisztika alapján történik, mely a neurális hálózat által becsült valószínűségi eloszlások szórásából, illetve abból adódik, hogy az egyes kontrollpontok mennyire következnek az alakmodell szerint a többi kontrollpontból. Amennyiben konfidencia alapú súlyozást kikapcsoljuk (H egységmátrix), úgy a 4.11 található eredményeket kapjuk a hiperparaméter optimalizáció elvégzése után. Jól látható, hogy konfidenciaértékek elhagyásával a modell teljesítménye jelentősen csökken a privát adathalmazon, ahol a bemeneti felvételek lényegesen különböző eloszlásból származnak, mint azon minták, amikre a javítópontokat becslő neurális hálózat tanítva lett. A jövőben a modell egy lehetséges továbbfejlesztési iránya lehet H meghatározására egyéb heurisztikák vizsgálata, esetlegesen matematikailag precízebb eljárás levezetése.

	JSRT		Privát adathalmaz	
	Dice Score	Hausdorff	Dice Score	Hausdorff
$l: 10^{-3.409}, \kappa: 0.559, k_1: 0.285, k_2: 0.108$	0.939 ± 0.026	$10.36 \text{ mm} \pm 4.84 \text{ mm}$	0.876 ± 0.067	$18.95 \text{ mm} \pm 9.09 \text{ mm}$
$l: 10^{-2.668}, \kappa: 1.537, k_1: 0.147, k_2: 0.060$	0.941 ± 0.026	$10.24 \text{ mm} \pm 5.05 \text{ mm}$	0.884 ± 0.063	$18.70 \text{ mm} \pm 9.85 \text{ mm}$
$l: 10^{-2.491}, \kappa: 2.344, k_1: 0.196, k_2: 0.143$	0.940 ± 0.024	$10.18 \text{ mm} \pm 4.91 \text{ mm}$	0.886 ± 0.062	$18.31 \text{ mm} \pm 10.02 \text{ mm}$
$l: 10^{-2.805}, \kappa: 1.208, k_1: 0.043, k_2: 0.209$	0.940 ± 0.027	$10.29 \text{ mm} \pm 5.21 \text{ mm}$	0.882 ± 0.064	$18.75 \text{ mm} \pm 9.63 \text{ mm}$
$l: 10^{-2.762}, \kappa: 3.029, k_1: 0.017, k_2: 0.031$	0.940 ± 0.027	$10.35 \text{ mm} \pm 5.28 \text{ mm}$	0.882 ± 0.062	$20.42 \text{ mm} \pm 10.43 \text{ mm}$
$l: 10^{-2.892}, \kappa: 1.054, k_1: 0.080, k_2: 0.214$	0.941 ± 0.027	$10.21 \text{ mm} \pm 5.13 \text{ mm}$	0.881 ± 0.064	$18.83 \text{ mm} \pm 9.53 \text{ mm}$
$l: 10^{-2.176}, \kappa: 0.989, k_1: 0.148, k_2: 0.187$	0.938 ± 0.024	$10.43 \text{ mm} \pm 4.75 \text{ mm}$	0.881 ± 0.064	$17.61 \text{ mm} \pm 8.24 \text{ mm}$
$l: 10^{-1.846}, \kappa: 1.314, k_1: 0.299, k_2: 0.221$	0.934 ± 0.026	$10.86 \text{ mm} \pm 4.66 \text{ mm}$	0.872 ± 0.066	$18.34 \text{ mm} \pm 7.33 \text{ mm}$

4.11. táblázat. Hiperparaméter optimalizáció eredménye a konfidencia alapú súlyozás kikapcsolása után.

4.2.9. Neurális háló által prediktált görbe újraparaméterezésének hatása

Az optimalizációs eljárás folyamán az alakmodellel nem közvetlenül a neurális hálózat predikcióit kívánjuk rekonstruálni, hanem annak egy lokális lineáris újraparaméterezését a kontrollpontok mentén. A lépés motivációja, hogy az összes javítópont predikciója egy lépésben történik egyazon neurális háló által, ezáltal fennáll a lehetősége, hogy a neurális hálózat egy implicit alakmodellt tanuljon a pontok egymáshoz képesti viszonyára, mely nem teljesen konzisztens az ASM-es modellel, így nehezítheti az optimalizációt. A reinterpoláció célja ezen implicit alakmodell hatásának kiküszöbölése. Ha ezen újraparaméterező lépést kikapcsoljuk (azaz az optimalizációs feladatban $\Phi(\Xi)$ helyett Ξ -t használjuk), úgy a hiperparaméter optimalizáció elvégzése után a 4.12 táblázatban található eredményeket kapjuk. Látható, hogy a modell teljesítménye kis mértékben visszaesik a komponens elhagyásának hatására, mely implikálja az implicit alakmodell létezését.

	JSRT		Privát adathalmaz	
	Dice Score	Hausdorff	Dice Score	Hausdorff
$l: 10^{-2.329}, \kappa: 2.024, k_1: 0.179, k_2: 0.098$	0.939 ± 0.025	10.27 mm $\pm$ 4.39 mm	0.892 $\pm$ 0.042	16.63 mm $\pm$ 6.35 mm
$l: 10^{-3.562}, \kappa: 3.112, k_1: 0.185, k_2: 0.095$	0.917 ± 0.021	13.26 mm $\pm$ 3.31 mm	0.869 ± 0.061	22.32 mm $\pm$ 10.04 mm
$l: 10^{-1.594}, \kappa: 1.553, k_1: 0.289, k_2: 0.190$	0.889 ± 0.030	16.76 mm $\pm$ 5.19 mm	0.820 ± 0.051	23.69 mm $\pm$ 5.61 mm
$l: 10^{-3.119}, \kappa: 2.713, k_1: 0.031, k_2: 0.019$	0.940 $\pm$ 0.030	10.38 mm $\pm$ 5.48 mm	0.879 ± 0.062	22.21 mm $\pm$ 10.50 mm
$l: 10^{-3.974}, \kappa: 2.468, k_1: 0.260, k_2: 0.193$	0.844 ± 0.035	22.85 mm $\pm$ 5.39 mm	0.776 ± 0.047	30.52 mm $\pm$ 5.54 mm
$l: 10^{-3.217}, \kappa: 3.461, k_1: 0.077, k_2: 0.258$	0.939 ± 0.029	10.55 mm $\pm$ 5.44 mm	0.879 ± 0.059	22.29 mm $\pm$ 10.28 mm
$l: 10^{-2.135}, \kappa: 2.945, k_1: 0.110, k_2: 0.122$	0.937 ± 0.025	10.52 mm $\pm$ 4.81 mm	0.890 ± 0.045	16.40 mm $\pm$ 6.02 mm
$l: 10^{-3.489}, \kappa: 2.742, k_1: 0.173, k_2: 0.081$	0.937 ± 0.027	10.93 mm $\pm$ 4.72 mm	0.878 ± 0.059	22.38 mm $\pm$ 10.30 mm

4.12. táblázat. Hiperparaméter optimalizáció eredménye, a neurális hálózat által prediktált pontok reinterpolációja nélkül.

4.2.10. Tisztán neurális baseline hálózat

Annak érdekében, hogy bemutassuk a modell alapú neurális megközelítés előnyeit a szívszegmentáció problémájában, megvizsgáltunk egy tisztán neurális megközelítést is, amely hasonlóan működik javasolt módszerhez, azonban ebben az esetben az explicit alakmodellt és a hozzá tartozó optimalizációs eljárást kihagyjuk. Ezt úgy valósítjuk meg, hogy ugyanazokat a neurális hálókat használjuk a landmark pontok lokalizálására és finomítására, mint a javasolt eljárásban, de csak az átlagos alakzatot használjuk az inicializálásnál a sarokpontokból a kezdeti alakzat becsléséhez, valamint a javító hálózat által becsült módosításokat egyéb regularizációs megfontolások nélkül elfogadjuk. A hasonló működésen felül azért is ezt az eljárást választottuk baseline modellként, mert továbbra is robusztus módon, sűrűségfüggvényként becsüli a landmark pontokat, nem úgy mint a szegmentálást pixel szintű osztályozásként értelmező megoldások. Amennyiben így végezzük el a modell kiértékelését a 4.13 táblázatban kapható eredményeket kapjuk.

	JSRT		Privát adathalmaz	
	Dice Score	Hausdorff	Dice Score	Hausdorff
Baseline modell	0.940 $\pm$ 0.031	10.69 mm $\pm$ 5.74 mm	0.869 $\pm$ 0.075	24.66 mm $\pm$ 11.09 mm

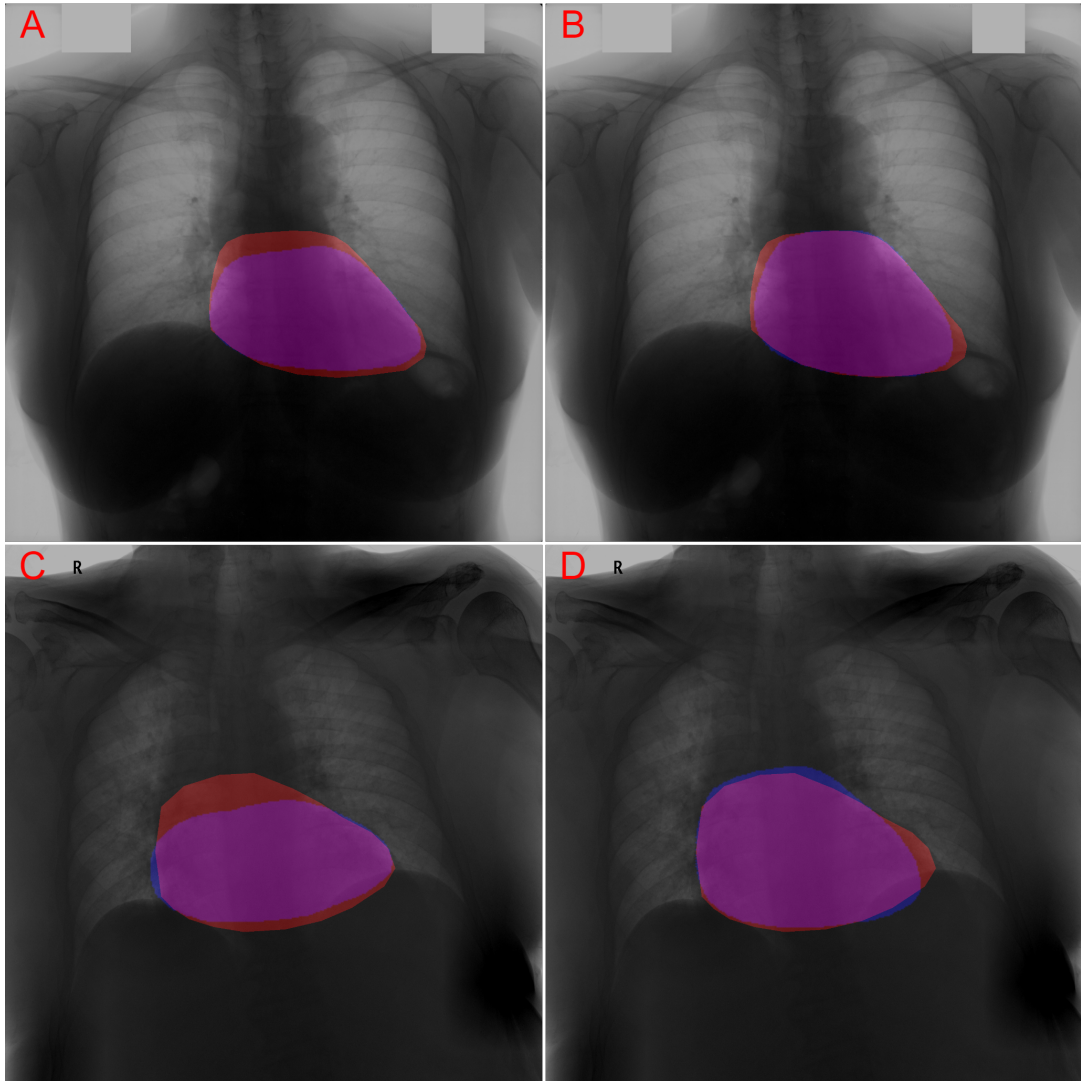
4.13. táblázat. Baseline szegmentációs modell eredményei.

4.3. Eredmények

Az egyes aspektusokban legjobban teljesítő paraméterkonfigurációk és a tisztán neurális baseline modell eredményeit a 4.14 táblázatban foglaltuk össze, valamint példa szegmentációk az optimalizált és baseline modell közötti kvalitatív különbségekre 4.2 ábrán láthatók. Összességében elmondható, hogy a JSRT adathalmaz validációs partícióján minden modell gyakorlati szempontból megfelelően teljesít, látványos pontatlanságokat nem mutat, azonban az optimalizált megoldások általában még kisebb varianciával dolgoznak. A kevésbé homogén privát adathalmazon az optimalizáció által a teljesítmény jelentős javulása figyelhető meg, különösen a robusztusságot mérő Hausdorff távolságban, valamint az eredmények szórása is jelentősen csökken, amely az optimalizáció által a modell általánosítóképességének és robusztusságának jelentős növekedésére utal.

	JSRT		Privát adathalmaz	
	Dice Score	Hausdorff	Dice Score	Hausdorff
Baseline modell	0.940 ± 0.031	$10.69 \text{ mm} \pm 5.74 \text{ mm}$	0.869 ± 0.075	$24.66 \text{ mm} \pm 11.09 \text{ mm}$
$l: 10^{-2.591}, \kappa: 2.316, k_1: 0.271, k_2: 0.107$	0.939 ± 0.028	$10.32 \text{ mm} \pm 5.02 \text{ mm}$	0.895 ± 0.036	$16.96 \text{ mm} \pm 6.71 \text{ mm}$
$l: 10^{-2.447}, \kappa: 3.035, k_1: 0.240, k_2: 0.227$	0.940 ± 0.025	$10.21 \text{ mm} \pm 4.82 \text{ mm}$	0.894 ± 0.039	$16.59 \text{ mm} \pm 6.50 \text{ mm}$
$l: 10^{-1.955}, \kappa: 2.955, k_1: 0.253, k_2: 0.015$	0.936 ± 0.025	$10.65 \text{ mm} \pm 4.67 \text{ mm}$	0.885 ± 0.047	$16.46 \text{ mm} \pm 5.10 \text{ mm}$

4.14. táblázat. Az egyes metrikák szerinti legjobb modellek teljesítményének összehasonlítása a baseline szegmentáló modellel.



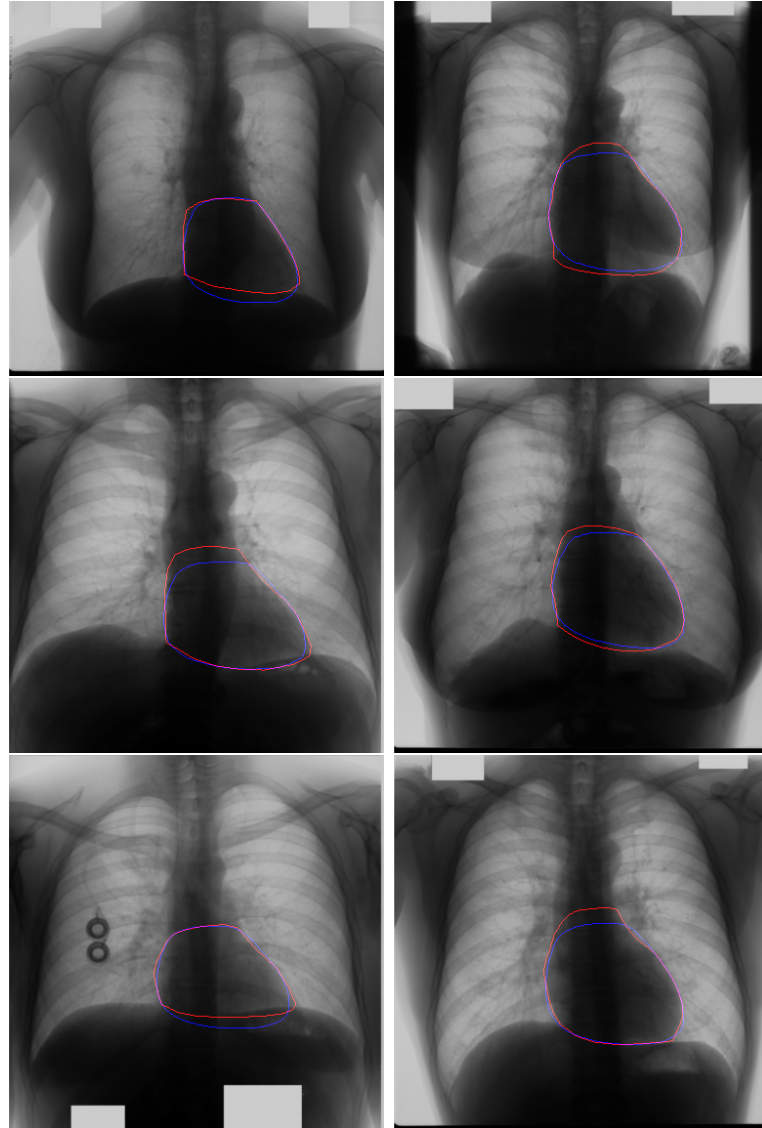
4.2. ábra. Az ábrán az (A) és (C) képeken a baseline szegmentálás eredménye látható egy-egy felvételre a JSRT és a privát adathalmazokból, míg a (B) és (D) képeken javasolt módszerrel végzett szegmentálás eredménye látható. Az eredmény szegmentációk a kék színcsatornán, míg a ground truth maszkok a piros csatornán láthatók.

5. fejezet

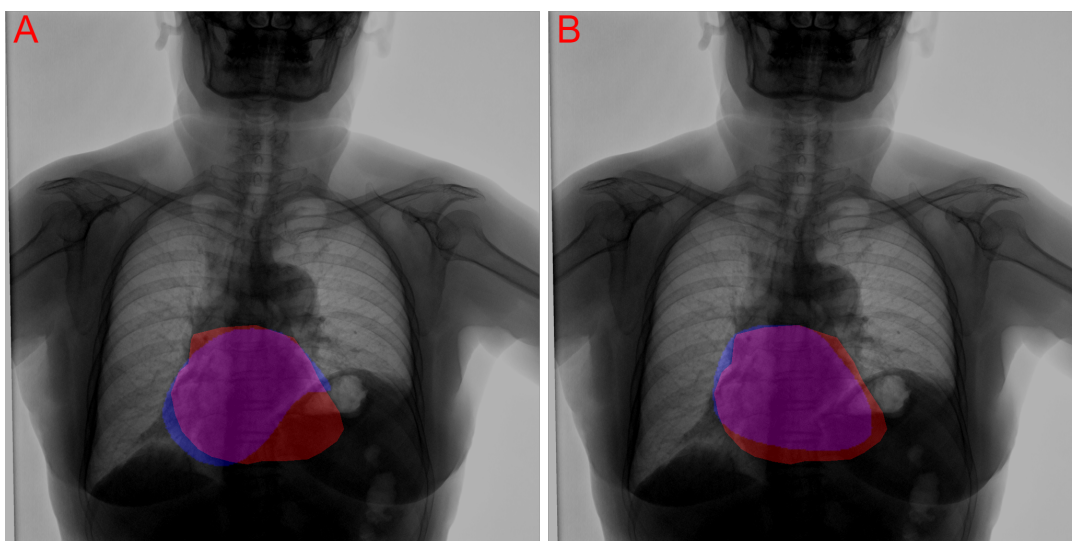
Összefoglalás

Összességében elmondható, hogy a javasolt módszer a hiperparaméterek megfelelő beállítása mellett minden aspektusában legalább olyan jó, mint a baseline módszer. Bár a különbség a JSRT adathalmazon jellemzően nem szignifikáns sem kvantitatívan sem kvalitatívan (5.1 ábra) (mivel itt mindkét modell megfelelően képes általánosítani a rendelkezésre álló adatokból). A nagyobb varianciájú privát adathalmazon már jelentős javulást sikerült elérni a tisztán neurális megközelítéshez képest (A Dice error nagyságrendileg 20%-al, a Hausdorff távolság pedig 30%-al csökken, miközben az elért eredmények varianciája is nagyjából 50%-al csökken mindkét esetben), továbbá bizonyos felvételeken csak az optimalizált módszer eredménye volt elfogadható, ami az általunk javasolt és implementált eljárás jó általánosító képességét is mutatja. A baseline modell pontosságát határozottan meghaladó eredmények ellenére további fejlesztési lehetőségek is adódnak, mint például más heurisztikák alkalmazása a H konfidenciamátrix becslésére, vagy nem izotropikus skálázás alkalmazása a változó oldalarányú célterületek jelenleginél precízebb kezelésére a regisztrációs lépések során. Az eljárás ablációja folyamán, bár nem minden komponens bizonyult elengedhetetlennek a kvantitatív eredmények szempontjából, azonban azok használata robusztussági tartalékot biztosít a még nehezebb problémákra. Ezen tartalékok szükségességének indoka, hogy bár a privát adathalmaz nagyobb diverzitású, azonban ez sem tartalmazta a felismerést leginkább nehezítő elváltozásokat (pl.: erősen aszimmetrikus rekeszív, súlyos tüdőgyulladás, folyadékkal megtelt lebenyek, stb.), így kvalitatívan további összehasonlításokat végeztünk a VinDr-CXR [18] adathalmazon. Ennek eredményeként megállapítottuk, hogy a javasolt módszer továbbra is robusztusan működik, ellentétben a baseline megoldással, amely például aszimmetrikus rekeszív esetén gyakran hibázik. Egy ilyen, szakértő által annotált példa a 5.2 ábrán látható.

Összefoglalva az eredményeket arra jutunk, hogy a javasolt megoldás nagy pontossággal, robusztusan, az adathalmazok között jól általánosítva képes a szívárnyék szegmentációjára alacsony mintaszám és homogén tanító adathalmaz mellett is. Érdemes továbbá kiemelni, hogy a dolgozatban javasolt módszer könnyen adaptálható más szegmentációs problémákra is, ezáltal robusztus és jól általánosító eljárásokat készíthetünk, mely során lehetőség nyílik ezen megoldásokat további domain-specifikus kényszerek beépítésével finomítani.



5.1. ábra. A szegmentációk Dice Score metrika minősítése szerinti kvalitatív elemzése a JSRT adathalmazon. A ground truth maszkok körvonalait a vörös, a modell által szegmentáltakét pedig a kék színcsatornán láthatjuk. A bal oldali oszlopban található felvételeken a szegmentált maszkok DSC-je a $[0.930, 0.935]$ tartományba esik, ami a baseline és az optimalizált modellek átlagos teljesítményétől elmarad, míg a jobb oldali oszlopban a $[0.945, 0.950]$ tartományba, amely meghaladja a vizsgált modellek átlagos teljesítményét. Látható, hogy egyik esetben sem történik szignifikáns anatómiai tévesztés, valamint a prediktált és ground truth maszkok közötti különbség elhanyagolható.



5.2. ábra. A szegmentációs modellek teljesítményének összehasonlítása egy minta felvételen a VinDr-CXR adathalmazból, amelyet egy szakértő manuálisan annotált. Az (A) képen a baseline modell predikcióját láthatjuk, míg (B) a javasolt módszer eredményét mutatja. Az eredmény szegmentációk a kék színcsatornán, míg a ground truth maszkok a piros csatornán láthatók.

Irodalomjegyzék

- [1] H. C. Kim, „Epidemiology of cardiovascular disease and its risk factors in korea,” *Global Health amp; Medicine*, vol. 3, p. 134–141, June 2021.
- [2] R. Jagannathan, S. A. Patel, M. K. Ali, and K. M. V. Narayan, „Global updates on cardiovascular disease mortality trends and attribution of traditional risk factors,” *Current Diabetes Reports*, vol. 19, June 2019.
- [3] R. C. Woodruff, X. Tong, S. S. Khan, N. S. Shah, S. L. Jackson, F. Loustalot, and A. S. Vaughan, „Trends in cardiovascular disease mortality rates and excess deaths, 2010–2022,” *American Journal of Preventive Medicine*, Nov. 2023.
- [4] H. Lee, J. Cho, D. W. Shin, S.-P. Lee, S.-S. Hwang, J. Oh, H.-K. Yang, S.-H. Hwang, K. Y. Son, S. H. Chun, B. Cho, and E. Guallar, „Association of cardiovascular health screening with mortality, clinical outcomes, and health care cost: A nationwide cohort study,” *Preventive Medicine*, vol. 70, p. 19–25, Jan. 2015.
- [5] M. J. Oude Wolcherink, C. M. Behr, X. G. L. V. Pouwels, C. J. M. Doggen, and H. Koffijberg, „Health economic research assessing the value of early detection of cardiovascular disease: A systematic review,” *PharmacoEconomics*, vol. 41, p. 1183–1203, June 2023.
- [6] M. Amini, F. Zayeri, and M. Salehi, „Trend analysis of cardiovascular disease mortality, incidence, and mortality-to-incidence ratio: results from global burden of disease study 2017,” *BMC Public Health*, vol. 21, Feb. 2021.
- [7] D. Pan, P. Pellicori, K. Dobbs, J. Bulemfu, I. Sokoreli, A. Urbinati, O. Brown, S. Sze, A. S. Rigby, S. Kazmi, J. M. Riistama, J. G. F. Cleland, and A. L. Clark, „Prognostic value of the chest x-ray in patients hospitalised for heart failure,” *Clinical Research in Cardiology*, vol. 110, p. 1743–1756, Mar. 2021.
- [8] A. H. Dallal, C. Agarwal, M. R. Arbabshirani, A. Patel, and G. Moore, „Automatic estimation of heart boundaries and cardiothoracic ratio from chest x-ray images,” in *Medical Imaging 2017: Computer-Aided Diagnosis* (S. G. Armato and N. A. Petrick, eds.), SPIE, Mar. 2017.
- [9] S. Candemir, S. Jaeger, W. Lin, Z. Xue, S. Antani, and G. Thoma, „Automatic heart localization and radiographic index computation in chest x-rays,” in *SPIE Proceedings* (G. D. Tourassi and S. G. Armato, eds.), SPIE, Mar. 2016.

- [10] B. van Ginneken, M. B. Stegmann, and M. Loog, „Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database,” *Med. Image Anal.*, vol. 10, pp. 19–40, Feb. 2006.
- [11] A. A. Novikov, D. Lenis, D. Major, J. Hladuvka, M. Wimmer, and K. Buhler, „Fully convolutional architectures for multiclass segmentation in chest radiographs,” *IEEE Transactions on Medical Imaging*, vol. 37, p. 1865–1876, Aug. 2018.
- [12] B. Maga, „Chest x-ray lung and heart segmentation based on minimal training sets,” 2021.
- [13] M. Frid-Adar, A. Ben-Cohen, R. Amer, and H. Greenspan, *Improving the Segmentation of Anatomical Structures in Chest Radiographs Using U-Net with an ImageNet Pre-trained Encoder*, p. 159–168. Springer International Publishing, 2018.
- [14] O. Gómez, P. Mesejo, O. Ibáñez, A. Valsecchi, and O. Cerdón, „Deep architectures for high-resolution multi-organ chest X-ray image segmentation,” *Neural Computing and Applications*, Oct. 2019.
- [15] M. Kholiavchenko, I. Sirazitdinov, K. Kubrak, R. Badrutdinova, R. Kuleev, Y. Yuan, T. Vrtovec, and B. Ibragimov, „Contour-aware multi-label chest x-ray organ segmentation,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 15, p. 425–436, Feb. 2020.
- [16] J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K.-i. Komatsu, M. Matsui, H. Fujita, Y. Kodera, and K. Doi, „Development of a digital image database for chest radiographs with and without a lung nodule,” *American Journal of Roentgenology*, vol. 174, pp. 71–74, 01 2000.
- [17] A. Tumay, D. Hadhazi, and G. Hullam, „Heart segmentation on pa chest x-ray images by model-based deep learning approach,” in *2024 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pp. 1–6, 2024.
- [18] H. Q. Nguyen, K. Lam, L. T. Le, H. H. Pham, D. Q. Tran, D. B. Nguyen, D. D. Le, C. M. Pham, H. T. T. Tong, D. H. Dinh, C. D. Do, L. T. Doan, C. N. Nguyen, B. T. Nguyen, Q. V. Nguyen, A. D. Hoang, H. N. Phan, A. T. Nguyen, P. H. Ho, D. T. Ngo, N. T. Nguyen, N. T. Nguyen, M. Dao, and V. Vu, „Vindr-cxr: An open dataset of chest x-rays with radiologist’s annotations,” 2020.
- [19] E. Catmull and R. Rom, „A class of local interpolating splines,” in *Computer Aided Geometric Design* (R. E. BARNHILL and R. F. RIESENFELD, eds.), pp. 317–326, Academic Press, 1974.
- [20] T. Cootes, C. Taylor, D. Cooper, and J. Graham, „Active shape models-their training and application,” *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

- [21] M. A. Fischler and R. C. Bolles, „Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” in *Readings in Computer Vision* (M. A. Fischler and O. Firschein, eds.), pp. 726–740, San Francisco (CA): Morgan Kaufmann, 1987.
- [22] O. Ronneberger, P. Fischer, and T. Brox, „U-net: Convolutional networks for biomedical image segmentation,” 2015.
- [23] G. Huang, Z. Liu, L. van der Maaten, and K. Weinberger, „Densely connected convolutional networks,” 07 2017.
- [24] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, „Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends in Machine Learning*, vol. 3, pp. 1–122, 01 2011.