БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ

ИНСТИТУТ ПО ОПТИЧЕСКИ МАТЕРИАЛИ И ТЕХНОЛОГИИ
"Акад. Йордан Малиновски"

АВТОРЕФЕРАТ

на дисертация
за присъждане на образователна и научна степен "доктор"

Направление 4.1. „Физически Науки:, специалност „Физика на вълновите процеси"

Дълбокообучаващи модели за реконструиране и анализ на съдови изображения , получени чрез оптична кохерентна томография с променлива дължина на вълната (SS-OCT)

Мариам Викар

Научни ръководители:
доц. д-р Виолета Маджарова (ИОМТ-БАН, България)
проф. д.н. Елена Стойкова (ИОМТ-БАН, България)
д-р Ердем Шахин (Университет на Тампере, Финландия)
проф. д-р Атанас Гочев (Университет на Тампере, Финландия)

София, 2025

Дисертацията е обсъдена и одобрена за защита на заседание на Колоквиума на Института по оптични материали и технологии -Българска академия на науките(ИОМТ-БАН), проведено на …… ...

Дисертацията се състои от декларация за оригиналност, списък с основни символи и съкращения, пет глави, включително уводната глава, основни приноси, списък с публикациите и презентациите на автора, свързани а дисертационния труд, списък с цитати на публикации на автора и библиография на източниците. Дисертационният труд съдържа 128 страници, 33 фигури и 12 таблица.

Защитата на дисертацията ще се проведе на ……. в ……. в зала ….. на блок 109 на ИОМТ-БАН в открито заседание на научното жури в състав:

Списък с ключови термини и основни съкращения

СЪКРАЩЕНИЕ	ЗНАЧЕНИЕ
ADD	Допълнителен слой
AMD	Възрастово-обусловена макулна дегенерация
AO-OCT	Оптична кохерентна томография с адаптивна оптика
ASPP	Атрус пространствен пирамидален пуллинг (pooling - операция за обобщаване)
BN	Нормализация по партиди
CNN	Конволюционни невронни мрежи
CVD	Хронични венозни заболявания
DL	Дълбоко обучение
DSP	Цифрова обработка на сигнали
ESDM	Ефективен модел за сегментация и потискане на шум
FD	Фурие област
FDML	Заключване на модове във Фурие пространството
FFT	Бърза Фурие трансформация
$\boldsymbol{FOV_{axial}}$	Аксиално зрително поле
$\boldsymbol{FOV_{lateral}}$	Латерално зрително поле
FWHM	Пълна ширина при половин максимум
GSPS	Гига извадки в секунда
GVD	Дисперсия на груповата скорост
MCME	Смес от експерти с много-мащабни конволюции
NA	Числова апертура
NAION	Предна исхемична оптична невропатия
NDFT	Неравномерна дискретна Фурие трансформация
OCT	Оптична кохерентна томография
PPM	Пирамидален пуллинг (pooling) модул
PSF	Функция на разсейването на точков източник
PSNR	Пиково съотношение сигнал–шум
RNFL	Слой от нервни влакна на ретината
SD-OCT	Оптична кохерентна томография в спектралната област
SNR	Отношение сигнал/шум
SSIM	Индекс на структурна прилика
SS-OCT	Оптична кохерентна томография с източник с промянлива дължина на вълната
TD	Времева област
TNode	Томографско намаляване на спекъла с нелокални средни
Up	Слой за увеличаване на резолюцията
WNNM	Минимизиране на претеглените ядрени норми

Глава 1 Въведение

Оптичната кохерентна томография (OCT) представлява неинвазивен метод за образна диагностика, който получава изображения на биологични образци като генерира обемни данни с микрометрова резолюция дълбочина на проникване от няколко милиметра. Пълният потенциал на ОКТ все още остава недостатъчно изследван за преодоляване на предизвикателствата в редица биомедицински области. С бързото развитие на методите за дълбоко обучение (DL) автоматизацията на задачите в образната диагностика отбелязва значителен напредък, водещ до по-ефективни работни процеси. Усъвършенстваната биомедицинска образна диагностика с дълбоко обучение се развива непрекъснато, за да предложи автоматизирани методи, способни да извличат смислена информация чрез бърз анализ на големи обеми данни, дори в реално време. Ръчната интерпретация на подобни данни е силно времеемка и трудоемка, особено в клинични условия с висок поток от пациенти. Автоматизираните методи предлагат високо ниво на стандартизация и възпроизводимост, докато резултатите, генерирани от човека, често варират. Освен това, някои фини детайли и характеристики могат лесно да бъдат пропуснати от човешкото око.

Въпреки съществения напредък в биомедицинските изследвания, постигнат чрез използването на най-съвременни методи за образна диагностика, множество предизвикателства в различни направления продължават да бъдат нерешени. Областта на съдовите изследвания е една от тези, в които е необходим изключително прецизен и безразрушителен метод за образна диагностика с цел надежден анализ. Подробното изследване на вените е от критично значение за разбирането на тяхната функция, проектирането на ефективни венозни заместители и за по-задълбочено изучаване на биомеханичните аспекти на венозните заболявания. Сегментационните карти са ключови за разбирането на формата на вените и за извличане на смислена информация от тези данни. Съществуващите водещи DL-базирани модели за сегментация страдат от висока изчислителна сложност, обусловена от големия брой параметри. Това ограничава тяхната приложимост в клинични условия. Допълнително предизвикателство представлява разработването на интегриран модел, способен както на сегментация, така и на оценка на физични характеристики (дебелина на стените на вените). Такъв модел би позволил едновременно висикосемантична сегментация на разширени вени и прецизно определяне на дебелината на венозните слоеве. Настоящите водещи подходи обикновено разглеждат двете задачи поотделно или правят компромис с ефективността.

Съществува също значителен потенциал за усъвършенстване на процеса по обработка на ОКТ данните, с цел подобряване на реконструирането и качеството на изображенията, намаляване на изчислителната сложност и ускоряване на времето за обработка. За извличане на дълбочинно-разрешена информация типичните системи Swept Source OCT (SS-OCT) извършват линеаризиране на сигнала по вълновото число като фундаментална стъпка в обработката. Този процес изисква допълнителни хардуерни ресурси или увеличава общата изчислителна тежест. С цел справяне с високите изчислителни изисквания и едновременно генериране на изображения с висока достоверност, настоящата дисертация анализира директно суровите ОКТ данни, които са нелинейни по вълновото число, без извършване на линеаризиране на сигнала по вълново число. В тази насока се изследват мултидисциплинарни подходи - комбиниращи обработка на сигнали, машинно обучение и вълнова оптика - с цел преодоляване на ключови предизвикателства и предлагане на иновативни изчислителни методи за системите SS-OCT.

Целта на дисертацията е да се отговори на следните изследователски въпроси:

RQ1. Колко точно може DL модел да оцени картите за сегментация на венозните слоеве, като същевременно поддържа ниска сложност чрез минимален брой параметри, използвайки обемни данни, придобити чрез SS-OCT?

RQ2. Как може да се разработи интегриран модел, който едновременно осигурява високо точни карти за сегментация – надеждно идентифицирайки истински положителни и отрицателни области, минимизирайки фалшивите детекции – и предоставя точни и последователни измервания на дебелината върху различни SS-OCT обемни набори данни на разширени вени?
RQ3. Как може да се разработи DL мрежа за реконструиране на SS-OCT изображения директно от данни, линейни по дължината на вълната, като се заобиколи k-линеаризацията, намалят се хардуерните разходи и изчислителните изисквания, като същевременно се поддържа качество на изображението, сравнимо със стандартния SS-OCT процес?
RQ4.? Как може да се разработи усъвършенствана DL мрежа за реконструиране, която оптимизира данните както в пространствената, така и във Фурие областта, генерирайки висококачествени изображения с намален спекъл шум и подобрени показатели Пиково съотношение сигнал–шум (PSNR) и Индекс на структурна прилика (SSIM) в сравнение с конвенционалната обработка на сурови SS-OCT данни?
RQ5. Възможно ли е да се разработи DL-базирано решение с ниска изчислителна сложност като алтернатива на стандартната процедура за реконструиране, което постига висококачествени изображения, позволявайки визуализация в реално време?

В обобщение, настоящата дисертация изследва SS-OCT като перспективен образен метод за оценка на меки съдови структури като вените и анализ на техните механични характеристики. Синергията между OCT и методите за дълбоко обучение демонстрира значителен потенциал за надеждна рамка за реконструиране, способна да отговори на посочените изследователски въпроси. Разработените в настоящата дисертацията методи показват висока ефективност и производителност, като така допринасят за развитието съвременни методи в областта на OCT.
Получените резултати са представени в Глава 3 (ИВ1 и ИВ2), Глава 4 (ИВ4–ИВ5), а Глава 5 предоставя заключенията. Глава 2 е посветена на литературен обзор на основните области на приложение и добре установените алгоритми.

Глава 2 Литературен обзор и исторически контекст

OCT представлява триизмерна (3D) образна техника, която използва нискокохерентна интерферометрия за генериране на изображения в дълбочина с микрометрова разделителна способност. От своето създаване през 90-те години на XX век [1] насам, OCT се е утвърдила като незаменим метод в медицинската образна диагностика. Въпреки че основното ѝ приложение е в офталмологията, обхватът ѝ бързо се разширява в различни клинични направления - дерматология [2], онкология [3] и редица други медицински области [4] - [10]. OCT изображенията съдържат профили в дълбочина на отражателната способност на определени точки в изследвания образец [11]. Те се използват за конструиране на напречни 3D обеми. Тъй като показателите за пречупване, поглъщането и разсейването варират между слоевете на образеца, това се проявява под формата на различни интензитетни пикове в интерференчния профил [12].

Системите за OCT могат да се разделят основно системи във времева (TD) и Фурие област (FD) [11]. В TD-OCT профил в дълбочин на отражателната способност се получава чрез механично сканиране на референтното огледало. При FD-OCT този профил се извлича чрез записване на интерференчния сигнал като функция на дължината на вълната на източника на светлина и последваща Фурие трансформация [13], [14]. Докато TD-OCT системите предлагат потенциални предимства по отношение на цена и по-дълбоко проникване [15], [16], FD-OCT осигуряват значително по-високи скорости на запис [17], по-добро отношение сигнал/шум (SNR) и по-висока аксиална разделителна способност [18]. FD-OCT системите могат да се подразделят на SS-OCT и SD-OCT (OCT в спектралната област). Те са допринесли значително за увеличаване на скоростта на запис [19], като са се развили от първоначалните TD-OCT системи с около 2 A-скана в секунда до няколко милиона A-скана в секунда при

водещите съвременни SS-OCT системи [13]. Поради това системата, която е във фокуса на тази дисертация, е SS-OCT.

За запис на данните в настоящата работа беше използвана комерсиална система на Optores GmbH, снабдена с FDML лазерен източник, работещ при централна дължина на вълната 1309 nm, спектрална ширина 100 nm и честота на сканиране 1.6 MHz [20], [21]. В SS-OCT системите сигналът се регистрира последователно във времето чрез точков детектор, формирайки спектрални интерферограми. Чрез прилагане на Фурие трансформация върху този сигнал се реконструира пълният профил в дълбочина (аксиална структура), използвайки целия спектър на сигнала за формиране на изображението.

Напредъкът в областта на цифровата обработка на сигнали (DSP) позволява по-бързи реализации, подобрявайки изчислителните разходи, обработката в реално време, надеждността и мащабируемостта в рамките на OCT обработката. Основните стъпки при обработката на данни [22] в SS-OCT системите включват: запис на сурови интерферограми, премахване на DC компонента, k-линеаризация, компенсация на дисперсията, прилагане на филтри-прозорци, едномерна FFT, логаритмично мащабиране на сигнала, потискане на шума и визуализация. Тъй като сигналът в SS-OCT системите се дискретизира равномерно по дължината на вълната, а не по вълновото число, за да се извлече от спектралните интерферограми информацията по дълбочина, е необходима Фурие трансформация. За тази цел данните трябва да бъдат линейни по вълновото число (k). Следователно се извършват прекалибриране и k-линеаризация в k-пространството, за да се осигури реконструкция с висока достоверност, обикновено чрез интерполационни техники за изчисляване на новите данни [23]–[25]. Тези техники обаче могат да доведат до проблеми като повишен шум, артефакти и увеличена изчислителна сложност, което може да повлияе негативно върху обработка в реално време. Освен това, методите често разчитат на функция за ремапинг, извлечена чрез независимо калибриране, и могат да изискват допълнителна хардуерна поддръжка. Хардуерни решения като k-часовници [26] често срещат затруднения, особено при скорости над 1,3 гига-извадки в секунда (GSPS). Използването на двуканален запис в MEMS е ефективно [27], но увеличава сложността на системата. Лазерни методи [28], използващи акинетични лазери за линеаризация по вълново число, също показват добри резултати, но високата им цена ограничава комерсиалното приложение.

Последният напредък в областта на машинното зрение, движен от развитието на методологии за дълбоко обучение, значително трансформира множество аспекти на обработката на изображения: детекция на характеристики, подобрение на качеството, потискане на шум, класификация и реконструиране на изображенията в широк спектър от приложения – от здравеопазване до биомедицински изследвания. Тези модели не само могат да произвеждат висококачествени, визуално ясни реконструкции, но и да предоставят диагностично значими и ясни за интерпретация изображения. Водещите съвременни DL модели, които се прилагат за различни задачи по обработка на изображения, често използват архитектури като U-Net [29], ResNet [30], SegNet [31], PSPNet (Pyramid Scene Parsing Network) [32], Deeplab3+ [33] и с механизми за внимание [34]. Макар че тези архитектури значително са допринесли за напредъка в областта на обработката на изображения, те имат ограничения при прилагане към OCT данни – било поради висока изчислителна сложност, било поради ниска точност, произтичаща от трудности при улавянето на дългообхватни зависимости и други фактори. Въпреки това, бъдещите разработки в DL архитектурите са насочени към удовлетворяване на изискванията на специфичните приложения и източници на изображения.

Глава 3 Анализ на SS-OCT изображения на вени с дълбоко обучение (DL)

Изследванията в тази глава се отнасят до публикации на автора P1-P2. Те се фокусират върху OCT-базирано изследване на съдови структури в човешкото тяло, физическите им

характеристики и свързаните с вените проблеми, като например разширените вени. Предложени са методи за оценка на формата и дебелината на разширените вени, използвайки DL. В литературата съществуват няколко модела, които могат да извършват сегментация за физически анализ на вените, но е силно желателно да има оптимизиран модел с подобрена изчислителната сложност чрез минимизиране на броя на използваните параметри, заедно с високо точни резултати от сегментирането, както е представено в публикация I. Тази глава разглежда подробно как директното изчисляване на дебелината от семантичните карти от сегментирането може да бъде погрешно. DL моделът, предложен в публикация II, преодолява този проблем, използвайки метод, базиран на трансферно обучение, по уникален начин и позволява много прецизна оценка на дебелината на вените.

3.1. Въведение във венозната система и свързани с нея заболявания

3.1.1 Вените и тяхното значение

Човешката кръвоносна система се състои от сърце и съдове, които пренасят кръв, кислород и основни хранителни вещества до различни части на тялото. Артериите пренасят наситена с кислород кръв и се разделят на по-фини съдове, наречени капиляри. Вените са отговорни за пренасянето на близо 70% от общия обем на кръвта, а сърдечната дейност е силно зависима от вените за своята функционалност. Те могат да бъдат засегнати от няколко заболявания, като например хронични венозни заболявания, често срещани като разширени вени. Следователно, механиката на вените трябва да бъде изучавана и анализирана, за да се разберат по-добре факторите, които променят свойствата на вените при болните пациенти. Разбирането на промените, като промени в състава на стените, скованост и др., е от решаващо значение за изясняване на прогресията на заболяването и разработване на специфични за вените иновативно лечение.

За всички често използвани механични тестове, анализът на параметрите се основава на геометрията на тестваните образци. Тези параметри могат да включват оценка на 3D формата, дебелината, обиколката, дължината, площта на напречното сечение и др. Основните променящи се параметри за анализ са формата и размерът. Това изискване прави изключително важно да има визуален 3D модел на формата на тестваните съдови структури и да се оценят физическите свойства с висока прецизност и точност. С напредъка в областта на образната диагностика, базираните на 3D модели изображения могат лесно да бъдат изградени за съдовите структури, за да се улови тяхната морфология. Това ни позволява да обработваме обемите и да ги визуализираме в различни равнини под различни ъгли. Точната сегментация отваря възможности за оценка на размери като диаметър, дебелина и площ на напречното сечение. Целта на тази дисертация беше да се получат ex-vivo венозни обеми, използвайки SS-OCT система, за да се анализират, проучат и измерят няколко физически характеристики на вените при болни пациенти, страдащи от разширени вени. В настоящата дисертация, DL-базирани методи се използват за извличане на пълна информация под формата на сегментационни карти и стойности за дебелината на стените от тези обемни сканирания, описани в разделите по-долу.

3.1.2 База данни с изображения на вени: дизайн, запис и обработка

3D модели на изображенията на вените са създадени с помощта на SS-OCT метод, за да предоставят пространствено организирана структурирана информация и да допълнят DL моделите с морфологични детайли. За създаването им е осъществено сътрудничество с Университетска болница „София Мед“, София, България, където е извършена хирургична интервенция под ръководството на проф. д-р Димитър Николов, ръководител на катедра „Съдова хирургия“. Етично одобрение за това изследване е получено от Етичната комисия на Българската академия на науките (разрешение № 1-44/11.06.2021). Всички експерименти са проведени в съответствие с принципите, изложени в Декларацията от Хелзинки (1975 г.), изменена през 2013 г.

В това проучване образците бяха получени чрез процедурата за отстраняване на разширените вени от болните пациенти. Всеки пациент беше подготвен съгласно стандартния хирургичен протокол. Извлечените вени бяха консервирани с помощта на физиологичен разтвор по преценка на лекаря. В рамките на 24 часа след хирургичната процедура вените бяха използвани за събиране на обемни данни. Пробата беше поставена в стерилна петриева паничка за прецизно изобразяване, като се избягват замърсявания. По време на процедурата по събиране пробата беше навлажнена с разтвора, за да се предотврати сухота на повърхността. Чрез внимателно поддържане на контролните условия бяха запазени физичните и оптичните характеристики на вената, за да се извърши прецизно ОСТ получаване на изображения.
Данните за вените са записани с помощта на системата SS-OCT от Optores GmbH [20],[21], описана в Глава 2. Наборът от данни се състои от пет различни обема, съдържащи 1300 B-скана, получени от четири пациента. Подробности относно пола и възрастта на пациентите участници и броя на обемите, получени от всеки пациент, могат да бъдат намерени в Таблица 3.1. Всяко B-сканиране, обхващащо 1024 A-скана, впоследствие беше изрязано и преоразмерено до 512 × 256 пиксела, за да се фокусира върху областта от интерес и да се отговори на изчислителните ограничения по време на обработката. Изображенията са под формата на B-сканиране и са засегнати от шум, главно от спекъл. Следователно, премахването на шума се извършва с помощта на минимизиране на претеглената ядрена норма (WNNM) [35].

Таблица 3.1 Описание на пациентите от ОСТ базата данни

Пациент	Пол	Възраст	Обени	Вена	B-скан (във всеки обем)
P1	Мъж	57	1	Разширена вена	(200)
P2	Мъж	68	2	Разширена вена	(650,200)
P3	Мъж	33	1	Разширена вена	(100)
P4	Жена	38	1	Разширена вена	(150)

3.2. Сегментиране и измерване на дебелината на стените на вените

Този раздел демонстрира формулировката на проблема за сегментиране на вени, последвана от представения метод за извършване на сегментиране на венозни слоеве чрез оптимизиране на броя на параметрите. Подходът за сегментиране има за цел да раздели вътрешните и външните граници. Освен това, моделът се използва за точно измерване на дебелината като ключов физически параметър, необходим за разбирането на функционалността на вените или техните изкуствени заместители. Лекарите могат да очертаят слоевете и структурите чрез ръчно сегментиране на всеки срез от ОСТ обема. Това обаче отнема много време и изисква обучен експерт. Съществува допълнително предизвикателство в приложенията в реално време. С наличието на техники за сегментиране в биомедицинско изобразяване, управлявани от изкуствен интелект, тази задача може да се изпълни ефективно с помощта на подходящ модел. Разглеждат се и различни съвременни подходи, базирани на DL, за извършване на пикселна сегментация и за получаване на стойността на дебелината чрез регресионен анализ.

U-Net [29] е широко признат за ефективната си сегментация в случай на малък набор от данни, но неговата изключително сложна архитектура, включваща милиони параметри, налага значително натоварване на хардуерните ресурси. Намаляването на броя на параметрите позволява по-бърз извод, по-малък размер на модела и използване на по-малко памет при обработка на големи партиди или изображения с по-висока резолюция по време на обучение. Друго ключово предизвикателство, засягащо сегментацията в архитектури, базирани на енкодер-декодер, е директното увеличаване на резолюцията (upsampling) на

региони, получени от максималното обединяване, извършено от страната на енкодера, които притежават ограничено рецептивно поле.

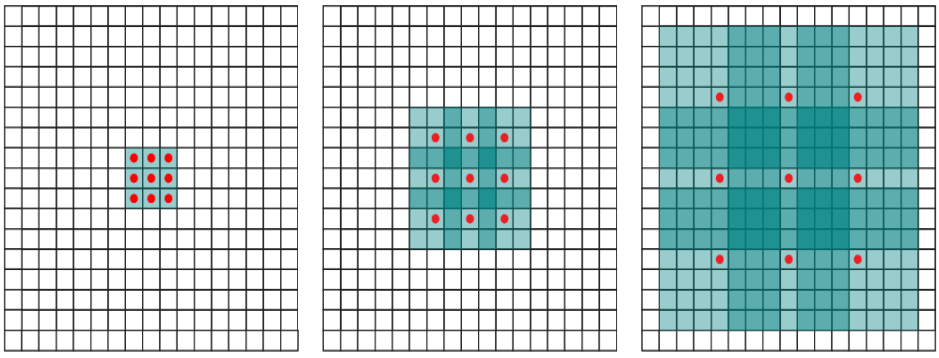
За да се създаде високо оптимизиран DL модел, въвеждаме усъвършенствана архитектура, базирана на енкодер-декодер, както е представено в Публикация I. Тя е проектирана да намали сложността чрез минимизиране на броя на параметрите, като същевременно гарантира точна пикселна сегментация на вените. Това се постига чрез въвеждане на паралелен мостов блок, вграден с атрус и дълбочинно разделими конволюционни слоеве. Моделът е допълнен и с остатъчни връзки [30], за да се разреши проблемът с изчезващите градиенти, често срещани в по-дълбоките архитектури. Такива връзки между различните мрежови слоеве улесняват безпроблемния поток на информация към следващите слоеве. Този подобрен поток от информация е постигнат чрез навлизането на съпоставяния на характеристиките от един слой към следващите слоеве чрез идентичностни съпоставяния. Математически, конволюцията може да бъде представена с помощта на входни данни X∈ $Z^{H\times W\times D}$, където H, W, и D означават размерите на картата на характеристиките и W_g е филтър с размер (k^f, l), както следва:

$$\sigma_C\left(X, W_g\right)_{(i,j)} = \sum_{k^f,l,m}^{K^f,L,M} X(k^f, l, m) * W_g(i + k^f, j + l, m), \quad (3.1)$$

В уравнение (3.1), ***m*** представлява каналите на входа, а позицията на картата на входните характеристики е маркирана с индекси ***(i, j).*** Атрус или разширена конволюция с коефициент на дилатация ***r*** може да се изрази като:

$$\alpha\left(X, W_g\right)_{(i,j)} = \sum_{k^f,l}^{K^f,L} X(i + rk^f, j + rl) * W_g(k^f, l), \quad (3.2)$$

Систематичната дилатация чрез атрус конволюция позволява експоненциално разширяване на рецептивното поле, като същевременно се запазват както разделителната способност, така и покритието. Това улеснява интегрирането на контекстуална информация на множество мащаби. Процесът представлява семплиране, при което честотата на семплиране се изразява чрез стойността на дилатацията. Използването на дилатация позволява извличане на връзки от по-високо ниво, като спомага за намаляване на загубите в точността, предизвикани от процесите на обединяване и увеличаване на резолюцията. На Фигура 3.1 е показана дилатираната конволюция с различни рецептивни полета и стойности на дилатация.



Фигура 3.1 Дилатация или атрус конволюция: (a) 1-дилатирана конволюция с рецептивно поле 3 × 3; (b) 2-дилатирана конволюция с рецептивно поле 7 × 7; (c) 4-дилатирана конволюция с рецептивно поле 15 × 15.“

Основната разлика между стандартната конволюция и дълбочинно-разделимата конволюция (depthwise separable convolution) е в начина, по който конволюционното ядро взаимодейства с входните канали. Дълбочинно-разделимата конволюция, показана на Фигура 3.2, се състои от две стъпки: първата включва независимо прилагане на пространствена конволюция към всеки канал на входа, а втората представлява 1×1 поточкова конволюция, която проектира изходните канали от дълбочинната конволюция в ново пространството на каналите.

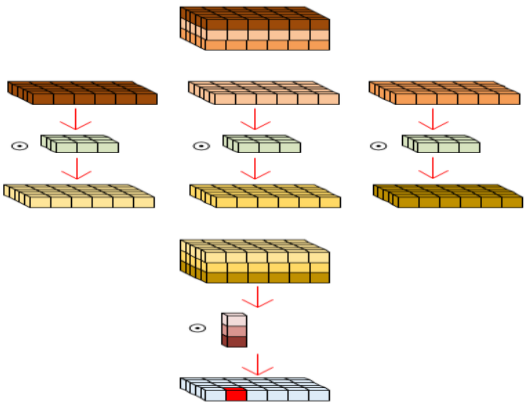
Двустъпковият процес, включващ поточкова (pointwise) и дълбочинна (depthwise) конволюция, е описан съответно в уравнения (3.3) и (3.4), както следва:

$$P(X, W_g)_{(i,j)} = \sum_m^M X(m) * W_g(i, j, m), \tag{3.3}$$

$$\partial(X, W_g)_{(i,j)} = \sum_{k^f,l}^{K^f,L} X(k, l) * W_g(i + k^f, j + l). \tag{3.4}$$

Дълбочинно-разделимата конволюция може да бъде записана като:

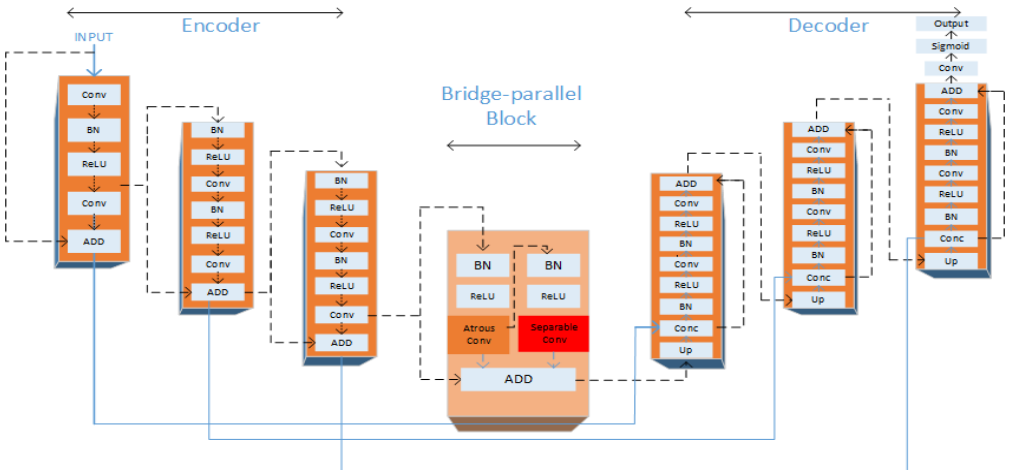
$$\delta(X_p, X_{(d)}, W_g)_{(i,j)} = P_{(i,j)}\left(X_p, \partial_{(i,j)}(X_{(d)}, W_g)\right). \tag{3.5}$$



Фигура 3.2 Дълбочинно-разделима конволюция

3.2.1 Оптимизиране на сегментирането на изображения на вени

За дизайна на ефективен модел за сегментация е от съществено значение оптимизацията чрез намаляване на сложността - една от цели на настоящата дисертация. В тази работа се предлага модел, представляващ енкодер-декодер, изграден върху архитектурата U-Net [29]. Моделът е обозначен като Opto-UNet [П I] и е предложен в Публикация I. Той включва паралелен свързващ-блок (bridge block), който използва атрус и разделими конволюции. Това интегриране подобрява извличането на пространствено обхватни и отчетливи карти, което води до по-добра ефективност в сравнение със съвременните водещи модели. Допълнително, мрежата е оптимизирана чрез използване на дълбочинно-разделима конволюция, което минимизира броя на параметрите и значително намалява сложността на модела.



Фигура 3.3 Блок схема на Opto-UNet модела за сегментиране на изображения на разширени вени (П.I)

Блок-схемата на Opto-UNet е показана на Фигура 3.3. Трите основни части, подчертани тук, са енкодерът, декодерът и паралелният bridge-блок. Такива по-дълбоки мрежи предлагат предимството на повишена точност, но по време на обучението могат да претърпят деградация на ефективността, основно поради често срещания проблем с изчезващите градиенти. За да бъде преодолян този проблем, в предложената мрежа са включени резидуални блокове, които осигуряват допълнителен трансфер на информация между

последователните слоеве в дълбоките мрежи, както е показано на Фигура 3.3 с черни прекъснати линии. Всеки от трите блока от страната на енкодера включва различни слоеве: конволюция (Conv), batch нормализация (BN), Изправена линейна функция (Rectified Linear Unit (ReLU)) и слой за събиране (ADD). В допълнение към тези слоеве, трите блока от страната на декодера съдържат и слоеве за увеличаване на резолюцията (Up) за разширяване на картите на признаците чрез операция по увеличаване на резолюцията, както и слой за конкатенация (Conc) за реализиране на skip-свързвания между енкодера и декодера, отбелязани със сини линии. Ключовият елемент на предложения модел е паралелният bridge -блок, разположен в центъра на мрежата, както е показано на Фигура 3.3. Тук картите на признаците са в най-компактен вид, като същевременно запазват най-важната контекстуална информация. Този блок включва атрус конволюция (оранжев блок) и дълбочинно-разделима конволюция (червен блок), както и BN и ReLU функции. Картите на признаците на изхода от паралелния bridge-блок произхождат от два източника: едни, които улавят пространствена информация, и други, които включват както пространствени, така и дълбочинни характеристики. Това позволява на модела да усвоява разнообразни признаци едновременно. Блок-схемата показва, че изходите от двата нововъведени слоя (атрус и разделима конволюция), заедно с операцията identity mapping, се комбинират в последния слой чрез събиране. Изходът от паралелния bridge-блок се насочва към декодерните блокове. Декодерите извършват разширяване на тези карти на признаците и използват информация от енкодерните блокове чрез skip-свързвания, за да постигнат прецизна локализация на структурните детайли. Всички конволюции (Conv) в енкодера, декодера и паралелния bridge-блок използват двумерно ядро с размер 3×3, със stride 1×1 и padding, зададен на „same“. Стойността на дилатацията за атрус конволюцията е 2.

Данните за обучение на мрежата се състоят от входни изображения и еталонни (ground truth) изображения, получени от FDML SS-OCT системата на Optores GmbH, описана в Глава 2. Поради хардуерни ограничения, входните изображения представляват изрязани и преоразмерени B-сканове с размер 512×256 пиксела, извлечени от оригинални OCT B-сканове с размер 1024×1152 пиксела. Интензитетите на всички входни изображения са нормирани в диапазона от 0 до 1. Общият брой изображения, използвани за обучение на мрежата, е 800 B-скана, разделени на обучаващи, валидационни и тестови съответно в съотношение 600, 100 и 100. Еталонни изображенията бяха получени чрез ръчно анотиране на B-сканoвете за съответната област на интерес под ръководството на проф. д-р Димитър Николов от Катедра по съдова хирургия, който участва и в извличането на вените, описано в раздел 3.1. Определени области от венозния слой бяха изключени от сегментационните маски поради предизвикателства при точното разграничаване на граничните пиксели. Тези еталонни изображения могат да бъдат представени като $y_i \in \{0,1\}$ за всяка пикселна стойност, при общ брой N^p пиксела. За функция на загуба е използвана Dice-загубата, извлечена от коефициента на Sørensen–Dice. Тя се прилага за минимизиране на разликата между предсказания изход и еталонния изход (ground truth). Тази функция е добре известна със своята ефективност при справяне с проблема за дисбаланс между класовете в задачи за сегментация. В конкретния случай, тъй като преднният план заема значително по-малка площ в сравнение с фона, е използвана именно Dice-загубата. Dice коефициентът може да бъде изразен по следния начин:

$$DCC = \frac{\sum_{i=1}^{N^p} p_i y_i + \epsilon}{\sum_{i=1}^{N^p} p_i + y_i + \epsilon} + \frac{\sum_{i=1}^{N^p} (1-p_i)(1-y_i) + \epsilon}{\sum_{i=1}^{N^p} 2 - p_i - y_i + \epsilon}, \quad (3.6)$$

Тук, в уравнение (3.6) за DCC, $p_i \in [0,1]$ представлява предсказаната стойност на пиксел от изображението, а $\epsilon = 1$ е фактор за изглаждане, използван за избягване на деление на нула в знаменателя. DCC съдържа два члена - единият за фона, а другият за обекта, но в изчисленията се взема предвид само загубата, отнасяща се за обекта. Загубата по Dice-коефициент $Loss^1$ може да бъде записана по следния начин:

$$Loss^1 = 1 - DC' \quad (3.7)$$

Тук DC' в уравнение (3.7) включва само първия член на DC от уравнение (3.6), като се отчита единствено загубата за областта на фона. Предложеният модел Opto-UNet беше изпълнен върху графичен ускорител NVIDIA GeForce RTX 3060 с 12 GB графична памет и 64 GB RAM. Обучението и тестването бяха проведени с използване на рамката TensorFlow 2.0, като се приложи оптимизаторът RMSprop с коефициент на обучение (learning rate) 0.0001 и загубата по Dice-коефициент ($Loss^1$), при размер на партидата (batch size) от 16.

3.2.2 Експериментални резултати и дискусия

За анализ на ефективността на предложения модел Opto-UNet, бяха проведени експерименти върху записаните венозните данни. Получените резултати от предложения модел са представени след систематични оценки и сравнени с водещи архитектури, а именно U-Net [29], SegNet [31] и PSPNet [32]. Всички тези модели бяха обучени при идентични условия и в същата изчислителна среда като предложения Opto-UNet, с цел да се гарантира коректност на сравнението. Сравнението с други сегментационни модели се основава на ключови показатели за оценка, включително точност (Acc), пресичане върху обединение (IoU), Dice-коефициент на сходство (DSC), чувствителност (SEN) и специфичност (SPE), които са математически дефинирани по следния начин:

$$\text{Acc}=\frac{TP+TN}{TP+TN+FP+FN}, \tag{3.8}$$

$$\text{IoU}=\frac{TP}{TP+FP+FN}, \tag{3.9}$$

$$\text{SEN}=\frac{TP}{TP+FN}, \tag{3.10}$$

$$\text{SPE}=\frac{TN}{TN+FP}, \tag{3.11}$$

$$\text{DSC}=\frac{2*SEN*SPE}{SEN+SPE}. \tag{3.12}$$

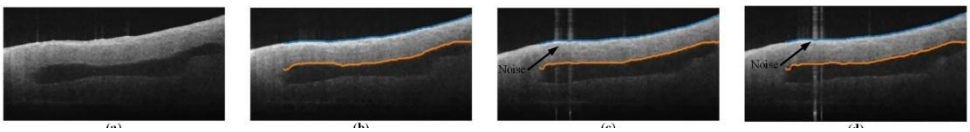
В уравнения (3.8)–(3.12) TP, TN, FP и FN обозначават броя пиксели, класифицирани съответно като истински положителни, истински отрицателни, фалшиво положителни и фалшиво отрицателни за резултатите, получени от модела. Таблица 3.2 представя обобщение на различните модели и тяхната ефективност върху набора от данни за разширени вени, което позволява директно сравнение между Opto-UNet и водещите съвременни архитектури. Като цяло моделът Opto-UNet постига най-добри резултати по всички метрики, подчертани в удебелен шрифт в Таблица 3.2. Той достига точност на сегментация 0.9830, значително по-висока от тази на SegNet и съпоставима с резултатите на U-Net. Освен това броят на използваните параметри е съществено по-нисък в сравнение с всички останали модели. От таблицата ясно личи, че SegNet [31] показва чувствително по-слаби резултати, особено по метриките SEN и IoU, поради високия дял на фалшиви детекции. За разлика от него, предложеният модел демонстрира устойчиво подобрение по всички показатели спрямо U-Net [29], като същевременно осигурява съществено повишение на ефективността чрез

Таблица 3.2 Сравнение на резултатите от сегментиране за разширени вени (P.I)

Модел	Acc	SEN	SPE	DSC	IoU	PARAMETERS
U-Net	0.9791	0.8346	0.9968	0.9085	0.8210	31.04M
SegNet	0.8868	0.1340	0.9031	0.2333	0.0245	29.45M
PSPNet	0.9619	0.8142	0.9813	0.8899	0.7127	48.70M
Opto-UNet	0.9830	0.8425	0.9980	0.9136	0.8395	8.54M

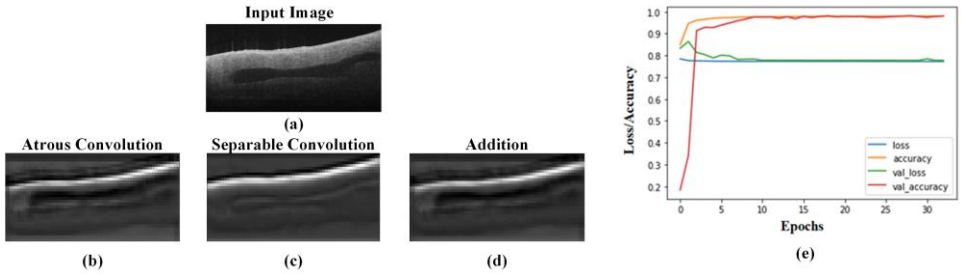
използването на приблизително 3.5 пъти по-малък брой параметри. За илюстриране на

резултатите, на Фиг. 3.4 е показано (a) оригиналното оразмерено B-scan изображение, генерирано от ОСТ системата, както и (b), (c) и (d) – три представителни скана, при които горната и долната граница са очертани съответно в син и оранжев цвят.



Фигура 3.4 (a) B-скан за разширени вени от SS-OCT (b),(c),(d) резултати от сегментазия с Opto-UNet (P.I)

Стрелките, показани на Фиг. 3.4(c) и (d), са използвани за илюстриране на точността на сегментационния модел, когато B-scan изображенията са повлияни от артефакти. Следва да се отбележи също, че фоновата област е засегната от спекъл шум дори след изглаждане на изображенията чрез прилагане на WNNM филтриране [35]. Моделът Opto-UNet като цяло демонстрира изключително висока прецизност и ефективно се справя с шума. За да се подчертае влиянието на ключовите функции, реализирани в рамките на паралелния мостов блок чрез интегриране на (i) атросна (atrous) конволюция, (ii) разделима (separable) конволюция и (iii) операция по събиране, на Фиг. 3.5(b), (c) и (d) са представени резултатите от изходните канали за съответните слоеве. Тези визуализации осигуряват поглед върху работата на Opto-UNet, като демонстрират начина, по който входното изображение навлиза през различните слоеве на паралелния мостов блок. Предимството от интегрирането на тези конволюционни слоеве за извличане на структурни детайли както на пространствено, така и на дълбочинно ниво ясно личи във картата на Фиг. 3.5(d), получена след операцията по събиране. Допълнително, на Фиг. 3.5(e) е представена динамиката на промяната в загубата (loss) и точността по време на обучението и валидирането на модела спрямо броя епохи. С увеличаването на броя на епохите обучителната загуба (loss в Фиг. 3.5(e)) постепенно намалява и се стабилизира около 15-та епоха. Наблюдава се сходна тенденция и при кривата на точността, което свидетелства за бърза конвергенция на модела Opto-UNet.



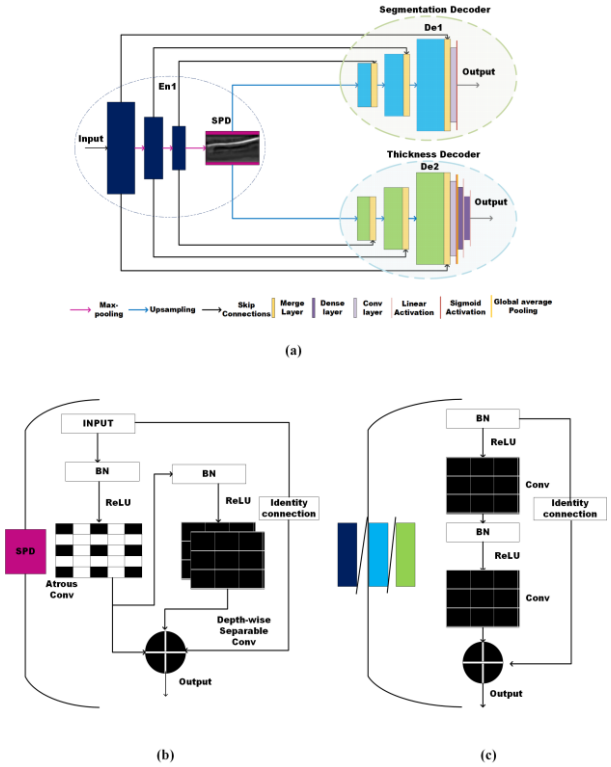
Фигура 3.5 (a) Оригиналния B-скан на вена от SS-OCT; (b),(c),(d) сегментации от Opto-UNet, (e) криви на загубата и точността по време на обучението и валидирането на Opto-UNet (P.I)

В обобщение, по-добри резултати бяха постигнати чрез интегрирането на слоеве с паралелния мостов блок, комбинирано с използването на резидуални връзки и прилагането на загуба, базирана на Dice-коефициента, в предложения модел Opto-UNet. Намаляването на броя параметри, спрямо водещите съвременни модели, е съществено. Като цяло предложеният модел демонстрира бърза конвергенция и висока точност, което го прави обещаващо решение за сегментация на разширени вени. Въпреки това, като едно от първоначалните експериментални изследвания в рамките на дисертацията, основното ограничение на работата, представена в Публикация I, е свързано с обхвата на набора от данни, включващ изображения само от един пациент, диагностициран с разширени вени. Подходът ефективно съхранява разделянето на информацията по дълбочинни канали на различни пространствени позиции, което води до по-информативна и по-дълбока мрежа. В изследването се разглежда и оценката на дебелината на вените. Тя може да бъде директно изчислена чрез сегментационната карта, получена от различните модели за сегментация [29]–[33], [36]. Въпреки това този подход може да доведе до натрупване на грешки поради

потенциални неточности в картите. Затова беше предложен двоен предсказващ модел, базиран на самообучаващо трансферно учене, за едновременно предсказване на дебелината и сегментационната карта. Подходът с трансферно обучение бе използван за дообогатяване на признаците, извлечени по време на първоначалната оптимизация за сегментация. Този ансамблов модел е способен да предсказва с висока точност стойностите на дебелината наред със сегментационните карти, като използва последователно обучение с една енкодерна и две декодерни архитектури. В Раздел 3.3 е представен предложеният модел за едновременна сегментация и оценка на дебелината върху по-голям набор от данни..

3.3. Сегментиране и измерване на дебелина с използване на трансферно обучение

За да се подобри разбирането на характеристиките на вените, разработихме автоматизиран модел, предназначен за едновременна сегментация и оценка на дебелината, който може да изчислява формата и дебелината на областта на интерес. Този ансамблов модел е наречен TREE (TransfeR learning-based approach for thicknEss Estimation). Той следва модифицирана архитектура енкодер-декодер, при която един енкодер служи като основа, разклонявайки се към два декодера. Декодерите се специализират съответно в предсказване на сегментационната карта и стойностите на дебелината. Архитектурата на мрежата и експерименталните детайли са описани по-долу. На Фиг. 3.6 е представен моделът TREE, който следва дизайн с един енкодер и два декодера. Основната архитектура на енкодер-декодер тук е вдъхновена от **U-Net [29]**. Кондензационният (bottle neck) модул между енкодера и декодера в U-Net [29] съдържа високоразредни структурни детайли и играе



Фигура 3.6 (a) Блок схема на предложения модел, показваща общата архитектура на енкодера и двата декодери, (b) модула Sensitive Pattern Detector, (c) блоковете на енкодера или декодера, показващи слоевете и потока на информацията. (P.II)

ключова роля за генериране на точни предсказания.

Предложеният модел **TREE** използва уникално тези признаци, за да улесни трансфера на детайли от задачата за сегментация към оценката на дебелината. Моделът се състои от три блока: един енкодер и два декодера. Декодерът, наречен **De1** (показан на Фиг. 3.6(a)), произвежда сегментационната карта чрез семантична сегментация на пикселите, докато другият декодер, наречен **De2** (показан на Фиг. 3.6(a)), оценява дебелината на вените.

Енкодерът, наречен **En** (показан на Фиг. 3.6(a)), е свързан към двата декодера, създавайки две отделни разклонения: **En-De1** и **En-De2**. Изходът на енкодера преминава през кондензационен модул (показан в розово на Фиг. 3.6(a)), който обслужва и двата декодера. Този модул подобрява производителността на модела, като се фокусира върху областта на интерес във всеки B-scan.

Както се вижда на Фиг. 3.6(a), енкодерът (**En**) и двата декодера (**De1, De2**) съдържат модули, маркирани съответно с тъмносини, светлосини и светлозелени правоъгълни контури. Тези модули са допълнително разяснени на Фиг. 3.6(c), за да се демонстрира структурата на слоевете в тях. Всички тези модули следват процеса: (i) нормиране чрез слой Binary Normalization (BN), последвана от ReLU активиране; (ii) две последователни конволюционни (Conv) операции с междинни стъпки чрез BN и ReLU слоеве; (iii) добавяне на изходната карта на признаците към входната карта чрез идентични връзки, известни като резидуални връзки [30], както е показано на Фиг. 3.6(c). Тези връзки подпомагат мрежата да постигне стабилно обучение с нарастването на броя на слоевете при по-дълбоки модели, като главно адресират проблема с изчезващите или експлодиращи градиенти.

В модела е включен блок, наречен Sensitive Pattern Detector (SPD), разположен в точката на свързване между енкодера и декодерните разклонения, показан като розов контур на Фиг. 3.6(a) и (b). В енкодера (En) всеки модул е последван от слой за максимално обединяване (max pooling), който намалява размера на картите на признаците чрез използване на максималната стойност за дадена област. След преминаване през SPD блока, картата на признаците се обогатява с глобална информация и преминава към декодера за разширяване на картите чрез увеличаване на резолюцията. Слоевете за максимално обединяване и увеличаване на резолюцията (upsampling) са предназначени за извличане на детайли на множество резолюции. Глобалните и локални детайли от страните на енкодера и декодера съответно се обединяват, за да се улесни потокът на информация при еднаква резолюция на картите.

Следвайки математическото определение за стандартна конволюция, дадено в уравнение (3.1), можем да представим общата обработка на блока на енкодера/декодера по следния начин:

$$Y_{en} = \sigma_C(ReLU(BN(\sigma_C(ReLU(BN(X)), W_g), W_g)(ReLU(BN(X)), W_g). \quad (3.13)$$

За да се получи глобалната структурна информация, определяща границите или слоевете, изходът от всеки блок на енкодера се намалява чрез слой за максимално обединяване. След преминаване през блоковете на енкодера, изходът достига до кондензационен модула, който съдържа специална комбинация от атрусна (степен на дилатация = 2) и дълбочинно-разделима конволюция, както е описано по-горе (Публикация I).

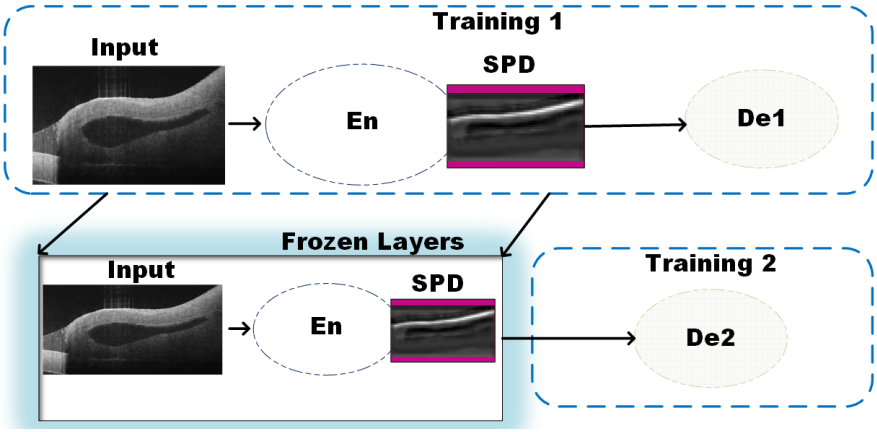
Стандартната конволюция поставя ограничения върху механизма на извличане на признаци поради ограниченото си рецептивно поле. Това води до ограничаване на обмена на информация до малка област, дефинирана от размера на ядрото (обикновено 3×3). Това представлява предизвикателство, когато сегментационните маски включват пространствени зависимости на дълги разстояния. Вените проявяват удължени структури, които изискват улавяне на разширени зависимости. За решаване на този проблем се използват дълбочинно-разделими и атрусни конволюции, които извличат по-дълбоки и дългосрочни карти на признаците, позволявайки на модела по-ефективно да научава корелациите.

Както е показано на Фиг. 3.6, **SPD** модулът съдържа **BN** и **ReLU** функции за активация заедно с дълбочинно-разделими и атрусни слоеве, разположени паралелно. В контраст с кондензационния модул от Публикация I, дълбочинно-разделимата конволюция тук е допълнена със същата степен на дилатация (dilation rate) = 2, както при атрусната конволюция, за да се запази прозрачността и плътността на двата слоя при интеграцията чрез операцията по добавяне.

След като се получи интегриран изход от конволюционните операции в последния слой, крайната карта на признаците се формира чрез комбиниране с резидуална връзка, за да се подобри представянето на признаците, както е показано на Фиг. 3.6.
Следвайки уравнение (3.5), общото изражение за SPD блока може да се формулира като:

$$Y_{bb} = \alpha\big(ReLU\big(BN(X)\big), W\big) + \delta'\big(ReLU\big(BN(X)\big), W\big) + \varphi(BN(X), W). \quad (3.14)$$

В уравнение (3.14) символите α, δ', φ представят съответно атрусната конволюция, дълбочинно-разделимата конволюция (със степен на дилатация = 2) и резидуалната операция. Глобалните признаци, извлечени с помощта на **SPD** блока, улесняват процеса на обучение както за сегментация, така и за оценка на дебелината на варикозните вени. За постигане на точност на ниво пиксел, тези карти на признаците се увеличават чрез увеличаване на резолюцията, осигурявайки прецизна сегментация и оценка на дебелината.
Блоковете на декодерите **De1** и **De2** са показани на Фиг. 3.6(a) и (c) с цветове синьо и зелено съответно. В **De1** крайните слоеве след трите основни блока на декодера извършват конволюция и **Sigmoid** активация за генериране на сегментационни карти. За разлика от това, както е показано на Фиг. 3.6(a), в **De2** блоковете са последвани от комбинация от конволюция, **Sigmoid**, **Global Average Pooling** и плътни слоеве, предназначени за получаване на оценка на дебелината. Важно е да се отбележи, че всички конволюции в енкодер и декодер блоковете, включително SPD модула, използват 2D ядро с размер 3×3, с **stride = 1×1** и padding, запазващ същите пространствени размери. За улавяне на ключовите признаци, необходими за сегментация и оценка на дебелината, обучението на предложен модел **TREE** се извършва

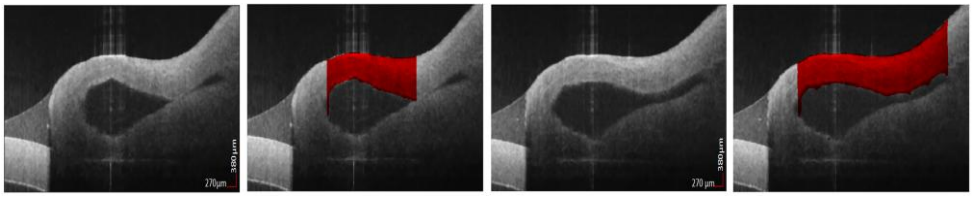


Фигура 3.7 (a) Блокова схема, показваща фазата на обучение на предложения метод, включваща Енкодер En и съответно двата Декодера De1, De2 (P.II)

стъпка по стъпка, както е показано на Фиг. 3.7. Първо моделът се обучава за сегментационната карта чрез разклонение 1, включващ **En** и **De1**. Оценката на дебелината се извършва след това чрез обучение на второто разклонение, включващо **En** и **De2**. Последователното обучение използва сегментационната карта на ниво пиксел, идваща от SPD и En, която преминава през слоевете на втория декодер (**De2**) за ефективно научаване на детайлите, необходими за прецизна оценка на дебелината.
Във втората фаза на обучение се тренира само декодер **De2** да извliqa нови признаци за точна оценка на дебелината. В този етап теглата на енткодера **(En)** и SPD се наследяват от първоначалното обучение, предназначено за сегментация.
Освен отличителната архитектура на предложения модел, ние въвеждаме и иновативен метод за обучение на двата декодера по последователен начин. Първоначално комбинацията от първия енкодер и декодер се обучава да предсказва сегментацията, а след това вторият декодер се тренира изключително за оценка на дебелината. По този начин се използва подходът на **transfer learning**, който подпомага оценката на дебелината чрез използване на характеристиките, вече научени по време на обучението за сегментация.

3.3.1 Експериментални резултати и дискусия

В този раздел илюстрираме представянето на предложения модел TREE и го сравняваме с водещите съвременни модели, главно U-Net [29], PSPNet [32], RESUNET [36], DeepLab3+ [33] и SegNet [31], за задачите сегментация и оценка на дебелината на варикозни вени.
Използваният за това изследване датасет включва 5 обема, съдържащи общо 1300 изображения (B-scans), получени от 4 различни пациенти, както е описано в Раздел 3.1. Размерът на изображенията след изрязване и преоразмеряване е 512 × 256, с цел покриване на областта на интерес и съобразяване с изчислителните ограничения. Изображенията са обработени чрез метод за намаляване на шума [35] и след това нормирани в диапазона [0,1].
Основният датасет, съдържащ четири тома, разделен на случаен принцип на тренировъчен, валидирационен и тестов комплект в съотношение 70%, 20% и 10% съответно. Петият обем (наричан independent-test set) съдържащ 200 изображения е случайно избран, за да се оцени способността на модела за генерализация. Този обем не е използван по време на обучение или валидиране, а единствено за представяне на независимите тестови резултати.
Всяко B-scan изображение представлява различни вени, демонстриращи разнообразни форми и размери. За увеличаване на разнообразието в датасета по време на обучението са приложени геометрични трансформации като транслация и отразяване. Изображенията са транслаирани с 20% от тяхната височина и ширина, като отразяването е извършено както хоризонтално, така и вертикално.
Сегментационните карти, служещи като еталонни изображения и показани на Фиг. 3.8, са



Фигура 3.8 (а) Оригиналните изображения със сегментационната карта, показана с червено върху тях (P.II)

ръчно очертани под ръководството на експерт по съдова хирургия проф. д-р Димитар Николов от Катедра „Съдова хирургия“, споменат и в Раздел 3.1. Поради ограниченията, свързани с дълбочинното проникване на използваната SS-OCT система, долната повърхност остава скрита. Следователно в настоящата работа се разглежда само горният 2D слой на вената.
След това се изчислява еталонната дебелината за всеки ред по осевата посока, формирайки т.нар. карта на дебелината.

Системата, използвана за експеримента, е оборудвана с процесор AMD Ryzen 7, графичен процесор NVIDIA GeForce RTX 3060, 32 GB RAM и SSD с капацитет 1TB, като се използва Keras API в рамките на TensorFlow. Оптимизаторът, използван по време на обучението, е Adam с ниво на учене 0.0001. Размерът на батча е 16, като за обучението са използвани максимум 200 епохи. За да се осигури справедливо сравнение, тези параметри на обучението са прилагани последователно за всички водещи модели. В оптимизационния процес на предложения модел TREE, се използват две различни функции за загубите, съответстващи на две различни задачи. Първоначално ансамбълът енкодер (En) и декодер (De1) се обучават чрез минимизация на Binary Cross-Entropy Loss (Loss1) за прецизно извършване на сегментацията на венозните слоеве, математически представена като:

$$Loss1 = -\frac{1}{N_b}\sum_{i=1}^{N_b}(Y_i^g.\log\left(\mathrm{p}(\hat{Y}_i)\right) + \left(1 - Y_i^g\right).\log\left(1 - p(\hat{Y}_i)\right) \qquad (3.15)$$

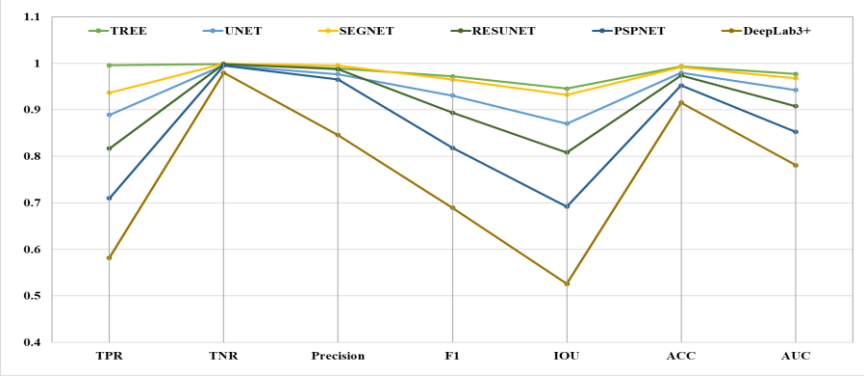
В уравнение (3.15) N_bе общият брой пиксели, $\hat{Y}_i$ е предсказаната вероятност за i-тия пиксел, а Y_i^g е еталонният етикет за i-тия пиксел. Тази функция на загубата (Loss1) се оптимизира между изхода на ансамбъла (En-De1) и ръчно анотираната маска (ground truth), чийто пиксели са класифицирани като 0 или 1, представляващи съответно фон и обект. Следващата фаза на обучението, както е описано в Раздел 3.3 и показано на Фиг. 3.7, включва обучението на

ансамбъла (En-De2) с използване на Mean Squared Error (MSE), наречен Loss2, даден в уравнение (3.16):

$$Loss2 = \frac{1}{N^C}\sum_{c=1}^{N^C}(Y_c - X_c)^2. \quad (3.16)$$

Тук N^C, Y_c и X_c представляват общия брой колони в изображението, еталона и предсказаната дебелина (в пиксели) за c-тата колона, съответно. По време на тази втора фаза теглата на блока En заедно с модула SPD са фиксирани. Следователно изходът от SPD блока, базиран на теглата, научени по време на оптимизацията за сегментация, служи като вход за De2 по време на неговата оптимизация.

За оценка на представянето и за сравнение с водещите методи се използват следните показатели: Accuracy, Area Under Curve (AUC), Intersection over Union (IoU), TNR, Recall, Precision и F1 score. Тези метрики се използват, за да се сравнят резултатите на предложен



Фигура 3.9 Графика на метриките за сегментация. (P.II)

модел TREE с водещите модели U-Net [29], SegNet [31], RESUNET [36], PSPNet [32] и DeepLab3+ [33]. Всички модели са обучавани при идентични условия като предложените, за да се осигури справедлива оценка. Резултатите от сегментацията са визуализирани чрез линейни графики на Фиг. 3.9 за обективно сравнение. Моделът TREE превъзхожда водещите методи по всички показатели, с изключение на метриката Precision, където само моделът SegNet [31] показва леко преимущество.

Дебелината беше изчислена за еталонните изображенията и за водещите методи по следния начин. Първо, за обучението на модела TREE за оценка на дебелината се използват еталонни сегментационни карти, очертани под надзора на експерт, за да се оцени стойността на дебелината за всяка колона, попадаща в областта на интерес. За водещите методи U-Net [29], SegNet [31], RESUNET [36], PSPNet [32] и DeepLab3+ [33], дебелината се изчислява от техните изходни сегментационни маски. Пикселите, очертаващи двата слоя на областта на интерес, т.е. горния и долния слой (бели пиксели с пикселна стойност y=1), се отделят от сегментационната карта. Матричният масив $\boldsymbol{M_{I\times J}}$ се използва тук, за да представи сегментационната карта, където всяка единица на картата се означава с $\boldsymbol{m_{ij}}$, като **i** и **j** съответно представляват реда и колоната. Нека C_j представлява колона от матрицата на изображението, а E^{kl} да бъде едномерна матрица, представяща границата на областта на интерес, изразена по следния начин:

$$C_j = [m_{1j}, m_{1j}, \ldots\ldots\ldots\ldots m_{I\times j}],^T \quad (3.17)$$

$$E^{kl} = [e_1, e_2, \ldots\ldots\ldots\ldots e_n]. \quad (3.18)$$

В уравнение (3.18) $kl \in$ {Layer1, Layer 2} и координатите на двата слоя L1 и L2 в колоната C_j могат да се запишат като:

$$B^{L1} = \arg\min \text{ i } C_j, \quad (3.19)$$

$$B^{L2} = \arg\max \text{ i } C_j. \quad (3.20)$$

Стойността на дебелината (в пиксели) от координатите на колоните, определени в уравнения (3.19) и (3.20), можем да изчислим като:

$$N = |B^{L1} - B^{L2}|. \tag{3.21}$$

Дебелината на вената е представена като вектор с размери 1×n, където n варира в зависимост от пространственото покритие на областта на интерес в рамките на вената. За да се оцени връзката между предсказаните и действителните стойности на дебелината, изчислени чрез уравнение (3.21), върху тестовия набор от данни са изчислени показатели за грешка. Те

Таблица 3.3 Грешки (пиксели) при определяне на дебелината на вените, изразени чрез мерките MSE, MAE, Bias, SD and RMSE (P.II)

Метод	MSE	MAE	BIAS	SD	RMSE
DeepLab3+	29.21	4.459	-2.155	4.957	5.405
U-Net	14.36	3.115	0.794	3.706	3.790
PSPNet	56.900	6.4701	-4.010	6.388	7.543
RESUNET	34.224	4.7930	2.154	5.439	5.850
SegNet	7.388	2.3684	-2.276	1.485	2.718
TREE	2.409	1.109	0.184	1.541	1.552

включват: Mean Square Error (MSE), Mean Absolute Error (MAE), Bias, Standard Deviation (SD) и Root Mean Square Error (RMSE). Чрез анализ на тези стандартни показатели за грешка при дебелината се оценява точността и надеждността на предсказанията на предложения модел TREE съпоставен с водещите методи U-Net [29], SegNet [31], RESUNET [36], PSPNet [32] и DeepLab3+ [33], като референтна стойност се използва еталонна маската. Резултатите от тези грешки са представени в Таблица 3.3.

Наблюдава се, че предложеният модел TREE демонстрира висока точност при оценката на дебелината, показвайки минимални грешки по различните показатели. Докато стандартното отклонение (SD = 1.541) е сравнимо с SegNet [31] (SD = 1.485), моделът превъзхожда SegNet [31] по всички останали показатели за грешка. Той значително превъзхожда PSPNet [32] и RESUNET [36], особено при сравнение на стойностите на MSE, Bias и MAE. За да се определи измеримата дебелина от стойностите в пиксели, дебелината се изчислява в милиметри (mm), използвайки следната формула:

$$\boldsymbol{T} = \frac{\boldsymbol{\mu}}{n_{st}} \times N_{pt} \times r_f \tag{3.22}$$

Тук в уравнение (3.22), μ представлява средната стойност на пикселното разстояние (9.61 μm) а показателят на пречупване на венозните тъкани е n_{st}= 1.38. Броят на пикселите, участващи в оценката на дебелината, е N_{pt}, а факторът за преоразмеряване е r_f=2, тъй като изображението е преоразмерено (по вертикалната ос) преди да бъде използвано като вход за модела. Дебелината за всички модели, включително TREE, се изчислява чрез уравнение (3.22) и се сравнява с дебелината по еталон (Mask). Усреднените стойности на дебелината за тестовия набор от данни са представени в Таблица 3.4. Тази еталонна дебелина е извлечена аксиално от маски, анотирани от експерт, независимо от какъвто и да е сегментационен модел, което гарантира висока точност. Резултатите показват, че TREE предоставя

Таблица 3.4 Дебелина (mm) като резултат от сегментацията от моделите DeepLab3+ [33], U-Net [29], PSPNet [32], RESUNET [36], SegNet [31]; ground truth (mask) и дебелина, получени от модела TREE (P.II)

Метод	**DeepLab3+**	**U-Net**	**PSPNet**	**RESUNET**	**SegNet**	**MASK**	**TREE**
Дебелина (mm)	0.7486	0.7076	0.7744	0.6887	0.7502	**0.7168**	**0.7160**

изключително прецизни оценки на дебелината. Освен стойностите на дебелината,

представяме и броя на параметрите, използвани от всеки модел. Лесно може да се заключи, че TREE съдържа най-малък брой параметри, благодарение на намаляването на броя блокове от четири на три (както в енкодера, така и в декодера), в сравнение с оригиналния U-Net [29] и други подобни архитектури като RESUNET [36]. Следователно, благодарение на приноса на модула SPD, възможностите на модула се увеличават, дори при по-малък брой слоеве.

За да се оцени **генерализируемостта** на предложения модел, неговата производителност бе

Таблица 3.5 Сравнение на точността за независимия test set (P.II)

Метод	DeepLab3+	U-Net	PSPNet	RESUNET	SegNet	TREE
Accuracy	0.8325	0.8796	0.8471	0.8668	0.8295	**0.9242**

допълнително тествана върху 200 B-скана, получени от петия обем от данни. Този обем съдържа нов венозен образец, който не е бил включен в обучителната или валидационната фаза. Резултатите от сегментацията на **TREE** бяха сравнени с други модели върху този независим тестов набор от данни. Както е обобщено в **Таблица 3.5**, моделът демонстрира превъзходна производителност спрямо останалите методи, което потвърждава способността му за генерализация.
За да се оцени въздействието на **transfer learning**, чрез внедряване на архитектура с два декодера и двуфазно обучение на модела, бе проведено **ablation study**. Моделите бяха обучени върху обема от венозни данни, следвайки същата експериментална процедура, описана по-горе. Бяха изследвани пет различни модела:

1. **En-De1 Model** – Моделът за сегментация (En-De1) се обучава първо, последван от допълнителни три плътни слоя с активация (както в De2), за да се определи дебелината на вената.
2. **En-De1-De2 Model** – В този подход и двата декодера (De1 и De2) се обучават едновременно заедно с енкодера, но без използване на transfer learning.
3. **En-De2 Model** – Този модел предсказва само дебелината на вената, без сегментация, като се обучават единствено енкодерът и De2.
4. **En-De2-De1 Model** – Моделът En-De2 от предишната конфигурация се използва, като теглата на енкодера са фиксирани. След това се обучава декодерът за сегментация (De1), позволявайки първо на модела да научи оценка на дебелината на вената преди сегментацията.
5. **NO-SPD Model** – Вариация на TREE, при която модулът **SPD** е заменен с конвенционални bottleneck bridge слоеве, подобни на архитектурата на U-Net [29]. Atrous и depthwise свивките са заменени със стандартни свивки, като същевременно се запазват размерът на ядрото, стъпката и структурата на слоевете, както при TREE.

За тези модели, производителността се сравнява чрез метриките точност (accuracy) и MSE за сегментация и предсказание на дебелината, съответно, върху тестовия набор от данни. Изчисляването на средната дебелина се извършва чрез намиране на средната стойност на

Таблица 3.6 Ablation study за сегментация и дебелина. (P.II)

Метод	En-De1	En-De1-De2	En-De2	En-De2-De1	NO-SPD	TREE
Точност (Accuracy)	0.9910	0.9282	-	0.6636	0.9514	**0.9937**
MSE (column)	0.1140	0.3554	58	58	0.2651	**0.0537**

редовия вектор, съдържащ дебелината на всяка колона за всеки B-скан. Тази средна дебелина се използва допълнително за изчисляване на MSE между еталона и предсказаните стойности, както е показано в Таблица 3.6. Изчислявайки средната грешка по колоните, се получава

представа за точността на оценките на дебелината. Това позволява по-дълбоко разбиране на влиянието на transfer learning, двата декодера и последователното обучение, предложени в настоящата работа.
Първият модел En-De1 е базиран на същата сегментационна рамка като TREE, водейки до сходни резултати. Основната разлика обаче се проявява в стойността на MSE за дебелината, тъй като не е приложен transfer learning. При модела En-De1-De2 едновременното обучение ограничава способността на модела да предсказва както сегментация, така и дебелина. Междувременно моделът En-De2 показва значителни грешки при оценката на дебелината и не успява да постигне целта за получаване едновременно на сегментационна карта и дебелина от една мрежа.
Моделът, обучен последователно (En-De2-De1), където първо се предсказва дебелината, а след това се извършва сегментация, също демонстрира по-слаба производителност. Започването на обучението с оценка на дебелината не се превръща ефективно в точна пикселно-базирана сегментация, тъй като този подход не успява да улови глобалния пространствен контекст, необходим за сегментацията, което прави transfer learning неефективен.
Накрая, тестването на модела чрез заместване на модула SPD, както бе споменато по-горе, показва, че вариантът NO-SPD също демонстрира по-слаба производителност в сравнение с TREE. Превъзходната производителност на TREE, както е показано по-горе, се дължи основно на стратегията за transfer learning, докато включването на модула SPD допълнително подобрява точността.

Глава 4 Рамка за дълбоко обучени за реконструкцип на изображения от SS-OCT

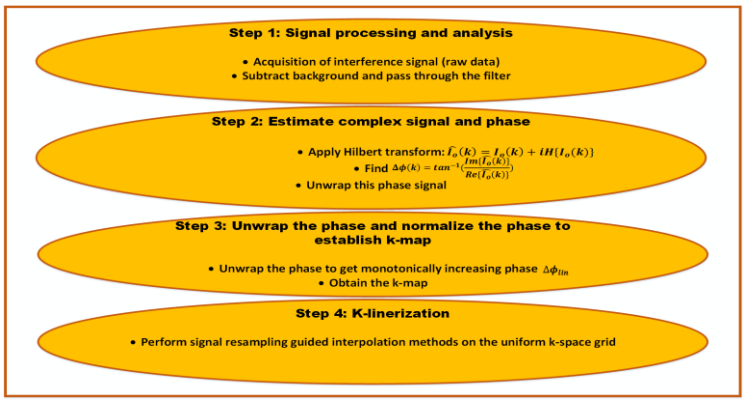
Изследванията, представени в Четвърта глава, се основават на публикации P III и P IV на автора. Досега в дисертацията фокусът бе поставен върху биомедицинското придобиване на данни за разширени вени и съответните методи на базата на дълбоко обучение за автоматизация на сегментацията и измерването на дебелината. Междувременно бе установено, че OCT системите страдат от понижено качество на изображението поради спекъл шум и усложнения, произтичащи както от скъп хардуер, така и от софтуерни методи за обработка. В тази глава се представя общ преглед на процеса на реконструкция на изображения в SS-OCT, базиран на данни, линейни по дължината на вълната. Основният акцент е поставен върху разработването на високооптимизирани и ефективни DL модели, способни да реконструират висококачествени изображения в SS-OCT системи при същевременно намалена време на обработка.

4.1 Реконструкция на данни, линейни по дължината на вълната

SS -OCT системите записват данни в спектралната област чрез лазер, който бързо променя дължината на излъчваната светлина във времето, т.е. извършва "преминаване" (*sweeping*) през определени дължини на вълните. Тази техника кодира информацията за дълбочинно-зависимата отражателна способност на изследвания образец като спектър във Фурие-областта. Полученият профил обикновено се нарича *depth profile*, тъй като отразява структурните детайли на различни дълбочини.

За да бъде извлечена тази информация от Фурие областта в A-сканове (пространствен домейн), интерференчният сигнал, кодиран по дължина на вълната, се подлага на обратна Фурие трансформация (IDFT). Типично ограничение при тези системи е необходимостта интерференчният сигнал да бъде линеен в k-областта (wavenumber domain) преди прилагането на Фурие трансформацията. Поради това, преди трансформацията сигналът в линеен по дължината на вълната се преобразува в сигнал в k-пространството чрез процес, известен като k-калибриране (*k-mapping*).

В Фиг. 4.1 е показана калбрационната процедура [22], която има за цел да изясни концепциите за калибрация и k-линеаризация, критично важни за реконструкцията на изображението и постигане на оптимална аксиална резолюция. Методът започва от записа на сурови данни, съдържащи интерференчен сигнал, който е линеен по отношение на дължината на вълната. Следва премахване на фиксирания шум чрез фоново изваждане и филтриране, осигуряващо коректно фазово извличане при Хилбертовата трансформация. За получаване



Фигура 4.1 Диаграма процедурата за линеаризиране по k

на комплексен аналитичен сигнал се прилага Хилбертова трансформация върху реалния сигнал (Стъпка 2, Фиг. 4.1). Определя се и се разгъва (*unwrap*) фазата, за да се получи монотонно нарастваща фаза. Информацията от фазата се използва за извличане на k-карта, отразяваща линейността в k-пространството. Линеаризацията по k се извършва чрез интерполация за ресемплиране на сигнала върху равномерна k-мрежа. Последната стъпка е прилагане на **IDFT** върху линеаризирания в k-пространството сигнал за получаване на **A-скан**.

За справяне с трудностите, свързани с калибрацията и k-линеаризацията, често се използват методи за ресемплиране чрез различни интерполации. Последните години са предложени и хардуерни решения като алтернативи на конвенционалните софтуерни техники – напр. оптични k-часовници, двуканално придобиване, акинетични лазери и др. Въпреки това тези методи страдат от недостатъци: високи разходи, сложност и неефективна обработка, което често води до неточности.

Когато изследваните образци проявяват динамично поведение, вътрешното движение може да предизвика артефакти при опит за намаляване на спекъл шума чрез осредняване. Тази времева флуктуация налага внимателен подбор на метода за обработка и съобразяване с времевите ограничения, тъй като и двете са решаващи за висока точност и надеждност на ОСТ-базирана система. Предложените в тази дисертация DL-базирани техники предоставят обещаващи решения за:

- преодоляване на проблема със спекъл шума,
- избягване на осредняването като метод за намаляване на спекъл шума (невъзможно при динамични образци),
- елиминиране на необходимостта от скъп хардуер за линеаризация,
- намаляване на времевата сложност при обработка на данни, линейни по дължината на вълната.

4.2 Формулиране на проблема

За моделиране на профила на отражателната способност в SS-OCT система, базирана на интерферометър на Michelson, интензитетът I_o може да бъде изразен в практически сценарий по следния начин:

$$I_o = I_{DC} + I_{CC} + I_{AC} + N_G(t). \tag{4.1}$$

В уравнение (4.1), $N_G(t)$ представлява адитивен бял гаусов шум, а I_{DC}, I_{CC}, и I_{AC} обозначават съответно DC, крос-корелационен и автокорелационен компоненти. Отражателната способност по дълбочина е кодирана в крос-корелационния член, който може да бъде развит по следния начин:

$$I_{CC} = S\big(k(t)\big)\left[\frac{\rho}{2}\sum_{n=1}^{N}\sqrt{R_{rm}R_{Sn}}\left(\cos\big(k(t)(z_R - z_{Sn})\big)\right)\right]. \quad (4.2)$$

В уравнение (4.2), $S(k(t))$ е спектралната форма на източника на светлина, R_{rm} и R_{sn} представят отражението от референтното огледало (на дълбочина z_R) и съответно от nth частица на образеца (на дълбочина z_{Sn}), N е броят на частиците, и ρ е отклика на детектора. Профилът на отражателната способност на образеца може да бъде възстановен чрез прилагане на обратна дискретна Фурие трансформация (IDFT) - $\boldsymbol{\mathcal{F}^{-1}}$ в k-пространството, както е дадено уравнение (4.3):

$$i_z = \boldsymbol{\mathcal{F}^{-1}}\{I_{CC}\}, \quad (4.3)$$

$$i_z = \frac{1}{2}\sum_{i\neq 1}^{\infty}\sqrt{R_{rm}R_{Sn}}(\alpha(z_R - z_{Sn}) + \alpha(-(z_R - z_{Sn}))). \quad (4.4)$$

Тук, в уравнение (4.4), $\alpha(z)$ е DFT двойката на $S(k)$. Тъй като IDFT предполага еквидистантни точки на дискретизация, необходимо е спектралният интерференчен сигнал да бъде записан с равномерно разпределение в k-пространството, за да се гарантира точна реконструкция. За разлика от това, записаните спектрални данни обикновено са дискретизирани равномерно по дължината на вълната във времеви интервал −Δt/2 до Δt/2, показвайки линейна зависимост от времето, както е дадено в:

$$\boldsymbol{\lambda} = \beta t + \boldsymbol{\lambda}_0. \quad (4.5)$$

В уравнение (4.5), β е скоростта на сканиране на източника, $\boldsymbol{\lambda}$ е дължината на вълната в интервала от $\boldsymbol{\lambda}_{\min}$ до $\boldsymbol{\lambda}_{max}$, а $\boldsymbol{\lambda}_0$ представлява централната дължина на вълната. Ако използваме $\boldsymbol{\lambda} = 2\pi/\boldsymbol{k}$, получаваме следния израз:

$$\mathrm{t} = \frac{1}{\beta}(\boldsymbol{\lambda} - \boldsymbol{\lambda}_0) = \frac{2\pi}{\beta}\left(\frac{1}{\boldsymbol{k}} - \frac{1}{k_0}\right). \quad (4.6)$$

Чрез разлагане в ред на Тейлър на уравнение (4.6), изразът може да бъде преформулиран като:

$$t = \frac{2\pi}{\beta}\left[-\frac{1}{k_0}\left(\frac{\boldsymbol{k}}{k_0} - 1\right) + \frac{1}{k_0}\left(\frac{\boldsymbol{k}}{k_0} - 1\right)^2 - \frac{1}{k_0}\left(\frac{\boldsymbol{k}}{k_0} - 1\right)^3 + \cdots\right]. \quad (4.7)$$

В допълнение, за източници на светлина, като например FDML лазери [20], зависимостта между дължината на вълната и времето показва синусоидален характер, описан математически като:

$$\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}_0 + \frac{\Delta\boldsymbol{\lambda}}{2}(\sin(2\pi f_t t)), \quad (4.8)$$

където $\Delta\boldsymbol{\lambda}$ е спектралната ширина на източника, а f_t е честотата на настройка на Fabry-Perot (FP) филтъра в източника на светлина.

От горните формулировки може да се заключи, че фундаменталната характеристика на ОСТ системите се състои в дълбочинно-кодираните спектрални данни. Както показват уравнение (4.7) и уравнение (4.8), спектралният сигнал демонстрира нелинейна зависимост от вълновото число. Наличието на значими членове от по-висок ред в разлагане в ред на Тейлър налага прилагането на калибрация и k-линеаризация, за да се осигури точна генерация на B-сканове. В допълнение, както бе обсъдено в главата „Предистория“ на дисертацията, ОСТ системите използват кохерентен източник на светлина, което води до явлението спекъл. Спекъл шумът влошава контраста на изображението и прикрива фините морфологични особености. Следователно, прилагането на методи за потискане на шума е абсолютно необходимо.

Публикации III и IV са насочени към високо достоверни реконструкции на изображения чрез ефективно потискане на шума, без компромис по отношение на пространствената разделителна способност или структурната цялост, като се използват сурови данни. В Публикация III е представен метод за директна реконструкция на висококачествени изображения от сурови ОСТ данни линейни по дължина на вълната, като се елиминира необходимостта от k-линеаризация посредством включване на CNN-базиран модел, разгледан в раздел 4.3.
За да се повиши потенциалът на тези модели, в Публикация IV е предложена мрежа, която изследва данните едновременно в пространствената и във честотната област. Този подход е мотивиран от факта, че FD-OCT системите се основават на фундаменталната теорема на Вайнер–Хинчин, която формулира връзката между пространствената и Фурие-областта. Съгласно тази теорема, спектралната плътност на мощността на детектирания интерференционен сигнал е математически свързана с неговата автокорелационна функция чрез преобразуване на Фурие. Именно тази зависимост във Фурие-пространството се използва от FD-OCT системите, които регистрират спектрално разделени интерференционни сигнали. Последните се обработват чрез обратно дискретно преобразуване на Фурие (IDFT) за реконструкция на дълбочинно-резолвирани профили на отражателната способност в пространствената област.

Вдъхновени от този основополагащ принцип, ние разработихме двумрежова рамка, която интегрира обработката в пространствената и Фурие-областта с цел подобрена реконструкция на ОСТ изображенията. Подраздел 4.4 от тази глава представя метода, при който първо се обработват суровите данни в пространствена област, след което следва тяхната обработка в честотната област чрез CNN-базирана мрежа.

4.3 Реконструкция на изображения от суровите SS-OCT данни

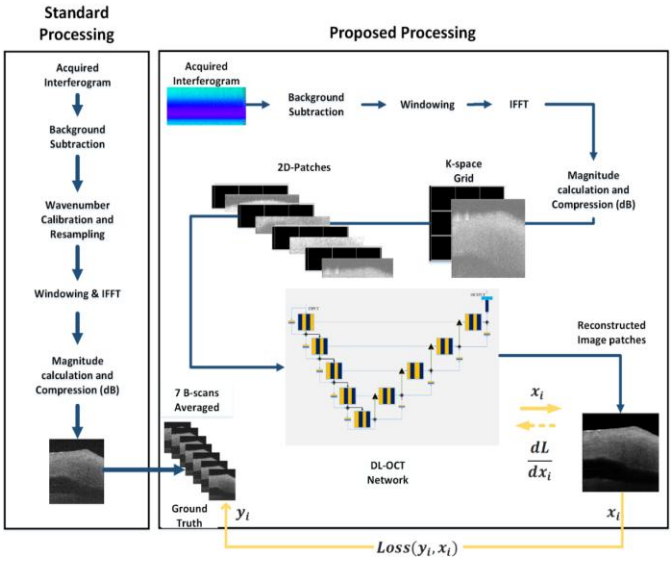
В този раздел обобщаваме работата, представена в Публикация III, в която се предлага DL модел за генериране на ОСТ B-сканове чрез използване на интерферометричен сигнал, линеен в дължина на вълната. Както беше обсъдено по-горе, ако интерферометричния сигнал, който е нелинеен в k-пространството, бъде трансформиран чрез IDFT в пространствения домейн, получените изображения губят своята рязкост и структурни детайли. Предложеният модел е способен да обработва директно сигнала, линеен в дължината на вълната (λ-пространство), като по този начин елиминира необходимостта от процес на k-линеаризация. Моделът използва DL-базирана архитектура, представляваща модифициран U-Net [29], в който са внедрени механизми за внимание [34] и остатъчни връзки [30]. Освен това, рамката е насочена към проблемите, свързани с влошаването на качеството на изображенията поради наличието на спекъл шум.

4.3.1 Методология на рамката за реконструкция

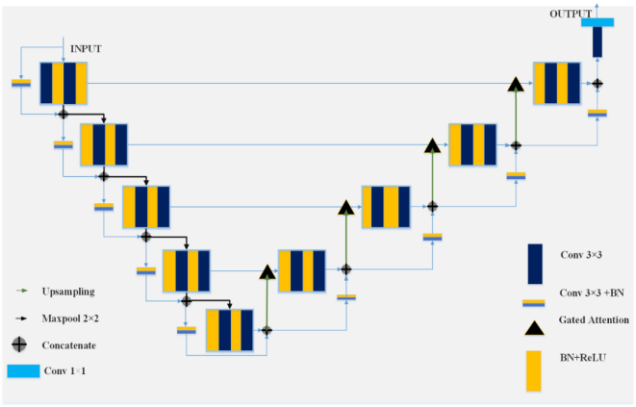
Предложената рамка, наречена WAVE-UNET, е представена на Фиг. 4.2 заедно със стандартната процедура за обработка, включваща k-линеаризация. Тази рамка първоначално обработва суровия интерферометричен сигнал, записан в нелинейното *k*-пространство, който съответства на единично B-сканиране. Целта е да се реконструират профили по дълбочина и съответните изображения, които обаче са замъглени поради нелинейността на сигнала. Предварителната обработка на суровия сигнал в λ-пространството се извършва на няколко етапа:

1. извличане на спектъра на фона без поставяне на образец и последващо изваждане на този спектър за елиминиране на спектралните изкривявания, породени от източника;
2. прилагане на прозореца на Hann с цел потискане на спектралното изтичане (*spectral leakage*);
3. едномерно обратно дискретно преобразуване на Фурие (IDFT) по вертикалната ос на всяко B-сканиране;

4. изчисляване на амплитудата на получения комплексен сигнал, трансформиране в логаритмична скала (dB) и елиминиране на излишната част на спектъра поради конюгатна симетрия.



Фигура 4.2 Обработка на ОСТ сигнала стъпка по стъпка (Standard Processing), предложена WAVE-UNET рамка за реконструиране на ОСТ изображения (Proposed Processing). (P.III)



Фигура 4.3 DL-базиран модел, използван в WAVE-UNET с модифициран U-Net with резидуални връзки и блокове с управлявано (gated) внимание. (P.III)

Крайният резултат от този процес е дълбочинно-кодирано ОСТ изображение, получено директно от линейни измервания в λ-пространството. Това изображение се означава тук като λ-пространствено изображение. Изображенията, реконструирани от данни, неравномерно дискретизирани в *k*-пространството, по своята същност са с ниско качество поради артефакти от замъгляване, както бе обсъдено по-горе. Тези нискокачествени изображения, илюстрирани на Фиг. 4.2, впоследствие се подават като вход към предложената рамка WAVE-UNET.

WAVE-UNET е базирана на DL-модел, вдъхновен от архитектурата U-Net [29], усъвършенстван с attention механизми [34] и резидуални връзки [30], за подобряване на представянето на признаците. Attention-модулите са поставени на всяка връзка за заобикаляне на слой (skip-connection), която обединява енкодерните и декодерните блокове на съответните нива, както е детайлно представено на Фиг. 4.3.

За да се интегрира предварителното знание за вълновите числа (*wavenumbers*) в обучителния модел, е въведен специален слой, наречен слой за споделяне на вълнови числа (Wavenumber Sharing, WS). Този слой кодира предварително дефинирана мрежа от вълнови числа в интервала от 4.62×10^6 rad/m до 4.99×10^6 rad/m, разпределени върху 1152 пространствени позиции, съответстващи на точките в A-line сканирането. Поставен като вторичен канал, WS-слоят се наслагва върху входното λ-пространствено изображение

(нискорезолюционно B-сканиране), формирайки комбинирана двуканална входна матрица, където първият канал съдържа интензитетните стойности от λ-пространственото изображение, а вторият канал – кодираната решетка от вълнови числа.
Дължината на вълната обхваща интервала $(\boldsymbol{\lambda}_{min}, \boldsymbol{\lambda}_{max})$, който се разделя на P^p точки, като s е дискретизационната точка. Всяка точка в λ-пространството може да бъде представена във формата на уравнение (4.9) и уравнение (4.10) по-долу:

$$\lambda = \boldsymbol{\lambda}_{min} + \frac{s}{P^p - 1}(\boldsymbol{\lambda}_{max} - \boldsymbol{\lambda}_{min}) \tag{4.9}$$

$$k = \frac{2\pi}{\lambda} = \frac{2\pi}{\boldsymbol{\lambda}_{min} + s(\boldsymbol{\lambda}_{max} - \boldsymbol{\lambda}_{min})/(P^p - 1)} \tag{4.10}$$

Тази интеграция е визуално представена като мрежа в k-пространството на Фиг.4.2. Те са разделени на четири двумерни блока, като всеки блок съдържа двата слоя, т.е. изображението в λ-пространство и информацията за вълновото число. Тези блокове служат като вход към предложената архитектура енкодер-декодер, като всяка страна на мрежата (енкодер/декодер) се състои от четири резидуални блока. По време на обучението моделът обработва батчове от тези двумерни блокове, като генерира съответните изходни блокове x_i, докато входните дании преминава през модела. Този двуслоен подход с patch-базиран вход позволява на мрежата да научи по-информирано представяне, използвайки както пространствената интензивност, така и свързания спектрален диапазон. Разликата между генерирания изход x_i и очаквания изход, т.е. еталона y_i, се минимизира с цел да се ръководи процесът на обучение на предложената мрежа.

За оптимизация се използва оптимизаторът Adam, който итеративно актуализира параметрите на модела. Както е илюстрирано на Фиг.4.2, плътните жълти стрелки представят директното (forward) разпространение на информацията, докато прекъснатите стрелки показват обратно (backward) разпространение на градиентите. Всички резидуални единици се състоят от слоеве за бинарна нормализация (BN), конволюции (convolutional operations) и ReLU функции за активация, интегрирани с резидуални връзки.

Енкодер модулът постепенно намалява пространствените размери чрез операции по максимално обединяване, докато декодерът реконструира обратно същото пространствено разрешение чрез слоеве за увеличаване на резолюцията (upsampling). Механизмът за внимание (attention) между блоковете на енкодера и декодера по връзките за заобикаляне на слой подпомага повишаването на точността при реконструкцията на изображението, като потиска нерелевантни характеристики [34].

4.3.2 Експерименти и анализ на WAVE-UNET

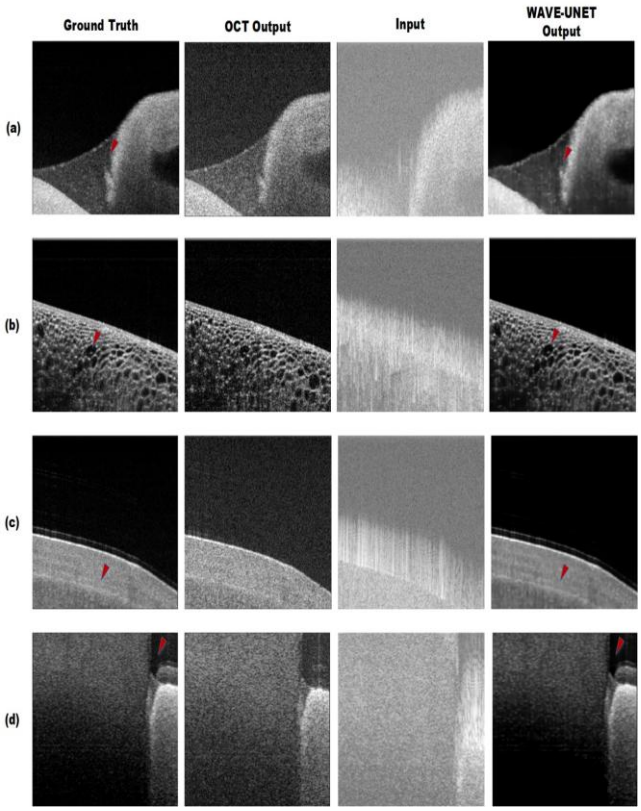
В този раздел представяме експериментите, проведени за анализ на производителността на рамката WAVE-UNET, базирана на набор от данни, получени от MHz SS-OCT система с FDML източник на Optores GmbH, описана в Глава 2. Общият набор от данни се състои от 3000 изображения. За всеки запис са използвани 5 различни образци, като всеки от тези обеми съдържа по 600 изображения и съответните сурови данни, линейни в λ-пространство. Използваните образци включват лимон, череша, човешки пръст, човешки зъб и разширена вена. Суровият входни данни са в λ-пространство, представени като матрица с размери 2304 × 1024. Тази матрица преминава през предварителна обработка на изображението, както е описано в предходния раздел. Процесът включва преобразуване във Фурие-пространство, като поради наличието на конюгатна симетрия, причиняваща излишък от данни, използвахме 1152 × 1024 пиксела за извличане на изображението в λ-пространство. По-нататък, поради ограничения на паметта, регионът на интерес за всяко изображение беше ограничен до размер 1152 × 512.
След това се формират patch-ове с размери 288 × 512 × 2, чрез обединяване на матрицата с вълнови числа (wavenumber) с изображението в λ-пространство и разделяне на B-scan-а (с размер 1152 × 512) на четири части по вертикала. Всеки patch се стандартизира чрез изваждане на средната стойност и последващо деление на стандартното отклонение. Наборът

от данни, съдържащ общо 3000 изображения, бе разделен на тренировъчен, валидиращ и тестов поднабори с цел ефективно обучение и настройка на хиперпараметри. По-конкретно, тренировъчният набор съдържа 70%, валидиращият – 20%, а тестовият – оставащите 10% от данните.

За еталонните изображения за обучението на мрежата, бяха използвани B-scan изображения, произведени от OCT системата на Optores след усредняване на седем последователни скана. Това усредняване беше извършено с цел намаляване на спекъл шума, като се взимаше предвид съхраняването на анатомичната цялост и предотвратяване на деградация на фини детайли, които обикновено се наблюдават при прекомерно усредняване. Тези изображения преминаха през вградените процедури на OCT системата (Optores GmbH), включително k-линеаризация и калибрация.

По време на фазата на обучение моделът премина през 150 епохи. Размерът на батча беше 12 patch-а, всеки с размери 288 × 512 × 2. Параметрите на модела бяха оптимизирани чрез оптимизатора ADAM със скорост на обучение (learning rate) 0.0001. Цялата рамка бе изпълнена на изчислителна система, оборудвана с 64 GB RAM и графичен процесор NVIDIA RTX 3090.

За количествено оценяване на разликата между предсказаните patch-ове изображения, генерирани от модела WAVE-UNET, и съответните еталонни patch-ове се използва функцията за загуба Средно статистическа грешка (MSE). Фиг. 4.4 представя изходните patch-ове изображения с размери 288 × 512 пиксела, съответстващи на (a) вена, (b) череша, (c) пръст на ръка и (d) зъб. Колоните са организирани както следва: колона 1 – еталонни изображения, колона 2 – OCT изходно изображение, колона 3 – входно изображение, колона



Фигура 4.4 OCT B-скан patches от (a) вена, (b) череша, (c) ръка и (d) зъб, Колони 1 до 4 показват съответно: (1) Ground Truth (left), (2) OCT изходно изображение, (3) входно изображение и (4) WAVE-UNET реконструирано изображение.

4 –WAVE-UNET реконструирано изображение. От резултатите, показани на Фиг. 4.4, е видно, че моделът WAVE-UNET демонстрира висококачествени реконструкции в сравнение с оригиналните замъглени λ-space изображения (входа) и съответния OCT изход, произведен от системата на Optores. Областите, маркирани с червени стрелки, допълнително подчертават способността на WAVE-UNET да възстановява фини структурни детайли. Тези наблюдения съвкупно потвърждават възможността на предложената DL-базирана рамка да реконструира

висококачествени OCT изображения с редуциран спекъл шум, като същевременно опростява процеса на реконструкция в сравнение с традиционния OCT pipeline, който разчита на обширно усредняване.

Таблица 4.1 Comparison of images generated by SS-OCT system and WAVE-UNET (P.III)

Metric		VEIN	LEMON	CHERRY	TOOTH	HAND
PSNR	I/P	17.06	15.41	15.56	14.57	14.24
	OCT	21.94	19.18	14.11	17.56	19.65
	WAVEUNET	25.50	27.34	19.21	19.75	22.81
SSIM	I/P	0.05	0.008	0.05	0.09	0.05
	OCT	0.21	0.04	0.03	0.08	0.03
	WAVEUNET	0.59	0.51	0.41	0.29	0.46
MSE	I/P	1.31	1.30	1.33	0.72	1.12
	OCT	0.42	0.54	1.86	0.36	0.32
	WAVEUNET	0.18	0.08	0.57	0.21	0.15

За количествена оценка на представянето, в настоящата работа се използват следните метрики: Structural Similarity Index Metric (SSIM), Peak Signal to Noise Ratio (PSNR) в dB и MSE, за по-добър анализ на резултатите върху различни образци. В този контекст, еталонното изображение служи като референтен стандарт за изчисляване на тези метрики за оценка. Той също така позволява количествено да се разбере разликата между входното изображение (raw λ-space image), изхода на OCT системата (OCT) и изхода на предложената рамка (WAVE-UNET), както е представено в Таблица 4.1 за 10% от patch-овете изображения, избрани случайно от всеки тестов набор, а именно: вена, лимон, череша, зъб и ръка.
Предложеният подход превъзхожда SS-OCT система на Optores по отношение на качеството на изображението за всички измерени параметри. Наблюдава се значително подобрение от около 4 dB в PSNR за реконструирания изход в сравнение с изхода от OCT системата за наборите лимон, вена и череша. По същия начин, SSIM показва подобрение за всички образци. Важно е също да се отбележи, че за образците от череша, зъб и ръка измерените SSIM стойности на суровите λ-изображенията надвишават тези на обработените с OCT изображения. Това се дължи на факта, че OCT изображенията са значително деградирани поради спекъл шума. Следователно, стойностите на SSIM сами по себе си може да не служат като напълно надежден индикатор за качество на изображението, особено при изображения, засегнати от спекъл шум. Подобна тенденция се наблюдава и при MSE стойностите, като наборът лимон показва най-малка грешка, а наборът череша – най-висока.
Тези сравнителни наблюдения обосновават, че предложеният метод демонстрира значително по-добро представяне в сравнение със стандартната обработка на OCT изображения чрез системата на Optores. Освен това, предложената рамка е подходяща и за приложения, изискващи бързи реконструкции, като постига време за извод от 90 секунди за обем, съдържащ 600 B-scan изображения, което е значително по-бързо в сравнение с традиционните 792 секунди, необходими за OCT системата на Optores (използвайки вградената обработваща верига), тествана на същата GPU-ускорена система, описана по-горе в този раздел.

4.4 Реконструиране на OCT изображения чрез оптимизиране ба данните във Фурие и пространствените области

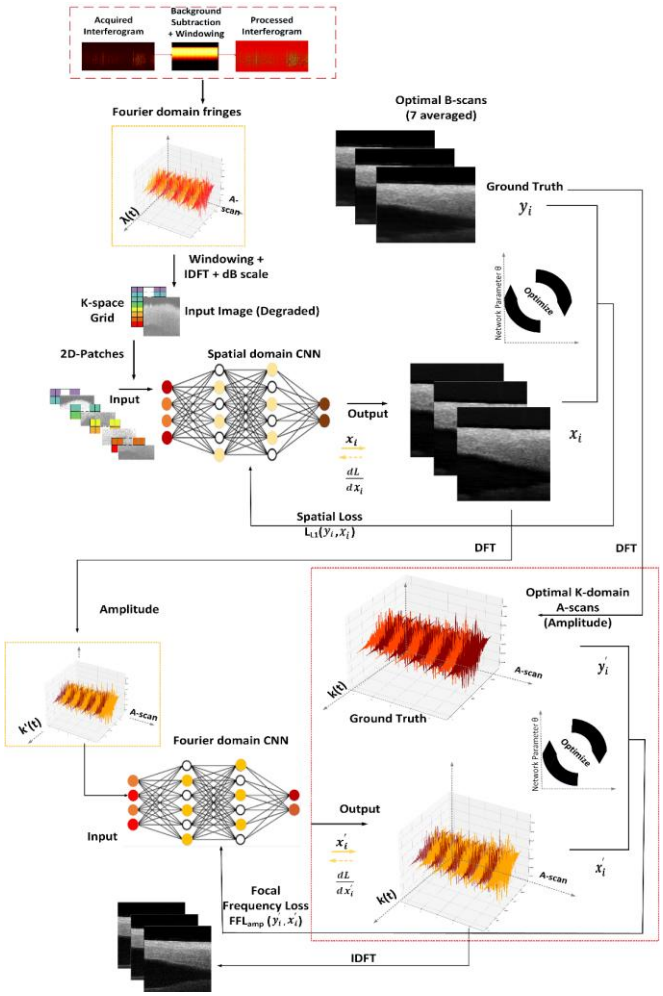
Този раздел представя работата, описана в Публикация IV, която се основава на метод за реконструкция на OCT изображения, отчитащ връзката между Фурие и пространствените области, формулирана чрез теоремата на Винер–Кинчин, разгледана подробно в раздел 4.2. Тази теоретична основа е интегрирана като физически основа с цел подобряване на процеса на обучение в предложената рамка, базирана на CNN.

Предложената рамка е реализирана чрез включването на две невронни мрежи, които извършват обработка на различни нива и се използват последователно. Първата мрежа, обозначена като SD-CNN (Spatial Domain Convolution Neural Network), както подсказва самото наименование, оптимизира в пространствената област. Втората мрежа, наречена FD-CNN (Fourier Domain Convolution Neural Network), извършва оптимизация във Фурие областта.
Мрежата SD-CNN стартира, като използва силно деградирани изображения, получени чрез прилагане на Фурие трансформация върху интерферометричния сигнал в λ-домейна, като входни данни. Нейната основна цел е възстановяване на сложните морфологични структури и потискане на нежелания шум. От своя страна FD-CNN е проектирана да функционира като мрежа за по-високостепенно прецизиране във Фурие областта, насочена към реконструкция на изображения с високо качество. Тя поставя особен акцент върху възстановяването на финитe структурни детайли, които не са били адекватно уловени от мрежата SD-CNN.
Тази стратегия, базирана на последователна оптимизация чрез двете мрежи, дава възможност на цялостната рамка да генерира B-scan изображения с отчетливо подобрено визуално качество.

4.4.1 Методология за реконструкция на ОСТ изображения чрез рамка, базирана на SD-CNN и FD-CNN

Предложената в Публикация IV рамка, включваща SD-CNN и FD-CNN за реконструкция на ОСТ изображения, е илюстрирана на Фиг.4.5. Тя използва суровите данни под формата на λ-пространствени изображения, получени след предварителна обработка. Тази обработка се извършва в същата последователност, както е описано в раздел 4.3.1, включваща: изваждане на фоновия спектър от записаните интерферограми, прилагане на Hann-прозорец, последвано от Фурие трансформация чрез IDFT и мащабиране на амплитудата. Дълбоката невронна мрежа, използвана в тази рамка, е показана на Фиг.4.6 (a) и описана в детайли по-долу. Моделът, приложен както в SD-CNN, така и във FD-CNN, е почти идентичен по архитектура, като се различава само по незначителни аспекти. Той е изграден върху базовата архитектура U-Net, включваща „тясно място“ (bottleneck), четири енкодерни и четири декодерни блока. За повишаване на ефективността са интегрирани механизъм за управлявано внимание (attention gating) и резидуални (residual) връзки. Както бе обсъдено по-рано, механизмът за внимание работи чрез потискане на нерелевантните характеристики, осигурявайки по-добра производителност. Енкодерният блок включва слоеве в следната подредба: Batch Normalization (BN), активационна функция Parametric ReLU (PReLU) и конволюционни (convolutional) слоеве. От своя страна декодерните блокове използват слоеве за увеличаване на резолюцията за увеличаване на разделителната способност, съчетани с конволюционни транспониращи слоеве (Conv Transpose), активация чрез PReLU и нормиране чрез BN. Междинните връзки за заобикаляне на слой (skip connections) между енкодера и декодера улесняват преноса на информация на съответните нива (резолюция). Въпреки това, те могат да пренасят и излишна информация на ниско ниво, която се филтрира чрез механизма за внимание, намалявайки активациите в областите с дублиращи се характеристики.

За обучение на мрежата са използвани еталонни изображения (y_i), генерирани от системата Optores ОСТ. Те са получени чрез усредняване на 7 последователни изображения от записаните обемни данни. Това е възможно благодарение на високата степен на сходство между съседни изображения. На Фиг.4.6(b) е показано как изглеждат ОСТ изображенията след усредняване на 5, 7 и 9 съседни скана (съответно в колони 2, 3 и 4), за образци от вена, лимон и череша, подредени по редове. Единичните B-scan изображения от различните образци показват значително наличие на спекъл шум.



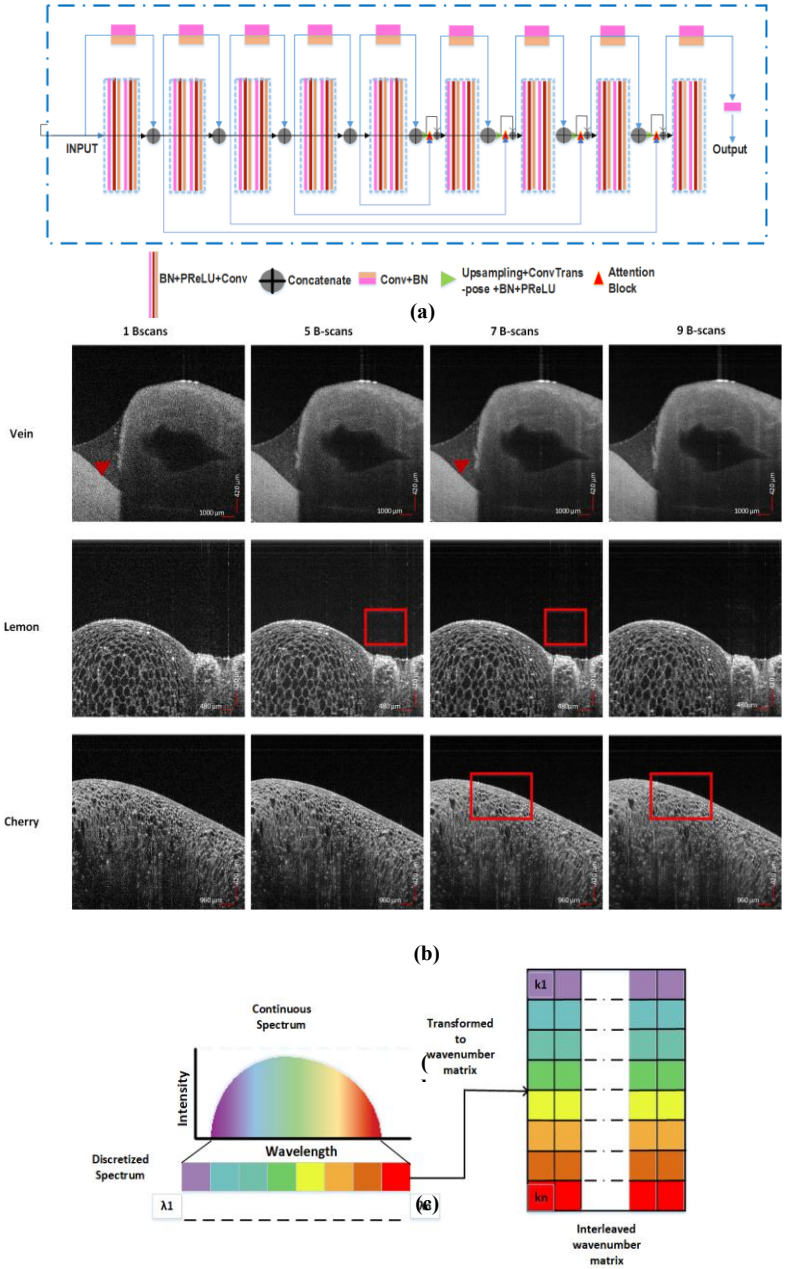
Фигура 4.5 Схема на предложената рамка, включваща SD-CNN FD-CNN (P.IV)

Визуалното сравнение разкрива: за венозния образец се наблюдава забележимо подобрение в структурните детайли (отбелязано с червена стрелка); при образеца от череша – постепенно увеличаване на свръхизглаждането с увеличаване броя на усреднените сканове; при образеца от лимон – съществено намаляване на фоновия шум, особено при усредняване на пет и седем скана (червена рамка). Като се отчита балансът между потискането на шума, риска от прекомерно изглаждане и запазването на латералната разделителна способност, усредняването на 7 B-скан изображения е възприето като еталонно (y_i) за настоящото изследване.

Като входни данни за SD-CNN се използва деградираното изображение, получено след обработка на суровите интерферометрични данни, описана по-горе. Информацията за обхвата на вълновите числа, съответстващи на записаните данни, е интегрирана в мрежата като помощен входен слой по време на обучението – обозначен като k-пространствена решетка (k-space grid) (Фиг.4.5 и Фиг.4.6(c)). Разделителната способност на тези входни λ-пространствени изображения е компрометирана, тъй като неравномерно разположените интерференчни ивици в λ-пространството се подлагат директно на IDFT. Това паралелно слоесто входно представяне позволява на SD-CNN да се обучи за реконструкция на OCT изображения. Точките, използвани за създаването на k-пространствената решетка, се изчисляват чрез уравнения (4.9) и (4.10), предоставяйки на мрежата информация за нелинейността в k-домейна. Тъй като идентични стойности на вълнови числа съществуват за всяко A-scan, колонният вектор се репликира във всички редове на входното изображение. Първоначално този вторичен слой е с размер (1152 × 256), но впоследствие се разделя на пачове с размер (288 × 256) за подобряване на сходимостта и за съобразяване с хардуерните

ограничения. Това разделяне е важно, тъй като тези пачове съответстват на аксиалните (дълбочинни) пикселни позиции във всяко A-scan чрез връзката на Фурие трансформацията.

Изходът на SD-CNN (x_i) се използва като вход за FD-CNN след подходяща обработка. Този вход се генерира чрез трансформиране на резултата на SD-CNN във Фурие домейна и прилагане на Дискретна Фурие трансформация (DFT) за извличане на амплитудния компонент. Еталонното изображение (y_i') (ground truth) за тази мрежа представлява амплитудата, получена чрез DFT върху еталонното изображения, описани по-горе. Всяко от 7-те изображения, използвани за създаването на референтния резултат, преминава през k-линеаризация преди IDFT и следователно съдържа приблизително равномерно разпределени



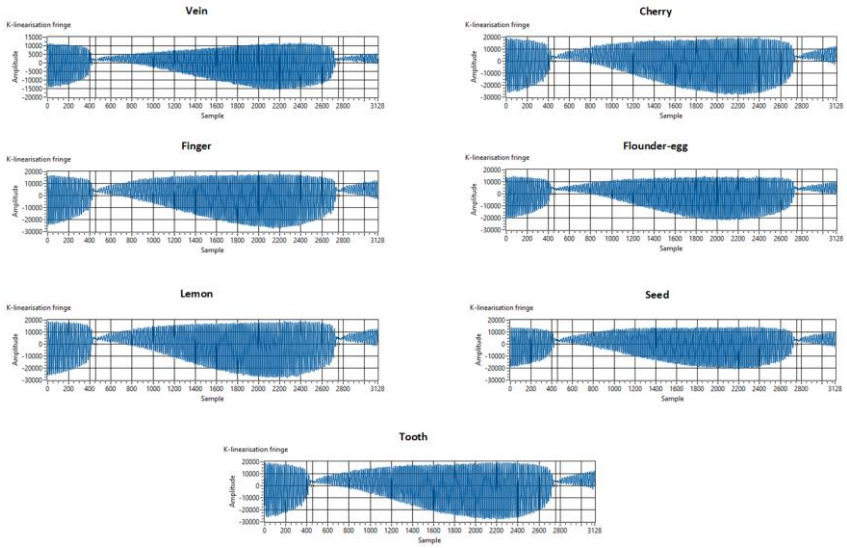
Фигура 4.6 (a) DL мрежа, използвана за FD-CNN и SD-CNN (b) 1 B-скан и изображения след усредняване на 5, 7, 9 B-скана на вена, лимон и череша *(c)* Слоя с вълнови числа за вход към SD-CNN (P.IV)

данни в k-домейна. С оглед на едномерната структура на зависимостите в A-scan, FD-CNN използва 1-D конволюционни ядра за извличане на характеристики. В процеса на обучение FD-CNN обработва амплитудните стойности във Фурие домейна и минимизира загубата, за да произведе равномерно разпределен спектър във вълново-пространствен домейн.

За крайното предсказание оптимизираният амплитуден изход x_i', генериран от FD-CNN, преминава през IDFT, докато съответната фазова информация се извлича от Фурие-преобразувания изход на SD-CNN.

4.4.2 Експерименти и анализ на рамката, базирана на SD-CNN и FD-CNN

Наборът от данни, използван в Публикация IV, е записан със системата Optores. За неговото създаване са използвани следните образци: вена, пръст, лимон, зъб, череша, хайвер от писия и семе (грах). Всеки B-скан от различните обеми се състои от 1024 A-скана, като всеки A-скан съдържа 2304 дълбочинни точки, съответстващи на записания суров спектър. В



Фигура 4.7 Analysis of k-linearization fringes across various sample volumes, including vein, finger, lemon, tooth, cherry, flounder egg, and pea seed. (P.IV)

края на обработващия конвейер както суровите данни, така и съответните B-сканове бяха изрязани до размери 2304 × 256, преди да бъде приложен IDFT, отчитайки ограниченията в паметта. След стъпката IDFT, описана в Раздел 4.4.1, поради огледалната симетрия се запазва само половината от сигнала, което води до краен размер 1152 × 256.

Суровите данни са директно достъпни чрез софтуерното приложение *MHz-OCT Processing* на системата Optores, което предоставя и индивидуални OCT сканове, както и 7-усреднени OCT изображения. Получените сурови данни и еталонните изображения бяха стандартизирани чрез изваждане на средната стойност и нормализиране спрямо стандартното отклонение.

За да се разработи високо генерализируема рамка, наборът от данни е проектиран така, че да включва образци, придобити при различни условия. Това обхваща вариации в зрителното поле, калибрационните модели и материалните характеристики. Латералното (x-y) зрително поле за отделните образци е следното: лимон (4 mm × 4 mm), вена (8 mm × 8 mm), череша (8 mm × 8 mm), зъб (4 mm × 4 mm), пръст (8 mm × 8 mm), семе (10 mm × 10 mm) и хайвер от писия (3 mm × 3 mm). На Фиг. 4.7 са показани различни k-линеаризационни интерференционни ивици за образците, които илюстрират различията в калибрационните схеми, използвани в началото на процеса на запис на данни чрез системата Optores OCT. Освен това, показателите на пречупване на избраните образци варират както следва: човешка венозна тъкан (1.3–1.4), лимон и череша (1.47), човешка кожа от пръст (1.42), човешки зъб (2.6–3.1), семе (1.5–1.7) и хайвер от писия (1.3–1.4). Това дава възможност за създаване на разнообразни спектрални модели, отговарящи на различни материални характеристики. Като цяло, такива вариации подпомагат обучението и оценяването на модела с цел по-висока адаптивност към разнообразието на набора от данни, включващ различни образци и параметри на придобиване, както и повишена устойчивост чрез включването на образци, засегнати от различна степен на спекъл шум. От седемте обемни набора от данни, описани

по-рано, пет — а именно лимон, вена, череша, зъб и пръст — съдържащи общо 3000 B-скана (по 600 от всеки обем), бяха случайно разпределени в поднабори за обучение, валидация и тестване в съотношение 70:20:10 за предложената рамка. Останалите два обема — семе (грах) и хайвер от писия — всеки съдържащ по 200 B-скана с размери 1152 × 256 пиксела, бяха използвани изключително за оценка на способността на модела да се генерализира, съгласно процедурата за независимо тестване. Тези два обема не бяха включвани по време на обучение или валидация.

Фазата на обучение започва с независимо трениране на SD-CNN. След завършване на това обучение, теглата на SD-CNN се фиксират, а FD-CNN се тренира, използвайки обработените изходни данни от SD-CNN, както е описано подробно в раздел 4.4.1. По време на фазата на за изводи (inference phase) SD-CNN първо обработва деградирано входно изображение, получено от неравномерно разпределена по вълново число информация, заедно с k-пространствената решетка, следвайки подхода, илюстриран на Фиг. 4.5. Върху изхода на SD-CNN се прилага DFT, за да се извлече амплитудата, която се подава като вход към FD-CNN. Предсказаният изход след това преминава през IDFT, за да се получи финалното OCT изображение. По този начин интегрираната рамка функционира чрез последователното прилагане на двата модела за реконструкция на изображението. За оптимизация на двата модела е използван Adam оптимизатор, със скорост на обучение 10^{-4}. SD-CNN достига сходимост след приблизително 200 епохи, докато FD-CNN изисква допълнително финно настройване и постига сходимост след около 400 епохи.

Използвани са два различни функционала на загуба: L_{L1} за SD-CNN и FFL за FD-CNN. По-конкретно, L_{L1} представлява средната абсолютна грешка, изчислена между изхода на SD-CNN (LR) и 7-усредненото еталонно изображение (HR), получено от системата Optores. Тази функция на загуба може да бъде записана (Уравнение 4.11), като пространствените индекси са означени с *i* и *j*::

$$L_{L1} = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}|HR - LR|, \tag{4.11}$$

FD-CNN използва Focal Frequency Loss (FFL) [150], изчисляван за изхода на модела ($F_\lambda(u,v)$) и съответното еталонно изображение ($F_k(u,v)$), където u и v представляват индексите на честотните коефициенти. Тази функция на загуба FFL може да се запише по следния начин::

$$FFL(u,v) = \frac{1}{UV}\sum_{u=0}^{U-1}\sum_{v=0}^{V-1} w(u,v)\,|F_k(u,v) - F_\lambda(u,v)|^2 \tag{4.12}$$

където тегловата матрица $w(u,v)$ се определя чрез:

$$w(u,v) = |F_k(u,v) - F_\lambda(u,v)|^\alpha \tag{4.13}$$

В уравнение (4.12) размерът на спектъра се означава чрез U и V, като тегловният фактор α=1. Минимизацията на грешката се прилага само върху амплитудната компонента, обозначена като FFL_{amp}, тъй като реконструкцията на B-scan изображения от суровите данни разчита изключително на спектралната амплитуда.

Рамката беше тествана върху тестовия набор от данни, описан по-горе, като се използваха следните показатели за оценка на резултатите от реконструкцията: MSE, SSIM, CNR_r (Contrast-to-Noise Ratio), β_s– параметър за гладкост и PSNR. CNR_r се изчислява математически за r-тата област чрез следната формула:

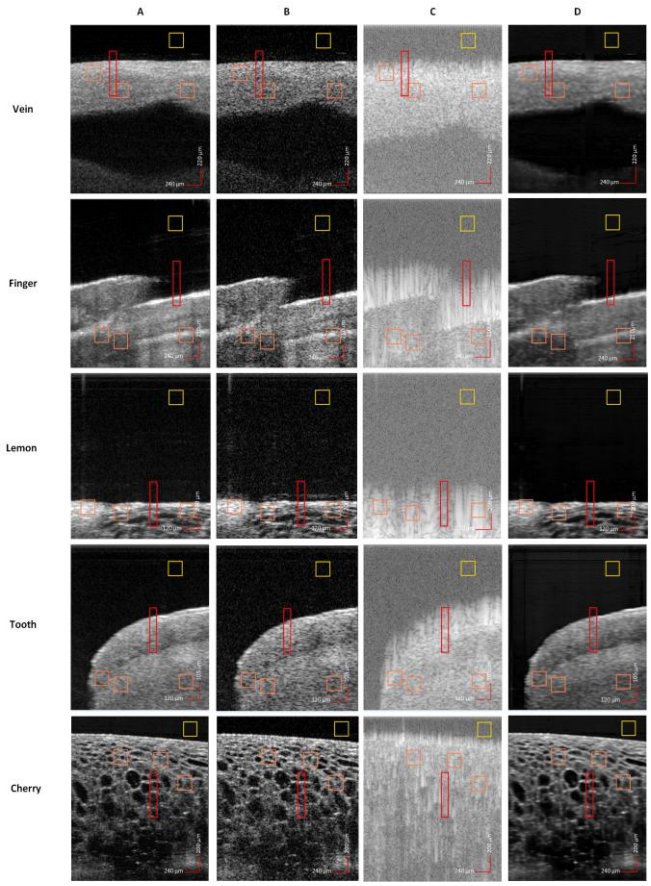
$$CNR_r = 10log_{10}\left(\frac{|\mu_f - \mu_b|}{\sqrt{{\sigma_f}^2 + {\sigma_b}^2}}\right), \tag{4.14}$$

където μ_f, и μ_b, а средните стойности на преден и заден план, а σ_f и σ_b а съответните стандартни отклонения за тези области. Параметърът за гладкост β_s се определя чрез:

$$\beta_s = \frac{\Gamma(I_{out}-\mu_{out}, I_{in}-\mu_{in})}{\sqrt{\Gamma(I_{out}-\mu_{out}, I_{out}-\mu_{out}).\Gamma(I_{in}-\mu_{in}, I_{in}-\mu_{in})}}, \tag{4.15}$$

където $\Gamma(I_1, I_2) = \sum_{i,j}[I_1(i,j).I_2(i,j)]$, а I_{in} и I_{out} са входните и изходните изображения. Анализът се извършва, като се сравняват деградираният вход, OCT изходът от системата Optores, изходът от предложената DL рамка и висококачествените еталонни изображения върху тестовия набор от данни.

Фиг. 4.8 илюстрира еталонното изображение (A), OCT изход без усредняване (B), деградирания суров вход за SD-CNN (C) и крайния изход от предложената DL рамка (D) за всички пет образци от тестовия набор. Крайният резултат от реконструкцията, постигнат чрез предложената рамка, демонстрира значително подобрение спрямо деградираните входове, като запазва фино структурните детайли в съответствие с еталонните изображения. Предложеният метод постига високо качество на изображението и намален шум в сравнение със стандартния OCT изход от системата Optores. Особено за образци като вена, зъб и пръст, които съдържат по-големи хомогенни региони, рамката ефективно потиска спекъл шума,



Фигура 4.8 Сравнение на B-сканове за 5 различни обема. (A) Ground truth (Усреднени 7 B-скана от OCT системата), (B) OCT система, (C) деградиран OCT, (D) резултат от предложената рамка. (P.IV)

като значително подобрява визуалното качество.

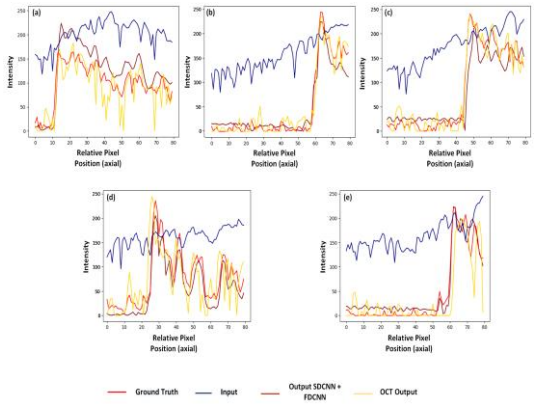
В Таблица 4.2 се сравнява предложената рамка чрез показателите PSNR, SSIM, CNR и β_s върху тестовия набор от данни. PSNR показва подобрение от 2 dB и 13 dB за предложената рамка спрямо резултатите, предоставени от OCT изхода (система Optores) и входа (деградиран вход за SD-CNN). Образците от лимон и череша, които съдържат по-фини структурни детайли в сравнение с другите образци, демонстрират относително високи стойности на PSNR, което ясно показва способността на рамката за висококачествена реконструкция на изображението. Сравнението по параметъра β_s показва подобрение в качеството на изображението, като метриката нараства от 0.87 до 0.93, подчертавайки способността на предложената рамка ефективно да потиска спекъл шума. Освен това, CNR беше изчислен чрез уравнение (4.14) за маркираните области, показани на Фиг. 4.8, като r = 3; оранжевите кутии се използват за преден план, а жълтата кутия за фон. Резултатите

показват, че моделът ефективно подчертава критичните структурни елементи на изображението спрямо шума за различни видове образци.

За оценка на качеството на изображението относно високочестотните компоненти като ръбове, се изследва вертикална колона пиксели, разположена в центъра на червената правоъгълна област, отбелязана на Фиг. 4.8. Тази оценка е представена чрез кривите на интензитета, показани на Фиг. 4.9 за всички пет категории образци: суров вход, еталонно изображение, стандартен OCT изход и крайният изход, произведен от предложения метод.

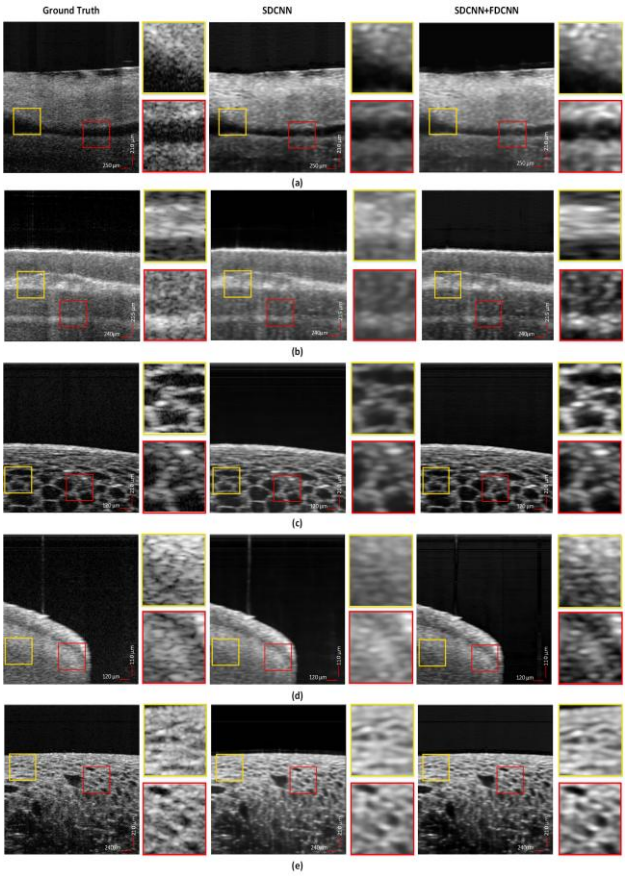
Таблица 4.2 Сравнение на показателите PSNR, SSIM, CNR, и β (P.IV)

Method	Dataset	PSNR	SSIM	CNR	β
Input	Overall	8.94	0.08	-	0.71
	Vein	12.76	0.14	4.44	0.68
	Finger	7.92	0.06	4.62	0.75
	Lemon	7.14	0.05	5.69	0.64
	Tooth	10.11	0.12	4.31	0.77
	Cherry	8.79	0.07	5.04	0.73
OCT Output	Overall	19.95	0.35	-	0.87
	Vein	21.20	0.22	5.21	0.88
	Finger	19.93	**0.45**	3.04	0.92
	Lemon	20.40	0.26	6.73	0.78
	Tooth	**22.11**	**0.56**	0.75	**0.93**
	Cherry	17.55	0.30	3.96	0.85
Proposed	Overall	**22.30**	**0.46**	-	**0.93**
	Vein	**21.98**	**0.43**	**8.71**	**0.93**
	Finger	**21.75**	0.42	**4.86**	**0.94**
	Lemon	**25.74**	**0.54**	**7.98**	**0.91**
	Tooth	21.61	0.32	**6.48**	0.92
	Cherry	**21.67**	**0.62**	**4.86**	**0.95**



Фигура 4.9 Сравнение на профилите на интензитетите (A.U.) за централната колона на червения правоъгълник, отбелязан на Фигура 4.8, като кривите са за (a) *вена*, (b) *пръст* (c) *лимон*, (d) *зъб* и (e) *череша*. (P.IV)

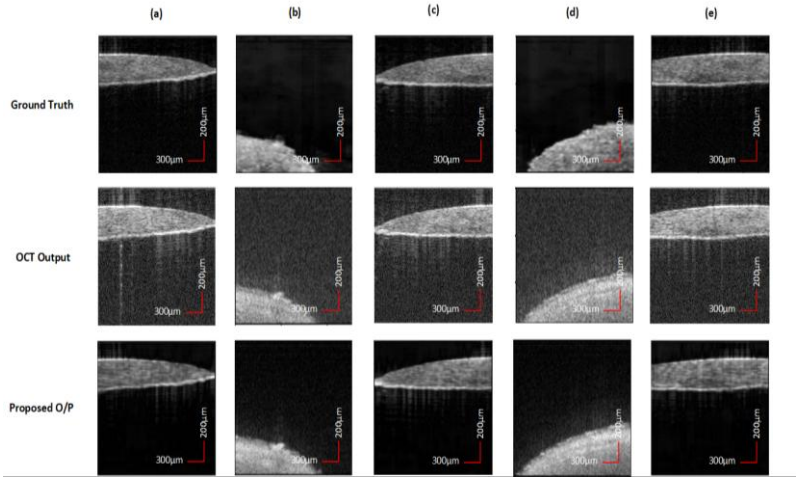
Оста x на графиките представлява относителен индекс на пикселите, като горният ред на маркирания правоъгълник е обозначен като позиция 1. Тези позиции са относителни и не съответстват на абсолютните координати на оригиналното изображение.



Фигура 4.10 Сравнение на SD-CNN и FD-CNN с еталонните данни. Образците са (a) *вена*, (b) *пръст*, (c) *лимон,* (d) *зъб* and (e) *череша*. (P.IV)

Кривата на еталонното изображение, показана в червено, е следвана от тъмночервената крива, която представя изхода на предложената рамка, което демонстрира висока достоверност на реконструкцията. ОСТ изходът от системата Optores показва повтарящи се колебания в интензитета, особено в равномерни или гладки области поради устойчив спекъл шум. Както се очаква, входните данни демонстрират значително отклонение на профила, доминирано от размазване и шум, поради нелинейното картографиране в спектъра на вълновия брой.

Ефективността на SD-CNN и FD-CNN се оцени чрез визуално сравнение на образците от тестовия набор – вена, пръст, лимон, зъб и череша (редове 1–5) – и съответните еталонни изображения, показани на Фиг. 4.10. Всяко изображение съдържа две маркирани кутии в



Фигура 4.11 Сравнение на резултатите от независимя тестови набор (a),(c),(e) *хайвер* and (b),(d) *семе(грах)* за еталони (ground truth), ОСТ реконструирани от предложения метод. (P.IV)

Таблица 4.3 Резултати за независим тестови набор (P.IV)

Volume	PSNR	SSIM
Flounder egg	22.09	0.45
Seed (pea)	21.53	0.42

жълто и червено, уголемени почти петкратно и поставени до оригиналите за по-добър анализ. Наблюдава се, че SD-CNN се справя добре при изображения с високо хомогенни региони, докато FD-CNN укрепва рамката чрез възможността си да реконструира високочестотни детайли като ръбове и граници. Цялата рамка осигурява последователна висококачествена реконструкция на изображенията благодарение на интегрирания физически базиран приоритет, който ръководи оптимизацията в пространствената и Фурие област.

Изпълнението е допълнително анализирано с независимия тестов набор от данни, съдържащ образци от семе (грах) и хайвер, състоящ се от 200 изображения, като се използват

Таблица 4.4 Аблационно изследване за оценка на приноса на отделните мрежи FD-CNN и SD-CNN при реконструиране на изображенията. (P.IV)

SD-CNN	FD-CNN	SD-CNN+FD-CNN	PSNR		SSIM	
			Avg	Std	Avg	Std
✓	-	-	20.81	2.64	0.42	0.11
-	✓	-	10.97	0.80	0.03	0.01
-	-	✓	22.30	2.51	0.46	0.08

стойности на PSNR и SSIM (Таблица 4.3). На Фиг. 4.11 се извършва сравнение между еталонното изображение, OCT изхода и изхода на предложената рамка за независимия тестов набор. Тези резултати потвърждават устойчивостта и адаптивността на рамката, показвайки силната ѝ способност да генерализира дори върху непознати данни.

За да се определят количествено индивидуалните приноси на SD-CNN и FD-CNN към цялостната рамка, беше проведено аблационно изследване (ablation study). В Таблица 4.4 са представени резултатите за отделните модели и комбинирания модел (SD-CNN + FD-CNN). От тях се вижда, че SD-CNN е основната мрежа, отговорна за реконструкцията на изображенията поради по-високите стойности на PSNR и SSIM. FD-CNN основно допълва производителността чрез маргинално подобрение на реконструкцията, тъй като проектирането и механизмът за извличане на характеристики на CNN позволяват постигане на по-висока точност в пространствената област, отколкото при приложение единствено във Фурие областта.

Времетраенето, свързано с различни междинни процеси в OCT системата – като калибрация, ресемплиране и FFT за един B-scan и усредняване за премахване на спекъл (7 изображения), е представено в Таблица 4.5. Това не отразява всички стъпки, включени в обработващия pipeline на Optores OCT системата. От измерванията на времетраенето става

Таблица 4.5 Сравнение на времетраене (P.IV)

Operations			Time (s)
Calibration			0.008
Resampling in K-space			0.17
FFT			0.05
Averaging (Speckle reduction)			0.07
Overall time Complexity -Volume (s)			
OCT system [30]	*Fourier Domain-CNN*	*Spatial Domain-CNN*	*Fourier Domain-CNN + Spatial Domain-CNN*
792	68.55	79.723	142.5

ясно, че ресемплирането изисква значително време от 0.17 секунди. В реални сценарии, индивидуалният B-scan е силно повлиян от спекъл шума, което налага използването на техники като усредняване на множество сканове или регистрационно-базирани подходи върху множество обеми за потискането му, увеличавайки изчислителното време за генериране на множество B-scans. Анализирано е и времетраенето за реконструкция на обем, съдържащ 600 изображения с размери 1152 × 1024. За тези експерименти, суровите данни са записани, а еталонните изображения са получени с намаляване на спекъл шума (усредняване върху 7 B-scans) чрез системата Optores. Предложената рамка обработва целия обем за 142.5 секунди, значително превъзхождайки комерсиалната система [20],[21], която изисква 792 секунди. При отделно изпълнение, компонентите FD-CNN и SD-CNN отнемат съответно 68.55 и 79.723 секунди. Подходът за усредняване на съседните сканове, използван от OCT системата [20],[21] за потискане на спекъл шума, води до по-гладки B-скан изходи, но увеличава изчислителното натоварване. За разлика от това, предложената рамка постига потискане на спекъл шума без необходимостта от множество B-scans, което значително ускорява обработката. Това ускорено и ефективно обработване позволява безпроблемна интеграция на OCT системите с модерни технологии, осигурявайки бърза и надеждна функционалност.

Глава 5 Заключения

Като модалност за биомедицинско визуализиране, Оптичната кохерентна томография (OCT) е постигнала значителни подобрения по отношение на автоматизацията чрез интегриране на подходи на дълбокото обучение (DL). Тези подходи имат способността да извличат сложни и многостепенни абстрактни характеристики, които са по-устойчиви и ефективни в сравнение с традиционните ръчно анотирани характеристики. В резултат на това те осигуряват висока точност на резултатите при ефективна обработка.

OCT системите са особено ефективни за визуализация на меки тъкани, позволявайки извършване на разнообразни измервания и улеснявайки анализа на техните основни физични характеристики. С микрометрова пространствена разделителна способност и милиметрова дълбочинна на проникване, OCT се оказва особено ценна при генериране на обемни данни за разширени вени, както беше разгледано в настоящата дисертация. Интеграцията на DL модели допълнително подобри разбирането на тези вени чрез възможност за точно измерване на дебелината на слоевете чрез невронни мрежи, специално оптимизирани за тези задачи.

В Глава 3 бяха представени ключовите компоненти, необходими за оценка на дебелината и автоматично прогнозиране на карти за сегментация от SS-OCT венозни изображения, елиминирайки нуждата от ръчно експертно въвеждане. Наборът от данни за тези цели беше събран от пациенти с разнообразни физически характеристики. Моделите за сегментация бяха оценявани по отношение на тяхната ефективност и производителност.

За да се преодолее предизвикателството на сложността на модела поради голям брой параметри, първият подход, беше разработен модела Opto-UNet, базиран на архитектурата U-Net [29]. Той предложи високо оптимизирана структура за точна сегментация на венозни изображения. Моделът беше тестван и сравнен с няколко съвременни метода, като моделът надмина другите по възможностите си за сегментация. Вторият модел, представен в Глава 3, TREE, целеше прогнозиране както на карти за сегментация, така и на стойности на дебелината от SS-OCT изображения. Моделът използва подход на трансферно обучение (transfer learning) и позволява първо да се научат картите за сегментация, които впоследствие улесняват по-ефективното измерване на дебелината. TREE интегрира специализиран SPD модул, който комбинира атросни и дълбочинно-сепарабилни конволюции (depthwise separable convolutions) в персонализирана конфигурация. Тази интегрирана рамка позволява извличане на пълна информация за форма и дебелина в една единна мрежа. Проведени бяха обширни експерименти за оценка на модела и количествено сравнение със съвременните

методи. Резултатите демонстрираха, че TREE надминава съществуващите методи по точност и прецизност както при сегментацията, така и при измерванията на дебелината на вените.

Реконструкцията на висококачествени изображения с ниска изчислителна сложност и потиснат спекъл шум е ключова за ефективната OCT визуализация. В Глава 4 бяха представени DL базирани модели, разработени за подобряване на реконструкцията на изображенията в OCT системи. Моделът WAVE-UNET генерира изображения директно от сурови данни, линейни по дължината на вълната – формат, който обикновено води до нискокачествени реконструкции при обработка във Фурие пространството. Вместо да се извършва k-линеаризация и ремапинг, DL моделите приемат суровите интерферограми като вход и ги обработват за генериране на B-scan изображения. Беше демонстрирано, че размазаните входни OCT изображения могат да бъдат използвани от невронната мрежа за генериране на B-скан с висок контраст. Реконструкционната рамка, включваща DL модели както в пространствената, така и във Фурие oblast, успешно генерира сложни морфологични структури, като значително потиска спекъл шума – два критични фактора в биомедицинското визуализиране. Тази рамка беше също така оптимизирана за ниско време на извличане на заключения, гарантирайки ефективно изпълнение. Предложените методи бяха стриктно оценени и визуализирани върху разнообразен набор от образци, включително вена, череша, лимон, пръст, зъб, яйце от камбала и грахово семе, всяка със специфични материални свойства. Количествените и качествените резултати последователно демонстрираха превъзходството на моделите спрямо съществуващите подходи.

Авторски приноси в дисертационната работа

Основните приноси на настоящата дисертация се намират в приложно-теоретичната област и служат като основа за разработване и оптимизация на DL модели за измерване на венозни структури, реконструкция на OCT изображения от данни, линейни в пространството на дължината на вълната, като същевременно допринасят за разбирането на физиката на вълновите процеси. В резултат на систематично теоретично и експериментално изследване бяха предложени нови решения за напреднала обработка на данни и разработени нови реализации на DL-базирани методи. Въз основа на горепосочените заключения, дисертацията се стреми да отговори на пет основни изследователски въпроса, идентифицирани като научни пропуски:

RQ1. Колко точно може DL модел да оцени картите за сегментация на венозните слоеве, като същевременно поддържа ниска сложност чрез минимален брой параметри, използвайки обемни данни, придобити чрез SS-OCT?

Тук беше интегрирана възможността на атросни и дълбочинно-сепарабилни конволюции и бе предложена енкодер-декодер архитектура, базирана на U-Net с резидуални връзки. Основният принос се състои в предложението на нов bridge-parallel блок, който минимизира броя на параметрите на модела, като същевременно осигурява висока точност на сегментацията на венозните слоеве.

RQ2. Как може да се разработи интегриран модел, който едновременно осигурява високо точни карти за сегментация – надеждно идентифицирайки истински положителни и отрицателни области, минимизирайки фалшивите детекции – и предоставя точни и последователни измервания на дебелината върху различни SS-OCT обемни набори данни на разширени вени?

Това бе постигнато чрез използване на трансферно обучение (transfer learning) в модел с единичен енкодер и два декодера, обучен специално за извличане на характеристиките, необходими за сегментация, последвано от измерване на дебелината. Кондензационният (Bottleneck) модулът съдържа компресирани и високо абстрактни представяния, научени по време на процеса на сегментация, които обхващат най-информативните, отличителни и семантично значими характеристики. Тези представяния се използват за измерване на

дебелината чрез трансферно обучение, позволявайки на модела да постигне висока точност и прецизност.

RQ3. Как може да се разработи DL мрежа за реконструиране на SS-OCT изображения директно от данни, линейни по дължината на вълната, като се заобиколи k-линеаризацията, намалят се хардуерните разходи и изчислителните изисквания, като същевременно се поддържа качество на изображението, сравнимо със стандартния SS-OCT процес?

Запазените интерферограми (линейни по дължината на вълната) бяха подложени на изваждане на фона, прозоречен филтър и Фурие трансформация, за да се получат нискокачествени размазани изображения. Тези размазани изображения се използваха като вход за реконструкционната мрежа, позволявайки да се заобиколи k-линеаризацията. Предложеният DL модел бе обучен с висококачествени еталонно изображения (ground truth), генерирани чрез усредняване по 7 B-скана, за минимизиране на спекъл шума. Моделът постигна висока производителност, осигурявайки превъзходно качество на изображенията в сравнение с OCT изображенията, генерирани от комерсиалната система Optores, без необходимост от допълнителен хардуер за линеаризация.

RQ4. Как може да се разработи усъвършенствана DL мрежа за реконструиране, която оптимизира данните както в пространствената, така и във Фурие областта, генерирайки висококачествени изображения с намален спекъл шум и подобрени показатели PSNR и SSIM в сравнение с конвенционалната обработка на сурови SS-OCT данни?

За получаване на висококачествени резултати, еталонните изображения бяха извлечени от SS-OCT системата. За подобряване на детайлите и структурите, DL моделът беше оптимизиран не само в пространствената област, но и във Фурие областта за подобряване на финитe детайли. Тази многодоменна оптимизация позволи рамката да запази високочестотни характеристики като ръбове, като същевременно потиска шума без прекомерно заглаждане.

RQ5. Възможно ли е да се разработи DL-базирано решение с ниска изчислителна сложност като алтернатива на стандартната процедура за реконструиране, което постига висококачествени изображения, позволявайки визуализация в реално време?

Предложен бе DL-базиран модел, който преодолява множество етапи от конвенционалната SS-OCT процедура за обработване на сигнала – от трансформиране на записаните интерферограми до финални OCT B-скан изображения, значително подобрявайки времетраенето. Единната DL мрежа постига ниско време за предсказване (inference) благодарение на предварително обучен модел с паралелни изчисления, batch processing и високо ниво на оптимизации.

В контекста на посочените изследователски въпроси е важно да се отбележат някои ограничения на представената работа. Моделът за сегментация и измерване на дебелината, предложен в Публикация II, използва набор от данни само от четирима пациенти. В бъдещи изследвания е необходимо да се включат и нормални вени и по-голям брой образци за оценка на устойчивостта на модела.

Публикации III и IV използват само един тип SS-OCT система – комерсиалната Optores GmbH с един лазерен източник, което ограничава оценката на способността за генерализация на модела. В настоящата работа бе използван методът за потискане на спекъл шума чрез 7-кратно усредняване, като в бъдещи изследвания могат да се изучават по-напреднали методи за шумопотискане.

В перспектива, изследването може да бъде разширено чрез включване на повече вени от пациенти с различни заболявания, както и здрави индивиди, като се обърне внимание и на други венозни заболявания като Хронична венозна недостатъчност, за подобряване на способността на DL модела да генерализира и увеличаване на устойчивостта към вариабилност.

Що се отнася до реконструкцията на SS-OCT изображения от сурови данни, в бъдещи изследвания могат да се използват разнообразни OCT системи с различни светлинни

източници. Допълнително, включването на компенсиране на дисперсията в обработката може да подобри качеството на изображенията, а времевата сложност може да бъде още намалена чрез усъвършенствани техники за оптимизация и използване на модерни невронни мрежи.

Литература

[1] D. Huang et al., “Optical Coherence Tomography,” Science (80-.)., vol. 20, pp. 1–26, 1991..

[2] Sattler, Elke, Raphaela Kästle, and Julia Welzel. "Optical coherence tomography in dermatology." Journal of biomedical optics 18, no. 6 (2013): 061224-061224.

[3] Boone, Marc Alm, Sarah Norrenberg, G. B. E. Jemec, and Véronique Del Marmol. "Imaging of basal cell carcinoma by high-definition optical coherence tomography: histomorphological correlation. A pilot study." British journal of dermatology 167, no. 4 (2012): 856-864.

[4] Kermany, Daniel S., Michael Goldbaum, Wenjia Cai, Carolina CS Valentim, Huiying Liang, Sally L. Baxter, Alex McKeown et al. "Identifying medical diagnoses and treatable diseases by image-based deep learning." cell 172, no. 5 (2018): 1122-1131.

[5] Ahmed, Hanya, Qianni Zhang, Ferranti Wong, Robert Donnan, and Akram Alomainy. "Lesion Detection in Optical Coherence Tomography with Transformer-Enhanced Detector." *Journal of Imaging* 9, no. 11 (2023): 244.

[6] Wan, B., Clarisse Ganier, X. Du-Harpur, N. Harun, F. M. Watt, Rakesh Patalay, and M. D. Lynch. "Applications and future directions for optical coherence tomography in dermatology." British Journal of Dermatology 184, no. 6 (2021): 1014-1022.

[7] Zhou, Kang, Shenghua Gao, Jun Cheng, Zaiwang Gu, Huazhu Fu, Zhi Tu, Jianlong Yang, Yitian Zhao, and Jiang Liu. "Sparse-gan: Sparsity-constrained generative adversarial network for anomaly detection in retinal oct image." In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), pp. 1227-1231. IEEE, 2020.

[8] Sigsgaard, V., L. Themstrup, P. Theut Riis, J. Olsen, and G. B. Jemec. "In vivo measurements of blood vessels’ distribution in non-melanoma skin cancer by dynamic optical coherence tomography—a new quantitative measure?." Skin Research and Technology 24, no. 1 (2018): 123-128.

[9] Yang, Chin-Yi, Ja-Hon Lin, and Chien-Ming Chen. "Optical Coherence Tomography as a Diagnosis-Assisted Tool for Guiding the Treatment of Melasma: A Case Series Study." Diagnostics 14, no. 18 (2024): 2083.

[10] Lindert, Judith, Kianusch Tafazzoli-Lari, Ludger Tüshaus, Beke Larsen, Anna Bacia, Marie Bouteleux, Tina Adler et al. "Optical coherence tomography provides an optical biopsy of burn wounds in children—a pilot study." Journal of Biomedical Optics 23, no. 10 (2018): 106005-106005.

[11] Bouma, Brett E., Johannes F. de Boer, David Huang, Ik-Kyung Jang, Taishi Yonetsu, Cadman L. Leggett, Rainer Leitgeb et al. "Optical coherence tomography." Nature Reviews Methods Primers 2, no. 1 79, (2022).

[12] Drexler, Wolfgang, and James G. Fujimoto, eds. Optical coherence tomography: technology and applications. Springer Science & Business Media, 2008.

[13] De Boer, Johannes F., Rainer Leitgeb, and Maciej Wojtkowski. "Twenty-five years of optical coherence tomography: the paradigm shift in sensitivity and speed provided by Fourier domain OCT." Biomedical optics express 8, no. 7 (2017): 3248-3280.

[14] P. H. Tomlins, and R. K. Wang. "Theory, developments and applications of optical coherence tomography." Journal of Physics D: Applied Physics, 38, no. 15, 2519 (2005).

[15] Zhou, Shi-you, Chun-xiao Wang, Xiao-yu Cai, David Huang, and Yi-zhi Liu. "Optical coherence tomography and ultrasound biomicroscopy imaging of opaque corneas." Cornea 32, no. 4 (2013): e25-e30.

[16] Ramos, Jose Luiz Branco, Yan Li, and David Huang. "Clinical and research applications of anterior segment optical coherence tomography–a review." Clinical & experimental ophthalmology 37, no. 1 (2009): 81-89.

[17] Chen, Teresa C., Barry Cense, Mark C. Pierce, Nader Nassif, B. Hyle Park, Seok H. Yun, Brian R. White, Brett E. Bouma, Guillermo J. Tearney, and Johannes F. De Boer. "Spectral domain optical coherence tomography: ultra-high speed, ultra-high resolution ophthalmic imaging." Archives of ophthalmology 123, no. 12 (2005): 1715-1720.
[18] Leitgeb, R. A., W. Drexler, A. Unterhuber, B. Hermann, T. Bajraszewski, T. Le, A. Stingl, and A. F. Fercher. "Ultrahigh resolution Fourier domain optical coherence tomography." Optics express 12, no. 10 (2004): 2156-2165.
[19] Potsaid, Benjamin, Bernhard Baumann, David Huang, Scott Barry, Alex E. Cable, Joel S. Schuman, Jay S. Duker, and James G. Fujimoto. "Ultrahigh speed 1050nm swept source/Fourier domain OCT retinal and anterior segment imaging at 100,000 to 400,000 axial scans per second." Optics express 18, no. 19 (2010): 20029-20048.
[20] Wieser, Wolfgang, Benjamin R. Biedermann, Thomas Klein, Christoph M. Eigenwillig, and Robert Huber. "Multi-megahertz OCT: High quality 3D imaging at 20 million A-scans and 4.5 GVoxels per second." Optics express 18, no. 14 (2010): 14685-14704.
[21] Klein, Thomas, Wolfgang Wieser, Lukas Reznicek, Aljoscha Neubauer, Anselm Kampik, and Robert Huber. "Multi-mhz retinal oct." *Biomedical optics express* 4, no. 10 (2013): 1890-1908.
[22] Attendu, Xavier, Roosje M. Ruis, Caroline Boudoux, Ton G. Van Leeuwen, and Dirk J. Faber. "Simple and robust calibration procedure for k-linearization and dispersion compensation in optical coherence tomography." *Journal of biomedical optics* 24, no. 5 (2019): 056001-056001.
[23] [48] Wu, Tong, Zhihua Ding, Ling Wang, and Minghui Chen. "Spectral phase based k-domain interpolation for uniform sampling in swept-source optical coherence tomography." Optics express 19, no. 19 (2011): 18430-18439.
[24] Szkulmowski, Maciej, Maciej Wojtkowski, Tomasz Bajraszewski, Iwona Gorczyńska, Piotr Targowski, Wojciech Wasilewski, Andrzej Kowalczyk, and Czesław Radzewicz. "Quality improvement for high resolution in vivo javascript:void(0)images by spectral domain optical coherence tomography with supercontinuum source." Optics communications 246, no. 4-6 (2005): 569-578.
[25] Chen, Yuping, Hong Zhao, and Zhao Wang. "Investigation on spectral-domain optical coherence tomography using a tungsten halogen lamp as light source." Optical review 16, no. 1 (2009): 26-29.
[26] Liang, Kaicheng, Zhao Wang, Osman O. Ahsen, Hsiang-Chieh Lee, Benjamin M. Potsaid, Vijaysekhar Jayaraman, Alex Cable, Hiroshi Mashimo, Xingde Li, and James G. Fujimoto. "Cycloid scanning for wide field optical coherence tomography endomicroscopy and angiography in vivo." Optica 5, no. 1 (2018): 36-43.
[27] Zhang, Jason, Tan Nguyen, Benjamin Potsaid, Vijaysekhar Jayaraman, Christopher Burgner, Siyu Chen, Jinxi Li et al. "Multi-MHz MEMS-VCSEL swept-source optical coherence tomography for endoscopic structural and angiographic imaging with miniaturized brushless motor probes." Biomedical Optics Express 12, no. 4 (2021): 2384-2403.
[28] Lee, Hwi Don, Gyeong Hun Kim, Jun Geun Shin, Boram Lee, Chang-Seok Kim, and Tae Joong Eom. "Akinetic swept-source optical coherence tomography based on a pulse-modulated active mode locking fiber laser for human retinal imaging." Scientific reports 8, no. 1 (2018): 17660.
[29] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." In International Conference on Medical image computing and computer-assisted intervention, pp. 234-241. Cham: Springer international publishing, 2015.
[30] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016.
[31] Badrinarayanan, Vijay, Alex Kendall, and Roberto Cipolla. "Segnet: A deep convolutional encoder-decoder architecture for image segmentation." IEEE transactions on pattern analysis and machine intelligence 39, no. 12 (2017): 2481-2495.

[32] Zhao, Hengshuang, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. "Pyramid scene parsing network." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2881-2890. 2017.
[33] Chen, Liang-Chieh, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. "Encoder-decoder with atrous separable convolution for semantic image segmentation." In Proceedings of the European conference on computer vision (ECCV), pp. 801-818. 2018.
[34] Ashish, Vaswani. "Attention is all you need." Advances in neural information processing systems 30 (2017): I.
[35] Gu, Shuhang, Lei Zhang, Wangmeng Zuo, and Xiangchu Feng. "Weighted nuclear norm minimization with application to image denoising." *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2862-2869. 2014.
[36] Zhang, Zhengxin, Qingjie Liu, and Yunhong Wang. "Road extraction by deep residual u-net." *IEEE Geoscience and Remote Sensing Letters* 15, no. 5 (2018): 749-753.

Списък с публикации, включени в дисертацията

Publication I M. Viqar, V. Madjarova, V. Baghel, and E. Stoykova, "Opto-UNet: optimized UNet for segmentation of varicose veins in optical coherence tomography," in *Proc. 2022 10th Eur. Workshop on Visual Information Processing (EUVIP)*, pp. 1–6, IEEE, 2022.

Publication II M. Viqar, V. Madjarova, E. Stoykova, D. Nikolov, E. Khan, and K. Hong, "Transfer learning-based approach for thickness estimation on optical coherence tomography of varicose veins," *Micromachines*, vol. 15, no. 7, 2024.

Publication III M. Viqar, E. Sahin, V. Madjarova, E. Stoykova, and K. Hong, "WAVE-UNET: wavelength-based image reconstruction method using attention UNET for OCT images," in *Proc. Medical Imaging 2024: Physics of Medical Imaging*, vol. 12925, pp. 839–847, SPIE, 2024.

Publication IV M. Viqar, E. Sahin, E. Stoykova, and V. Madjarova, "Reconstruction of optical coherence tomography images from wavelength space using deep learning," *Sensors*, vol. 25, no. 1, 2024.

Списък на участията в научни форуми и мероприятия

Устни доклади:

1. 10th European Workshop on Visual Information Processing (EUVIP), Lisbon, Portugal, 11-14 September, 2022.
2. PLENOPTIMA Webinar Series 1, 2022.
3. Presentation at Tampere University, Secondments, January, 2023.
4. PLENOPTIMA Webinar Series 2, 2023.
5. HISTRATE conference on Advanced Composites under High Strain Rates Loading: A Route to Certification-by-Analysis, Istanbul, 5-6 June, 2024.
6. PLENOPTIMA Webinar Series 3, 2024.
7. European Light Field Imaging Workshop (ELFI), Sozopol, Bulgaria, 25-27 June, 2024.
8. Bulgarian-Korean Workshop, Photonics and Sensorics, Sofia, Bulgaria, 10-11 October, 2024.

Постерни доклади:

9. Digital Holography and 3-D Imaging, Cambridge, United Kingdom, 1–4 August 2022.
10. Materials, Methods & Technologies Conference, Burgas, Bulgaria, 19- 22 August, 2022.
11. PLENOPTIMA Training School 2: Machine and Deep Learning and Workshop 3: Research Communication, Mid Sweden University, Sundsvall, Sweden, 19-23 September, 2022.
12. Medical Imaging 2024: Physics of Medical Imaging, San Diego, California, United States of America, 18-22 February, 2024.
13. HISTRATE conference on Advanced Composites under High Strain Rates Loading: A Route to Certification-by-Analysis, Istanbul, Turkiye, 5-6 June, 2024.
14. International Conference on Advances in Material Science & Photonics (AMSP'25), Bansko, Bulgaria, 15 - 18 July, 2025.

Списък на цитатите на авторските статии

Viqar, Maryam, Violeta Madjarova, Vipul Baghel, and Elena Stoykova. "Opto-UNet: optimized UNet for segmentation of varicose veins in optical coherence tomography." In *2022 10th European Workshop on Visual Information Processing (EUVIP)*, pp. 1-6. IEEE, 2022.

Citations:

1. Rufer, Libor, Josué Esteves, Didace Ekeom, and Skandar Basrour. "Piezoelectric Micromachined Microphone with High Acoustic Overload Point and with Electrically Controlled Sensitivity." *Micromachines* 15, no. 7 (2024): 879.
2. Su, Junjia, Yihao Chen, Pengcheng Feng, Zhelong Jiang, Zhigang Li, and Gang Chen. "A high resolution and configurable 1T1R1C ReRAM Macro for Medical Semantic Segmentation." *IEICE Electronics Express* 21, no. 8 (2024): 20240071-20240071.
3. Mao, Rongzhi, Liangxu Xie, Xiaohua Lu, Jialu Pei, Xiaojun Xu, and Shan Chang. "Harnessing Multiple Level Features to Improve Segmentation Performance of Deep Neural Network: A Case Study in Magnetic Resonance Imaging of Nasopharyngeal Cancer." *IEEE Access* (2024).
4. Rajathi, V., A. Chinnasamy, and P. Selvakumari. "DUTC Net: A novel deep ulcer tissue classification network with stage prediction and treatment plan recommendation." *Biomedical Signal Processing and Control* 90 (2024): 105855.
5. Rathnam, KS Revanth Sai, V. Lokanadham, C. Vignesh, and Vijayakumar Peroumal. "Varicose Veins Disease Detection and Automated Treatment using Body Area Network." In *2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)*, pp. 1-6. IEEE, 2024.
6. Arunkumar, M., A. Mohanarathinam, and Kamalraj Subramaniam. "Detection of varicose vein disease using optimized kernel Boosted ResNet-Dropped long Short term Memory." *Biomedical Signal Processing and Control* 87 (2024): 105432.
7. Das, M. Ashwin, I. Anand, C. Nihal, Kamalraj Subramaniyam, and A. Mohanarathinam. "Early Detection and Prevention of Varicose Veins using Embedded Automation and Internet of Things." In *2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA)*, pp. 1476-1482. IEEE, 2023.

Viqar, Maryam, Violeta Madjarova, and Elena Stoykova. "Modified watershed approach for segmentation of complex optical coherence tomographic images." Materials, Methods & Technologies.

- Arrabiyeh, Peter A., Moritz Bobe, Miro Duhovic, Maximilian Eckrich, Anna M. Dlugaj, and David May. "Implementing low budget machine vision to improve fiber alignment in wet fiber placement." *Journal of Reinforced Plastics and Composites* (2024): 07316844241278050.
- Anzabi, Leila Chehreghani, Azizollah Shafiekhani, Elham Darabi, Mirosław Bramowicz, Sławomir Kulesza, Jamshid Sabbaghzadeh, and Shahram Solaymani. "Nanoscale 3D spatial analysis of FTO/ZnO/Ag-x films subjected to photocatalytic activity." *Scientific Reports* 15, no. 1 (2025): 7330.

Viqar, Maryam, Violeta Madjarova, Amit Kumar Yadav, Desislava Pashkuleva, and Alexander S. Machikhin. "Deep learning based segmentation of optical coherence tomographic images of human saphenous varicose vein." In *Digital Holography and Three-Dimensional Imaging*, pp. W2A-5. Optica Publishing Group, 2022.

- Cui, Weixin, Shan Lou, Wenhan Zeng, Visakan Kadirkamanathan, Yuchu Qin, Paul J. Scott, and Xiangqian Jiang. "BlurRes-UNet: A novel neural network for automated surface characterisation in metrology." *Computers in Industry* 165 (2025): 104228.
- MJ, Prasanna Kumar, and H. N. Prakash. "A Comprehensive Review on Varicose Veins and its Implications using AI Methods." *Grenze International Journal of Engineering & Technology (GIJET)* 10 (2024).

Благодарности

Този дисертационен труд е осъществен в рамките на програмата European Joint Doctoral Programme on Plenoptic Imaging (PLENOPTIMA) и е получил финансиране от the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie проект No 956770.