

УДК 004.852 + 616-018

ЗГОРТКОВА НЕЙРОННА МЕРЕЖА ДЛЯ СЕГМЕНТАЦІЇ СУДИН СІТКІВКИ ОКА

Добровська Людмила Миколаївна

Назарага Ярослав Русланович

nazaraha.yaroslav@iit.kpi.ua

кафедра біомедичної кібернетики

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»,
м. Київ, Україна

Анотація – Важливе значення для постановки діагнозу при різних офтальмологічних захворюваннях відіграє дослідження, моніторинг та оцінка судин сітківки ока. Ідентифікація конкретних об'єктів-патологій на зображеннях зводиться до розв'язання задач сегментації. Сегментація судин сітківки є ключовим кроком до точної візуалізації, діагностики захворювань ока, раннього лікування та планування хірургічного втручання. Саме тому важливою задачею є автоматизована сегментація судин сітківки.

Мета даної роботи полягала у розробці програмного застосунку для сегментації зображень судин сітківки ока з використанням машинного навчання у вигляді згорткової нейронної мережі.

Наразі найточнішими є нейромережеві методи сегментації, а саме методи на основі глибокого навчання.

База знімків, яка використовувалась для дослідження, була взята з загальнодоступного набору даних DRIVE, що надає еталонні сегментації (маски) для кожного зображення, для половини з яких застосовано аугментацію. Це надає змогу обчислити оцінки продуктивності моделі.

Під час першого етапу дослідження було встановлено, що наразі до найпоширеніших:

1) критеріїв, за якими кількісно можна оцінити якість сегментації, належать такі метрики: intersection over union; accuracy; precision; sensitivity; specificity; F1-score; dice coefficient; loss function;

2) мереж, здатних виконувати сегментацію зображень, належать такі: FCN, SegNet, U-Net, FC-Densenet, E-Net, Link-Net, RefineNet, PSPNet.

Для вирішення задачі сегментації зображень судин сітківки ока виконано порівняння вказаних мереж за точністю, перевагами та обмеженнями. Встановлено, що враховуючи точність, найкраще для вирішення вказаної задачі підходять мережі DeepLab, PSPNet, U-Net.

Другий етап дослідження полягав у розробці програмного застосунку (ПЗ), порівнянні та оцінюванні показників якості відомих систем сегментації судин сітківки та розробленого ПЗ.

Результатом дослідження є розробка ПЗ, який надає такі оцінки за метриками: accuracy=0.9452, sensitivity = 0.8991, specificity= 0.9468, dice= 0.8247. Ці показники якості отримано під час роботи *розробленого ПЗ*.

Ключові слова – штучний інтелект, нейронна мережа, сегментація судин сітківки ока, обробка зображень.

I. ВСТУП

Значна кількість інформації про оточення людини надходить через систему зору, тому вона є важливою частиною світосприйняття. Одним із найвагоміших елементів у цій системі є сітківка очей, яка відповідає за формування зображення та передачу нервових імпульсів у головний мозок. Наразі своєчасне діагностування патологій сітківки залишається істотною умовою успішної боротьби з хворобами очей. Важливе значення для постановки діагнозу при різних офтальмологічних захворюваннях відіграє дослідження, моніторинг та оцінка судин сітківки ока.

Ідентифікація конкретних об'єктів-патологій на зображеннях зводиться до

розв'язання задач сегментації. Сегментація судин сітківки є ключовим кроком до точної візуалізації, діагностики захворювань ока (таких як діабетична ретинопатія, глаукома тощо), раннього лікування та планування хірургічного втручання очних захворювань.

Актуальність роботи обумовлена тим, що у сучасному інформатизованому світі кількість офтальмологічних захворювань стрімко зростає, тому своєчасне діагностування проблем судин сітківки ока та розробка програмного застосунку на основі ефективних методів сегментації для подальшого використання отриманих даних для діагностики є надзвичайно важливими.

Мануальна сегментація судин сітківки (де процесом ідентифікації відхиленя

займається безпосередньо людина) є трудомісткою та виснажливою. На сьогодні для аналізу медичних зображень застосовуються методи машинного навчання з використанням нейронних мереж (НМ).

Останнім часом методи сегментації судин сітківки на основі глибокого навчання досягли значної ефективності.

II. МЕТА ДОСЛІДЖЕННЯ

Метою роботи було розробити програмний застосунок для сегментації зображень судин сітківки ока з використанням машинного навчання у вигляді згортової НМ та оцінити його продуктивність (оцінювання проводилося за метриками accuracy, precision, specificity, sensitivity та dice).

III. МАТЕРІАЛИ ТА МЕТОДИ

3.1. Поняття сегментації та її особливості

На сучасному етапі суспільного розвитку інформаційні технології (ІТ) проникли в усі сфери людської діяльності. Усе більшого значення набуває використання ІТ у сфері охорони здоров'я, завдяки чому медицині стали притаманні нові ознаки. Так, сьогодні складні дослідження в галузі медицини відбуваються за допомогою застосування обчислювальної техніки та спеціального програмного забезпечення, при цьому кількість різноманітних комп'ютерних алгоритмів та математичних методів невпинно зростає. Це пов'язано з необхідністю обробки, аналізу та систематизації нової інформації, отриманої під час різноманітних медичних досліджень.

Одним із напрямків, який отримав розвиток завдяки прискореному темпу удосконалення засобів комп'ютерної техніки, стала розробка теоретичних засад і практична реалізація систем розпізнавання та обробки зображень. Важливу роль у

задачах обробки медичних зображень, поряд із ресстрацією, класифікацією та локалізацією, відіграє сегментація.

Загалом, під *сегментацією* розуміють процес розділення цілого на окремі частини (сегменти). У випадку роботи з цифровими зображеннями, *сегментація* - це процес розділення цифрового знімка на кілька сегментів (множини пікселів або суперпікселів). При цьому зображення спрощується або змінюється його подання для полегшення подальшого аналізу [0]. У більшості випадків сегментація зображень використовується для виділення конкретних об'єктів та контурів на зображеннях.

Для встановлення правильного діагнозу з метою подальшого лікування використовуються різноманітні біомедичні зображення, такі як рентгенограми, ультразвукові, зображення цифрової мікроскопії, магнітно-резонансної томографії тощо.

У загальному випадку для довільних зображень не існує універсальних методів сегментації, а використовуються спеціалізовані методи, які для визначеного класу задач є найпоширенішими.

У роботі [0] Харлік Р. та Шапіто Л. виділили характеристики, властиві областям, отриманим після сегментації:

- сегменти зображення повинні бути однорідними щодо певної характеристики, наприклад, яскравості або текстури;
- сусідні сегменти повинні суттєво відрізнятися один від одного відносно обраної характеристики;
- середина області повинна бути суцільною та без великої кількості "дірок";
- межа кожної області повинна бути неперервною, простою та просторово точною.

Традиційні методи сегментації та їх класифікація. Відомі різні *підходи до класифікації алгоритмів сегментації*: Фу [0], Пала [0], Скарбека і Кошана [0], Лючеса

та Мітра [0]. Це пов'язано з тим, що наразі не існує повної та всеосяжної класифікації методів сегментації зображень, оскільки кожна із запропонованих так чи інакше не враховує деякі алгоритми, наприклад, авторські методики, що є не дуже відомими і поширеними.

Існуючі схеми класифікації методів сегментації базуються на таких критеріях: властивості областей, меж областей та точок, апріорні знання про об'єкти тощо. До інших критеріїв поділу алгоритмів сегментації відносяться наступні: тип зображення (бінарні, напівтонові, кольорові), вид обробки зображень під час сегментації (паралельна або послідовна) [7]. Але описані характеристики не завжди є однозначними. Деякі види сегментації (наприклад, порогова) можуть відбуватись як в послідовному, так і в паралельному режимах, а також працювати як з бінарними, так і напівтоновими зображеннями. Для уникнення неоднозначності в класифікації наявних алгоритмів сегментації у роботі [8] виділено такі критерії: на основі властивостей точок, властивостей областей і властивостей текстури, взаємозв'язків між точками, виділення границь. Варто зауважити, що ці традиційні методи сегментації не надають достатньої точності розпізнавання зображень. Найточнішими наразі є нейромереві методи сегментації, тобто методи на основі глибокого навчання.

3.2. Сучасні методи глибокого навчання для сегментації зображень

Глибоке навчання – це один із методів машинного навчання, в основі якого лежить застосування нейронних мереж. Завдяки розвитку глибокого навчання удосконалюються і методи сегментації зображень. У порівнянні з класичними методами сегментації алгоритми глибокого навчання, зазвичай, призводять до набагато кращих та точніших результатів, особливо в

умовах, що відрізняються від тих, у яких вони проходять тестування.

Основним принципом роботи глибокої НМ є послідовні лінійні операції, які поєднуються з нелінійними. Завдяки великій кількості вузлів у мережі та введенню нелінійних функцій активації такі структури можуть виявляти численні нелінійні закономірності в даних.

Зазвичай, в моделях глибокого навчання кожен шар супроводжується використанням нелінійної функції активації. Отже, вихід одного шару обчислюється наступним чином:

$$X_{out} = \varphi(wX_{input} + b),$$

де X_{input} - вхідне значення змінної; w - вага; b - зсув лінійної функції; φ - функція активації.

Вперше з метою семантичної сегментації були використані такі підходи машинного навчання як класифікатори TextonForest і Random Forest. А з появою згорткових нейронних мереж (CNN) у 1998 р. задачі обробки зображень вийшли на інший рівень розвитку.

Згорткові нейронні мережі мають двовимірну структуру та виконують аналіз зображень не повністю, а його окремих частин. Це забезпечується спеціальними згортковими та підвибірковими шарами, які зменшують розмір вхідного сигналу з кожним наступним шаром. Описаним способом утворюється певна «піраміда» шарів мережі, де кожен окремий шар складається з набору площин, для кожної такої площини використовується власний набір значень ваги. Згорткові нейронні мережі призводять до швидкого та ефективного навчання внаслідок меншої кількості налаштовуваних вагових коефіцієнтів.

Однією з перших спроб сегментації з використанням CNN була *латкова класифікація* [9]. За допомогою цього

методу класифікують кожен піксель зображення після обробки згортковими шарами певної частини зображення (латки), а потім рішення про віднесення до деякого класу приймається з використанням повнозв'язних шарів. У зв'язку зі збільшенням часу роботи та вимог до обчислювальних машин такий метод не набув поширення.

Зауважимо, що спочатку спроби використати CNN для задач сегментації були невдалими через застосування повнозв'язних та пулінгових шарів. Це призводило до обмежень на розмір вхідного зображення, бо сегментувати можна зображення однакового розміру. Застосування пулінгових шарів дозволяє брати більше ознак з більшої області зображення й об'єднувати їх, втративши інформацію про розташування ознак. Ці недоліки вперше усунув Джонатан Лонг (Jonathan Z. Long) у роботі [10].

Більшість мереж для сегментації формують шляхом поєднання двох складових - кодувальника та декодувальника: на першому етапі кодувальником визначається об'єкт на зображенні, а потім декодувальником - місце знаходження об'єкта.

Критерії визначення якості сегментації зображення

На сьогодні якість сегментації для більшості практичних завдань визначається як міра близькості двох зображень: сегментованого деяким експертом та сегментованого відповідними методами. Якість сегментації зображення на практиці визначається як на суб'єктивному (якісному) [11], так і на об'єктивному (кількісному) рівні.

До суб'єктивних критеріїв відносяться критерії візуального сприйняття, що отримуються під час експертизи певною групою експертів. Найпоширенішим є метод оцінок, який дозволяє оцінити такі

характеристики зображення; правильність виділення кольорів, контурів, областей тощо

До об'єктивних критеріїв належать критерії, отримані внаслідок деякого порівняння кількісних ознак сегментованого певним методом зображення та зображення, сегментованого експертом (так званого еталонного). Критерії такого типу часто використовуються в системах автоматизованого аналізу зображень, для визначення кількісного еквівалента якості сегментації [12]. До найпоширеніших критеріїв, за якими кількісно можна оцінити якість сегментації, належать метрики [13]: intersection over union (IOU); accuracy; precision; recall (sensitivity); specificity; F1-score; dice coefficient; loss function.

Введемо позначення [14]: P (positive) — кількість позитивних відповідей (результатів сегментації); N (negative) — скільки негативних відповідей.

Тепер подивимось на результат роботи мережі. Усі позитивні відповіді (мережа сказала "так", тобто правильно сегментувала) поділяють на: TP (true positive) — мережа відповіла "так" і вгадала, FP (false positive) — мережа відповіла "так" і помилилася. Усі негативні відповіді подають аналогічно: TN (true negative) — мережа відповіла "ні" і це було правильно, FN (false negative) — мережа відповіла "ні" і це помилка.

Intersection over Union (IOU) є метрикою оцінки, що використовується для визначення достовірності розташування обмежувачої рамки і є відношення площ обмежувальних рамок:

$$IoU = \frac{S(A \cap B)}{S(A \cup B)}$$

де A та B — передбачена рамка, що обмежує, і справжня рамка, що обмежує відповідно. IoU дорівнює «0», якщо обмежувальні рамки не перетинаються, і «1» у випадку ідеального накладання (рис. 1). При сегментації зображення область

може мати будь-яку форму (правильну чи неправильну) і не обов'язково має бути прямокутною. Це означає, що прогнози – це не обмежувальні рамки, а маски сегментації.

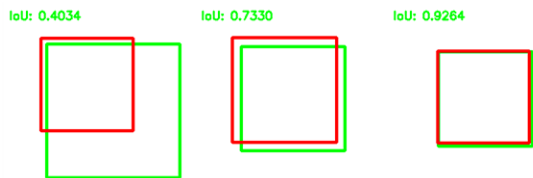


Рис.1. Вплив взаємного розміщення рамок на метрику IoU

У цьому випадку виконується попіксельний аналіз:

$$IoU = \frac{TP}{TP+FP+FN}$$

Для багатокласової сегментації середнє значення IoU зображення обчислюється шляхом обчислення IoU кожного класу та усереднення:

$$mIoU = \frac{1}{N} \sum_{i=1}^N IoU_i$$

Accuracy - це відношення правильно класифікованих елементів до всіх доступних елементів, обчислюється так:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

Precision - позитивне прогностичне значення (PPV) - це відношення між істинними позитивними та всіма елементами, класифікованими як позитивні:

$$Precision = \frac{TP}{TP+FP}$$

Recall (sensitivity) - справжнє позитивне значення (TPV) - це відношення справжніх позитивних до всіх позитивних елементів:

$$Recall = \frac{TP}{P} = \frac{TP}{TP+FN}$$

Specificity (selectivity or true negative rate - TNR) - це відношення справжніх негативних до всіх негативних елементів:

$$Specificity = \frac{TN}{N} = \frac{TN}{TN+FP}$$

Dice coefficient використовується для вимірювання схожості двох зображень, відображає подібність між передбачуваним

(сегментованим) зображенням та відповідним оригінальним зображенням. *Dice* обчислюють наступним чином:

$$Dice = \frac{2 * TP}{(TP + FP) + (TP + FN)}$$

IoU, *accuracy*, *precision*, *sensitivity*, *specificity*, *dice* $\in [0, 1]$, чим значення ближче до «1», тим краще.

F1-score - це комбінація Recall и Precision (середнє гармонійне):

$$F1 - score = 2 * \frac{Recall * Precision}{Recall + Precision} = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}}$$

Максимальнє значення F1-score ми отримаємо, якщо Recall та Precision достатньо далекі від нуля. За допомогою цієї метрики можна знайти компромісне рішення, фактично, між максимізацією TP за різними шкалами.

Loss function (функція втрат) використовується для розрахунку помилок між реальними й отриманими відповідями. Глобальною метою є мінімізація цієї помилки. Функція втрат одновимірна і може залежати від таких змінних, як вага та зсув. Вона оцінює, наскільки добре працює НМ. Серед відомих функцій втрат можна виділити такі: 1) квадратична (середньо-квадратичне відхилення); 2) кросс-ентропія; 3) експоненціальна (AdaBoost); 4) відстань Кульбака — Лейблера або приріст інформації.

Loss function $\in [0, \infty]$, чим значення ближче до «0», тим краще.

Тут TP, TN, FP та FN відповідають кількості правильних позитивних, правильних негативних, помилкових позитивних та помилкових негативних відповідей відповідно.

Порівняння методів сегментації на основі НМ. До найпоширеніших НМ, які можуть виконувати сегментацію зображень, належать FCN, SegNet, U-Net, FC-Densenet, E-Net, Link-Net, RefineNet, PSPNet. Для вирішення задачі семантичної сегментації, а саме сегментації зображень судин сітківки

ока потрібно обрати один з перелічених методів, який найкраще підходить для сегментації медичних зображень, тобто враховує специфіку вхідних даних, а також наявну кількість зображень для тренування, розмір зображень, обмеження на час навчання, наявні ресурси. Кожен з цих методів має свої особливості та переваги, які можна застосувати для різних задач сегментації зображень.

Для вибору найоптимальнішого методу потрібно враховувати розмір зображень, їхню кількість, сферу застосування, обчислювальні можливості персонального комп'ютера тощо. У табл. 1 подано порівняння НМ, призначених для семантичної сегментації зображень, за точністю (показник mIoU), перевагами та обмеженнями. З табл. 1 можна зробити висновок, що враховуючи точність, найкраще для сегментації медичних зображень підходять мережі *DeepLab*, *PSPNet*, *U-Net*.

Варто зауважити, що у медичній галузі дані пацієнтів є конфіденційними, вони захищені законами, через це пошук потрібних зображень, а також їх наявність у достатній кількості викликає певні проблеми. Тому однією з умов вибору методу глибокого навчання буде можливість працювати з базою зображень невеликого розміру.

Розглянуто метод умовного випадкового поля (conditional random field - CRF), який використовується як метод постобробки, що дозволяє покращити вихідні результати роботи згорткової НМ.

Умовне випадкове поле. Умовне випадкове поле — це метод дискримінаційного статистичного моделювання, який використовується, коли мітки класів для різних вхідних даних не є незалежними. Наприклад, у сегментації зображення мітка класу для пікселя також залежить від мітки його сусідніх пікселів.

Відмінністю умовного випадкового поля від дискретного класифікатора (що здійснює передбачення мітки окремого зразка без урахування «сусідніх»), є можливість враховувати контекст.

Таблиця 1

Порівняння нейромережових методів сегментації

Мережа	mIoU	Переваги	Обмеження
FCN-8	~ 57	Швидкодія, хороша продуктивність	Невисока точність сегментації, необхідність адаптації для сегментації
SegNet	~ 56.1	Швидкодія, висока продуктивність	Невисока точність сегментації, необхідність адаптації для сегментації
U-Net	~ 77.5	Гарні результати сегментації на невеликій кількості даних, швидкодія, висока точність	Невисока точність сегментації для деяких задач
FC-Dence Net	~ 66.9	Висока точність та якість карток сегментації на виході	Недостатня точність сегментації для деяких задач
Refine Net	~ 73.6	Висока точність	Невисока швидкодія
PSPNet	~ 81.2	Висока точність	Невисока швидкодія
Link-Net	~ 76.4	Працює в режимі реального часу	Потрібен потужний процесор
ENet	~ 58.3	Працює в режимі реального часу	Потрібен потужний процесор
Deep Labv3	~ 81.3	Висока точність	Невисока швидкодія

CRF складається з набору пікселів зображення, яким присвоєно мітку/значення на основі міток/значення їхніх прямих сусідів. CRF часто застосовуються для зменшення шуму та виділення границь об'єктів на сильно зашумлених картах сегментації.

Використання умовних випадкових полів в поєднанні з нейромережевими методами сприяє вирішенню проблем зашумлення.

Завдяки застосуванню методу CRF до вихідних даних глибокої згорткової НМ карта сегментації стає більш згладженою та отримує чіткішу межу між сегментами зображення. CRF часто розглядається як метод постобробки, що дозволяє ітеративно покращити вихідні результати роботи згорткової НМ.

Отже, поєднання методів CRF та архітектур глибокого навчання може дати перспективні результати для сегментації зображень судин сітківки ока.

Методи аугментації. У процесі навчання моделі сегментації, якщо вхідних даних недостатня кількість, використовуються різні стратегії збільшення їх кількості. Таке збільшення вибірки даних для навчання нейромережі через модифікацію існуючих вхідних даних називають аугментацією. Потім отримані зображення додаються до набору даних у процесі тренування [15].

Серед *методів аугментації* найпоширенішими є такі: віддзеркалення, гаусовий шум, тремтіння, масштабування, розмиття за Гаусом, обертання, афінна трансформація, випадкова гамма-корекція.

3.3. Аналіз останніх досліджень (приклад застосування методів розв'язання задач сегментації зображень судин сітківки ока)

Ознайомитися із прикладами застосування класичних методів розв'язання задач сегментації кровоносних судин, які

включають ручні фільтри, зокрема детектори лінії, можна у [0, 07], а методи розширення судин – у роботах [08, 09, 20].

Для автоматизованого аналізу зображень сітківки за останні десятиліття було запропоновано багато підходів та методів, наприклад:

- у роботі [0] описано метод сегментації кровоносних судин на зображеннях очного дна на основі дискримінаційно навченої повністю пов'язаної моделі на основі CRF, у якій комбінуються різні характеристики судин, а зображення сегментують за допомогою повністю пов'язаних умовних випадкових полів;

- у роботі [02] пропонується структура посилення градієнта для навчання фільтрів у керований спосіб;

- у роботах [0, 04] розроблені алгоритми, здатні покращувати тонкі структури за умови регресії судин сітківки;

- в [Ошибка! Источник ссылки не найден.] запропоновано метод суперпіксельної класифікації.

За останні роки нейромережеві методи глибокого навчання зробили надзвичайний прогрес у сфері комп'ютерного зору. Архітектура CNN спочатку була розроблена для задач природної класифікації зображень [0], і останні дослідження призвели до вражаючих результатів у вирішенні подібних завдань. CNN були використані для сегментації судин сітківки в [0] для класифікації особливостей плям.

Для сегментації судин сітківки з'явилася велика кількість вдосконалених алгоритмів, заснованих на U-Net. Наприклад,

- у роботі [0] пропонують U-Net-подібну модель із механізмом зваженої ваги та схемою пропуску з'єднань для вирішення проблем сегментації, пов'язаних із малими тонкими судинами, низькою дискримінаційною здатністю в області диска зорового нерва тощо;

- у роботі [30] поєднують фільтрування за Гаусом та U-Net з метою сегментації зображень судин сітківки;

- у роботі [**Ошибка! Источник ссылки не найден.**] сегментацію зображень судин сітківки здійснено за допомогою методів глибокого навчання в умовах невизначеності з використанням мінімаксного підходу;

- робота [**Ошибка! Источник ссылки не найден.**] присвячена вирішенню проблеми зашумлених міток при сегментації судин сітківки на основі НМ.

3.4. Практична частина

Розроблено програмний застосунок (ПЗ) для сегментації зображень судин сітківки ока. Для створення ПЗ використано мову програмування Python, архітектуру U-net та бібліотеку Keras.

U-net. Повнозгорткова нейромережа U-Net розроблена у 2015 р. у відділенні Computer Science Фрайбургського університету і є підвидом згорткових НМ, призначених для потреб сегментації зображень, насамперед медичних [**Ошибка! Источник ссылки не найден.**9]. Цей вид НМ відрізняється від відомих традиційних мереж тим, що просторова інформація (тобто розміщення у просторі кожного елемента згортки) враховується для кожного шару. Це можливе завдяки тому, що у мережі відсутні повнозв'язні шари. Такий підхід дозволяє також працювати із зображеннями будь-якого розміру, що неможливе для традиційних мереж через прямий зв'язок розміру зображення і кількості вхідних нейронів першого повнозв'язного шару. Нейромережі подібного типу можуть навчатися на невеликій кількості навчальних зображень та мають високу швидкості роботи.

Мережа U-Net побудована за принципом «кодувальник-декодувальник». Блок кодування є типовою згортковою НМ. Він складається з двох згорткових шарів з

вікном згортки 3×3 (3 пікселі на 3 пікселі), за кожним з яких розміщується шар ReLU та пулінговий шар з вікном 2×2 з функцією максимуму та кроком 2, після чого розмір зображення зменшується вдвічі. На кожному із кроків блоку кодування кількість каналів ознак подвоюється і після проходження зменшувального шляху зображення зменшується до розміру 32×32 з 512 каналами ознак. Кожен крок блоку декодування передбачає збільшення карти ознак, після чого здійснюється 1) згортка 2×2 , яка зменшує кількість каналів ознак удвічі, 2) конкатенація (об'єднання) зі зменшеною картою ознак, отриманою на відповідному кроці блоку кодування, та двома згортками 3×3 з шарами ReLU.

Об'єднання карт ознак симетричних блоків кодувальника та декодувальника необхідне для відновлення дрібнозернистої інформації.

На останньому кроці роботи мережі згортка 1×1 перетворює вектор ознак (складається з 64 елементів) у потрібну кількість класів. Всього в архітектурі U-Net використовується 23 шари згортки.

Інструменти для реалізації програмного застосунку. На сьогодні існує багато алгоритмів оптимізації для навчання нейромереж. Серед найвідоміших – алгоритм зворотного поширення похибки, який застосовується для мінімізації відхилень поточного значення вихідного сигналу НМ від потрібного. Недоліками цього алгоритму є можливість перенавчання у випадку вибору малої швидкості навчання або велика розбіжність при надто великій швидкості. Для усунення цих недоліків існують різні модифікації алгоритму, основна відмінність яких від класичного алгоритму полягає в тому, що для них використовується адаптивна швидкість навчання для кожної НМ, яка адаптується під час навчання [33].

У роботі використовується алгоритм

Adam, який показує кращі результати в порівнянні з іншими алгоритмами: кращу точність класифікації; достатньо простий для реалізації; ефективний у застосуванні; інваріантний до зміни масштабу градієнта по діагоналі; ефективний для завдань з великою кількістю параметрів.

Структура ПЗ. Під час виконання роботи було розроблено дві підсистеми: для навчання нейромережі (рис. 3) та для її використання у задачах сегментації зображень судин сітківки ока (рис. 4). Для навчання НМ модуль завантаження завантажує з файлового сховища зображення сітківки ока та відповідну маску для цього зображення. Після чого для зображення та маски змінюється розмір до заданого (512x512), далі вони подаються на вхід модуля аугментації, у якому виконуються однакові модифікації зображення та відповідної маски.

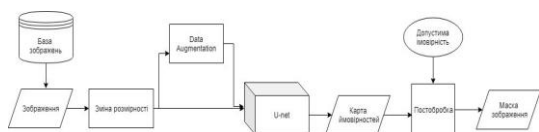


Рис. 3. Схема навчання моделі

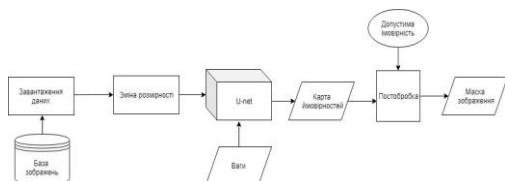


Рис. 4. Схема роботи системи в режимі передбачень

Після аугментації зображення обробляється мережею на базі архітектури U-net. Нейронна мережа генерує маску, яка має вигляд карти ймовірностей. Згенерована маска порівнюється з початковою вхідною маскою, обчислюються метрики, після чого нейромережа корегує вагу.

Робота підсистеми для передбачень (рис. 4) відрізняється від роботи підсистеми для навчання НМ.

Метою роботи підсистеми передбачень

є генерація маски (сегментоване зображення), тому на вхід подається лише зображення сітківки ока. Аугментація в цій підсистемі не проводиться.

Результатом роботи мережі є карта ймовірностей належності пікселя поданого вхідного зображення до шуканого сегменту. Після чого отримуємо сегментоване зображення.

Для покращення отриманих результатів початкової моделі сегментації, на етапі постобробки використано умовні випадкові поля, що дозволяє позбутися шумів і покращити зовнішній вигляд вихідного зображення.

Вхідні зображення. У роботі використано загальнодоступний набір даних DRIVE, який містить 40 зображень судин сітківки ока разом з відповідними масками. Перед використанням розмір зображень потрібно змінити за допомогою застосунку до потрібного. На вхід розробленої моделі повинні подаватися зображення однакового розміру. Крім того, час навчання НМ залежить від розміру зображень, оскільки при роботі із зображеннями великого розміру будуть виникати проблеми, пов'язані з нестачею оперативної пам'яті.

Аугментація. Застосування аугментації, тобто стратегії збільшення даних, стало необхідним при використанні нейромережі U-Net. Варто зауважити, що для практичних задач сегментації судин сітківки не дуже багато даних для навчання, що пов'язано насамперед із конфіденційністю особистої інформації кожного пацієнта. Застосування аугментації дозволяє збільшити кількість вхідних даних (розмір датасету), а це надає НМ нову додаткову інформацію для навчання. Крім того, застосування аугментації дозволяє компенсувати втрати точності внаслідок якості зображення та його розмірів. Також збільшення кількості даних допоможе нейромережі U-Net при використанні

операції об'єднання (pooling), оскільки модель погано опрацьовує інформацію на краях вхідного зображення.

У роботі застосовано такі методи аугментації: 1) віддзеркалення: вико-нується горизонтальне та вертикальне відображення для кожного зображення та його маски з базового навчального набору; 2) еластичне перетворення: зображення та маска були еластично перетворені з використанням аргументів $\alpha=120$, $\sigma=6$ та $\alpha_affine=3.6$; 3) розмиття сіткою (GridDistortion): зображення та маска були розмиті за допомогою розміщення пікселів на сітці; 4) оптичне спотворення (Optical Distortion): зображення та маска були змінені за допомогою оптичного спотворення. У задачах сегментації зображення та його маска повинні відповідати одне одному, тому методи аугментації застосовуються і для зображення, і для його маски.

Процес навчання моделі. Основними кроками тренування моделі є такі: підготовка даних, побудова моделі, налаштування зворотних викликів, навчання моделі. Підготовка даних передбачає завантаження тренувального та валідаційного наборів зображень і еталонних масок, далі виконується аугментація зображень тренувального набору для кращого розподілу даних.

При побудові моделі створюється модель UNet з архітектурою, яка складається з енкодера та декодера. Модель компілюється з використанням функції втрат, оптимізатора Adam із заданою швидкістю навчання та метриками.

Під час тренування модель буде навчатися, оновлюючи значення ваги, та зберігати найкращу модель. Ця модель буде мати параметри, які надали найкращий результат на тестовому наборі зображень (рис. 5).

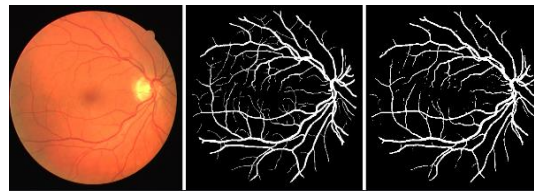


Рис. 5. Оригінальне зображення, маска, результат сегментації

Також після навчання буде сформований файл, у якому міститься історія тренування, включаючи значення втрат та метрик на кожній епосі тренування. Ці результати допомагають оцінити продуктивність моделі, її здатність до узагальнення та ефективність тренування. Можна проаналізувати журнали та результати метрик, а також використати збережену найкращу модель для подальших застосувань, таких як тестування на нових даних.

Серед основних можливостей програмного застосунку можна виділити: завантаження, обробка перегляд та сегментація зображень, обчислення метрик точності. Користувачі отримують результати сегментації у зручному форматі (зображення із розширенням .png). Якщо для зображення існує еталонна маска, то отримаємо результат сегментації (рис. 6) та будуть автоматично обчислені метрики, такі як точність (accuracy), чутливість (recall), специфічність (specificity) і коефіцієнт Серенсена (Dice coefficient). Ці метрики дозволяють оцінити якість сегментації, порівнюючи результати з еталонними масками.

У випадку відсутності еталонної маски для завантаженого зображення видається відповідне повідомлення про те, що метрики обчислені не будуть (рис. 7).

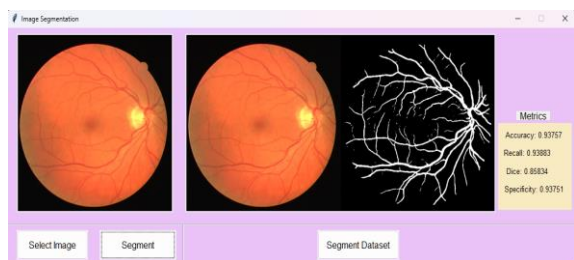


Рис. 6. Сегментація зображення



Рис. 7. Результат сегментації для зображення без маски (метрики не обчислені)

У такому випадку для зображення без маски сегментація відбувається без обчислення метрик (рис. 7). Також у ПЗ є можливість сегментувати весь датасет. Для цього потрібно натиснути кнопку «Segment Dataset» та обрати датасет. Після цього починається процес сегментації набору зображень.

IV. РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Порівняння застосованого методу з аналогами. Застосований метод оцінювався на публічному наборі даних DRIVE, який надає еталонні сегментації (маски) для кожного зображення. Це надає змогу обчислити оцінки продуктивності. У роботі оцінки проводяться за метриками accuracy, precision, specificity, recall (sensitivity) та dice. У табл. 2 подано результати порівняння застосувань різних методів сегментації зображень судин сітківки на наборі даних навчання DRIVE, результати взято з робіт [Ошибка! Источник ссылки не найден., Ошибка! Источник ссылки не найден., Ошибка! Источник ссылки не найден.-44].

У останньому рядку таблиці - показники якості, отримані під час виконання роботи

розробленого ПЗ.

Таблиця 2
Порівняння різних методів сегментації зображень судин сітківки

Methods	Accuracy	Sensitivity	Specificity	Dice
Human observer	0.947	0.7761	-	-
Staal	0.944	0.7194	-	-
Mendonca	0.945	0.7344	-	-
Marin	0.945	0.7067	-	-
Fraz	0.948	0.7406	-	-
Nguyen	0.941	0.7429	-	-
Zhao	0.948	0.7354	-	-
Melinscak	0.947	0.7276	-	-
Azzopardi	0.944	0.7655	-	-
Roychowdhury	0.949	0.7395	-	-
HED	0.944	0.7364	-	-
DeepVessel w/o S	0.934	0.7164	-	-
DeepVessel 2	0.95	0.7603	-	-
LadderNet	0.956	0.7856	0.981	0.82
IterNet	0.957	0.7791	0.983	0.82
Hrychaniuk, Nosovets	0.969	0.8341	0.983	0.83
Martsenyuk	-	0.7668	0.988	0.87
ПЗ	0.945	0.899	0.947	0.83

V. ОБГОВОРЕННЯ

Отже, у роботі оцінювання проводилося за метриками accuracy, precision, specificity, sensitivity та dice.

Результатом дослідження є розробка ПЗ, який надає такі оцінки: accuracy=0.9452, sensitivity = 0.8991, specificity= 0.9468, dice= 0.8247. Показники якості отримано під час виконання роботи розробленого ПЗ. Як видно з табл. 2, у деяких випадках метрики для застосованого у роботі методу кращі за метрики в аналогічних методах, а в деяких - гірші.

Отримані результати свідчить про хорошу роботу мережі на нових даних, і

можна говорити про конкурентоспроможність розробленого програмного застосунку на основі методів глибокого навчання та умовного випадкового поля.

V. ВИСНОВКИ

В результаті виконання даного дослідження був реалізований програмний застосунок на основі бібліотек Keras та Tensorflow з використанням мови Python. Для покращення якості сегментації застосовано метод умовного випадкового поля CRF, що використовує інформацію про сусідні пікселі для визначення меж судин. Його протестовано на загальнодоступному наборі даних DRIVE, що складається з 40 зображень, для половини з яких застосовано аугментацію.

Наведено метрики, отримані у результаті навчання мережі, а також метрики, отримані після етапу прогнозування, тобто роботи підсистеми передбачення. Для оцінювання якості роботи розробленого методу використовувались метрики accuracy, precision, specificity, recall (sensitivity) та dice. Виконано порівняння застосувань різних методів сегментації зображень судин сітківки. Отримані результати пояснюються хорошою роботою нейромережі в режимі передбачень та свідчать про конкурентоспроможність обраного методу. Точність роботи розробленого програмного застосунку можна покращити шляхом навчання нейромережі на більшій кількості навчальних даних.

Таким чином, у результаті поєднання методів глибокого навчання та CRF можна отримати конкурентоспроможну модель для сегментації зображень судин сітківки ока. Маємо, такі метрики: accuracy=0.9452, sensitivity = 0.8991, specificity= 0.9468, dice= 0.8247.

Оптимальний варіант реалізації ПЗ дає користувачу високу точність сегментації,

прийнятний час навчання мережі та непоганий функціонал і передбачає використання мови програмування Python, бібліотеки Keras та архітектури нейронної мережі U-Net.

Фінансування. Дане дослідження не отримувало зовнішнього фінансування.

Конфлікт інтересів. Автори заявляють про відсутність конфлікту інтересів.

Згода на публікацію. У роботі використано загальнодоступний набір даних DRIVE, який містить 40 зображень судин сітківки ока разом з відповідними масками. Усі пацієнти, що мають відношення до рукопису дали згоду на публікацію даної роботи.

ORCID ID та внесок авторів.

- 0000-0002-4055-6834 (D, E, F)
Ludmila Dobrovska
- 0009-0007-4395-8672 (A, B, C, D)
Yaroslav Nazaraha

ПЕРЕЛІК ПОСИЛАНЬ

- [1] Linda G. Shapiro and George C. Stockman. Computer Vision. New Jersey: Prentice-Hall, 2001. - 608 p. (pp. 279–325).
- [2] Haralick R.M, Shapiro L.G. Image Segmentation Techniques. *Computer Vision, Graphics and Image Processing*. 1985. Vol. 29. - P. 100-132.
- [3] Fu K., Mui J. A survey on image segmentation. *Pattern Recognition*. 1981. Vol. 13. - P. 3 - 16.
- [4] Pal N., Pal S. A survey on image segmentation techniques. *Pattern Recognition*. 1993. Vol. 26. - P. 1277 - 1294.
- [5] Skarbek W., Koschan A. Color Image Segmentation - A Survey. *Technischer Bericht: Technical University of Berlin*. 1994. - P. 94 - 32.
- [6] Lucchese L., Mitra S. Color Image Segmentation: A State-of-the-Art Survey, Image Processing, Vision, and Pattern Recognition. Proc. of the Indian National Science Academy (INSA-A). New Delhi, India: CRC, 2001. P. 207 - 221.
- [7] Zhang Y. Advances in Image And Video Segmentation. USA: IRM Press, 2006. - 457 p.
- [8] Березький О.М., Мельник Г.М., Батько Ю.М. Методи сегментації біомедичних зображень / Вісник Хмельницького національного університету. Технічні науки. 2010. №1. - С. 251 - 258.
- [9] Le Hou, Dimitris Samaras, Tahsin M. Kurc. Patch-Based Convolutional Neural Network for Whole Slide Tissue Image Classification. *Proc. IEEE Conference on Computer Vision and Pattern Recognition*. 2016. - P. 2424 – 2433. DOI: 10.1109/CVPR.2016.266/

- [10] Jonathan Long, Evan Shelhamer, Trevor Darrell. Fully Convolutional Networks for Semantic Segmentation. *Computer Vision and Pattern Recognition*. 2015. - DOI: 10.48550/arXiv.1411.4038/
- [11] Батько. Ю.М. Метод і алгоритми сегментації біомедичних зображень на основі попередніх розміток / Штучний інтелект. 2010. N 4.
- [12] Глушков Г.А. Комп'ютерна графіка. - Харків: Фоліо, 2000. - 500 с.
- [13] David Martin Ward Powers. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*. 2011. Vol. 2 (1). P. 37 – 63. DOI: 10.9735/2229-3981/
- [14] Метрики в машинному навчанні: precision, recall і не тільки. Веб-сайт. URL: <http://www.michurin.net/computer-science/ml-precision-recall.html/>
- [15] Гричанюк І.О., Носовець О.К. Аналіз методів аугментації даних для задач сегментації судин сітківки / *Young Scientist*. № 10 (98). 2021. DOI: 10.32839/2304-5809/2021-10-98-23/
- [16] Nguyen U.T., Bhuiyan A., Park L.A., Ramamohanarao K. An effective retinal blood vessel segmentation method using multi-scale line detection. *Pattern Recognition*. 2013. Vol. 46. Issue 3. P.703 – 715. DOI: 10.1016/j.patcog.2012.08.009/
- [17] Ricci, E., Perfetti, R. Retinal blood vessel segmentation using line operators and support vector classification. *IEEE Transactions On Medical Imaging*. 2007. Vol. 26. N 10. P. 1357 – 1365. DOI: 10.1109/TMI.2007.898551/
- [18] Fraz M.M., Remagnino P., Hoppe A., Uyyanonvara B., Rudnicka A.R., Owen C.G., Barman S.A. Blood vessel segmentation methodologies in retinal images - A survey. *Computer Methods and Programs in Biomedicine*. 2012. Vol.108, Issue 1, P. 407 – 433. DOI: 10.1016/j.cmpb.2012.03.009/
- [19] Soares J.V., Leandro J.J., Cesar R.M., Jelinek H.F., Cree M.J. Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification. *IEEE Transactions On Medical Imaging*. 2006. Vol. 25(9). P. 1214 – 1222. DOI: 10.1109/tmi.2006.879967/
- [20] Zhang B., Zhang L., Zhang L., Karray, F. Retinal vessel extraction by matched filter with first-order derivative of gaussian. *Comput. Biol. Med.* 2010. Vol. 40(4). P. 438 – 445. DOI: 10.1016/j.compbiomed.2010.02.008/
- [21] Orlando J., Blaschko M. Learning fully-connected CRFs for blood vessel segmentation in retinal images. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. 2014. Vol. 8673. P. 634 – 641. DOI:10.1007/978-3-319-10404-1_79/
- [22] Becker C., Rigamonti R., Lepetit V., Fua P. Supervised feature learning for curvilinear structure segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. 2013. Vol. 8149. P. 526 – 533. DOI: 10.1007/978-3-642-40811-3_66/
- [23] Gu L., Cheng L. Learning to boost filamentary structure segmentation. *IEEE International Conference on Computer Vision (ICCV)*: Proc. of the Conf., 2015. P. 639 – 647. URL: https://www.ece.ualberta.ca/~lcheng5/papers/GuChe_ICCV15.pdf/
- [24] Sironi A., Lepetit V., Fua P. Projection onto the manifold of elongated structures for accurate extraction. *IEEE International Conference on Computer Vision (ICCV)*: Proc. of the Conf., 2015. P. 316 – 324. URL: https://openaccess.thecvf.com/content_iccv_2015/html/Sironi_Projection_onto_the_ICCV_2015_paper.html/
- [25] Cheng J. et al. Superpixel Classification Based Optic Cup Segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. 2013. Vol. 8151. P. 421–428. DOI: 10.1007/978-3-642-40760-4_53/
- [26] Krizhevsky A., Sutskever I., Hinton G.E. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems 25 (NIPS)*. 2012. URL: https://papers.nips.cc/paper_files/paper/2012
- [27] Ganin Y., Lempitsky V. N4-fields: neural network nearest neighbor fields for image transforms. *ACCV*. 2014. DOI: 10.48550/arXiv.1406.6558/
- [28] Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. 2015. Vol 9351. P. 234 – 241. DOI: 10.1007/978-3-319-24574-4_28
- [29] Xiao X., Lian S., Luo Z., Weighted Res-UNet for highquality retina vessel segmentation. *9th Int. Conf. Inf. Technol. Med. Educ. (ITME)*: Proc. Of Conf., 2018. URL: <https://ieeexplore.ieee.org/abstract/document/8589312/>
- [30] Gao X., Cai Y., Qiu C. Retinal blood vessel segmentation based on the Gaussian matched filter and U-net. *CISP-BMEI*, 2017. URL: <https://ieeexplore.ieee.org/abstract/document/8302199>.
- [31] Martsenyuk V., Milian N., Milian R. The Minimax Approach Application For Segmentation Of Retinal Vessels. *Innovative Solutions in Modern Science*. 2020. Vol. 7(43). DOI: 10.26886/2414-634X.7(43)2020.13
- [32] Гричанюк І.О., Носовець О.К. Навчання нейронної мережі для сегментації судин сітківки за допомогою зашумлених міток / *Розвиток наукових міжгалузевих досліджень*: зб. матер. конф. 26-27 лист. 2021р. Вінниця, 2021. - с.73 - 76.
- [33] Добровська Л. М. Теорія та практика нейронних мереж: навч. посіб. /Л. М. Добровська, І. А. Добровська. – Київ: КПІ ім. Ігоря Сікорського (Вид-во «Політехніка»), 2015. – 396 с. – URL: <https://ela.kpi.ua/handle/123456789/49841>
- [34] Staal J., Abramoff M. and all. Ridge-based vessel segmentation in color images of the retina. *IEEE Trans. Med. Imaging*. 2004. Vol. 23 (4). P. 501 – 509. DOI:10.1109/TMI.2004. 825627/
- [35] Mendonca A., Campilho A. Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. *IEEE Trans. Med. Imaging*. 2006. Vol. 25 (9), P. 1200 – 1213. DOI: 10.1109/tmi.2006.879955/
- [36] Marin D, Aquino A., Gegundez M. E., Bravo J. A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features. *IEEE Trans. Med. Imaging*. 2011. Vol. 30(1). P. 146 – 158. DOI: 10.1109/TMI.2010.2064333
- [37] Fraz M., Remagnino P., Hoppe A., Uyyanonvara B., Rudnicka A., Owen C., Barman S. An ensemble classification-based approach applied to retinal blood vessel segmentation. *IEEE Trans. Biomed. Eng.* 2012. Vol. 59(9), P. 2538 – 2548.

DOI: 10.1109/TBME.2012.2205687/

[38] Nguyen U., Bhuiyan A., Park L., Ramamohanarao K. An effective retinal blood vessel segmentation method using multi-scale line detection. *Pattern Recogn.* 2013. Vol. 46 (3). P. 703–715. DOI: 10.1016/j.patcog.2012.08.009

[39] Zhao Y., Wang X., Wang X., Shih F. Retinal vessels segmentation based on level set and region growing. *Pattern Recogn.* 2014. Vol. 47 (7), P. 2437–2446. DOI: 10.1016/j.patcog.2014.01.006

[40] Melinscak M., Prentasac P., Loncaric S. Retinal vessel segmentation using deep neural networks. *International Conference on Computer Vision Theory and Applications: Proc. of Conf.*, 2015. P. 577–582. DOI: 10.5220/0005313005770582/

[41] Azzopardi G., Strisciuglio N., Vento M., Petkov N. Trainable COSFIRE filters for vessel delineation with

application to retinal images. *Med. Image Anal.* 2015. Vol. 19 (1), P. 46–57. DOI: 10.1016/j.media.2014.08.002/

[42] Roychowdhury S., Koozekanani D., Parhi K.: Iterative vessel segmentation of fundus images. *IEEE Trans. Biomed. Eng.* 2015. Vol. 62 (7), P. 1738–1749. DOI: 10.1109/TBME.2015.2403295/

[43] Xie S., Tu Z. Holistically-nested edge detection. *Int. Conf. on Computer Vision (ICCV 2015): Proc. of Conf.*, 2015. P. 1395–1403. URL: https://openaccess.thecvf.com/content_iccv_2015/html/Xie_Holistically-Nested_Edge_Detection_ICCV_2015_paper.html/

[44] Fu H., Xu Y., Lin S., Kee Wong D.W., Liu J. DeepVessel: Retinal Vessel Segmentation via Deep Learning and Conditional Random Field. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. 2016. Vol. 9901. DOI: 10.1007/978-3-319-46723-8_16/

УДК 004.852 + 616-018

A CONVOLUTIONAL NEURAL NETWORK FOR SEGMENTATION OF RETINA VESSELS

Ludmila Dobrovska

Yaroslav Nazaraha

nazaraha.yaroslav@lil.kpi.ua

National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”,
Kyiv, Ukraine

Abstract – Research, monitoring and assessment of retinal vessels play an important role in the diagnosis of various eye diseases. Identification of specific pathological objects in the images comes down to solving segmentation problems. Segmentation of retinal vessels is a key step towards accurate imaging, diagnosis of eye diseases, early treatment and surgery planning. This is the reason why the automated segmentation of retinal vessels is an important task.

The purpose of this work was to develop a software application for segmentation of images of retinal vessels using machine learning in the form of a convolutional neural network.

The neural network segmentation methods, namely methods based on deep learning, are the most accurate at the moment.

The image database used for the study was taken from the public dataset DRIVE, which provides reference segmentations (masks) for each image and for the half of these segmentations the augmentation is used. This enables to calculate the estimates of model performance.

During the first stage of research, it was established that currently the most common are:

- 1) criteria by which the quality of segmentation can be quantitatively assessed include the following metrics: intersection over union; accuracy; precision; sensitivity; specificity; F1-score; dice coefficient; loss function;
- 2) networks capable of performing image segmentation include the following: FCN, SegNet, U-Net, FC-Densenet, E-Net, Link-Net, RefineNet, PSPNet.

To solve the problem of retinal vessel image segmentation, a comparison of the indicated networks was performed by the accuracy, advantages and limitations. It was established that taking into account the accuracy, the DeepLab, PSPNet, U-Net networks are best suited for solving the described problem.

The second stage of the research consisted in the development of a software application, comparison and evaluation of quality indicators of known retinal vessel segmentation systems and the developed software.

The result of the research is the development of software that provides the following metrics: accuracy=0.9452, sensitivity=0.8991, specificity=0.9468, dice=0.8247. These quality indicators are obtained during the execution the developed software.

Keywords - artificial intelligence, neural network, segmentation of retinal vessels, image processing.