

**НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
ІНСТИТУТ ПРОБЛЕМ МАТЕМАТИЧНИХ МАШИН І СИСТЕМ**

ШКУРАТ ОКСАНА СЕРГІЇВНА



УДК 004.932.2: 004.932.72: 004.93'1

**МЕТОДИ ТА ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ОБРОБЛЕННЯ АРХІВНИХ
МЕДИЧНИХ ЗОБРАЖЕНЬ**

05.13.06 – інформаційні технології

Автореферат дисертації на здобуття наукового ступеня кандидата технічних наук

Київ – 2020

Дисертацією є рукопис

Робота виконана в Національному технічному університеті України «Київський політехнічний інститут імені Ігоря Сікорського» Міністерства освіти і науки України

Науковий керівник кандидат технічних наук, доцент
Сулема Євгенія Станіславівна,
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»
Міністерства освіти і науки України,
доцент кафедри програмного забезпечення
комп'ютерних систем

Офіційні опоненти: доктор технічних наук, професор,
заслужений діяч науки і техніки України,
лауреат Державної премії України в галузі науки і техніки
Павленко Петро Миколайович,
Національний авіаційний університет
Міністерства освіти і науки України,
завідувач кафедри механіки

доктор технічних наук, професор
Філатова Ганна Євгенівна,
Національний технічний університет
«Харківський політехнічний інститут»
Міністерства освіти і науки України,
професор кафедри обчислювальної техніки
та програмування

Захист відбудеться "02" грудня 2020 р. о 14⁰⁰ годині на засіданні спеціалізованої вченої ради Д 26.204.01 в Інституті проблем математичних машин і систем НАН України за адресою: 03187, м. Київ, проспект Академіка Глушкова, 42.

З дисертацією можна ознайомитись у бібліотеці Інституту проблем математичних машин і систем НАН України за адресою: 03187, м. Київ, проспект Академіка Глушкова, 42.

Автореферат розісланий "26" жовтня 2020 р.

Вчений секретар
спеціалізованої вченої ради



М.Г. Ієвлєв

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Візуалізація даних у вигляді зображень є важливим етапом медичних досліджень. Медична наука та практика дозволили накопичити протягом десятків років архіви нецифрових та цифрових медичних зображень значного обсягу, які становлять інтерес, що полягає у наступному: (1) автоматизований аналіз індивідуальних архівів медичних зображень для моніторингу стану здоров'я населення упродовж життя; (2) автоматизований аналіз загальних архівів медичних зображень лікувальних закладів для використання у медичних дослідженнях при пошуку специфічних випадків хвороб та аналізі динаміки їх розвитку з метою загального розвитку медичної теорії; (3) використання узагальнених архівів медичних зображень для навчання нейронних мереж при розробленні наступних поколінь медичних інформаційних систем. Можливості медичних діагностичних пристроїв та апаратно-комп'ютерних систем зумовлюють широкий діапазон видів зображень. Для їх інтерпретації, перетворення та дослідження застосовується спектр методів оброблення та аналізу.

Значний внесок у розвиток методів оброблення та аналізу зображень зробили зарубіжні та вітчизняні вчені: Gonzalez R. C., Woods R. E., Pratt W. K., Serra J., Duda R. O., Hart P. E., Путятін Є. П., Білинський Й. Й., Батько Ю. М. та ін. В галузі медичних інформаційних систем значний внесок зробили зарубіжні та вітчизняні вчені: Cimino J. J., Chen C. H., Амосов Н. М., Мінцер О. П., Коваленко О. С. та ін.

Сучасні методи отримання, аналізу, зберігання та передавання даних, в тому числі графічного характеру, зумовлюють створення медичних інформаційних систем, які дозволяють автоматизувати всі види медичної діяльності. Розрізняють адміністративні системи, експертні системи для діагностики, прогнозування та моніторингу, системи для лабораторних і діагностичних досліджень, навчальні системи тощо. Розвиток інформаційних технологій зумовлює створення медичних інформаційних систем, що інтегрують вищевказані системи. На сьогодні в Україні подібні інтегровані системи створюються та функціонують на базі одного медичного закладу або групи медичних закладів. Проте подібні системи можуть бути розширені до рівня загальнодержавної медичної інформаційно-комунікаційної системи. Долучення архівних медичних зображень до сучасних медичних знань приведе до розширення бази наукових досліджень, що сприятиме подальшому розвитку медичної теорії та практики. Водночас створення архівів медичних зображень, що включатимуть великі обсяги різнорідних графічних даних, вимагатиме нових підходів до їх оброблення. Тому розроблення та вдосконалення методів оброблення архівних медичних зображень є актуальною науковою задачею.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційне дослідження виконувалось на кафедрі програмного забезпечення комп'ютерних систем КПІ ім. Ігоря Сікорського відповідно до планів, програм і договорів науково-дослідної роботи «Розроблення та дослідження методів оброблення, розпізнавання, захисту та зберігання медичних зображень в розподілених комп'ютерних системах» (номер державної реєстрації 0117U004267) та міжнародного наукового проєкту «Future eHealth powered by 5G» (номер державної реєстрації 0118U100449).

Мета і завдання дослідження. Метою дисертаційної роботи є розроблення методів та інформаційної технології оброблення архівних медичних зображень для автоматизованого аналізу медичних даних у програмних та програмно-апаратних системах, що використовуються у галузі охорони здоров'я.

Досягнення зазначеної мети передбачає вирішення таких завдань:

- розроблення архітектури інформаційної технології оброблення архівних медичних даних;
- розроблення методу сегментації медичних зображень;
- розроблення методу контурного опису об'єктів медичних зображень;
- розроблення методу нечіткого пошуку об'єктів на медичних зображеннях;
- розроблення методу початкової класифікації цифрових та оцифрованих медичних зображень;
- удосконалення методу підвищення яскравості медичних зображень;
- удосконалення методу згладжування контурних кривих об'єктів зображень;
- програмна реалізація та дослідження ефективності розроблених методів оброблення зображень.

Об'єктом дослідження є процеси подання, аналізу, оброблення й пошуку медичних зображень та їх елементів.

Предметом дослідження є методи та технології оброблення медичних зображень у програмних і програмно-апаратних системах у галузі охорони здоров'я.

Методи дослідження. У теоретичному дослідженні використовуються положення теорії цифрового оброблення зображень. Опис цифрових та векторних зображень здійснюється методами лінійної алгебри та аналітичної геометрії. Для реалізації розроблених методів використовуються методи комп'ютерного моделювання.

Наукова новизна отриманих результатів визначається такими положеннями:

1. *Вперше* розроблено метод сегментації зображень, який, на відміну від існуючих, ґрунтується на аналізі статистичних показників вхідного масиву однорідних зображень, що дозволяє отримувати складові частини та контури зображень для масивів медичних зображень.

2. *Вперше* розроблено метод опису медичних зображень, який, на відміну від існуючих, ґрунтується на контурному представленні зображення та його елементів масивами координат «опорних» вершин, розмір яких задається залежно від вимог, які визначаються поставленою задачею пошуку та потрібною точністю відтворення, що дозволяє зменшити обсяг графічних даних, які зберігаються та оброблюються у медичній програмно-апаратній системі, у 2.71 ± 0.53 рази.

3. *Вперше* розроблено метод нечіткого пошуку векторних зображень, який, на відміну від існуючих, ґрунтується на визначенні областей припустимих значень координат вершин для двоетапного аналізу схожості зображення-кандидата та зображення-шаблону, що дозволяє підвищити точність пошуку на $20.23 \pm 2.14\%$.

4. *Удосконалено* метод підвищення яскравості медичних зображень, який ґрунтується на застосуванні нелінійних функцій до колірних компонент зображень

для моделі *HSV*, що дозволяє підвищити точність виявлення контурів на $31.83 \pm 5.81\%$.

5. *Удосконалено* метод згладжування контурних кривих об'єктів зображень, який ґрунтується на застосуванні нелінійних операцій перетворення координат контурних вершин, що дозволяє підвищити точність відтворення об'єктів зображення на $10.96 \pm 1.53\%$.

Практичне значення отриманих результатів. За результатами дисертаційного дослідження було створено інформаційну технологію оброблення архівних медичних зображень, що ґрунтується на безперервному накопиченні бази медичних зображень, отриманих різними засобами візуалізації, й дозволяє оброблення різнорідних медичних зображень з метою спрощення їх подальшого якісного аналізу лікарем. До складу інформаційної технології увійшли програмні засоби, що реалізують запропоновані методи оброблення зображень. Також практичну цінність має метод початкової класифікації оцифрованих та цифрових медичних зображень.

Теоретичні та практичні результати дисертаційного дослідження впроваджено у Національній медичній академії післядипломної освіти імені П.Л. Шупика МОН України (акт від 7.10.2019 та акт від 20.01.2020) та у ТОВ «Відео Інтернет Технології» (акт про впровадження від 20.12.2019).

Особистий внесок здобувача полягає в теоретичному обґрунтуванні отриманих результатів, їх експериментальній перевірці та дослідженні. Основні результати дисертаційного дослідження автором отримано самостійно. Наукові праці [7] – [8] та [11] є одноосібними. У публікаціях, написаних у співавторстві, здобувачеві належить: [1] – метод сегментації медичних зображень; [2] – метод контурного опису об'єктів зображень; [3] – критерії оцінки подібності зображень для задачі сортування послідовності медичних зображень; [4] – метод оброблення зображень комп'ютерної томографії (КТ) мозку для визначення морфометричних показників патологій; [5] – аналіз результатів застосування методів математичної морфології для зображень КТ мозку; [6] – архітектура інформаційної системи автоматизованого аналізу архівних медичних зображень; [9] – метод попередньої класифікації медичних зображень; [10] – метод підвищення яскравості медичних зображень.

Апробація результатів дисертації. Основні результати дисертації доповідались, обговорювались та отримали позитивну оцінку на таких конференціях: Міжнародна конференція «Health Technology Management» (м. Кишинів, 2016 р.); Міжнародна науково-практична конференція «Inżynieria i technologia Naukowa i Praktyczna Nauka światowa: problemy i innowacje» (м. Сопот, 2017 р.); Міжнародна наукова інтернет-конференція «Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення (випуск 23)» (м. Тернопіль, 2017 р.); Одинадцята наукова конференція магістрантів та аспірантів «Прикладна математика та комп'ютинг ПМК (II)» (м. Київ, 2018 р.); Восьма наукова конференція магістрантів та аспірантів «Прикладна математика та комп'ютинг ПМК» (м. Київ, 2016 р.); Всеукраїнська науково-методична відеоконференція з міжнародною участю «Актуальні питання дистанційної освіти та телемедицини 2014» (м. Запоріжжя, 2014 р.).

Публікації. За матеріалами дисертації опубліковано 11 друкованих праць (з них 3 роботи є одноосібними, 3 статті входять до наукометричної бази Scopus), серед яких 3 статті у наукових фахових виданнях ДАК України та 2 статті у закордонних наукових періодичних виданнях, що індексуються у міжнародних наукометричних базах, 6 публікацій в збірниках тез доповідей наукових конференцій.

Структура та обсяг дисертації. Дисертація складається зі вступу, чотирьох розділів, висновків, списку використаних джерел та додатків. Робота містить 119 сторінок основного тексту, 36 рисунків, 12 таблиць. Загальний обсяг дисертації становить 210 сторінок, список використаних джерел налічує 262 найменування. Дисертація містить 8 додатків, розміщених на 64 сторінках.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У вступі обґрунтована актуальність теми дисертаційної роботи; показано зв'язок роботи з науковими програмами, планами, темами; сформульовано мету і завдання наукових досліджень; визначено наукову новизну та практичне значення отриманих результатів; наведено відомості про публікації, апробацію і впровадження результатів роботи.

У першому розділі на основі аналізу літературних джерел розкрито стан проблеми, пов'язаної з необхідністю розроблення автоматизованої інформаційної технології оброблення медичних даних. Розглянуто сучасний стан, основні напрями та критерії розвитку галузі електронної охорони здоров'я у світі. Розглянуто призначення та перспективи застосування телемедицини в охороні здоров'я. Досліджено рівень впровадження телемедичних систем у розвинених країнах та в Україні. Проаналізовано сучасні підходи до інформатизації охорони здоров'я та подальші стратегії розвитку даного напрямку у світі та в Україні. Проаналізовано сучасні підходи до створення медичних інформаційних систем та розглянуто перспективи їх впровадження. Досліджено рівень впровадження медичних інформаційних систем у світі та в Україні. Проаналізовано сучасні медичні інформаційні системи в Україні, виявлено їх переваги та недоліки. Розглянуто способи підвищення якості рівня медичних послуг у сільській місцевості в Україні. Запропоновано інформаційну технологію оброблення архівних медичних даних (рис. 1) для створення напівавтоматизованих амбулаторій в Україні.

Запропонована інформаційна технологія охоплює процеси накопичення, документування та оброблення медичних даних. У рамках дисертаційної роботи розроблено методи оброблення графічних даних, зокрема, медичних зображень. Запропонована інформаційна технологія ґрунтується на створенні бази медичних даних для довготривалої безперервної медичної допомоги (Lifelong Care). Безперервне накопичення медичних зображень та формування відповідної діагностичної бази відбувається шляхом візуалізації результатів періодичних обстежень та долучення архівних зображень, що зберігаються на електронних та паперових носіях. Тому початковими процедурами запропонованої технології є

сканування зображення, початкове оброблення, класифікація та автоматизоване складання архіву зображень, а також інших графічних даних.

Наступними процесами запропонованої технології є сегментація зображень, виділення ознак, опис та пошук складових об'єктів зображень. Ґрунтуючись на базі зображень, ці процеси дозволяють проведення інтерактивного автоматизованого моніторингу, а саме: ідентифікацію об'єктів дослідження, їх оброблення, виявлення відхилень, здійснення пошуку подібних випадків, візуалізацію результатів дослідження.

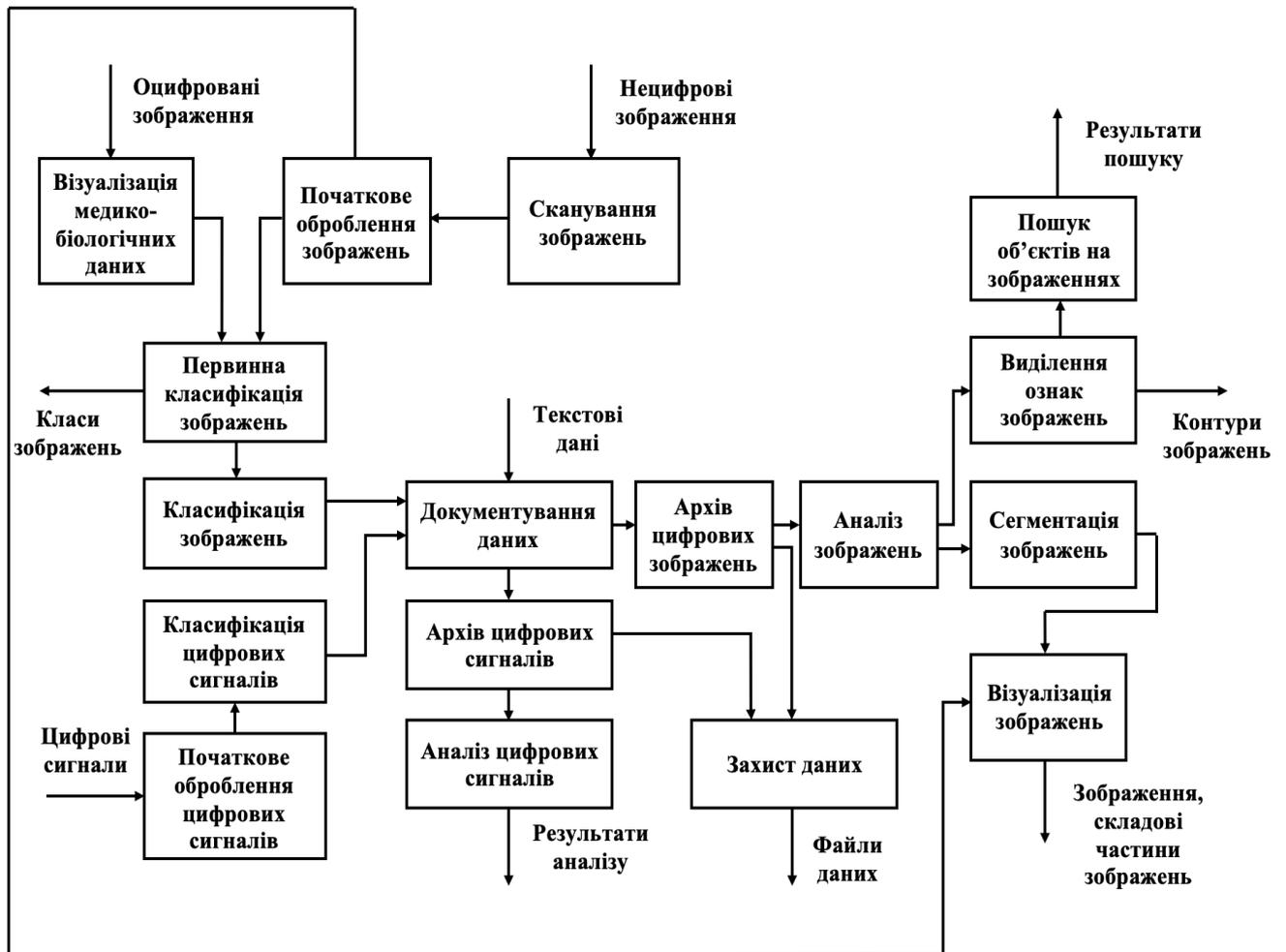


Рисунок 1 – Архітектура інформаційної технології оброблення архівних медичних даних

Використання інформаційної технології оброблення медичних даних підвищить ефективність та якість медичної допомоги, зокрема, в сільській місцевості, за рахунок впровадження запропонованої технології в телемедичні системи. Впровадження цієї інформаційної технології передбачає перерозподіл економічних та трудових ресурсів у галузі охорони здоров'я, оскільки значну частину рутинної роботи виконуватимуть програмно-апаратні системи медичної діагностики.

У другому розділі досліджено основні підходи до класифікації зображень та розроблено метод класифікації цифрових та оцифрованих медичних зображень.

Розглянуто основні градаційні перетворення для підвищення яскравості та контрасту зображень, а також розроблено метод підвищення яскравості медичних зображень.

Запропонований метод підвищення яскравості зображень ґрунтується на застосуванні градаційних перетворень до колірних компонент зображення. Даний метод є ефективним у випадках, коли зображення за візуальною оцінкою відображає відтінки одного тону та одночасно характеризується задовільним контрастом. Таким вхідним даним відповідають гістологічні зображення. Початковим етапом є перетворення графічних даних відповідно до колірного простору моделі *HSV*, результат якого описується повноколірним зображенням $Im(i, j) = (h, s, v)$ розміром $M \times N$, $i = 1, 2, \dots, M$, $j = 1, 2, \dots, N$. Наступним етапом є нелінійні перетворення компонент *Saturation* та *Value* відповідно до (1) – (2).

$$s'(i, j) = \left(\sin \left(\frac{s(i, j) - \min(s)}{\max(s) - \min(s)} \right) \right)^m, \quad (1)$$

$$v'(i, j) = \left(\frac{l(i, j) - \min(l)}{\max(l) - \min(l)} \right)^n, \quad (2)$$

де $l(i, j) = \log \left(\frac{v(i, j) - \min(v)}{\max(v) - \min(v)} + 1 \right)$;

\min , \max – операції пошуку мінімального та максимального значень;
 m , n – значення ступенів.

Повноколірне зображення $Im'(i, j) = (h, s', v')$ є результатом застосування методу (рис. 2а, 2б). Виділення контурів (рис. 2в, 2г) є оцінкою отриманих результатів в умовах, коли підвищення яскравості є початковим етапом деякого комплексу процедур для оброблення зображення.

Вхідними даними запропонованого методу класифікації є цифрові та оцифровані медичні зображення. Об'єктами зображень є клітини, тканини, судини, органи та системи органів. У таких умовах процедура класифікації характеризується значною кількістю початкових класів. Простір ознак складають такі властивості зображень: використаний метод візуалізації зображення (p_1); колірна гама зображення (p_2); розмір об'єкта дослідження (p_3); щільність однорідних об'єктів зображення (p_4); кількість однорідних об'єктів зображення (p_5).

Запропонований простір ознак визначається метаданими зображень (p_1, p_3), а також якісними (p_2) та кількісними властивостями (p_4, p_5). Визначення класу зображення відбувається за принципом порівняння відповідного вектора ознак з еталоном.

Вхідне зображення розміром $M \times N$ можна описати вектором у просторі моделі *RGB*, як $I(i, j) = (r, g, b)$, де r, g, b – інтенсивність складових кольору пікселя по компонентах кольору, $r, g, b \in [0; 255]$, i, j – координати пікселя на площині зображення, $i = 1, 2, \dots, M$, $j = 1, 2, \dots, N$. У випадках, коли $r = g = b$, вхідне зображення

є зображенням у відтінках сірого. Якщо $r \neq g \neq b$ одночасно та яскравість більшості пікселів обмежується діапазоном значень (наприклад, $r, g, b \in [150; 250]$ для гістологічних зображень лімфатичних вузлів), то вхідне зображення є зображенням з обмеженою колірною гамою, що є характерним для зображень мікроскопічних досліджень. В інших випадках вхідне зображення є повноколірним.

Щільність об'єктів зображень визначається послідовністю операцій виділення однорідних об'єктів (порогова обробка, операція ерозії) та обчисленням кількості таких об'єктів на одиницю площини. Даною ознакою (p_4) ідентифікуються мікроскопічні зображення, що належать різним класам (наведений приклад використовувався для ідентифікації зображень мікроскопічних досліджень лімфатичних вузлів та спинномозкової рідини).

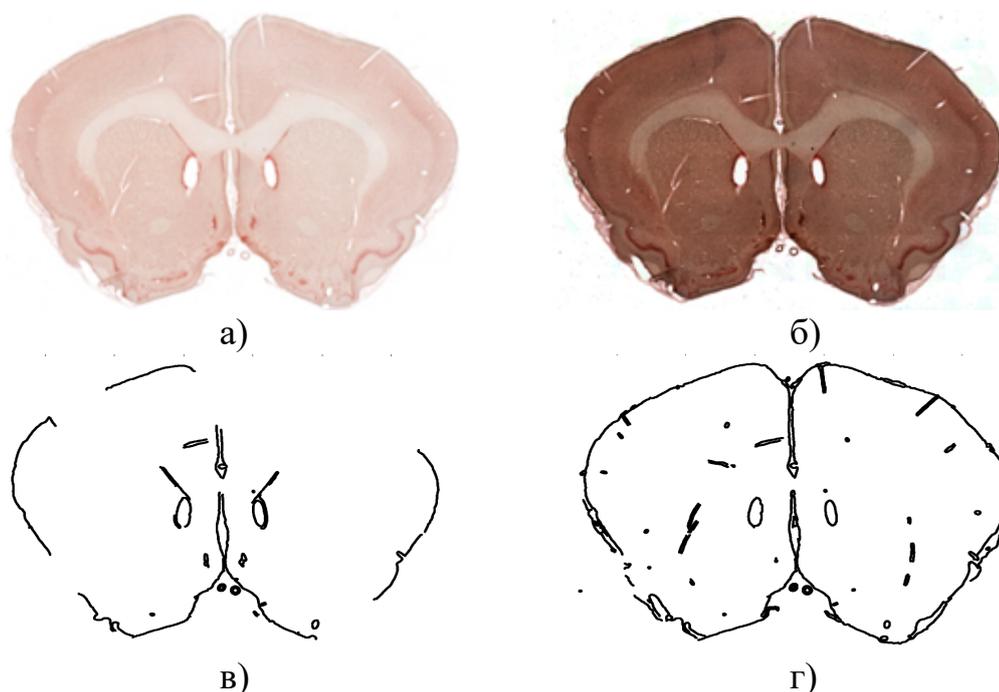


Рисунок 2 – Метод підвищення яскравості зображень: а) вхідне зображення; б) результат перетворення; в) виявлення контурів зображення (а) методом Canny; г) виявлення контурів зображення (б) методом Canny

Кількість однорідних об'єктів зображень визначається послідовністю операцій виділення таких об'єктів (порогова обробка, операції *close*, ерозії, інвертування та дилатації). Ці перетворення застосовуються для ідентифікації радіологічних зображень, що належать різним класам (наведений приклад використовувався для ідентифікації зображень комп'ютерної томографії легень).

Група запропонованих ознак застосовувалась для попередньої класифікації медичних зображень, що використовуються при діагностиці нервової та дихальної систем. Масив таких графічних даних включає 21 тип зображень. Кількість правильно класифікованих зображень при тестуванні ознаки p_4 для 100 зображень складає 76%, кількість зображень, що не потрапили до жодного класу, складає 5%, кількість неправильно класифікованих зображень – 19% (рис. 3а). При тестуванні

ознаки p_5 для 100 зображень кількість правильно класифікованих зображень складає 75%, кількість неправильно класифікованих зображень – 25% (рис. 3б).

У третьому розділі досліджено основні підходи до сегментації зображень, а також розроблено метод сегментації зображень, вхідними даними якого є масив однорідних зображень $\Omega = \{I_1, I_2, \dots, I_q, \dots, I_Q\}$, де I_q – q -те растрове зображення, $q = 1, 2, \dots, Q$, Q – кількість зображень масиву.

Однорідними можна вважати зображення, які містять однаковий об'єкт дослідження та отримані однакою методом візуалізації.

Class_True	Class_False	False	Class_name
[39]	[9]	[5]	'Клас нервової системи'
[37]	[10]	'-'	'Клас дихальної системи'

а)

Class_True	Class_False	Class_name
[38]	[13]	'Клас нервової системи'
[37]	[12]	'Клас дихальної системи'

б)

Рисунок 3 – Результати тестування ознак зображень: а) щільність однорідних об'єктів; б) кількість однорідних об'єктів

Для сегментації зображень масиву застосовується перетворення колірному простору відповідно до моделі XYZ та виділяються дані Z-компоненти. Результат перетворення описується масивом $\Omega' = \{I_1', I_2', \dots, I_q', \dots, I_Q'\}$.

Запропонований метод належить до методів сегментації, що ґрунтуються на знаходженні однорідних областей. Показник однорідності визначається статистичними параметрами масиву (3) – (4).

$$M_Z = \frac{1}{Q} \times \sum_{q=1}^Q \left(\frac{1}{M \times N} \times \sum_{i=1}^M \sum_{j=1}^N I_q'(i, j) \right), \quad (3)$$

$$D_Z = \frac{1}{Q} \times \sum_{q=1}^Q \left(\frac{1}{M \times N} \times \sum_{i=1}^M \sum_{j=1}^N \left(I_q'(i, j) - M_Z \right)^2 \right), \quad (4)$$

де $I_q'(i, j)$ – значення яскравості пікселів q -го зображення;

M, N – розміри I_q' зображення.

Центри сегментів є сталими та визначаються відповідно до формул (5) – (6).

$$T_+(h_1) = M_Z + h_1 \times D_Z, \quad (5)$$

$$T_-(h_2) = M_Z - h_2 \times D_Z, \quad (6)$$

де $T_+(h_1)$ – функція, що визначає «зростаючі» значення центрів, $h_1 = 1, 2, \dots, H_1$;
 $T_-(h_2)$ – функція, що визначає «спадаючі» значення центрів, $h_2 = 1, 2, \dots, H_2$;
 H_1, H_2 – максимальні значення змінних h_1, h_2 , відповідно.

Значення яскравості зображень масиву Ω' належать діапазону $[0; 1.089]$, отже, кількість сегментів визначається формулою (7).

$$C = \text{int}\left(\frac{M_Z}{D_Z}\right) + \text{int}\left(\frac{1.089 - M_Z}{D_Z}\right), \quad (7)$$

де int – операція відкидання дробової частини числа.

Тобто, $H_1 = \text{int}\left(\frac{M_Z}{D_Z}\right)$, $H_2 = \text{int}\left(\frac{1.089 - M_Z}{D_Z}\right)$. Таким чином, обмеженим є розмір масиву $L = \{l_c \mid l_c \in (T_+ \cup T_-), c \in [1; C]\}$, елементами якого є центри сегментів. Для більш точної сегментації масив L доповнюється такими центрами: $l_{C+1} = 0$, $l_{C+2} = M_Z$, $l_{C+3} = 1.089$.

Кількісним критерієм однорідності пікселів зображень масиву Ω' до центрів сегментів L є функція відстані (8). Щоб визначити сегменти, до яких належать пікселі зображень, необхідно визначити мінімальне значення цієї функції.

$$D(c) = \sqrt{(I_q'(i, j) - l_c)^2}, \quad (8)$$

де $D(c)$ – ступінь подібності пікселя зображення I_q' до центру сегмента l_c ;

l_c – значення c -го елемента масиву L , $c = 1, 2, \dots, C+3$.

Отже, піксель зображення I_q' з координатами (i, j) належить до сегмента p , якщо $\min(D) = D(p)$, $p \in [1; C+3]$.

Перерозподіл значень пікселів зображень масиву Ω' відповідно до номерів сегментів можна представити масивом індексних зображень $\Omega_s = \{S_1, S_2, \dots, S_q, \dots, S_Q\}$, елементи якого описуються двовимірними матрицями $S_q(i, j) = c$, $c \in [1; C+3]$. Для візуальної оцінки таких результатів застосовується палітра кольорів, де значення пікселів зображень S_q є індексами палітри $\text{map}(c) = (r_c, g_c, b_c)$.

Зображення S_q можна представити набором зображень $W = \{B_1, B_2, \dots, B_c, \dots, B_{C+3}\}$, що відповідають сегментам вхідного зображення. Ці сегменти візуалізують складові області, контури, фон, шумові компоненти зображення I_q' . Зображення B_c є бінарними зображеннями та визначаються відповідно до (9).

$$B_c(i,j) = \begin{cases} 1, S_q(i,j) = c \\ 0, S_q(i,j) \neq c \end{cases} \quad (9)$$

Для виділення контурних точок складових областей спочатку застосовується операція ерозії (10) для зображень B_c , а потім розраховується різниця між B_c та попереднім результатом (11).

$$E_c(i,j) = \begin{cases} 1, \sum_{a=-1}^1 \sum_{b=-1}^1 B_c(i+a, j+b) = 9 \\ 0, \sum_{a=-1}^1 \sum_{b=-1}^1 B_c(i+a, j+b) < 9 \end{cases} \quad (10)$$

$$K_c(i,j) = B_c(i,j) - E_c(i,j) = \begin{cases} 0, B_c(i,j) = E_c(i,j) \\ B_c(i,j), B_c(i,j) \neq E_c(i,j) \end{cases} \quad (11)$$

Тестовими даними для запропонованого методу сегментації є зображення комп'ютерної томографії (КТ) легень (рис. 4а). Результатом застосування є індексне зображення (рис. 4б) та масив зображень сегментів (рис. 5), які візуалізують фон (рис. 5а), складові частини зображення (рис. 5 б-в, е-з) та контури (рис. 5 г-д). Також запропонованим методом виділяються контури зображень складових частин.

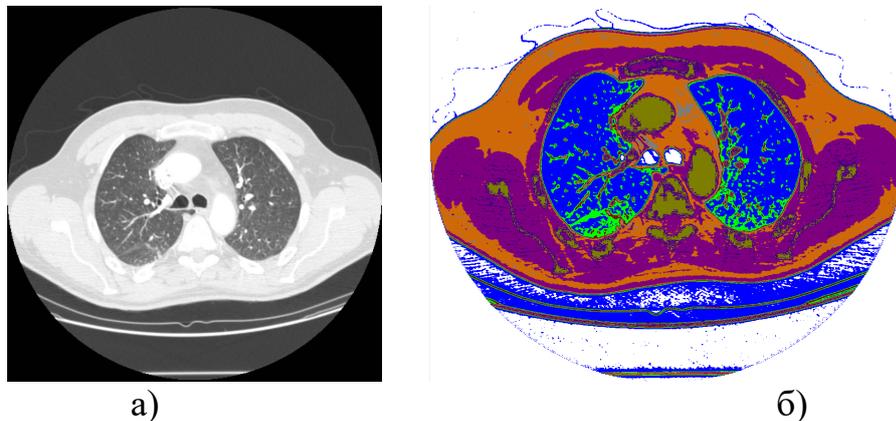


Рисунок 4 – Сегментація зображення: а) приклад зображення вхідного масиву даних; б) індексне зображення

У четвертому розділі досліджено основні підходи до виділення ознак зображень, зокрема, контурів та пошуку зображень, а також розроблено методи опису та пошуку об'єктів на зображеннях.

Метод опису ґрунтується на контурному представленні об'єктів зображень у вигляді масивів впорядкованих вершин з (x,y) координатами, що утворюють неперервні контурні криві. Вхідними даними є відеодані, що візуалізують деякі процеси. Ці процеси характеризуються зображеннями незмінних фонів та

зображеннями об'єктів спостереження, які рухаються та/або переміщуються. Також вхідними даними можуть бути статичні медичні зображення в умовах наявності бази даних з порівняльної та патологічної анатомії.

Відеодані можна представити мультифреймовим масивом. Елементами такого масиву є повноколірні зображення у просторі колірної моделі *RGB*.

Простір колірної моделі *RGB* вхідних зображень перетворюється відповідно до *YCbCr* моделі, а результати перетворення нормуються відповідно до (12).

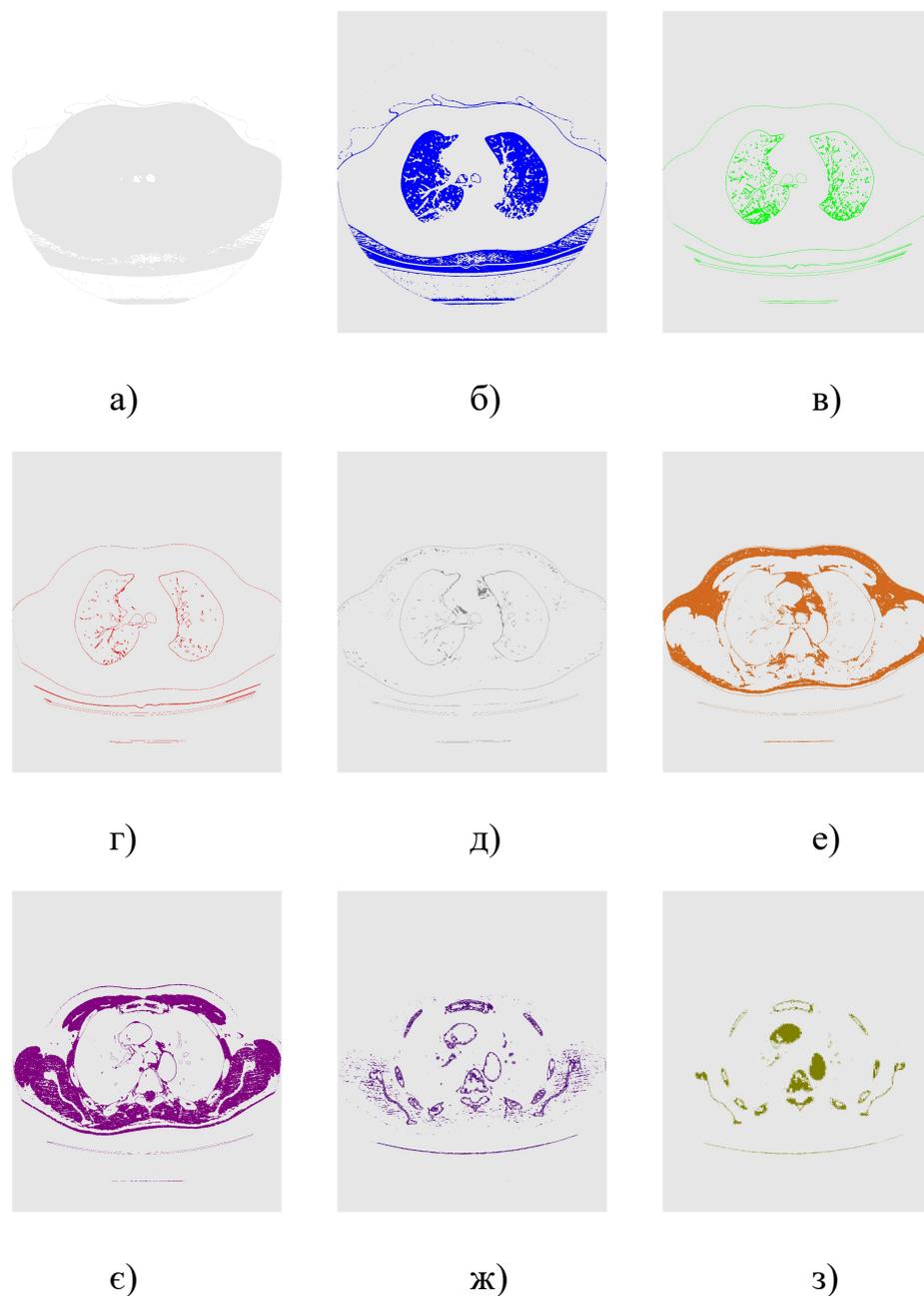


Рисунок 5 – Масив сегментів зображення КТ легень

$$\begin{bmatrix} Y' \\ Cb' \\ Cr' \end{bmatrix} = \begin{bmatrix} \sin\left(\frac{Y}{255}\right) \\ \sin\left(\frac{Cb}{255}\right) \\ \sin\left(\frac{Cr}{255}\right) \end{bmatrix}, \quad (12)$$

де Y' , Cb' , Cr' – результати нормування;

$$Y' \in [0.0627; 0.7966], \quad Cb', Cr' \in [0.0627; 0.8083], \quad Y \in [16; 235], \quad Cb, Cr \in [16; 240].$$

Колірні вектори (13) є значеннями яскравості пікселів зображення незмінного фону I_b та тестового зображення I_t , яке перевіряється на присутність об'єктів спостереження:

$$\begin{aligned} I_b(i, j) &= (b_1, b_2, b_3) \\ I_t(i, j) &= (t_1, t_2, t_3) \end{aligned}, \quad (13)$$

де i, j – координати пікселів зображень розміром $M \times N$;

$$i = 1, 2, \dots, M, \quad j = 1, 2, \dots, N;$$

$$b_1, t_1 \in [0.0627; 0.7966], \quad b_{2,3}, t_{2,3} \in [0.0627; 0.8083].$$

Виявлення об'єкта спостереження відбувається шляхом застосування логічних операцій, зокрема, логічного заперечення. Для її визначення у випадку повноколірного зображення використовується матриця V , множина значень якої не перетинається з множинами значень зображень I_b та I_t . Отже, значення елементів матриці V можуть належати діапазону $(0.8083; 1]$. Операція логічного заперечення виконується відповідно до (14).

$$\begin{aligned} I_b^n(i, j) &= V(i, j) - I_b(i, j) = (b_1^n, b_2^n, b_3^n) \\ I_t^n(i, j) &= V(i, j) - I_t(i, j) = (t_1^n, t_2^n, t_3^n) \end{aligned}. \quad (14)$$

де b_k^n, t_k^n – значення яскравості пікселів зображень, які є результатом операції логічного заперечення, $k = 1, 2, 3$.

Виявлення об'єкта спостереження виконується згідно з (15). Зображення $I_M(i, j)$ є контрастним растровим зображенням. Пікселі фону представлені векторами, які відповідають чорному кольору у просторі моделі $YCbCr$, а пікселі об'єкта спостереження – іншим кольорам моделі.

$$I_M(i, j) = \left| \max(b_k(i, j), t_k^n(i, j)) - \max(b_k^n(i, j), t_k(i, j)) \right|. \quad (15)$$

Вхідними даними для етапу виявлення контурів методом Differential Equation Interpolant (DEI) є дані Y -компоненти зображення $I_M(i, j)$. Отримані контурні криві візуалізують межі об'єктів спостереження, їх елементів та описуються різною

кількістю вершин. Нехай деяка контурна крива q (16) зображення $I_M(i, j)$ візуалізує форму площі, яку займає об'єкт спостереження у двовимірному просторі:

$$\text{Contour}(q) = \begin{pmatrix} x_1 & x_2 & \dots & x_{i-1} & x_i \\ y_1 & y_2 & \dots & y_{i-1} & y_i \end{pmatrix}, \quad (16)$$

де x_i, y_i – координати вершин контурної кривої q , $i=1,2,\dots,\theta$;
 θ – кількість вершин контурної кривої q , $q=1,2,\dots,Q$;
 Q – кількість контурних кривих зображення $I_M(i, j)$.

Наступним кроком є перерозподіл контурних вершин відносно початку координат. Даний перерозподіл вершин для q кривої (16) виконується за $(i = \beta, \beta+1, \dots, \theta, 1, 2, \dots, \beta-1)$ або проти $(i = \beta, \beta-1, \dots, 1, \theta, \theta-1, \beta+1)$ годинникової стрілки, якщо вершина з координатами (x_β, y_β) є найближчою до точки $O(0;0)$.

Отримані контурні криві містять «надлишкові» вершини, які візуалізуються у вигляді петель, хвилеподібних нерівностей та вказують на наявність шумів. Для зменшення шуму застосовується запропонований метод згладжування контурних кривих як етап постоброблення процедури виявлення контурів.

Для кожного кроку методу згладжування визначаються масиви даних із постійним розміром $1 \times W$, що містять x та y координати вершин контурної кривої. Наприклад, для деякої вершини K із координатами (x_1, y_1) та $W=5$ контурної кривої q (16) масиви таких даних дорівнюють $\omega_x = (x_1, x_2, x_3, x_4, x_5)$ та $\omega_y = (y_1, y_2, y_3, y_4, y_5)$. Наступним етапом є перерозподіл значень масивів та виділення медіани. Наприклад, масиви ω_x та ω_y після сортування дорівнюють $\omega_x' = (x_2, x_3, x_4, x_5, x_1)$ та $\omega_y' = (y_5, y_1, y_2, y_3, y_4)$, де x_4 та y_2 – медіани. Відповідно початкові координати вершин контурних кривих змінюються, а координати вершини K становитимуть (x_4, y_2) .

Наступним етапом є сортування даних, яке дозволяє виявити «опорні» та малоінформативні вершини контурних кривих. Початковим кроком сортування є визначення кількості відрізків або сегментів, що в сукупності відображають деякий контур. Наступний крок – формування шаблонного вектора, за значеннями якого визначаються «опорні» вершини контуру.

Контур об'єкта спостереження описується виразом $L = \{l_p(x, y) \mid p \in [1; P], l_p \in \text{Contour}(q)\}$, де P – кількість сегментів, які формують контур об'єкта спостереження. Діапазон значень шаблонного вектора обмежується розміром вхідних зображень та визначається відповідно до (17). Значення вектора T є незмінними для кожного сегмента l_p .

$$T = \left\{ 1, 1 + D, 1 + 2 \times D, \dots, 1 + \text{int} \left(\frac{N-1}{D} \right) \times D \right\}, \quad (17)$$

де D – параметр деталізації, $0 < D < N$;
 N – розмір зображення $I_M(i, j)$.

Процедура сортування виконується для значень x координати вершин кожного сегмента l_p . Для першого сегмента контуру L ці значення дорівнюють $l_{1x} = (x_1, x_2, \dots, x_{\theta/p})$. Визначення x координат «опорних» вершин контуру L відбувається шляхом знаходження для кожного сегмента l_{px} , $p = 1, 2, \dots, P$ та вершин вектора T значення $\min \left(\sqrt{(T_h - l_{px})^2} \right)$, де $h = 1, 2, \dots, 1 + \text{int} \left(\frac{N-1}{D} \right)$. Координата y для «опорної» вершини залишається незмінною та вилучається з l_{py} . У такий спосіб формується масив сегментів $L' = \{l_p'(x, y) \mid p \in [1; P], l_p' \in \text{Contour}(q)\}$, які містять вже «опорні» вершини початкової контурної кривої. Після того, як вершина потрапляє в масив сегментів L' , вона видаляється з масиву L .

Особливістю процедури виявлення «опорних» вершин є можливість описати контури будь-яких об'єктів спостереження рівною кількістю вершин, яка розраховується відповідно до (18). Це дозволить ефективно вирішувати задачі аналізу зображень, оскільки подальші операції відбуватимуться на однорідно представлених даних.

$$S = P \times \left(\text{int} \left(\frac{N-1}{D} \right) + 1 \right). \quad (18)$$

Для подання масиву сегментів L' строго впорядкованим набором «опорних» вершин, які послідовно з'єднуються неперервними кривими, для вершин сегментів l_p' застосовується процедура сортування $\left(\sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2} \right) \rightarrow \min$, де x_i, y_i – координати вершин сегментів l_p' , $i = 1, 2, \dots, 1 + \text{int} \left(\frac{N-1}{D} \right)$, $p = 1, 2, \dots, P$.

З метою уникнення надлишкових процедур оброблення початковими вершинами є перші вершини сегментів l_p . Коли масив впорядкованих вершин $L'' = \{l_p''(x, y) \mid p \in [1; P], l_p'' \in \text{Contour}(q)\}$ буде сформований, ці вершини необхідно видалити. Перша та остання вершини масиву L'' дублюються.

Пошук схожих зображень ґрунтується на початковому визначенні областей нечіткості для контурних вершин. Область нечіткості є деяким еліпсом із центром у такій вершині та радіусами $\mu\tilde{x}$, $\mu\tilde{y}$. Вхідними даними для пошуку є масиви вершин

із нечіткими координатами \tilde{x} , \tilde{y} та точними координатами x , y . Вершини з координатами \tilde{x} , \tilde{y} описують зображення-шаблони, відповідно до яких необхідно знайти схожі зображення. Вершини з координатами x , y описують зображення-кандидати, які є схожими до зображень-шаблонів у випадку потрапляння x , y в область нечіткості. Для пошуку схожих векторних зображень застосовуються два етапи. Початковий пошук характеризується невеликим обсягом даних та значною областю нечіткості. Обсяг даних визначається кількістю вершин, що дозволяють коректно описати об'єкт на зображенні (18). Кінцевий пошук здійснюється для більшого обсягу даних та з меншою областю нечіткості. Для визначення області нечіткості вершин розраховуються радіуси еліпсів (19).

$$\begin{aligned}\mu_{\tilde{x}}^{(1)} &= \lambda_{\tilde{x}} \times D_{\tilde{x}}^{(1)} \\ \mu_{\tilde{y}}^{(1)} &= \lambda_{\tilde{y}} \times D_{\tilde{y}}^{(1)}\end{aligned}\quad (19)$$

$$\begin{aligned}\text{де } \lambda_{\tilde{x}} &= \text{int}\left(\frac{M_{\tilde{x}}^{(1)}}{D_{\tilde{x}}^{(1)}}\right), \quad \lambda_{\tilde{y}} = \text{int}\left(\frac{M_{\tilde{y}}^{(1)}}{D_{\tilde{y}}^{(1)}}\right); \\ M_{\tilde{x}}^{(1)} &= \frac{1}{S_1} \times \sum_{i=1}^{S_1} \tilde{x}_i, \quad M_{\tilde{y}}^{(1)} = \frac{1}{S_1} \times \sum_{i=1}^{S_1} \tilde{y}_i, \quad \tilde{x}_i, \tilde{y}_i \in [0;1]; \\ D_{\tilde{x}}^{(1)} &= \frac{1}{S_1} \times \sum_{i=1}^{S_1} (M_{\tilde{x}}^{(1)} - \tilde{x}_i)^2, \quad D_{\tilde{y}}^{(1)} = \frac{1}{S_1} \times \sum_{i=1}^{S_1} (M_{\tilde{y}}^{(1)} - \tilde{y}_i)^2; \\ S_1 &\text{ – кількість вершин для початкового етапу пошуку.}\end{aligned}$$

Для виконання кінцевого етапу пошуку область нечіткості зменшується та дорівнює $\mu_{\tilde{x}}^{(2)} = (\lambda_{\tilde{x}} - \sigma) \times D_{\tilde{x}}^{(2)}$, $\mu_{\tilde{y}}^{(2)} = (\lambda_{\tilde{y}} - \sigma) \times D_{\tilde{y}}^{(2)}$, де $D_{\tilde{x}}^{(2)}$ та $D_{\tilde{y}}^{(2)}$ – значення дисперсій для S_2 вершин, $\sigma \in [1; \min(\lambda_{\tilde{x}}; \lambda_{\tilde{y}}) - 1]$. Ступінь схожості зображень-кандидатів до зображень-шаблонів для двох етапів пошуку визначається добутком значень схожості вершин, які описують об'єкти на зображенні (20). У випадку схожих вершин $\varepsilon \leq R_i < 1 + \varepsilon$, у протилежному випадку $R_i = 0$, у випадку співпадіння $R_i = 1 + \varepsilon$.

$$R = \prod_{i=1}^S R_i = \prod_{i=1}^S \max\left(0; 1 + \varepsilon - \left(\frac{(x_i - \tilde{x}_i)^2}{(\mu_{\tilde{x}}^{(\tau)})^2} + \frac{(y_i - \tilde{y}_i)^2}{(\mu_{\tilde{y}}^{(\tau)})^2}\right)\right), \quad (20)$$

де R_i – показник схожості i -ої вершини, $i = 1, 2, \dots, S$;

S – кількість контурних вершин, $x_i, y_i \in [0;1]$, $\tilde{x}_i, \tilde{y}_i \in [0;1]$, $\tau = 1, 2$;

ε – деяке невелике додатне число (наприклад, $\varepsilon = 0.0001$).

Перший етап пошуку застосовується для виявлення несхожих зображень ($R = 0$) або зображень із незначною схожістю ($R \rightarrow 0$). Другий етап застосовується

для виявлення найбільш схожих зображень. Кількісний показник ступеня схожості зображень R однозначно визначає такі категорії зображень-кандидатів: «однакові», «схожі», «незначно схожі» та «несхожі». Ступінь несхожості зображень можна оцінити за кількістю виконання умов $R_i = 0$.

Тестовими даними для методу контурного опису об'єктів є зображення КТ серця. Результати застосування даного методу показано на рис. 6. Для пошуку масив тестових даних складався зі 100 зображень-кандидатів КТ серця та містив зображення-шаблон Im_72 (рис. 6). Кількість контурних вершин для першого етапу пошуку становила 949, для другого – 1369. Тестові зображення є схожими до шаблону Im_72 зі значеннями показника R відповідно до даних таблиці 1.

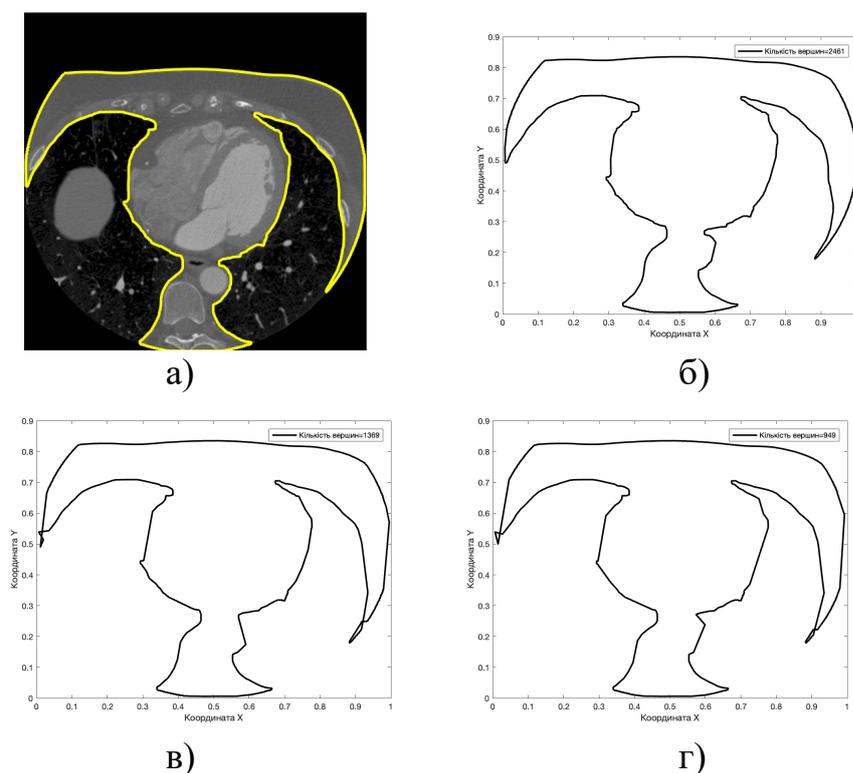


Рисунок 6 – Контур зображення та кількість вершин: а) початковий контур, кількість вершин – 3291; б) 2461; в) 1369; г) 949

Таблиця 1 – Результати визначення схожості для другого етапу пошуку

Назва файлу	Показник R						
Im_71	1.02e-05	Im_78	2.99e-35	Im_85	6.20e-44	Im_92	0
Im_72	1.15	Im_79	1.01e-22	Im_86	1.01e-51	Im_93	0
Im_73	6.34e-17	Im_80	5.53e-40	Im_87	4.64e-68	Im_95	2.07e-67
Im_74	1.24e-11	Im_81	3.64e-34	Im_88	4.13e-62	Im_96	9.14e-60
Im_75	3.67e-12	Im_82	8.73e-35	Im_89	5.97e-110	Im_97	0
Im_76	2.25e-18	Im_83	5.62e-48	Im_90	1.47e-88		
Im_77	4.64e-16	Im_84	9.05e-66	Im_91	1.37e-67		

Параметри нечіткості для першого етапу пошуку дорівнювали $\mu_{\tilde{x}}^{(1)} = 0.4944$, $\mu_{\tilde{y}}^{(1)} = 0.4607$. Для другого етапу пошуку параметри нечіткості становили $\mu_{\tilde{x}}^{(2)} = 0.4362$, $\mu_{\tilde{y}}^{(2)} = 0.3721$, $\sigma = 1$.

ВИСНОВКИ

У дисертаційній роботі запропоновано розв'язання актуального наукового завдання розроблення та удосконалення інформаційних рішень для оброблення медичних зображень. У рамках дисертаційного дослідження отримані наукові результати:

1. Запропоновано архітектуру інформаційної технології оброблення архівних медичних даних, яка, на відміну від існуючих, ґрунтується на безперервному накопиченні бази медичних зображень, отриманих різними засобами візуалізації, та дозволяє оброблення різнорідних медичних зображень з метою спрощення їх подальшого якісного аналізу лікарем.

2. Вперше розроблено метод сегментації медичних зображень, що ґрунтується на аналізі статистичних показників масиву однорідних зображень та дозволяє отримати складові частини та контури зображень шляхом утворення однорідних областей без урахування просторової близькості елементів зображення. Запропонований метод доцільно застосовувати для автоматизованого визначення відхилень на об'єктах зображень, а також подальшого їх дослідження лікарем. Метод застосовується для сегментації медичних зображень будь-якого типу, зокрема, КТ легень, МРТ головного мозку, ультразвукових досліджень серця, черевної порожнини, мікроскопічних досліджень тканин, клітин та ін.

3. Вперше розроблено метод опису зображень, що ґрунтується на контурному представленні зображення та його елементів масивами координат «опорних» вершин, розмір яких є вхідним параметром методу та може змінюватись залежно від вимог, які визначаються поставленою задачею пошуку та потрібною точністю відтворення. У порівнянні з методом Differential Equation Interpolant, запропонований метод дозволяє зменшити обсяг графічних даних, які зберігаються та оброблюються у медичній програмно-апаратній системі, у 2.71 ± 0.53 рази для описання форми площі, яку займає об'єкт спостереження. Запропонований метод застосовується для контурного опису медико-біологічних об'єктів довільної форми на медичних зображеннях, зокрема, КТ серця, голови, легень, МРТ головного мозку, мікроскопічних досліджень тканин та ін.

4. Вперше розроблено метод нечіткого пошуку векторних зображень, який, на відміну від існуючих, ґрунтується на визначенні областей припустимих значень координат вершин зображень-шаблонів. Запропонований метод пошуку дозволяє кількісно визначити ступінь схожості зображень-кандидатів та зображень-шаблонів. Запропонований метод дозволяє підвищити точність пошуку на $20.23 \pm 2.14\%$ у порівнянні з методом пошуку зображень за змістом. Запропонований метод

застосовується для пошуку медико-біологічних об'єктів довільної форми на медичних зображеннях, зокрема, КТ серця, голови та ін.

5. Розроблено метод класифікації оцифрованих та цифрових медичних зображень, що ґрунтується на використанні еталонних зображень, описаних векторами у параметричному просторі, який визначається такими ознаками: метод візуалізації та колірна гама зображення, розмір об'єкта дослідження, щільність та кількість однорідних об'єктів на зображенні. Метод класифікації доцільно використовувати на початковому етапі оброблення медичних зображень для створення груп однорідних даних з метою вторинної класифікації нейромережевими технологіями. Запропонований метод може застосовуватись для класифікації великих наборів графічних даних, зокрема для 21 типу зображень, що відображають процес діагностики нервової та дихальної систем.

6. Удосконалено метод підвищення яскравості медичних зображень, що ґрунтується на застосуванні комплексу нелінійних функцій до колірних компонент зображень для моделі *HSV*. Запропонований метод дозволяє підвищити точність виявлення контурів на $31.83 \pm 5.81\%$ у порівнянні з методом лінійного розтягування. Запропонований метод застосовується для зображень мікроскопічних досліджень тканин та клітин.

7. Удосконалено метод згладжування контурних кривих об'єктів зображень, що ґрунтується на застосуванні нелінійних операцій перетворення координат контурних вершин, що дозволяє пригнічувати перешкоди на медичних зображеннях, представлених у векторному форматі. Запропонований метод дозволяє підвищити точність відтворення об'єктів зображення на $10.96 \pm 1.53\%$ у порівнянні з методом медіанної фільтрації. Запропонований метод застосовується для згладжування контурних кривих об'єктів довільної форми на медичних зображеннях, зокрема, зображень КТ серця, мікроскопічних досліджень тканин та ін.

8. Створено програмне забезпечення для дослідження ефективності розроблених методів оброблення медичних зображень. Практичну цінність отриманих результатів підтверджено трьома актами впровадження.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Наукові праці, в яких опубліковані основні наукові результати дисертації:

1. Shkurat O., Sulema Y., Suschuk-Sliusarenko V., Dychka A. Image Segmentation Method Based on Statistical Parameters of Homogeneous Data Set. *Advances in Intelligent Systems and Computing*. 2020. Vol. 902. P. 271–281.

2. Shkurat O. S., Sulema Ye. S., Dychka A. I. Complicated Shapes Estimation Method for Objects Analysis in Video Surveillance System. *KPI Science News*. 2018. N 3. P. 53–62.

3. Hu Z., Dychka I., Sulema Y., Valchuk Y., Shkurat O. Method of Medical Images Similarity Estimation Based on Feature Analysis. *International Journal of Intelligent Systems and Applications*. 2018. Vol. 10, N 5. P. 14–22.

4. Shkurat O., Solomin A. Methods of Processing Medical Images for the Detection of Pathologies in Craniofacial Surgery. *Eastern-European Journal of Enterprise Technologies*. 2015. Vol. 3, N 2 (75). P. 35–41.

5. Шкурят О. С., Соломин А. В. Морфологическая фильтрация полутонных изображений черепно-лицевой области. *Проблеми інформаційних технологій*. 2015. № 1 (017). С. 209–212.

Наукові праці, які засвідчують апробацію матеріалів дисертації:

6. Sulema Y. S., Shkurat O. S. Information System for Archival Medical Images Automated Processing. *Health Technology Management (HTM-2016): Book of Abstracts of the 3rd International Conference (Chisinau, October 6 – 7, 2016)*. Chisinau: Pontos, 2016. P. 72.

7. Шкурят О. С. Интерактивный процесс кластеризации у автоматизованных системах анализа медицинских изображений. *Inżynieria i technologia Naukowa i Praktyczna Nauka światowa: problemy i innowacje: збірник тез доповідей міжнар. наук.-практ. конф. (м. Сопот, 31 жовтня 2017 р.)*. Варшава, 2017. С. 28–30.

8. Шкурят О. С. Идентификация медицинских изображений. *Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення (випуск 23): збірник тез доповідей міжнар. наук. інтернет-конф. (м. Тернопіль, 17 жовтня 2017 р.)*. Тернопіль, 2017. С. 51–53.

9. Сулема Є. С., Шкурят О. С. Спосіб попередньої класифікації архівних медичних зображень. *Прикладна математика та комп'ютинг (ПМК 2018): збірник тез доповідей одинадцятої наук. конф. магістрантів та аспірантів (м. Київ, 14 – 16 листопада 2018 р.)*. Київ: КПІ ім. Ігоря Сікорського, Вид-во «Політехніка», 2018. С. 261–267.

10. Sulema Y. S., Shkurat O. S. Histological Image Contrast Enhancement Technique Based on HSV Colour Model. *Прикладна математика та комп'ютинг (ПМК 2016): збірник тез доповідей восьмої наук. конф. магістрантів та аспірантів (м. Київ, 20 – 22 квітня 2016 р.)*. Київ: Просвіта, 2016. С. 265–269.

11. Шкурят О. С. Медична інформаційна система щелепно-лицевої хірургії. *Актуальні питання дистанційної освіти та телемедицини 2014: збірник тез доповідей всеукраїн. наук.-метод. відеоконф. з міжнар. участю (м. Запоріжжя, 16 – 17 жовтня 2014 р.)*. Запоріжжя: ЗДМУ, 2014. С. 151–152.

АНОТАЦІЯ

Шкурят О. С. Методи та інформаційна технологія оброблення архівних медичних зображень. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології. – Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

МОН України, Інститут проблем математичних машин і систем НАН України. – Київ, 2020.

Робота присвячена розв'язанню актуального наукового завдання розроблення та удосконалення інформаційних рішень для оброблення медичних зображень.

Запропоновано інформаційну технологію автоматизованого оброблення медичних даних, зокрема, архівів медичних зображень для забезпечення довготривалої безперервної медичної допомоги. Розроблено нові методи оброблення медичних зображень, зокрема, метод сегментації однорідних медичних зображень, метод контурного опису об'єктів зображень та метод нечіткого пошуку об'єктів на зображеннях. В запропонованих методах сегменти зображень описуються масивом зображень фону, складових частин та контурів, об'єкти зображень – впорядкованим набором координат контурних вершин. Розроблено метод класифікації оцифрованих та цифрових медичних зображень, а також удосконалено метод підвищення яскравості медичних зображень та метод згладжування контурних кривих об'єктів зображень.

Ключові слова: інформаційна технологія, аналіз зображень, оброблення зображень, сегментація, класифікація, пошук зображень.

АННОТАЦІЯ

Шкурат О. С. Методы и информационная технология обработки архивных медицинских изображений. – Квалификационная научная работа на правах рукописи.

Диссертация на соискание научной степени кандидата технических наук по специальности 05.13.06 – информационные технологии. – Национальный технический университет Украины «Киевский политехнический институт имени Игоря Сикорского» МОН Украины, Институт проблем математических машин и систем НАН Украины. – Киев, 2020.

Работа посвящена разрешению актуального научного задания разработки и усовершенствования информационных решений для обработки медицинских изображений.

Предложена информационная технология автоматизированной обработки медицинских данных, в частности, архивов медицинских изображений для обеспечения долговременной непрерывной медицинской помощи. Разработаны новые методы обработки медицинских изображений, в частности, метод сегментации однородных медицинских изображений, метод контурного описания объектов изображений и метод нечеткого поиска объектов на изображениях. В предложенных методах сегменты изображений описываются массивом изображений фона, составных частей и контуров, объекты изображений – упорядоченным набором координат контурных вершин. Разработан метод классификации оцифрованных и цифровых медицинских изображений, а также усовершенствованы метод повышения яркости медицинских изображений и метод сглаживания контурных кривых объектов изображений.

Ключевые слова: информационная технология, анализ изображений, обработка изображений, сегментация, классификация, поиск изображений.

ABSTRACT

Shkurat O. S. Methods and information technology of archival medical image processing. – Qualification scientific work as a manuscript.

Ph.D. thesis in specialty 05.13.06 – information technologies. – National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute» of the Ministry of Education and Science of Ukraine, Institute of Problems of Mathematical Machines and Systems of the National Academy of Sciences of Ukraine. – Kyiv, 2020.

The work is devoted to solving the actual scientific problem of developing and improving information decisions for processing and analysis of medical images.

The modern transformation instruments of medical data, in particular, medical information systems were considered, their advantages and disadvantages were detected. The advantages of using the archival images in the modern medical researches were defined. The necessity of creating the new approaches to medical data processing was shown. For the first time, the information technology for automated processing of medical data, in particular, medical image archives was proposed. The components of the proposed information technology for archival medical image processing are the preprocessing, classification and segmentation of images, as well as the feature extraction and object retrieval on images.

For the first time, the medical image segmentation method, which, unlike the existing ones, is based on analysis of the statistical rates for input array of homogeneous images, was developed. The input array of homogeneous images is not limited by the element number and the data type. The segment number and the values of their centers are the a priori information of the proposed method. The result of applying the proposed method are index images of complex segments, as well as the corresponding arrays of segment images visualized the background, image parts and contours. The practical segmentation problem for image arrays of computed tomography (CT) of the lungs, magnetic resonance imaging (MRI) of the brain, ultrasound of the heart, the abdominal cavity and microscopic images of tissues, etc., was considered. The implementing the proposed method to patients' long-term monitoring systems allows to automatically detect the deviations on images.

For the first time, the medical image description method, which, unlike the existing ones, is based on contour representation of the image and its elements by the arrays of the coordinates of the reference vertices, was developed. The size of the arrays is set according to the requirements which are defined by the given task of the retrieval and the desired reproduction accuracy. The segment number of contour and the value of the template vector are the main parameters of the proposed method that define the accuracy of the object reproduction, as well as the appropriate data volume. The application of the proposed method allows to reduce the graphical data volume stored and processed in the medical software and hardware systems by 2.71 ± 0.53 times compared to interpolation

method of Differential Equation Interpolant (DEI). The practical problem of contour description for video images and static medical images, in particular, CT of the heart, the brain, MRI of the brain and microscopic images of tissues, etc. was considered.

For the first time, the fuzzy retrieval method of the objects on images, which, unlike the existing ones, is based on defying regions of the permissible values of the vertex coordinates for two-stage analysis of the similarity for the candidate images and the template images, was developed. The first stage of the retrieval is characterized by the larger region of the fuzziness and less data than the second stage, and is applied to quantitatively detect such image categories as «dissimilar» images or «slightly similar» images. To quantitatively detect the «similar» images or the «identical» images, the second stage of retrieval is applied. The similarity degree of the candidate images to the template images is defined by the product of the similarity values of the vertices that describe the image objects. The application of the proposed method allows to increase the retrieval accuracy in $20.23 \pm 2.14\%$ compared to the content-based image retrieval method. The practical problem of the object retrieval on images, in particular, CT of the brain, the heart, etc., was considered.

The method of the primary classification for the digitized and digital medical images, which is based on applying the templates, was developed. To create the homogeneous data groups for the secondary classification by neural networks technology the proposed method is applied. The practical problem of image array classification, which includes 21 elements and reflects the diagnosis process of the nervous and respiratory systems, was considered.

The brightness enhancement method for medical images, which is based on applying the complexes of the non-linear functions to colour components of the images for the *HSV* model, was improved. The application of the proposed method allows to increase the contour detection accuracy on $31.83 \pm 5.81\%$ compared to the linear stretch method. The practical problem of the brightness enhancement, in particular, microscopic images, was considered.

The smoothing method of the contour curves for the image objects, which is based on applying the non-linear operations to contour vertex coordinates for the noise suppression on the medical images presented in a vector format, was improved. The practical problem of smoothing the contour curves of the objects on images, in particular, CT of the heart, the brain and microscopic images, etc., was considered. The application of the proposed method allows to increase the reproduction accuracy of image objects on $10.96 \pm 1.53\%$ compared to the median filtering method.

Keywords: information technology, image analysis, image processing, segmentation, classification, image retrieval.