

АЛГОРИТМ ПОРІВНЯННЯ ДВОХ ЗРАЗКІВ ДЛЯ АНАЛІЗУ МЕДИЧНИХ ЗОБРАЖЕНЬ ШЛЯХОМ СПІВСТАВЛЕННЯ ПАТЕРНІВ

Шулигін Д.Е., магістр
shulygin.d@gmail.com

Настенко Є.А., к.т.н., д.б.н., проф.
nastenko.e@gmail.com

м. Київ, Україна

Кафедра біомедичної кібернетики
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

Реферат – Метою роботи була розробка та реалізація нового алгоритму порівняння двох медичних зображень паренхіматозних органів шляхом виокремлення в них однакових та унікальних патернів та наступного їх аналізу. Для цього було розглянуто найбільш розповсюджені алгоритми порівняння зображень, серед яких були строге порівняння, порівняння нечітких пікселів та порівняння гістограм. Оскільки найбільш розповсюджені алгоритми порівняння зображень мають недоліки, які роблять неможливим їх використання в аналізі медичних зображень було створено програмний додаток для порівняння двох медичних зображень, який реалізує алгоритм порівняння зображень шляхом співставлення патернів, а також використовує для попередньої обробки алгоритм зменшення кількості відтінків сірого. Для розробки було використано фреймворк .NET та мову програмування C#. Після аналізу отриманих результатів роботи програмного додатку на зображеннях легень у станах норми, пневмонії, COVID-19, COVID-19 з пневмонією, а також печінки у нормі та при цирозі було з'ясовано, що при такому підході попередня обробка зображень шляхом зменшення кількості відтінків сірого є необхідним компонентом програми, а також що можливо отримати задовільні результати навіть при обробці зображень, які містять додаткову інформацію про оточуючі тканини, але для найкращого результату на вхід програми бажано подавати зображення, на яких міститься виключно текстура досліджуваних органів. Також результати роботи створеного додатку можуть бути використані для створення тренувальних вибірок для навчання нейронних мереж та інших класифікаційних алгоритмів.

Ключові слова – обробка зображень, аналіз текстури, системи підвищення ефективності, .NET, C#.

I. ВСТУП

Пандемія КОВІД-19 зумовила підвищення вимог до обробки медичних зображень [5]. Наявність постковідних ускладнень, розвиток приєднаних патологій, наприклад, пневмоній різного генезу, викликала потребу в отриманні якомога більшої та деталізованої інформації, необхідної для прийняття медичних рішень щодо діагнозу та тактики лікування [6,7]. Актуальність роботи полягає у створенні нового засобу для підтримки прийняття медичних рішень стосовно діагнозу пацієнтів шляхом порівняння медичних зображень, наприклад норми та патології, або однієї патології з іншою. Новизна одержаного програмного забезпечення полягає у тому, що за допомогою даного засобу можливо виділити ті частини зображення, які містять унікальну патологію, наприклад, пухлину [8], порівнявши зразок зі зразком норми, або відокремити ознаки однієї патології від супутньої іншої [1,2], або видалити певні

фрагменти зображення, які не мають діагностичної цінності. Результат роботи такої програми може більш чітко вказати лікарю, на якій частині зображення слід зосередити увагу для більш безпомилкового прийняття діагностичних рішень.

II. МЕТА ДОСЛІДЖЕННЯ

Метою роботи було створення нового алгоритму порівняння двох зображень шляхом виокремлення в них однакових та унікальних патернів та наступного їх аналізу.

Найбільш розповсюдженими є 3 алгоритми порівняння зображень:

1. Строге порівняння - має велику швидкість виконання, але зображення мають бути абсолютно ідентичними для отримання позитивного результату

2. Порівняння нечітких пікселів. Цей алгоритм дозволяє виявляти подібність двох зображень, незважаючи на можливу невелику

відмінність деяких пікселів. , але не дає змоги знаходити часткову подібність двох зображень.

3. Порівняння гістограм. Спрямоване на порівняння кольорових профілів, припускає деяку похибку, але абсолютно не враховує взаємне розташування пікселів[7].

Оскільки найбільш розповсюджені алгоритми порівняння зображень мають недоліки, які роблять неможливим їх використання в аналізі медичних зображень паренхіматозних органів, створення нового алгоритму, який порівнював би зображення за характеристиками текстури, є доцільним[1].

III. МАТЕРІАЛИ І МЕТОДИ

Суть розроблюваного алгоритму полягає в послідовному порівнянні двох зображень за патернами розміром $M \times N$ пікселів. Шляхом порівняння двох медичних зображень паренхіматозних органів в різних станах (наприклад, легені в нормі та при COVID-19) завдяки пошуку і відмічанню усіх патернів, які зустрічаються на обох поданих зображеннях, на зображенні з наявною патологією будуть помічені ті області зображення, що співпадають з нормою.

Створення нового засобу для підвищення якості постановки діагнозу має сенс оскільки клінічна діагностика пневмонії є особливо складною у пацієнтів літнього віку, бо симптоми менш специфічні, а декомпенсовані супутні захворювання (наприклад, серцева недостатність та гостре загострення хронічної обструктивної хвороби легень) можуть ще більше погіршити клінічну картину. Пацієнти похилого віку, як правило, повідомляють про симптоми менш точно через порушення пам'яті або марення. Нарешті, частота альтернативних діагнозів пневмонії (наприклад, тромбоемболія легеневої артерії, рак легенів, нереспіраторний сепсис тощо), а також поширеність хронічних респіраторних захворювань (наприклад, хронічна обструктивна хвороба легень та інтерстиціальна хвороба легенів) також зростає у пацієнтів літнього віку[8,9].

В результаті такої обробки лікар зможе одразу визначити масштаби ураження та окремо аналізувати області з наявною патологією. В рамках договору про співробітництво, ДУ «Національний інститут фізіатрії та пульмонології ім. Ф.Г. Яновського НАМН України» було надано анонімізовану підбірку

отриманих за допомогою комп'ютерної томографії (КТ) медичних зображень легень пацієнтів у чотирьох станах: здорові, пневмонія, COVID-19, COVID-19 та пневмонія.

Для дослідження достовірності роботи алгоритму з бази зображень обиралися зрізи, які найяскравіше демонструють відмінності між легеньми здорового пацієнта та хворого на COVID-19. Було обрано по п'ять перерізів (слайсів) здорових легень та при COVID-19. Для кожного зі зрізів було виділено по шість зображень, на яких представлена виключно текстура легень.

На рисунку 1 наведено приклади отриманих зображень текстури легень в нормі та при COVID-19.

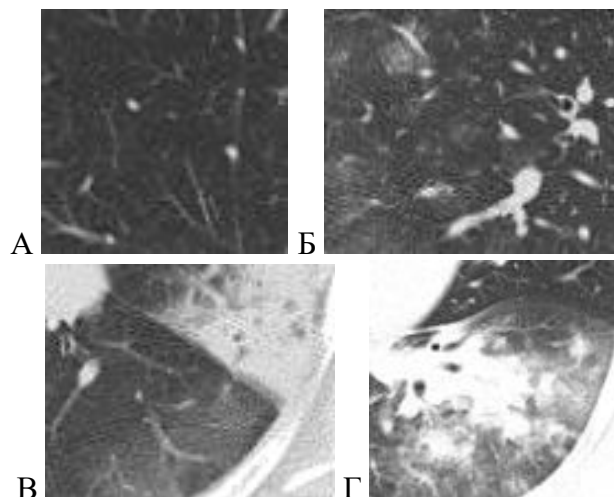


Рис. 1: А – текстура здорових легень; Б – текстура легень при COVID-19; В - текстура легень при COVID-19 та пневмонії; Г - текстура легень при пневмонії.

Для реалізації програмного додатку було обрано наступне програмне забезпечення:

1. Фреймворк .NET – це безкоштовна кросплатформна платформа з відкритим вихідним кодом для створення сучасних, масштабованих і високопродуктивних настільних, веб-, хмарних і мобільних додатків. .NET включає такі компоненти: фреймворки та бібліотеки додатків, стандарт .NET, компоненти часу виконання, компілятори мов, мови програмування – C#, F# та Visual Basic, інструменти — Visual Studio, Visual Studio для Mac, Visual Studio Code та інтерфейс командного рядка (CLI) — це інструменти, які використовуються для створення, тестування та розгортання програм .NET[10].

2. Мова програмування C# – сучасна мова програмування загального призначення, яку можна використовувати для виконання широкого кола завдань і цілей, які охоплюють різні професії. C# в основному використовується в платформі Windows .NET, хоча її можна застосувати до платформи з відкритим вихідним кодом. Ця надзвичайно універсальна мова програмування — це об'єктно-орієнтована мова програмування (ООП), яка не дуже поширена — і наразі досить нова, але вже стрімко набирає популярності[11].

3. Графічна підсистема Windows Presentation Foundation – це фреймворк інтерфейсу користувача, який не залежить від роздільної здатності дисплею та використовує векторний механізм візуалізації, створений для використання переваг сучасного графічного обладнання. WPF надає повний набір функцій розробки додатків, які включають розширену мову розмітки додатків (XAML), елементи керування, прив'язування даних, макет, 2D і 3D графіку, анімацію, стилі, шаблони, документи, медіа, текст і типографіку. WPF є частиною .NET, тому можна створювати програми, які включають інші елементи API .NET[12].

Через велику кількість відтінків сірого на вихідних зображеннях потрібно дослідити необхідність попередньої обробки поданих зображень.

Оскільки це може стати проблемою для пошуку співпадинь було вирішено зменшити кількість доступних відтінків сірого на етапі обробки зображень[13]. Налаштування параметру кількості відтінків сірого покладається на користувача, таким чином він зможе досягти найкращого можливого результату.

Для зменшення кількості відтінків сірого виконуються наступні дії:

1. Розраховується відстань різниці між сусідніми відтінками за заданою користувачем кількістю відтінків сірого за формулою $difference = 256 / (shades - 1)$, де $shades$ – кількість відтінків сірого;

2. Виконується перебір матриці, в якому для кожної точки проводяться такі операції:

a. $rest = i \% difference$, де i – значення в точці, $rest$ – залишок від ділення значення в точці на відстань між сусідніми відтінками

b. Якщо $rest \geq difference/2$, тоді

$$i = i + (difference - rest)$$

Інакше

$$i = i - rest$$

На вхід програми мають подаватися два зображення у відтінках сірого (від 0 до 255). При успішному завантаженні зображень вони мають переводитися у вигляд двовимірних масивів, розмір яких співпадає з розміром зображень.

Наступним кроком користувачем задається розмірність патерна, за яким будуть відшукуватися співпадиння. Патерн має прямокутну форму та може бути розміром $M \times N$, де M – це ширина, а N – висота.

Після завантаження зображень утворюється ще пара масивів, до яких буде записано маски зображень. Їх розмір буде відрізнятися від розміру зображень в залежності від заданого користувачем розміру патерна.

Коли всі необхідні параметри задано, починається пошук усіх можливих патернів на обох зображеннях. Для цього потрібно перебрати подані зображення з метою пошуку патернів.

Патерни визначаються як однакові коли усі точки в однакових позиціях мають однакове значення. На зображенні нижче продемонстровано приклади двох пар патернів. У випадку А патерни визначаються як однакові, у випадку Б – як різні:

А	0	128	0	128
	64	255	64	255
Б	0	128	0	64
	64	255	128	255

Рис. 2: А – пара патернів, які буде визначено співпадаючими; Б – пара патернів, яку буде визначено як різні

Якщо патерн раніше не зустрічався, проводиться його пошук на обох зображеннях. Знайдені результати мають зберігатися як структура, до якої входять:

- Інформація про патерн
- Позиції кожного патерну на обох зображеннях
- Частота, з якою патерн зустрічається на кожному зображенні.

Після того, як усі патерни знайдено, відбуватиметься заповнення масок. В масках зберігаються дані булевого типу, які вказують, чи є дана частина спільною для обох зображень. Для досягнення цього послідовно

перебираються усі знайдені патерни і, якщо вони були присутні на обох зображеннях в маску записується значення "true".

Маючи заповнені маски в графічний інтерфейс користувача виводяться додаткові зображення, які накладаються на завантажені користувачем.

Результатом роботи даного програмного застосунку мають стати зображення, на яких завдяки накладеній масці можна більш досконально дослідити наявність патології на одному з зображень у випадку якщо порівнювалася норма та патологія.

До переваг такого алгоритму можна віднести:

- Виявлення уражених тканин на зображеннях.
- Діагностика дифузних захворювань на ранніх етапах.
- Здатність працювати з зображеннями, отриманими з різних джерел завдяки можливості гнучкого налаштування таких параметрів, як розмір патерна та кількість відтінків сірого на зображеннях.

Серед недоліків можна відмітити

- Низьку чутливість без попередньої обробки зображень.
- Дещо низьку швидкість виконання, але її можна підвищити шляхом оптимізацій обчислень та застосуванням багатопотокового виконання.

На рисунку 3 наведені тестові зображення, на яких перевіряється коректність роботи алгоритму пошуку співпадінь на зображеннях.

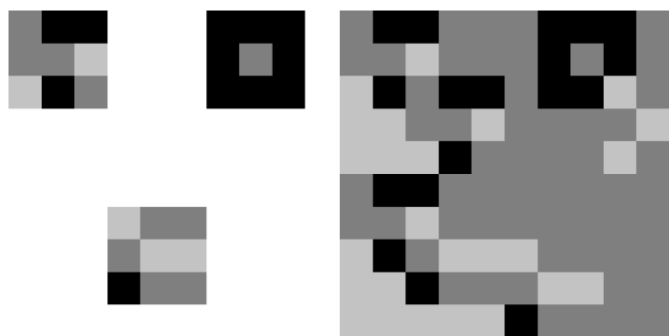


Рис. 3: Вихідні тестові зображення

На рисунок 4 продемонстровано, які були знайдені співпадіння на двох тестових зображеннях при пошуку за патерном розміром 3x3.

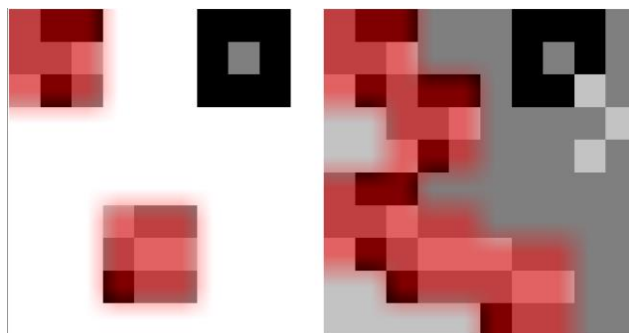


Рис. 4: Зображення зі співпадіннями, відміченими відтінками червоного кольору

IV. РЕЗУЛЬТАТИ

З рисунку 5 видно, що без попередньої обробки співпадінь за таким алгоритмом на зразках майже не було знайдено, при тому що використовувався патерн розмірності лише 2x2.

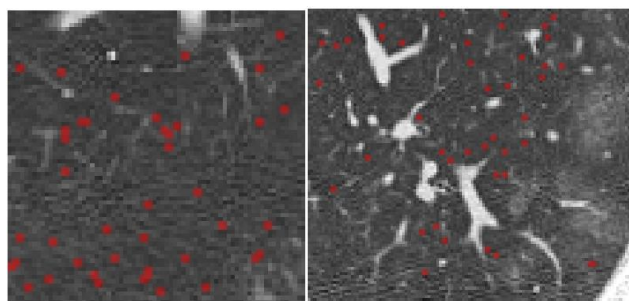


Рис. 5: Без попередньої обробки, розмір патерна 2x2

Для рисунку 6 при пошуку співпадінь було зменшено кількість відтінків сірого до 17, в результаті чого навіть зі збільшеним розміром патерну (3x4) було знайдено достатньо співпадінь щоб можна було знайти області інтересу на другому зразку.

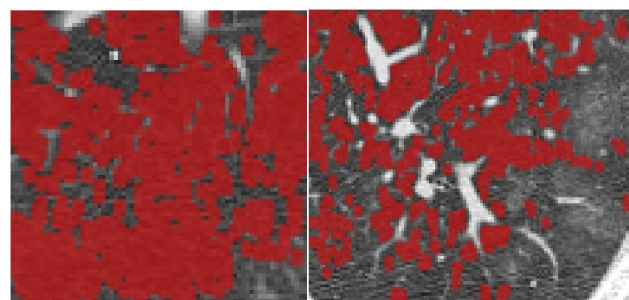


Рис. 6: Зі зменшенням кількості відтінків сірого, розмір патерна 3x4

Оскільки зображення, які подаються на вхід, від початку мають правильно встановлені параметри яскравості та контрастності, виконувати нормалізацію немає необхідності. Для перевірки коректності роботи програмного додатку на вхід було подано два зображення легень, отриманих за допомогою комп'ютерної

томографії. На зображенні 7а - легені здорового пацієнта, на зображенні 7б - легені пацієнта, хворого на COVID-19.

а

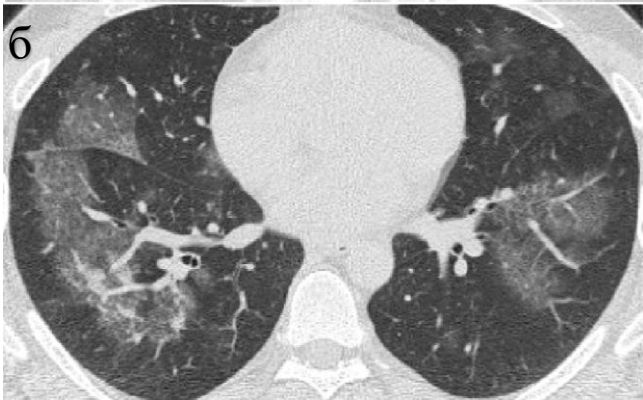
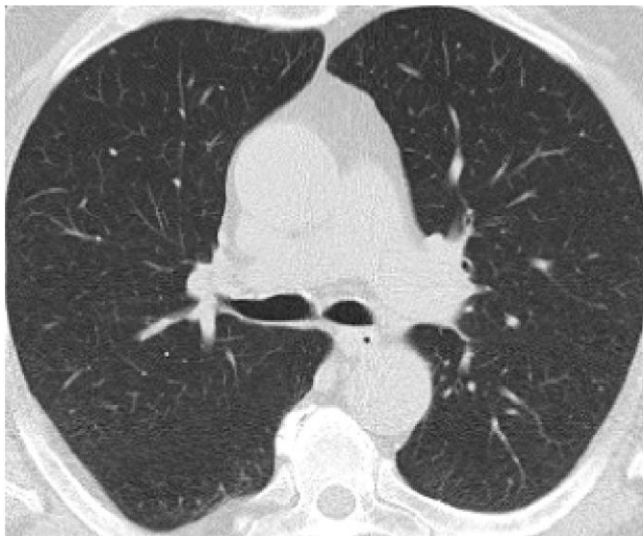


Рис. 7: А) легені здорового пацієнта, б) легені пацієнта, хворого на COVID-19.

Для порівняння зображень, результати якого наведені на рисунку 8, було встановлено такі параметри:

- Кількість відтінків сірого зменшено до 9;
- Розмір патерна 4x4.

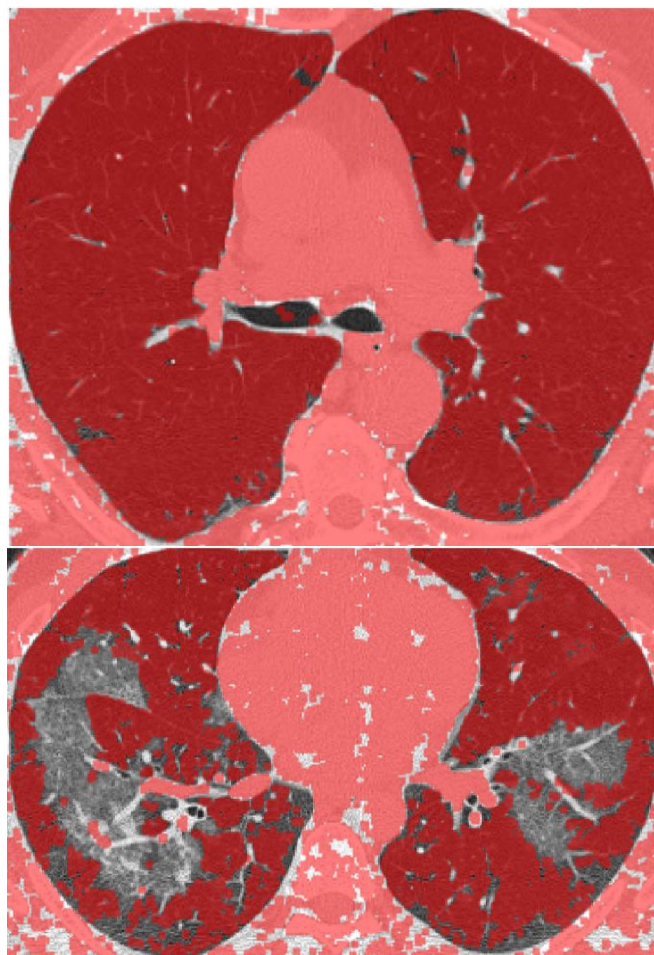


Рис. 8: Результат порівняння зображень КТ легень в нормі та при COVID-19.

З рисунку 8 видно, що навіть при безпосередньому порівнянні двох зображень легень (у нормі та при COVID-19) без вирізання частин зображення, що містять лише текстуру легень, на другому зображенні чітко видно області легень, які було уражено інфекцією.

Для рисунків 9 та 10 використовувалося зменшення кількості відтінків сірого до 17, розмір патерна становив 3x4.

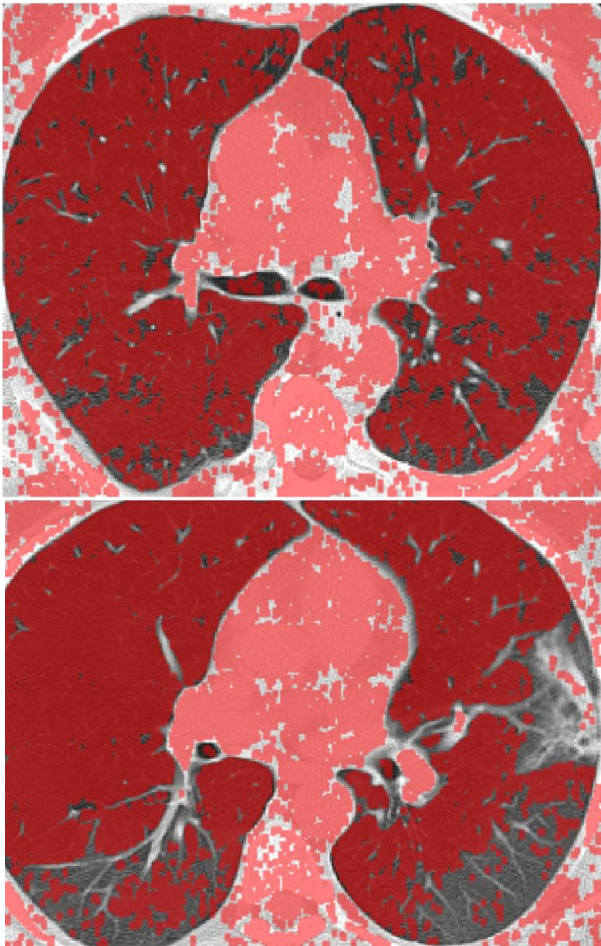


Рис. 9: Результат порівняння зображень КТ легень в нормі та при COVID-19 з пневмонією

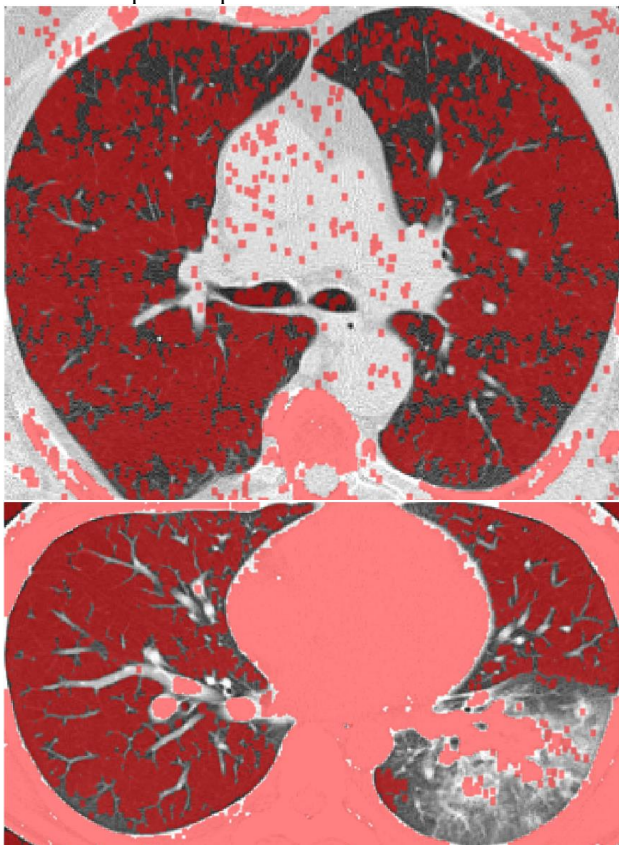


Рис. 10: Результат порівняння зображень КТ легень в нормі та при пневмонії

З рисунків 9 та 10 видно, що при тому ж безпосередньому порівнянні без вирізання текстури найбільш уражені області починають виділятися як характерні для норми. Це трапляється оскільки тканини, що зазнали фіброзу, мають світліший відтінок, через що становляться схожі на оточуючі легені тканини.

Перевірка працездатності алгоритму порівняння зображень пройшла успішно, як можна бачити за прикладом на рисунку 4. Усі співпадіння між зображеннями було підсвічено червоним, при цьому, патерни, які частково перетинаються, також коректно знаходяться і виділяються.

За рисунками 5 та 6 видно, що використання алгоритму зменшення кількості відтінків сірого дає змогу знаходити між зображеннями значно більшу кількість співпадінь, ніж без нього. При цьому, при правильному виборі розміра патерна можна отримати доволі наочно виділені уражені області.

Як можна бачити за рисунками 8, 9, 10, на вхід програмного додатку можна подавати навіть зображення, на яких наявні не тільки тканини органів, які нас цікавлять, а й оточуючі тканини. При цьому уражені хворобою ділянки на зображеннях все ще коректно виділяються. Лише при значному ступені фіброза найбільш уражені тканини можуть виділятися як характерні для норми через набування схожих кольорових характеристик. Цієї проблеми можна запобігти у випадку подання на вхід програми зображень, які містять лише текстуру досліджуваного органа.

V. ЗАКЛЮЧЕННЯ

1. В процесі розробки програмного додатку було виявлено, що в комбінації з попередньою обробкою зображення у вигляді зменшення кількості відтінків сірого алгоритм порівняння зображень дає змогу чітко виділяти на медичних зображеннях паренхіматозних органів зони, в яких наявна патологія.

2. Створене програмне забезпечення здатне підвищити точність та спростити діагностику дифузних захворювань різних систем організму завдяки автоматичному знаходженню на (зображеннях областей, які по своїм характеристикам не відповідають нормі).

3. Розроблений алгоритм дозволяє виокремити неінформативні ознаки зображення,

наприклад, здорові тканини, та спростити прийняття лікарем діагностичних заключень.

4. Розроблений алгоритм дозволяє автоматично створювати тренувальні вибірки для навчання нейронних мереж та інших класифікаційних алгоритмів.

ПЕРЕЛІК ПОСИЛАНЬ

- [1] Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, et al. Radiomics: the facts and the challenges of image analysis. *Eur Radiol Exp*. 2018 Nov 14;2(1):36. doi: 10.1186/s41747-018-0068-z.
- [2] Chen, H.J., Mao, L., Chen, Y. *et al.* Machine learning-based CT radiomics model distinguishes COVID-19 from non-COVID-19 pneumonia. *BMC Infect Dis* **21**, 931 (2021). <https://doi.org/10.1186/s12879-021-06614-6>
- [3] Philippe Lambin, Emmanuel Rios-Velazquez, Ralph Leijenaar, Sara Carvalho, Ruud G.P.M. van Stiphout, et. al). Radiomics: Extracting more information from medical images using advanced feature analysis. *European Journal of Cancer*.2012; (48): 441-446.
- [4] Xu X, Wang H, Guo Y, Zhang X, Li B, Du P, Liu Y and Lu H Study Progress of Noninvasive Imaging and Radiomics for Decoding the Phenotypes and Recurrence Risk of Bladder Cancer. *Frontiers in Oncology*. 11:704039. doi: 10.3389/fonc.2021.704039
- [5] Nicolas G., Christophe M., Max S., Jérôme S., Virginie P. Computed tomography scan contribution to the diagnosis of community-acquired pneumonia. *Curr Opin Pulm Med*. 2019 May;25(3):242-248. doi: 10.1097/MCP.0000000000000567
- [6] Pneumonia Symptoms and Diagnosis [Internet]. July 30, 2021 Available from: <https://www.lung.org/lung-health-diseases/lung-disease-lookup/pneumonia/symptoms-and-diagnosis>
- [7] Thomas L. Understanding Different Image Comparison Algorithms [Internet]. Feb 2021 Available from: <https://www.froglogic.com/blog/understanding-different-image-comparison-algorithms/>
- [8] Weinberger, S; Cockrill, B; Mandel, J. Principles of pulmonary medicine (Seventh ed.) December 26, 2017 177-183 pp.
- [9] Marino P. Mastering C# and .NET Framework. 2016, Packt Publishing
- [10] Joseph A., Eric J. C# 8.0 in a Nutshell. 2020, O'Reilly Media, Inc.
- [11] Windows Presentation Foundation documentation [Internet] 2021. Available from: <https://docs.microsoft.com/en-us/dotnet/desktop/wpf>
- [12] Shulyhin D.E., Nastenko E.A. Classification of norm / pathology in diffuse liver diseases according to the texture of ultrasound images with a reduced number of shades of gray. 2020, Biomedical Engineering and Technology

АЛГОРИТМ СРАВНЕНИЯ ДВУХ ОБРАЗЦОВ ДЛЯ АНАЛИЗА МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ ПУТЁМ СОПОСТАВЛЕНИЯ ПАТТЕРНОВ

Шулыгин Д.Э., бакалавр

shulygin.d@gmail.com

Настенко Е.А., Ph.D., D.Sc., prof.

nastenko.e@gmail.com

м. Київ, Україна

Кафедра биомедицинской кибернетики

Национальный технический университет Украины

«Киевский политехнический институт имени Игоря Сикорского»

Реферат – Целью работы была разработка и реализация нового алгоритма сравнения двух медицинских изображений паренхиматозных органов путем выделения в них одинаковых и уникальных паттернов и их последующего анализа. Для этого были рассмотрены наиболее распространенные алгоритмы сравнения изображений, среди которых были строгое сравнение, сравнение нечетких пикселей и сравнение гистограмм. Поскольку наиболее распространенные алгоритмы сравнения изображений имеют недостатки, которые делают невозможным их использование в анализе медицинских изображений, было создано программное приложение для сравнения двух медицинских изображений, которое реализует алгоритм сравнения изображений путем сопоставления паттернов, а также использует для предварительной обработки алгоритм уменьшения количества оттенков серого. Для разработки были использованы фреймворк .NET и язык программирования C#. После анализа полученных результатов работы программного приложения на изображениях легких в состояниях нормы, пневмонии, COVID-19, COVID-19 с пневмонией, а также печени в норме и циррозе было установлено, что при таком подходе предварительная обработка изображений путем уменьшения количества оттенков серого является необходимым компонентом программы, а также возможно получить удовлетворительные результаты даже при обработке изображений, содержащих дополнительную информацию об окружающих тканях, но для наилучшего результата на вход программы желательно подавать изображения, на которых содержится исключительно текстура исследуемых органов. Также результаты работы созданного приложения могут использоваться для создания тренировочных выборок для обучения нейронных сетей и других классификационных алгоритмов.

Ключевые слова – обработка изображений, анализ текстуры, системы повышения эффективности, .NET, C#.

ALGORITHM OF COMPARISON OF TWO SAMPLES FOR ANALYSIS OF MEDICAL IMAGES BY PATTERN COMPARISON

Shulyhin D.E., Bachelor student
shulygin.d@gmail.com

Nastenko E.A., Ph.D., Doctor of Biological Sciences, senior researcher
nastenko.e@gmail.com

Kiev, Ukraine

Department of Biomedical Cybernetics
National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”
Kiev, Ukraine

Abstract – *The aim of the work was to develop and implement a new algorithm for comparing two medical images of parenchymal organs by isolating identical and unique patterns in them and their subsequent analysis. For this, the most common image comparison algorithms were considered, among which were strict comparison, fuzzy pixel comparison and histogram comparison. Since the most common image comparison algorithms have drawbacks that make them impossible to use in the analysis of medical images, a software application for comparing two medical images was created that implements an image comparison algorithm by matching patterns, which also uses a grayscale reduction algorithm for preprocessing. .NET framework and C# programming language were used for development. After analyzing the results of the operation of the software application on images of the lungs in normal states, pneumonia, COVID-19, COVID-19 with pneumonia, as well as the liver in normal conditions and cirrhosis, it was found that with this approach, preliminary image processing by reducing the number of shades of gray is necessary component of the program, and it is also possible to obtain satisfactory results even when processing images contain additional information about the surrounding tissues, but for the best result, it is advisable to supply images containing only the texture of the organs that are under study to the program input. Also, the results of the created application can be used to create training samples for training neural networks and other classification algorithms.*

Keywords – *image processing, texture analysis, efficiency improvement systems, .NET, C#.*