

УДК 004.9 + 616-079.4

КЛАСИФІКАТОР ДИФЕРЕНЦІАЛЬНОЇ ДІАГНОСТИКИ АУТОІМУННОГО ГЕПАТИТУ ТА ХВОРОБИ ВІЛЬСОНА НА ОСНОВІ АНАЛІЗУ УЛЬТРАЗВУКОВИХ ЗОБРАЖЕНЬ ПЕЧІНКИ

Іванченко А.С.¹, студентка
alina142536@gmail.com

Бовсуновська К.С.¹, старший викладач
bm-kbs-fbmi@iit.kpi.ua

Дикан І.М.², д.м.н., чл.-кор.НАМН України, професор
irinadykan@gmail.com

Тарасюк Б.А.², к.т.н., доцент
btarasjuk13@gmail.com

Павлов В.А.¹, доц., к.т.н.
pavlov.vladimir264@gmail.com

Настенко Є.А.^{1,3}, с.н.с., д.б.н., к.т.н.
nastenko.e@gmail.com

¹Кафедра біомедичної кібернетики
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
м. Київ, Україна,

²ДУ "Інститут ядерної медицини та променевої діагностики НАМН України"
м. Київ, Україна,

³Національний інститут серцево-судинної хірургії імені М.М. Амосова,
м. Київ, Україна

Реферат: *Проблематика.* При інтенсивному моніторингу профілактичного огляду пацієнтів у медичних закладах первинної ланки найбільш зручно по ультразвуковим зображенням діагностувати лише наявність чи відсутність фіброзних змін печінки. Подібний підхід є найбільш ефективним при профілактиці захворювань, при цьому алгоритм класифікації визначає лише наявність патології, а уточнення діагнозу, ступінь ураження вже знайденої патології може визначатися в подальшому у спеціалізованих медичних закладах висококваліфікованим лікарем діагностом. Однак, розробка автоматизованих систем підтримки рішень при диференціації клінічно схожих захворювань завжди є актуальною задачею в медичній практиці. Однією з таких задач є диференціація аутоімунного гепатиту і хвороби Вільсона.

Мета. Розробити діагностичний алгоритм класифікації аутоімунного гепатиту і хвороби Вільсона за результатами аналізу ультразвукових зображень печінки.

Методика реалізації. Дані для виконання дослідження надано Інститутом ядерної медицини та променевої діагностики НАМН України – 9 знімків УЗД стосуються хворих на аутоімунний гепатит, 20 знімків пацієнтів з хворобою Вільсона. Об'єктами класифікації є області інтересу, що було виділено на ультразвукових зображеннях медичними фахівцями. Для збільшення об'єму навчальної вибірки та підвищення якості системи класифікації застосовано аугментацію одержаних зображень. В результаті для навчання (навчальна вибірка) та верифікації (тестова вибірка) було одержано загалом 600 областей інтересу (150 для аутоімунного гепатиту і 450 для хвороби Вільсона). Виходячи з припущення, що відмінності у характеристиках зображень класів знаходяться у відмінностях їх текстур в роботі розраховані текстурні ознаки на основі частот зустрічаємості патернів бінарного шаблону відтінків сірого. Для побудови класифікатора застосовано алгоритм Random Forest.

Результати дослідження. Загальна вибірка з 600 областей інтересу була розбита випадковим чином на навчальну (80%) і тестову (20%). Одержано модель класифікатора алгоритмом Random Forest з показниками якості класифікації на навчальній вибірці: точність - 100%, чутливість - 1, специфічність - 1, F-score -1, та на тестовій вибірці: точність 90,8% , чутливість 0.767 , специфічність – 0,956, F-score – 0,873.

Висновки. Запропоновано ефективний підхід для вирішення задачі автоматичної диференційної діагностики аутоімунного гепатиту та хвороби Вільсона. На основі текстурних ознак та алгоритму випадкового лісу була отримана високоякісна модель класифікації

Ключові слова – диференціальна діагностика, аутоімунний гепатит, хвороба Вільсона, ультразвукова діагностика, аугментація зображень, патерни, локальні бінарні шаблони, *Random Forest*.

I. ВСТУП

Пропонована робота продовжує цикл спільних робіт колективу кафедри біомедичної кібернетики Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» та фахівців ДУ "Інститут ядерної медицини та променевої діагностики НАМН України". Дані роботи розглядають можливості застосування алгоритмів аналізу ультразвукових зображень для визначення наявності патологій та ступеню уражень при дифузних захворюваннях печінки.

Класифікація патологій за медичними зображеннями являється однією з найбільш "трендових" задач біомедичної кібернетики і машинного навчання [1]. На відміну від класичних завдань класифікації багатомірних об'єктів (наприклад, визначити за клінічними характеристиками пацієнта, чи є у нього певне захворювання), медичні зображення не характеризуються напрямку ніякими ознаками в тому вигляді, який був би придатний для автоматичних алгоритмів класифікації (логістична регресія, дерево прийняття рішень, k -найближчих сусідів і т.д.). Тому для класифікації зображень вибирають один з двох шляхів: застосовуються різні класифікаційні алгоритми відповідно до специфіки зображень із застосуванням в якості ознак різноманітні текстурні характеристики [2-3], або ж застосують так звані «згорткові нейронні мережі» (CNN – Convolutional Neural Networks) [4-5], на вхід яких подаються значення інтенсивностей відтінків сірого пікселів зображення, що дає можливість виділити ознаки згортковими шарами, а глибинні шари виступають в ролі класифікатора.

Для одержання високих результатів класифікації за CNN необхідно мати великий обсяг даних навчальної вибірки (Big Data). При наявності вибірки малих обсягів, альтернативою є різноманітні класифікаційні алгоритми на основі ознак текстурного аналізу. В нашому випадку обсяг даних відповідає більш ефективному застосуванню класифікаційних алгоритмів типу Random Forest, та виділенню ознак

текстури на основі матриць суміжності та довжин відтінків сірого.

II. ПОСТАНОВКА ЗАДАЧІ

Попередні роботи даного напрямку було направлено на вирішення задач автоматизованого діагностування норма - патологія печінки по УЗ зображенням при дифузних захворюваннях: гепатози, гепатити, фібрози та цирози [6]. Результати реалізовано в системі підтримки прийняття діагностичних рішень для фахівців ДУ "Інститут ядерної медицини та променевої діагностики НАМН України" що передана в дослідну експлуатацію. Дана робота розглядає можливість створення автоматизованої системи диференціальної діагностики автоімунного гепатиту та хвороби Вільсона.

Математична формалізація задачі:

Передбачається наявність $K=2$ класів зображень (областей інтересу ОІ в цих зображеннях).

Кожен клас є нескінченна множина об'єктів (зображень ОІ) Z_i^* , $i = 1, \dots, K$. Передбачається, що $Z_i^* \cup Z_j^* = \emptyset$. Нам ці класи задаються скінченими навчальними підмножинами з кількістю об'єктів n_i , $i = 1, \dots, K$, представленими зразками ОІ z_{ij} , $j = 1, \dots, n_i$. Кожен зразок z_{ij} являє собою УЗ зображення ОІ печінки. На підставі заданих навчальних підмножин Z_i , $i=1, \dots, K$ необхідно запропонувати механізм найкращої класифікації об'єктів z_{ij}^* з Z_i^* , $i = 1, \dots, K$. в обраному класі алгоритмів класифікації.

III. ОГЛЯД ОСТАННІХ ДОСЛІДЖЕНЬ

Сенс текстурного аналізу полягає в тому, що початкове двовимірне зображення аналізується у якості відповідної двовимірної матриці GM відтінків сірого, розміром $m \times n$ (де m – кількість рядків матриці, n – кількість стовпців). Значення елементів матриці, по суті, є кодами відтінків сірого пікселів зображення, які варіюються від 0 до 255. У випадку сірошкальних

зображень, 0 позначає чорний колір пікселя, а 255 – білий колір. Всі інші значення є відтінками сірого [7]. Дану інтерпретацію можна представити на рис. 1.

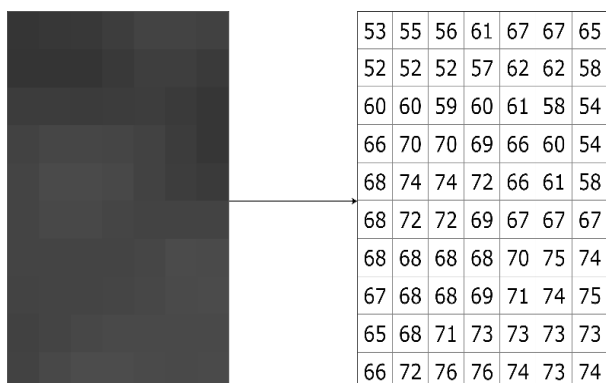


Рис. 1. Приклад зображення та відповідної матриці GM відтінків сірого. Джерело: [16]

Текстурний аналіз дає можливість отримати з подібних матриць цінну для класифікації інформацію. Такою інформацією можуть бути статистичні характеристики [9]: середнє арифметичне значення матриці, середньоквадратичне відхилення, дисперсія і т.д. у текстурному аналізі дані характеристики називаються текстурними ознаками першого порядку. Вони використовуються в різних дослідженнях по класифікації зображенні за ознаками текстурного аналізу, зокрема в роботах [10-13]. Текстурні ознаки другого порядку виходять з матриці суміжності відтінків сірого (GLCM – Gray-Level Co-Occurrence Matrix) [14] і матриці довжин пробігу відтінків сірого (GLRLM – Grey-Level Run length Matrix) [15].

Ознаки, що одержуємо з GLCM [16], називають *характеристиками Хараліка*. Кожна подібна ознака характеризує певні взаємовідношення між групами сусідніх пікселів. На особливості текстури можуть вказувати контраст зображення (великі відмінності між сусідніми пікселями) або ентропія (впорядкованість розподілу рівнів сірого в зображенні), тощо. Нижче показано, як визначаються особливості текстури:

- *Автокореляція:*

$$\sum_{i=1}^N \sum_{j=1}^N (i \cdot j) p(i, j)$$

i, j - тут і далі, рівні інтенсивностей в парі пікселів суміжності, $i, j = 0, \dots, N=255$ – кількість дискретних рівнів інтенсивності на зображенні; p – ймовірність появи пари з даним поєднанням інтенсивностей пікселів у зображенні.

- *Кластерний відтінок* - міра асиметрії та однорідності GLCM:

$$\sum_{i=0}^N \sum_{j=0}^N (i + j - 2\mu)^3 p(i, j)$$

μ – очікування інтенсивності рівня сірого p ,

- *Кластерна популярність*:

$$\sum_{i=0}^N \sum_{j=0}^N (i + j - 2\mu)^4 p(i, j)$$

- *Контраст*: відображає кількість локальних варіацій рівня сірого в зображенні;

$$\sum_{i=0}^N \sum_{j=0}^N (i - j)^2 p(i, j)$$

- *Кореляція*: вимірює лінійну залежність рівнів сірого від рівнів сусідніх пікселів; вона забезпечує міру, аналогічну методам автокореляції.

$$\sum_{i=0}^N \sum_{j=0}^N \frac{(i \cdot j) p(i, j) - \mu^2}{\sigma_i \sigma_j}$$

σ_i – стандартне відхилення по рядку, σ_j – стандартне відхилення по стовпцю.

- *Різниця ентропії*: вимірює безлад, пов'язаний з розподілом різниці рівнів сірого зображення.

$$- \sum_{k=0}^{N-1} p_{i-j}(k) \log p_{i-j}(k)$$

Де $p_{i-j}(k) = \sum_{i=1}^N \sum_{j=1}^N p(i, j)$, $k = |i - j|$ для $k = 0, 1, \dots, N - 1$, i та j – рівні інтенсивності відтінків сірого, N – кількість дискретних рівнів інтенсивності на зображенні, $p(i, j)$ – значення відносних частот відтінків сірого.

- *Різниця дисперсії*: міра неоднорідності, яка надає більшу вагу різним парам рівнів інтенсивності, які більше відхиляються від середнього.

$$\sum_{k=0}^{N-1} (k - \mu_{i-j})^2 p_{i-j}(k)$$

де μ_{i-j} – середня інтенсивність p_{i-j} .

- *Енергія*: вимірює однорідність (або впорядкованість) розподілу рівнів сірого зображення; зображення з меншою кількістю рівнів сірого мають більшу однорідність.

$$\sum_{i=1}^N \sum_{j=1}^N p(i,j)^2$$

- *Ентропія*: вимірює ступінь хаотичності інтенсивностей пікселів у зображенні; ентропія (приблизно) обернено корелює з однорідністю; зображення з великою кількістю рівнів сірого мають велику ентропію

$$\sum_{i=1}^N \sum_{j=1}^N p(i,j) \log p(i,j)$$

- *Однорідність*: вимірює гладкість розподілу рівнів сірого зображення; однорідність (приблизно) обернено корелює з контрастом - якщо контраст невеликий, зазвичай однорідність велика.

$$\sum_{i=1}^N \sum_{j=1}^N \frac{p(i,j)}{1 + (i-j)^2}$$

- *Інформаційна міра кореляції 1*: оцінює кореляцію між розподілами ймовірностей i та j (кількісна оцінка складності текстури), використовуючи взаємну інформацію

$$\frac{HXY - HXY1}{\max(HX, HY)}$$

Де $HX = - \sum_{i=1}^N p_x(i) \log_2(p_x(i) + \epsilon)$
є ентропією p_x ;

$HY = - \sum_{j=1}^N p_y(j) \log_2(p_y(j) + \epsilon)$ є ентропією p_y ;

$HXY = - \sum_{i=1}^N \sum_{j=1}^N p(i,j) \log_2(p(i,j) + \epsilon)$
є ентропією $p(i,j)$;

$HXY1 = - \sum_{i=1}^N \sum_{j=1}^N p(i,j) \log_2(p_x(i)p_y(j) + \epsilon)$;
 $p_x(i)$ – граничні ймовірності рядків; $p_y(j)$ – граничні ймовірності стовпців. ϵ – довільне мале додатне число $\sim 2.2 \cdot 10^{-16}$

- *Інформаційна міра кореляції 2*: оцінює кореляцію між розподілами ймовірностей i та j (кількісно оцінюючи складність текстури).

$$\sqrt{1 - \exp[-2(HXY2 - HXY)]}$$

де $HXY = - \sum_{i=1}^N \sum_{j=1}^N p(i,j) \log_2(p(i,j) + \epsilon)$
є ентропією $p(i,j)$;

$HXY2 = - \sum_{i=1}^N \sum_{j=1}^N p_x(i)p_y(j) \log_2(p_x(i)p_y(j) + \epsilon)$;

- *Обернена різниця*: міра локальної однорідності зображення;

$$\sum_{i=1}^N \sum_{j=1}^N \frac{p(i,j)}{1 + |i-j|}$$

- *Максимальна ймовірність*:
 $\max_{i,j} p(i,j)$

- *Сума середніх*: вимірює середнє значення розподілу суми рівнів сірого зображення.

$$\sum_{k=2}^{2N} k p_{i+j}(k)$$

Де $p_{i+j}(k) = \sum_{i=1}^N \sum_{j=1}^N p(i,j)$, $k = i + j$ для $k = 2, 3, \dots, 2N$

- *Сума ентропій*: вимірює ступінь хаотичності, пов'язаний з розподілом суми рівнів сірого зображення.

$$- \sum_{k=2}^{2N} p_{i+j}(k) \log p_{i+j}(k)$$

- *Сума квадратів*: вимірює дисперсію (відносно середнього) розподілу рівнів сірого.

$$\sum_{i=1}^N \sum_{j=1}^N (i - \mu)^2 p(i, j)$$

• *Сума дисперсій*: вимірює дисперсію (відносно середнього) розподілу суми рівнів сірого зображення.

$$\sum_{k=2}^{2N} (k - \mu_{i+j})^2 p_{i+j}(k)$$

Де $p_{i+j}(k) = \sum_{i=1}^N \sum_{j=1}^N p(i, j)$, $k = i + j$ для $k = 2, 3, \dots, 2N$, μ_{i+y} – середня інтенсивність p_{i+j} .

Всі вищезазначені ознаки можна використовувати в якості характеристик зображень для їх класифікації.

Також були розглянуті [8, 13]:

1. Характеристики спектра зображення засновані на GLCM [6].
2. Статистичні характеристики, отримані з GLRLM [6].
3. Оптимальні ансамблі пар відтінків сірого.
4. Коефіцієнти моделей просторової розгортки.
5. Коефіцієнти моделей ковзного вікна.

Більш детально характеристики описані в роботі [8]. Ідея полягає в тому, що існують різні способи отримання інформації зображень для їх класифікації. У статті розглянуто варіант розробки власного апарату для визначення ознак із УЗ зображень печінки.

IV. ОПИС КЛІНІЧНИХ ДАНИХ

Для вирішення задачі були використані дані Інституту ядерної медицини та променевої діагностики. Всього для дослідження використано: 9 зображень УЗД печінки з аутоімунним гепатитом і 20 зображень печінки з хворобою Вільсона. Приклад УЗД представлений на рис. 2.

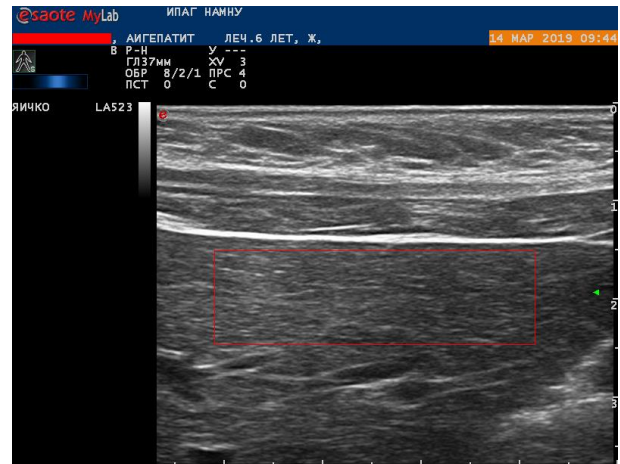


Рис. 2. Приклад зображення печінки на УЗД

Всі надані зображення УЗД є анонімними, і особиста інформація пацієнтів не може потрапити до сторонніх осіб.

На відібраних зображеннях УЗД лікарями позначені *області інтересу* (ОІ) в якості зон, обмежених червоними лініями. Дані ОІ є найбільш інформативними зонами на УЗД, для класифікації патології. Таким чином фахівцями були виділені зображення і помічені ОІ були сегментовані для отримання об'єктів класифікації. Всього виділено 10 ОІ зображень для аутоімунного гепатиту і 30 - для хвороби Вільсона.

Таким чином формуємо завдання бінарної класифікації, де першим класом є клас аутоімунного гепатиту, а другим класом – хвороба Вільсона. Завдання класифікації є дизбалансним, оскільки відношення об'єктів двох класів становить 1 до 3. Це буде враховано при подальшому дослідженні.

V. АУГМЕНТАЦІЯ ЗОБРАЖЕНЬ

Через те, що для класифікації надано досить мало зображень (40 об'єктів), пропонується застосувати техніку аугментації (збільшення) даних зображень [17], яка полягає у штучному збільшенні розміру навчального набору даних шляхом створення модифікованих версій зображень. Навчання моделей машинного навчання на більшій кількості даних призводить до отримання більш точних моделей, а методи аугментації можуть створювати варіації зображень, що поліпшує здатність моделей

узагальнювати навчальний процес на нові зображення.

Для збільшення вибірки були застосовані 2 основні операції аугментації:

1. Переворот (горизонтальний і вертикальний) - рис. 3.

2. Ротація, або поворот – на 45° і -45° - рис. 4.

Ці операції дозволили збільшити вибірку даних в 5 разів.

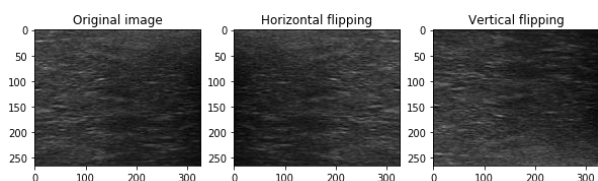


Рис. 3. Приклади перевертання зображення

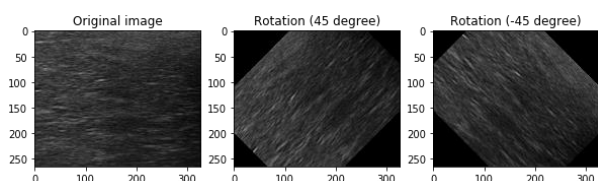


Рис. 4. Приклади повороту зображення

VI. ГІСТОГРАМНЕ ПЕРЕТВОРЕННЯ ЗОБРАЖЕНЬ

Доповнимо деякі деталі процесу одержання клінічних даних. Зображення УЗД бралися фахівцями при різному рівні контрастності, який індивідуально налаштовувався під кожного пацієнта. Через це, оригінальні значення пікселів ОІ, обраних в якості об'єктів класифікації, лежать в різних інтервалах. Для запобігання проблем з налаштуванням класифікатора автори застосували методи гістограмного перетворення зображень, для приведення їх значень в один інтервал. Були використані 2 популярних метода:

1. *max-min* нормалізація [18]:

$$I'(x,y) = \frac{I(x,y) - \min}{\max - \min} \quad (1)$$

min та *max* – мінімальне та максимальне значення яскравості з діапазону рівнів яскравості зображення, $I(x,y)$ – піксель зображення.

2. Еквалізація [19]:

$$I'(x,y) = \min + \frac{(\max - \min)}{N} \cdot \sum_{i=\min}^k n_i \quad (2)$$

N – загальна кількість пікселів зображення; *min* та *max* – мінімальне та максимальне значення яскравості з діапазону рівнів яскравості зображення; k – рівень яскравості який дорівнює значенню яскравості $I(x,y)$; n_i – загальна кількість пікселів, що мають i -й рівень яскравості.

Використання даних методів жодним чином не впливає на текстуру зображень (вони залишаються незмінними), таким чином, їх додавання в загальну вибірку не впливає на фінальні результати. До них також були використані вищеописані операції аугментації, таким чином загальна вибірка зображень склалася з: 150 об'єктів аутоімунного гепатиту і 450 об'єктів хвороби Вільсона.

VII. ОТРИМАННЯ ТЕКСТУРНИХ ОЗНАК

Для того, щоб виділити текстурні ознаки з областей інтересу, було запропоновано застосувати операцію локального бінарного шаблону LBP – Local Binary Pattern. Це підхід до текстурного аналізу, який позначає пікселі, визначаючи їх межі близькості, створюючи бінарний (двійковий) результат [20].

Оператор LBP був представлений з урахуванням особливостей текстури, застосований до сітки 3x3, де центральне значення позначає поріг. Код LBP сформульований наступним чином: порогові значення масштабуються за відповідними вагами пікселів. Оскільки сусідство визначається вісьмома пікселями, створюється набір з 256 дискретних міток, що відповідає рівню сірого відносного порогу та іншим пікселям сітки. Оператор LBP здатний перетворити зображення в матрицю цілочисельних рівнів, що визначають детальну локальну структуру зображення (рис. 5). Ці рівні, або їх похідні статистичні оператори, такі як гістограма, є параметрами, що використовуються для

створення аналітики зображень.

Приклад	Порогове перетворення	Ваги розрядів
6 4 3	1 0 0	1 2 4
8 6 2	1 0 0	128 8
9 8 7	1 1 1	64 32 16

Рис. 5. Ілюстрація функції оператора LBP

В якості ознак пропонується використовувати нормалізовані частоти кожного патерну для всіх зображень.

VIII. РЕЗУЛЬТАТИ КЛАСИФІКАЦІЇ

Для класифікації ОІ було застосовано алгоритм випадковий ліс. Даний метод був обраний, оскільки дозволив одержувати найкращі результати класифікації в попередніх роботах класифікації норма-патологія печінки у порівнянні з іншими алгоритмами класифікації, що дає підставу твердити про нього як основного конкурента методам глибокого навчання.

Для отримання моделей вибірку даних було поділено на тренувальну (80%) і тестову (20%). Для оцінки моделей використовувалися 4 метрики: точність, чутливість, специфічність і F-оцінка (F-score). Результати продемонстровані в табл. 2.

Таблиця 1 Результати класифікації випадкового лісу

Метрика оцінки	Навчальна вибірка	Тестова вибірка
Точність	100%	90.8%
Чутливість	1	0.767
Специфічність	1	0.956
F-score	1	0.873

IX. ВИСНОВКИ

В запропонованому дослідженні розроблено класифікатор для диференційної діагностики захворювань печінки аутоімунний гепатит та хвороба Вільсона по знімках ультразвукового дослідження (УЗД) з високою якістю класифікації (90.8% точності на тестовій вибірці). Для розрахунку класифікатора був застосований алгоритм Random Forest, версії та розвинуті концепції якого [13] є, по досвіду авторів, в даний час єдиним реальним конкурентом мережевих методів моделювання, проте

являється більш швидкодійним. Високий результат дослідження дозволило отримати використання аугментації зображень, гістограмного перетворення та розрахунків текстурних ознак, як частот наявності локальних бінарних шаблонів у зображеннях.

ПЕРЕЛІК ДЖЕРЕЛ

- Gavali, P., Banu, J. S. Deep Convolutional Neural Network for Image Classification on CUDA Platform: Deep Learning and Parallel Computing Environment for Bioengineering Systems. Elsevier, 2019.
- Armi, L., Fekri-Ershad, S. Texture image analysis and texture classification methods - A review / 2019. 29 p.
- Cai, J.-H., He, Y., Zhong, X.-L., et al. Magnetic Resonance Texture Analysis in Alzheimer's disease. Academic Radiology. 2020. Vol. 27, No. 12. pp. 1774–1783.
- Yamashita, R., Nishio, M., Do, R. K. G., et al. Convolutional neural networks: an overview and application in radiology. Insights into Imaging. 2018. Vol. 9, No. 4. pp. 611–629.
- Gu, J., Wang, Z., Kuen, J., et al. Recent advances in convolutional neural networks. Pattern Recognition. 2018. Vol. 77. pp. 354–377.
- Настенко, Є. А., Дикан, І. М., Тарасюк, Б. А., та ін. Класифікація станів печінки при дифузних захворюваннях на основі статистичних показників текстури ультразвукових зображень та МГУА. Індуктивне моделювання складних систем. 2019. Vol. 11. С. 54–66.
- Castro, N. de, Garrido-Vizuetе, M. A., Robles, R., et al. Contrast in greyscales of graphs. Journal of Combinatorial Optimization. 2020. Vol. 39, No. 3. pp. 874–898.
- Настенко, Є. А., Павлов, В. А., Носовець, О., та ін. Застосування текстурного аналізу у вирішенні задачі класифікації медичних зображень. Біомедична інженерія і технологія. 2020. Vol. 4. С. 69–82.
- Lei, Y., Wu, Z. Time series classification based on statistical features. EURASIP Journal on Wireless Communications and Networking. 2020. Vol. 2020, No. 1. P. 46.
- Humeau-Heurtier, A. Texture Feature Extraction Methods: A Survey. IEEE Access. 2019. Vol. 7. pp. 8975–9000.
- Tong, J., Zhao, Y., Zhang, P., et al. MRI brain tumor segmentation based on texture features and kernel sparse coding. Biomedical Signal Processing and Control. 2019. Vol. 47. pp. 387–392.
- Настенко, Є. А., Янковий, І. О. Класифікатор стану печінки у дітей з патологією гепатобіліарної системи за текстурними статистиками ультразвукового дослідження. Біомедична інженерія і технологія. 2019. Vol. 2. С. 15–23.
- Nastenko I., Maksymenko V., Galkin A., Pavlov V., Nosovets O., Dykan I., Tarasiuk B., Babenko V., Umanets V., Petrunina O. Klymenko D. (2021) Liver Pathological States Identification with Self-organization Models Based on Ultrasound Images Texture Features. In: Shakhovska N., Medykovsky M.O. (eds) Advances in Intelligent Systems and Computing V. CSIT 2020. Advances in Intelligent Systems and

Computing, vol 1293. Springer, Cham. https://doi.org/10.1007/978-3-030-63270-0_26

14. Benco, M., Hudec, R., Kamencay, P., et al. An Advanced Approach to Extraction of Colour Texture Features Based on GLCM. *International Journal of Advanced Robotic Systems*. 2014. Vol. 11, No. 7. P. 104.

15. Sohail, A. S. M., Bhattacharya, P., Mudur, S. P., et al. Local relative GLRLM-based texture feature extraction for classifying ultrasound medical images: 2011 24th Canadian Conference on Electrical and Computer Engineering(CCECE), IEEE, May.11. pp. 001092–001095.

16. Löfstedt, T., Brynolfsson, P., Asklund, T., et al. Gray-level invariant Haralick texture features. *PLOS ONE*. 2019. Vol. 14, No. 2. P. e0212110.

17. Shorten, C., Khoshgoftaar, T. M. A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*. 2019. Vol. 6, No. 1. P. 60.

18. Kociołek, M., Strzelecki, M., Obuchowicz, R. Does image normalization and intensity resolution impact texture classification? *Computerized Medical Imaging and Graphics*. 2020. Vol. 81. P. 101716.

19. Xiong, J., Yu, D., Wang, Q., et al. Application of Histogram Equalization for Image Enhancement in Corrosion Areas. *Shock and Vibration*. 2021. Vol. 2021. pp. 1–13.

20. Elias, S. J., Hatim, S. M., Hassan, N. A., et al. Face recognition attendance system using Local Binary Pattern (LBP). *Bulletin of Electrical Engineering and Informatics*. 2019. Vol. 8, No. 1. pp. 239–245.

REFERENCES

1. Gavali P, Banu JS. Deep Convolutional Neural Network for Image Classification on CUDA Platform. In: *Deep Learning and Parallel Computing Environment for Bioengineering Systems* [Internet]. Elsevier; 2019. p. 99–122. DOI: 10.1016/B978-0-12-816718-2.00013-0.

2. L. Armi and S. Fekri-Ershad, “Texture image analysis and texture classification methods - A review,” arXiv. p. 29, 2019. doi: 10.1016/B978-0-12-816718-2.00013-0.

3. J.-H. Cai et al., “Magnetic Resonance Texture Analysis in Alzheimer’s disease,” *Acad. Radiol.*, vol. 27, no. 12, pp. 1774–1783, Dec. 2020, doi: 10.1016/j.acra.2020.01.006.

4. R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, “Convolutional neural networks: an overview and application in radiology,” *Insights Imaging*, vol. 9, no. 4, pp. 611–629, Aug. 2018, doi: 10.1007/s13244-018-0639-9.

5. J. Gu et al., “Recent advances in convolutional neural networks,” *Pattern Recognit.*, vol. 77, pp. 354–377, May 2018, doi: 10.1016/j.patcog.2017.10.013.

6. Ie. A. Nastenکو et al., “GMDH Classification for Diffuse Liver Disease States Based on Statistical Features of the Ultrasound Images Texture” [in Ukrainian], *Induktyvne modeliuвання skladnyh system*, vol. 11, pp. 54–66, 2019.

7. N. de Castro, M. A. Garrido-Vizuete, R. Robles, and M. T. Villar-Liñán, “Contrast in greyscales of graphs,” *J. Comb. Optim.*, vol. 39, no. 3, pp. 874–898, Apr. 2020, doi: 10.1007/s10878-020-00528-w.

8. Ie. A. Nastenکو et al., “Texture Analysis Application in Medical Images Classification Task Solving” [in Ukrainian], *Biomedichna injeneriya i tekhnologiya*, vol. 4, pp. 69–82, Dec. 2020, doi: 10.20535/2617-8974.2020.4.221876.

9. Y. Lei and Z. Wu, “Time series classification based on statistical features,” *EURASIP J. Wirel. Commun. Netw.*, vol. 2020, no. 1, p. 46, Dec. 2020, doi: 10.1186/s13638-020-1661-4.

10. A. Humeau-Heurtier, “Texture Feature Extraction Methods: A Survey,” *IEEE Access*, vol. 7, pp. 8975–9000, 2019, doi: 10.1109/ACCESS.2018.2890743.

11. J. Tong, Y. Zhao, P. Zhang, L. Chen, and L. Jiang, “MRI brain tumor segmentation based on texture features and kernel sparse coding,” *Biomed. Signal Process. Control*, vol. 47, pp. 387–392, Jan. 2019, doi: 10.1016/j.bspc.2018.06.001.

12. Ie. A. Nastenکو and I. O. Yankovyi, “The classifier of a state of a child via texture statistics of the liver ultrasound” [in Ukrainian], *Biomedichna injeneriya i tekhnologiya*, vol. 2, pp. 15–23, Nov. 2019, doi: 10.20535/2617-8974.2019.2.184714.

13. Nastenکو I, Maksymenko V., Galkin A., Pavlov V., Nosovets O., Dykan I., Tarasiuk B., Babenko V., Umanets V., Petrunina O. Klymenko D. (2021) Liver Pathological States Identification with Self-organization Models Based on Ultrasound Images Texture Features. In: Shakhovska N., Medykovskyy M.O. (eds) *Advances in Intelligent Systems and Computing V. CSIT 2020. Advances in Intelligent Systems and Computing*, vol 1293. Springer, Cham. https://doi.org/10.1007/978-3-030-63270-0_26

14. M. Benco, R. Hudec, P. Kamencay, M. Zachariasova, and S. Matuska, “An Advanced Approach to Extraction of Colour Texture Features Based on GLCM,” *Int. J. Adv. Robot. Syst.*, vol. 11, no. 7, p. 104, Jul. 2014, doi: 10.5772/58692.

15. A. S. M. Sohail, P. Bhattacharya, S. P. Mudur, and S. Krishnamurthy, “Local relative GLRLM-based texture feature extraction for classifying ultrasound medical images,” in 2011 24th Canadian Conference on Electrical and Computer Engineering(CCECE), May 2011, pp. 001092–001095, doi: 10.1109/CCECE.2011.6030630.

16. T. Löfstedt, P. Brynolfsson, T. Asklund, T. Nyholm, and A. Garpebring, “Gray-level invariant Haralick texture features,” *PLoS One*, vol. 14, no. 2, p. e0212110, Feb. 2019, doi: 10.1371/journal.pone.0212110.

17. C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *J. Big Data*, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1186/s40537-019-0197-0.

18. M. Kociołek, M. Strzelecki, and R. Obuchowicz, “Does image normalization and intensity resolution impact texture classification?,” *Comput. Med. Imaging Graph.*, vol. 81, p. 101716, Apr. 2020, doi: 10.1016/j.compmedimag.2020.101716.

19. J. Xiong et al., “Application of Histogram Equalization for Image Enhancement in Corrosion Areas,” *Shock Vib.*, vol. 2021, pp. 1–13, Jan. 2021, doi: 10.1155/2021/8883571.

20. S. J. Elias et al., “Face recognition attendance system using Local Binary Pattern (LBP),” *Bull. Electr. Eng. Informatics*, vol. 8, no. 1, pp. 239–245, Mar. 2019, doi: 10.11591/eei.v8i1.1439.

КЛАССИФИКАТОР ДИФФЕРЕНЦИАЛЬНОЙ ДИАГНОСТИКИ АУТОИММУННОГО ГЕПАТИТА И БОЛЕЗНИ ВИЛЬСОНА НА ОСНОВЕ АНАЛИЗА УЛЬТРАЗВУКОВЫХ ИЗОБРАЖЕНИЙ ПЕЧЕНИ

Иванченко А.С.¹, студентка
alina142536@gmail.com

Бовсуновская К.С.¹, старший преподаватель
bnk-bks-fbmi@lil.kpi.ua

Дыкан И.М.², д.м.н., чл.-кор.НАМН Украины, профессор
irinadykan@gmail.com

Тарасюк Б.А.², к.т.н., доцент
btarasjuk13@gmail.com

Павлов В.А.¹, доц., к.т.н.
pavlov.vladimir264@gmail.com

Настенко Е.А.^{1,3}, с.н.с., д.б.н., к.т.н.
nastenko.e@gmail.com

¹Кафедра биомедицинской кибернетики

Национальный технический университет Украины

«Киевский политехнический институт имени Игоря Сикорского»

²ГУ "Институт ядерной медицины и лучевой диагностики НАМН Украины

Национальный институт сердечно-сосудистой хирургии им. Н.М. Амосова,
г. Киев, Украина

Реферат:*Проблематика.* При интенсивном мониторинге профилактического осмотра пациентов в медицинских учреждениях первичного звена наиболее удобно по ультразвуковым изображениям диагностировать только наличие или отсутствие фиброзных изменений печени. Подобный подход является наиболее эффективным при профилактике заболеваний, при этом алгоритм классификации определяет только наличие патологии, а уточнение диагноза, степень поражения уже найденной патологии может определяться в дальнейшем в специализированных медицинских учреждениях высококвалифицированным врачом-диагностом. Однако, разработка автоматизированных систем поддержки решений при дифференциации клинически похожих заболеваний всегда является актуальной задачей в медицинской практике. Одной из таких задач является дифференциация аутоиммунного гепатита и болезни Вильсона.

Цель. Разработать диагностический алгоритм классификации аутоиммунного гепатита и болезни Вильсона по результатам анализа ультразвуковых изображений печени.

Методика реализации. Данные для выполнения исследования предоставлены Институтом ядерной медицины и лучевой диагностики НАМН Украины - 9 снимков УЗИ касаются больных аутоиммунным гепатитом, 20 снимков пациентов с болезнью Вильсона. Объектами классификации являются области интереса, выделенные на ультразвуковых изображениях специалистами-медиками. Для увеличения объема обучающей выборки и повышения качества системы классификации была применена аугментация полученных изображений. В результате для обучения (обучающая выборка) и верификации (тестовая выборка) было получено всего 600 областей интереса (150 для аутоиммунного гепатита и 450 для болезни Вильсона). Исходя из предположения, что различия в характеристиках изображений классов находятся в различиях их текстур в работе рассчитаны текстурные признаки на основе частот встречаемости паттернов бинарного шаблона оттенков серого. Для построения классификатора использован алгоритм Random Forest.

Результаты исследования. Общая выборка из 600 областей интереса была разбита случайным образом на учебную (80%) и тестовую (20%). Получена модель классификатора алгоритма Random Forest с показателями качества

классификации на обучающей выборке: точность - 100%, чувствительность - 1, специфичность - 1, F-score 1, и на тестовой выборке: точность 90,8%, чувствительность 0.767, специфичность - 0,956 F-score - 0,873.

Выводы. Предложен эффективный подход для решения задачи автоматической дифференциальной диагностики аутоиммунного гепатита и болезни Вильсона. На основе текстурных признаков и алгоритма случайного леса была получена высококачественная модель классификации.

Ключевые слова - дифференциальная диагностика, аутоиммунный гепатит, болезнь Вильсона, ультразвуковая диагностика, аугментация изображений, паттерны, локальные бинарные шаблоны, Random Forest.

UDC 004.9 + 616-079.4

CLASSIFIER FOR DIFFERENTIAL DIAGNOSIS OF AUTOIMMUNE HEPATITIS AND WILSON'S DISEASE BASED ON ANALYSIS OF LIVER ULTRASONIC IMAGES

*Alina Ivanchenko*¹, student
alina142536@gmail.com

*Kateryna Bovsunovska*¹, Senior teacher
bm-bks-fbmi@iit.kpi.ua

*Irina Dykan*², Doctor of Medicine Sciences,
Corresponding member of the National Academy
of Sciences of Ukraine, Professor,
irinadykan@gmail.com

*Boris Tarasiuk*², Doctor of Medicine Sciences
btarasyuk13@gmail.com

*Volodymyr Pavlov*¹, Ph.D. of Technical Sciences, Assistant professor
pavlov.vladimir264@gmail.com

Ievgen Nastenka^{1,3}, Doctor of Biological Sciences,
Ph.D. of Technical Sciences, Professor
nastenko.e@gmail.com

¹Department of Biomedical Cybernetics
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

²SI "Institute of Nuclear Medicine Diagnostic Radiology
of National Academy of Medical Sciences of Ukraine"

³National M. Amosov Institute of Cardiovascular Surgery,
Kyiv, Ukraine

Abstract: Problems. When monitoring intensive prophylactic examination of patients in primary care medical institutions, it is most convenient to diagnose only the presence or absence of fibrotic changes in the liver using ultrasound images. This approach is the most effective in disease prevention, while the classification algorithm determines only the presence of pathology. Clarification of the diagnosis, the degree of damage to the already found pathology can be determined later in specialized medical institutions by a highly qualified diagnostician. However, the development of automated decision support systems for differentiating clinically similar diseases is always an urgent task in medical practice. Differentiating autoimmune hepatitis and Wilson's disease is one such challenge.

Purpose. To develop a diagnostic algorithm for the classification of autoimmune hepatitis and Wilson's disease using the results of liver ultrasound imaging analyzes.

Implementation technique. The data for the study were provided by the Institute of Nuclear Medicine and Radiological Diagnostics of the National Academy of Medical Sciences of Ukraine. 9 ultrasound images concern patients with autoimmune hepatitis, 20 images of patients with Wilson's disease. Classification objects are areas of interest identified on ultrasound images by medical specialists. To increase the volume of the training sample and improve the quality of the classification system, the augmentation of the obtained images was applied. As a result, only 600 regions of interest were obtained for

training (training set) and verification (test set) (150 for autoimmune hepatitis and 450 for Wilson's disease). Based on the assumption that the differences in the characteristics of the images of the classes are in the differences in their textures. Texture features are calculated based on the frequency of occurrence of patterns of a binary grayscale template. The Random Forest algorithm was used to build the classifier.

Research results. The total sample of 600 areas of interest was randomly divided into study (80%) and test (20%). A model of the classifier of the Random Forest algorithm was obtained with the classification quality indicators on the training set: accuracy - 100%, sensitivity - 1, specificity - 1, F-score 1, and on the test set: accuracy 90.8%, sensitivity 0.767, specificity - 0.956 F -score - 0.873.

Conclusions. An effective approach is proposed for solving the problem of automatic differential diagnosis of autoimmune hepatitis and Wilson's disease. Based on texture features and a random forest algorithm, a high-quality classification model was obtained.

Key words - *differential diagnosis, autoimmune hepatitis, Wilson's disease, ultrasound diagnostics, image augmentation, patterns, local binary patterns, Random Forest.*