

ГЕНТОШ ЛЕСЯ

Національний університет «Львівська політехніка»

<https://orcid.org/0000-0002-4957-1512>e-mail: lesia.i.mochurad@lpnu.ua

МОЧУРАД ЮЛІАННА

Львівський національний медичний університет імені Данила Галицького

<https://orcid.org/0009-0008-4486-7139>e-mail: yuliannamochurad@gmail.ua

ВАСИЛАШКО ДАРИНА

Національний університет «Львівська політехніка»

<https://orcid.org/0009-0009-4068-1094>e-mail: daryna.vasylashko.shi.2022@lpnu.ua

ПАРАЛЕЛЬНИЙ АНАЛІЗ МЕДИЧНИХ ТЕРМОГРАФІЧНИХ ЗОБРАЖЕНЬ ЗА ДОПОМОГОЮ МЕТОДУ ОПОРНИХ ВЕКТОРІВ ТА НАЇВНОГО БАЙЄСІВСЬКОГО КЛАСИФІКАТОРА

Система автоматизованого аналізу медичних термограм має широкий спектр застосування, охоплюючи такі галузі як медицина, спорт, військова технологія та промисловість. У медицині термографія є важливим інструментом для діагностики різноманітних захворювань, включаючи рак грудей, суглобів та інших органів, завдяки її здатності відображати відмінності у температурі, які можуть свідчити про патологічні процеси. У спортивній сфері термографія використовується для контролю та прогнозування травм, допомагаючи зменшити ризик пошкоджень та покращити ефективність тренувань. У військовій технології термографія використовується для виявлення терористів або підозрілих осіб на території, а також для локалізації поранених у зоні бойових дій, забезпечуючи швидке та точне реагування. У даній роботі пропонується паралельний підхід для автоматичної обробки термографічних зображень та виявлення можливих патологій за допомогою ансамблю методів класифікації. Цей ансамбль складається з методу опорних векторів (SVM) та наївного байєсівського класифікатора (NB), а для ухвалення кінцевого рішення використовується метод голосування. Отримано високий показник $F1\text{-Score}=0.93$, що свідчить про ефективність розробленої системи. Для прискорення часу тренування моделі використана паралелізація з використанням бібліотеки `joblib` та її функції `parallel_backend`, що призвело до показника прискорення близького до двох. Це значно підвищує продуктивність системи та забезпечує швидку обробку великих обсягів даних. Обговорено також подальші рекомендації та визначено цілі для наступних кроків, включаючи розширення функціональності системи шляхом навчання моделі на різних типах захворювань, покращення алгоритмів обробки даних та використання більш складних моделей машинного навчання для отримання більш точних результатів.

Ключові слова: ансамбль голосування, розпаралелення, медичні термограми, метод PCA, класифікація.

HENTOSH LESIA

Lviv Polytechnic National University

MOCHURAD YULIANNA

²Danylo Halytsky Lviv National Medical University

VASYLASHKO DARYNA

Lviv Polytechnic National University

PARALLEL MEDICAL THERMOGRAPHIC IMAGE ANALYSIS USING A SUPPORT VECTOR MACHINE AND A NAIVE BAYESIAN CLASSIFIER

The system for automated analysis of medical thermograms has a wide range of applications, covering such fields as medicine, sports, military technology, and industry. In medicine, thermography is an important tool for diagnosing a variety of diseases, including breast, joint, and other organ cancers, due to its ability to display temperature differences that may indicate pathological processes. In the sports industry, thermography is used to monitor and predict injuries, helping to reduce the risk of injury and improve training efficiency. In military technology, thermography is used to detect terrorists or suspicious persons on the ground, as well as to localize the wounded in a combat zone, ensuring a quick and accurate response. In this paper, we propose a parallel approach to automatically process thermographic images and detect possible pathologies using an ensemble of classification methods. This ensemble consists of a support vector machine (SVM) and a naive Bayesian classifier (NB), and a voting method is used to make the final decision. A high $F1\text{-Score}$ of 0.93 was obtained, which indicates the effectiveness of the developed system. To speed up the training time of the model, parallelization was used using the `joblib` library and its `parallel_backend` function, which led to an acceleration factor close to two. This significantly improves the system's performance and ensures fast processing of large amounts of data. Further recommendations and goals for the next steps were also discussed, including expanding the functionality of the system by training the model on different types of diseases, improving data processing algorithms, and using more complex machine learning models to obtain more accurate results.

Keywords: voting ensemble, parallelization, medical thermograms, PCA method, classification.

Постановка задачі

За останні кілька десятиліть термографія стала важливим методом діагностики, який забезпечує багато інформації про здоров'я людини та може бути використаний для ранньої діагностики різних захворювань. Система автоматизованого аналізу медичних термограм здатна забезпечити більш швидкий та точний аналіз, порівняно зі звичайним ручним аналізом, що дозволяє більш ефективно виявляти

захворювання та розпочати лікування у ранній стадії. Проте розробка такої системи є складним процесом, під час якого можуть виникнути різноманітні проблеми. Зокрема недостатня точність аналізу термографічних зображень: можуть виникнути проблеми з визначенням температурних значень на зображеннях, які можуть призвести до неточності діагностики захворювань. А також недостатня кількість даних: для розробки системи автоматизованого аналізу медичних термограм необхідно мати достатню кількість якісних даних для тренування моделі. Якщо кількість даних недостатня, то може виникнути проблема зі здатністю системи до точного визначення захворювань.

Дане дослідження спрямоване на вирішення вищезгаданих проблем. В результаті необхідно провести аналіз термограф різних захворювань. Створити систему автоматизованого аналізу медичних термограм, що сприятиме точному визначенню захворювань та допоможе медичним фахівцям у прийнятті клінічних рішень.

Нижче наведено формальне представлення задачі автоматизованого аналізу медичних термограм захворювань:

$$X_{train}, Y_{train} \rightarrow V(X_{train}, Y_{train}) \rightarrow M \rightarrow X_{test}, Y_{test} \rightarrow M(X_{test}) \rightarrow Y_{pred} \rightarrow E(Y_{test}, Y_{pred}),$$

де X_{train} - тренувальна множина вхідних даних ($X = \{x_1, x_2, \dots, x_m\}$, x_i - медична термограма, $i = 1, 2, \dots, m$), Y_{train} - тренувальна множина міток класів ($Y = \{y_1, y_2, \dots, y_m\}$, y_i - мітка класу для відповідної медичної термограми x_i , $i = 1, 2, \dots, m$), X_{test} - тестувальна множина вхідних даних, Y_{test} - тестувальна множина міток класів, $V(X_{train}, Y_{train})$ - алгоритм класифікації, який на вхід приймає дані для тренування ($V = \{v_1, v_2, \dots, v_n\}$, v_i - алгоритм класифікації, $i = 1, 2, \dots, n$), M - отримана натренована модель класифікації в результаті навчання алгоритмів класифікації VotingClassifier, $M(X_{test})$ - класифікація медичних термограм X_{test} з використанням натренованої моделі M для отримання передбачуваних міток Y_{pred} , Y_{pred} - отримані передбачення, $E(Y_{test}, Y_{pred})$ - оцінювання моделі на основі отриманих передбачень Y_{pred} та реальних значень Y_{test} .

Аналіз останніх джерел

Відомо, що термографія може бути корисною для виявлення раку грудей, оскільки ракові клітини зазвичай мають вищу температуру, ніж навколишні тканини [1]. Також останні дослідження показують, що термографія може бути корисною для виявлення інших захворювань, таких як артрит, діабет, пульсова ішемія та інші [2]. Крім того термографія може бути ефективною для моніторингу лікування раку грудей, оскільки зміни в термографічних зображеннях можуть вказувати на ефективність лікування та прогресію захворювання [3]. Одним із переваг термографії є те, що вона є неінвазивним методом, який не потребує введення радіоактивних або інших шкідливих речовин в організм пацієнта. Термографія може бути ефективним методом для моніторингу психічного стану пацієнтів, оскільки зміни в температурі можуть вказувати на стрес, депресію та інші психічні захворювання [4]. Термографія може бути корисною для виявлення захворювань, пов'язаних зі змінами температури внутрішніх органів, таких як гепатит та інші захворювання печінки.

У [5] запропоновано доступне та портативне рішення для отримання інфрачервоних зображень для дистанційного моніторингу здоров'я грудей. Запропонована система інтегрує модуль отримання інфрачервоного зображення, модуль скринінгу і модуль передачі. Оцінено продуктивність запропонованого рішення, яке продемонструвало точність класифікації 96,46%, що вважається конкурентоспроможним результатом порівняно з сучасними схемами. Проте експерименти було проведено лише на 71 пацієнті.

У статті [6] представлені зусилля Американської академії термоології (ААТ) зі стандартизації формату файлів цифрових зображень і зв'язку в медицині (DICOM) для медичної термоології. Також представлені спроби впровадження програмного забезпечення системи архівації та передачі зображень (PACS) для обробки великої кількості зображень для програм термографії з великими даними. У статті не наведено готового до використання програмного забезпечення чи методу.

Дослідження [7] завершується основним висновком про те, що для задачі прогнозування раку молочної залози на основі вихідного спектру з використанням навчальних даних 680 термограм досягнуто 95,8%. Поточний підхід продемонстрував значне покращення порівняно з раніше опублікованою точністю 93,30% із 50 термограмами. У роботі не проведено перевірку запровадженої технології для інших захворювань.

Робота [8] демонструє алгоритм SFDWT-SPS для об'єднання термографічних та мамографічних зображень для подальшого виявлення раку молочної залози. Важливо відзначити, що при порівнянні запропонованого способу з стандартними методами, було виявлено, що запропонований метод призвів до кращих візуальних і статистичних результатів. Проте автори все ж зазначають необхідність покращення параметрів запропонованого підходу.

Незважаючи на значний прогрес, досягнутий у дослідженнях у галузі аналізу медичних термограф, все ще існує необхідність у покращенні та вдосконаленні цих методів. Отже, **мета нашої роботи** полягає в розробці ефективного підходу щодо автоматичної обробки термографічних зображень та виявлення можливих патологій з високою точністю та швидкістю.

Виклад основного матеріалу

Для аналізу медичних термограм у роботі з метою покращення точності класифікації створено ансамбль класифікаторів SVM [9] і Naive Bayes [10], а для прийняття кінцевого рішення використано метод голосування. При цьому кожен класифікатор може бути навчений на різних ознаках термограм, або з

використанням різних параметрів, що може допомогти виявити різні аспекти класифікації. При використанні методу голосування для прийняття кінцевого рішення, кожен класифікатор повертає вектор ймовірностей для кожного класу. Ці вектори ймовірностей об'єднуються і середнє значення кожного класу обчислюється для кожної точки термограми. Кінцеве рішення приймається на основі найбільш високої середньої ймовірності.

Цей підхід може допомогти знизити ризик перенавчання (overfitting) і покращити загальну точність класифікації. Використання ансамблю може також допомогти зменшити ефект шуму в даних термограми і зробити класифікацію більш стійкою до змін в навчальних і тестових наборах даних. Для реалізації ансамблювання на основі голосування можна використати VotingClassifier [11] з бібліотеки scikit-learn. Формула для такого голосування виглядає наступним чином: $f(x) = \frac{1}{N} \cdot \sum f_i(x)$, де $f_i(x)$ - i -а функція класифікації, N - кількість класифікаторів.

Для покращення результатів та зменшення часу тренування ансамблю моделей, використовується паралелізація тренування ансамблю моделей SVM та Naive Bayes. Процес полягає в одночасному виконанні тренування кожної моделі на окремому ядрі процесора. Для цього можна використовувати різні підходи, наприклад, Python бібліотеку joblib або multiprocessing. В даному випадку ми використовуємо бібліотеку joblib із контекстним менеджером parallel_backend, що дозволяє запустити тренування на декількох ядрах процесора. Нижче наведено математичну формулу процесу ансамблювання двох моделей:

$$CombinedModel(X) = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^M w_j f_j^{(i)}(X),$$

де N – кількість моделей у ансамблі, M – кількість класифікаторів в кожній моделі, відповідно, у нашому випадку $N=2$, $M=1$. X – матричне представлення зображення медичної термограми, $f_j^{(i)}$ - функція алгоритму відповідного класифікатора (SVM чи Naive Bayes), w_j – значення ваг, які автоматизовано генеруються та оптимізуються під час виклику функції тренування моделі.

Паралелізація дозволяє зменшити час тренування і поліпшити результати. Швидкість тренування залежить від кількості процесорів, які використовуються, і розміру даних, які потрібно обробити. Більшість алгоритмів машинного навчання мають складність $O(n^2)$ або більше, що означає, що час тренування зростає дуже швидко зі збільшенням розміру даних. Таким чином, паралелізація дозволяє ефективно використовувати наявні ресурси та прискорити процес тренування. У нашій роботі підзадачами паралелізації є паралельне тренування моделей класифікації SVM та Naive Bayes, результати яких в подальшу будуть об'єднуватись за принципом Voting Classifier.

Паралельний алгоритм можна описати наступним чином:

- Розділити навчальний набір даних на рівні підмножини, які будуть розподілені між різними ядрами процесора. Якщо неможливо розділити дані на рівні частини, тоді дані розділяються на стільки рівних частин, скільки можливо, але при цьому до останньої частини додаються дані, які залишились. Таким чином усі дані будуть розподілені між ядрами. Такий принцип реалізується за допомогою додаткової перевірки довжини кожної підмножини: при довжині меншій за вказану - попередня визначена підмножина доповнюється цими даними, що залишились. У досліджуваній задачі відбувається розподіл даних на дві частини, оскільки відбувається паралельне тренування двох моделей, таким чином, якщо даних буде непарна кількість, то на одному з ядер процесора буде проводитись навчання моделі на основі підмножини, яка містить на одне зображення більше, ніж інша.
- Кожне ядро процесора отримує підмножину даних та одну з двох моделей для тренування (SVM або Naive Bayes).
- Для кожної моделі SVM та Naive Bayes тренується окремо на відповідному підмножині даних.
- Після закінчення тренування, навчені моделі для SVM та Naive Bayes повертаються до основної програми.
- Основна програма створює VotingClassifier, який містить навчені моделі для SVM та Naive Bayes.
- VotingClassifier може використовуватись для класифікації нових даних.

Цей підхід дозволяє ефективно використовувати обчислювальні можливості кожного ядра процесора, що дозволяє значно зменшити час тренування моделі та покращити її точність. Якщо використовуються k класифікаторів, а обробляються N записів з d ознаками, то складність ансамблю з паралельним тренуванням можна визначити як $O(k \cdot (N/k) \cdot d)$, що спрощується до $O(N \cdot d)$, що є складністю найбільш простого класифікатора з однією моделлю.

Для підготовки системи автоматичного аналізу медичних термограм використовуються теплові зображення, які можна знайти в базі даних, створеній для досліджень з інфрачервоним зображенням (DMR-IR). Ця база даних формується шляхом збору ІЧ-зображень із лікарні Університету UFF та може бути оприлюднена лише за згодою пацієнтів, які повинні підписати документ зі згодою на використання своїх зображень. На Рис. 1 продемонстровано термограми грудних клітин жінок.

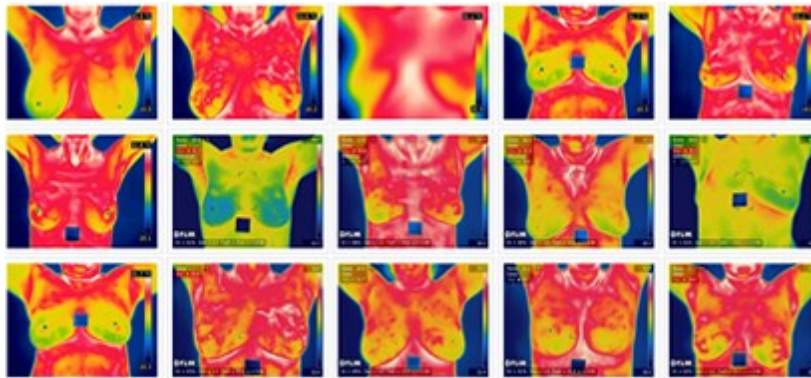


Рис. 1. Термограма грудних клітин жінок

Для даного дослідження було використано набір з 1000 фронтальних термограм з цієї бази даних, які були отримані з використанням ІЧ-камери FLIR SC-620 з роздільною здатністю 640×480 пікселів. Набір включав 500 нормальних і 500 аномальних зображень грудей різних розмірів та форм.

Першим кроком попередньої обробки даних було зміна розміру зображень до максимального значення 256×256 за допомогою функції "resize_image()". Це дозволило отримати зображення однакового розміру та зменшити кількість обчислень. Далі зображення були вирівняні до одновимірного масиву за допомогою функції flattening, що відноситься до перетворення багатовимірного масиву (у цьому випадку - зображення) в одновимірний масив, при цьому всі елементи зберігають своє значення, але в одному рядку замість матриці. Таким чином кожне зображення перетворюється в одновимірний масив, щоб мати змогу передати його в наступний етап обробки. Після цього всі перетворені зображення додаються до списку. Це необхідно, оскільки подальші методи обробки даних, які будуть застосовані до цих зображень, очікують вхідні дані в одному вимірі, а не у вигляді матриці. Наступним етапом було масштабування даних до середнього значення 0 та стандартного відхилення 1. Це зроблено для забезпечення стабільності та швидкості збіжності алгоритму навчання. Останнім кроком було застосування PCA [12] для зменшення розмірності даних перед навчанням моделі. Метод PCA визначає напрямки, в яких даний набір даних має найбільшу дисперсію, тобто напрямки, які найкраще описують даний набір даних. Потім метод зменшення розмірності вибирає підмножину з цих напрямків (головні компоненти), які найкраще описують дані. Отримані головні компоненти використовуються як нові ознаки, які передаються до моделі для навчання. В результаті попередньої обробки даних були отримані дані, які можна використовувати для тренування моделі класифікації.

У медичній термографії, F1-score може використовуватись для оцінки ефективності класифікації різних типів тканин. Наприклад, в дослідженнях з діагностикою раку грудей, F1-score може використовуватись для оцінки ефективності виявлення пухлин, діагностики захворювань молочних залоз та інших патологій. Під час експериментів, найкращим показником F1-Score було 0.9345678911. Таке значення є досить високим показником F1-Score, що свідчить про якісні результати моделі при класифікації медичних термограм.

Для перевірки системи було застосовано термограму грудної клітини з пацієнта, у якого немає проблем зі здоров'ям, та з хворого. Ці зображення не були включені до набору даних, що використовувався для тренування моделі. На рис. 2 зображено термограму, що використовувалась для оцінки стану здоров'я пацієнта з нормальною термограмою, відповідно результати роботи системи для цього випадку представлені на рис. 3.

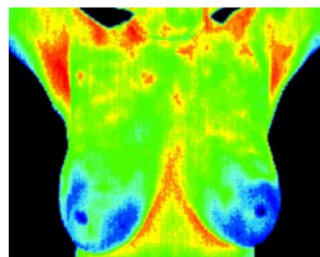


Рис. 2. Термограма здорового пацієнта

```
Predictions for image C:/Users/Iryna.Dosiak/PycharmProjects/Medical/extra_data/healthy_person.jpg
Probability that patient is healthy: 91.56%
Probability that patient is sick: 8.44%
```

Рис. 3. Результати роботи системи із термограмою здорового пацієнта

На Рис. 4 зображено термограму, що використовувалась для оцінки стану здоров'я пацієнта з аномальною термограмою, а на рис. 5 відображено результати.

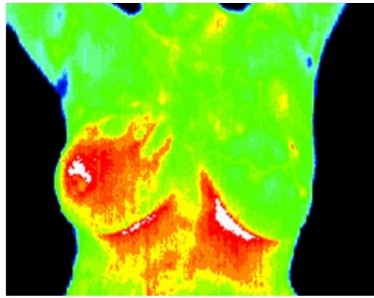


Рис. 4. Термограма хворого пацієнта

```
Predictions for image C:/Users/Iryna.Dosiak/PycharmProjects/Medical/extra_data/sick_person.jpg
Probability that patient is healthy: 9.05%
Probability that patient is sick: 90.95%
```

Рис. 5. Результати роботи системи із термограмою хворого пацієнта

Під час нашого дослідження ми застосовували паралелізацію з метою прискорення тренування моделі. Для цього ми використовували функцію `parallel_backend` з бібліотеки `joblib`, що дозволяє запускати тренування моделі з використанням багатопроцесорності.

У результаті проведеного дослідження ми підтвердили, що використання паралелізації дійсно дозволяє зменшити час тренування моделі. Проте конкретні результати залежать від конфігурації системи та параметрів моделі. Наприклад, у нашому випадку час тренування ансамблю з моделей SVM з перехресною перевіркою та NB з дефолтним набором параметрів без використання паралелізації складав близько 4 хвилини, що можна вважати прийнятним часом, оскільки набір тренувальних даних був невеликим.

Проте у випадку з більш об'ємними наборами даних, де час тренування може займати години або навіть дні, використання паралелізації може значно зменшити час, необхідний для тренування моделі. У дослідженні цієї роботи час тренування ансамблю з тими ж моделями та тією ж конфігурацією з використанням паралелізації становив 2.8 хв. Дослідивши час виконання тренування ансамблю моделей при послідовному підході та при паралельному, можна обчислити показник отриманого прискорення за наступною формулою:

$$S_p(n) = \frac{T_1(n)}{T_p(n)},$$

де $T_1(n)$ - часова складність послідовного виконання алгоритму, $T_p(n)$ - часова складність паралельного виконання алгоритму. Тоді $S_p(n) = 1.875$.

У проаналізованих джерелах було показано, що використання паралелізації при тренуванні кількох різних моделей для подальшого їх поєднання за принципом Voting Classifier може забезпечити прискорення близько 3-4 рази в порівнянні з послідовним запуском навчання. Що ще раз підтверджує ефективність такого підходу. У цій роботі було отримано прискорення майже у 2 рази, при цьому слід враховувати, що набір даних не є надто великим, тому отриманий результат можна вважати теж суттєвим.

Висновки

У роботі була розглянута проблема автоматизованого аналізу медичних термограм з метою класифікації зображень на основі ознак, що характеризують наявність захворювання. Для цього проведено дослідження та аналіз існуючих методів та алгоритмів для аналізу медичних зображень, проаналізовано набір даних, проведено попередню обробку та візуалізацію даних. Далі, було проведено тренування та налаштування ансамблю методів класифікації моделі, результати оцінено за допомогою метрики точності F1-Score. Для прискорення часу тренування моделей було використано розпаралелення. Запропонований підхід має перевагу у тому, що він не потребує значного обсягу вільного місця на комп'ютері та не схильний до перенавчання.

Проте, для того, щоб система розроблена на основі запропонованого паралельного підходу була більш гнучкою та здатною до застосування в різних медичних задачах, необхідно розширити її функціональність. Зокрема, доцільно натренувати модель на різних типах захворювань, а не лише на захворюваннях грудної клітини. Також, можна покращити алгоритми обробки даних та використовувати більш складні моделі машинного навчання, щоб отримати більш точні результати. Для забезпечення більшої ефективності та гнучкості системи, можна використовувати різні методи оптимізації, такі як автоматичний підбір гіперпараметрів та зменшення розмірності даних.

Література

1. Pavithra P.R. The Effect of Thermography on Breast Cancer Detection / P.R Pavithra, K.S. Ravichandran, K.R. Sekar, R. anikandan // A multi-faceted Review journal in the field of Pharmacy. – 2018. – vol. 9, No 1, 2018. – P. 10-16.
2. Ng E. Y. K. Advanced integrated technique in breast cancer thermography / E. Y. K. Ng, E. C. Kee //

Journal of Medical Engineering & Technology. – 2008. – Vol. 32, No. 2. – P. 103–114.

3. Schwartz R. G. Guidelines for breast thermography / R. G. Schwartz, M. Brioschi, J. Pittman, [et al.] // Pan American Journal of Medical Thermology. – 2015. – Vol. 2. – P. 26–34.

4. Alhammad M. Automated impact damage detection technique for composites based on thermographic image processing and machine learning classification / M. Alhammad, N. P. Avdelidis, C. Ibarra-Castanedo, [et al.] // Sensors. – 2022. – Vol. 22, No. 23.

5. Krishna S. An affordable solution for the recognition of abnormality in breast thermogram / S. Krishna, B. George // Multimedia Tools and Applications. – 2021. – Vol. 80, No. 18. – P. 28303–28328.

6. Gonzalez F. J. Artificial intelligence applications in medical thermography using pacs and dicom file formats / F. J. Gonzalez, R. G. Schwartz. – 2022. – ISBN 978-1-5106-5090-9.

7. Yadav S. S. Thermal infrared imaging based breast cancer diagnosis using machine learning techniques / S. S. Yadav, S. M. Jadhav // Multimedia Tools and Applications. – 2022. – Vol. 81, No. 10. – P. 13139–13157.

8. Sathy P. K. Breast cancer detection using bimodal image fusion: thermography and mammography images / P. K. Sathy, S. Shanthi, K. Anitha, [et al.] // Onkologia i Radioterapia. – 2022. – Vol. 16, No. 6.

9. Mochurad L. A new information system for road surface condition classification using machine learning methods and parallel calculation / L. Mochurad, A. Ilkiv, O. Kravchenko // International scientific journal «Computer systems and information technologies». – 2023. – № 1 – P. 53-61.

10. Yang F. -J. An Implementation of Naive Bayes Classifier / F. -J. Yang // 2018 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2018. – P. 301-306, doi: 10.1109/CSCI46756.2018.00065.

11. Khan M. A. Voting classifier-based intrusion detection for iot networks./ M. A. Khan, M. A Khan Khattk, S. Latif, A. A. Shah, M. Ur Rehman, W. Boulila, ... & J. Ahmad,. // In Advances on Smart and Soft Computing: Proceedings of ICACIn 2021. – 2022, Springer Singapore. – P. 313-328.

12. Bajwa I. S. PCA based image classification of single-layered cloud types / I. S. Bajwa, S. I. Hyder. – 2005. – 365–369 p.

References

1. Pavithra P.R. The Effect of Thermography on Breast Cancer Detection / P.R Pavithra, K.S. Ravichandran, K.R. Sekar, R. anikandan // A multi-faceted Review journal in the field of Pharmacy. – 2018. – vol. 9, No 1, 2018. – P. 10-16.

2. Ng E. Y. K. Advanced integrated technique in breast cancer thermography / E. Y. K. Ng, E. C. Kee // Journal of Medical Engineering & Technology. – 2008. – Vol. 32, No. 2. – P. 103–114.

3. Schwartz R. G. Guidelines for breast thermography / R. G. Schwartz, M. Brioschi, J. Pittman, [et al.] // Pan American Journal of Medical Thermology. – 2015. – Vol. 2. – P. 26–34.

4. Alhammad M. Automated impact damage detection technique for composites based on thermographic image processing and machine learning classification / M. Alhammad, N. P. Avdelidis, C. Ibarra-Castanedo, [et al.] // Sensors. – 2022. – Vol. 22, No. 23.

5. Krishna S. An affordable solution for the recognition of abnormality in breast thermogram / S. Krishna, B. George // Multimedia Tools and Applications. – 2021. – Vol. 80, No. 18. – P. 28303–28328.

6. Gonzalez F. J. Artificial intelligence applications in medical thermography using pacs and dicom file formats / F. J. Gonzalez, R. G. Schwartz. – 2022. – ISBN 978-1-5106-5090-9.

7. Yadav S. S. Thermal infrared imaging based breast cancer diagnosis using machine learning techniques / S. S. Yadav, S. M. Jadhav // Multimedia Tools and Applications. – 2022. – Vol. 81, No. 10. – P. 13139–13157.

8. Sathy P. K. Breast cancer detection using bimodal image fusion: thermography and mammography images / P. K. Sathy, S. Shanthi, K. Anitha, [et al.] // Onkologia i Radioterapia. – 2022. – Vol. 16, No. 6.

9. Mochurad L. A new information system for road surface condition classification using machine learning methods and parallel calculation / L. Mochurad, A. Ilkiv, O. Kravchenko // International scientific journal «Computer systems and information technologies». – 2023. – № 1 – P. 53-61.

10. Yang F. -J. An Implementation of Naive Bayes Classifier / F. -J. Yang // 2018 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2018. – P. 301-306, doi: 10.1109/CSCI46756.2018.00065.

11. Khan M. A. Voting classifier-based intrusion detection for iot networks./ M. A. Khan, M. A Khan Khattk, S. Latif, A. A. Shah, M. Ur Rehman, W. Boulila, ... & J. Ahmad,. // In Advances on Smart and Soft Computing: Proceedings of ICACIn 2021. – 2022, Springer Singapore. – P. 313-328.

12. Bajwa I. S. PCA based image classification of single-layered cloud types / I. S. Bajwa, S. I. Hyder. – 2005. – 365–369 p.