

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ  
імені ІГОРЯ СІКОРСЬКОГО»  
Приладобудівний факультет**

**Кафедра автоматизації та систем неруйнівного контролю**

До захисту допущено:

Завідувач кафедри

\_\_\_\_\_ Юрій КИРИЧУК

«\_\_» \_\_\_\_\_ 20\_\_ р.

**Дипломна робота**

**на здобуття ступеня бакалавра**

**за освітньо-професійною програмою «Роботизовані і автоматизовані  
системи неруйнівного контролю та діагностики»**

**спеціальності 151 «Автоматизація та комп'ютерно-інтегровані  
технології»**

**на тему: «Програмне забезпечення для аналізу ультразвукових  
зображень»**

Виконав (-ла):

студент (-ка) IV курсу, групи ПК-91

Андреев Сергій Михайлович \_\_\_\_\_

Керівник:

доцент, к.т.н., доцент

Галаган Роман Михайлович \_\_\_\_\_

Консультант з назва розділу:

Посада, науковий ступінь, вчене звання,

Прізвище, ім'я, по батькові \_\_\_\_\_

Рецензент:

доцент, к.т.н., доцент

Добролюбова Марина Валеріївна \_\_\_\_\_

Засвідчую, що у цій дипломній роботі  
немає запозичень з праць інших авторів  
без відповідних посилань.

Студент (-ка) \_\_\_\_\_

Київ – 2023 року

**Національний технічний університет України**  
**«Київський політехнічний інститут імені Ігоря Сікорського»**  
**Приладобудівний факультет**

**Кафедра автоматизації та систем неруйнівного контролю**

Рівень вищої освіти – перший (бакалаврський)

Спеціальність – **151 «Автоматизація та комп'ютерно-інтегровані технології»**

Освітньо-професійна програма **«Роботизовані і автоматизовані системи неруйнівного контролю та діагностики»**

ЗАТВЕРДЖУЮ

Завідувач кафедри

\_\_\_\_\_ Юрій КИРИЧУК

«\_\_\_» \_\_\_\_\_ 20\_\_ р.

**ЗАВДАННЯ**

**на дипломну роботу студенту**

**Андреєв Сергій Михайлович**

1. Тема роботи «Програмне забезпечення для аналізу ультразвукових зображень», керівник роботи Галаган Роман Михайлович, доцент, к.т.н., доцент, затверджені наказом по університету від «\_\_\_» \_\_\_\_\_ 20\_\_ р.  
№ \_\_\_\_\_

2. Термін подання студентом роботи \_\_\_\_\_

3. Вихідні дані до роботи

Система, що розробляється, і компоненти, що вміщуються в неї, повинні мати максимум до 5% похибки під час проведення діагностики. Система повинна мати модуль візуалізації ультразвукового зображення для інтерпретації висновків системи медичним співробітником. Система має бути стійкою до роботи з новими даними. Система має бути придатна до потенційного розгортання на хмарних сервісах.

4. Зміст роботи: Синдром полікістозних яєчників (СПКЯ), Ультразвукова діагностика, Використання машинного навчання в ультразвуковій діагностиці, Огляд розповсюджених методів комп'ютерного зору, Імплементация алгоритму комп'ютерного зору

5. Перелік ілюстративного матеріалу (із зазначенням плакатів, презентацій тощо)		завдання видав	завдання прийняв
Детальна діаграма алгоритму			
Ілюстрація обчислень згорткової нейронної мережі			
Ілюстрація результатів GRAD-CAM			
Опис даних і моделі			
Графіки та результати роботи моделі			

7. Дата видачі завдання \_\_\_\_\_

#### Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання етапів роботи	Примітка
1	Дослідження матеріалів про полікістоз	20.04.2023	
2	Дослідження діагностики полікістозу	30.04.2023	
3	Пошук даних для вирішення завдання	05.05.2023	
4	Дослідження знайдених ультразвукових зображень	07.05.2023	
5	Опис аналітичної частини	14.05.2023	
6	Проведення попередньої обробки даних	18.05.2023	
7	Проектування та реалізація системи класифікації	26.05.2023	
8	Проектування та реалізація системи візуалізації висновків системи	28.05.2023	
9	Оформлення та проведення висновків	30.05.2023	

Студент

Андреев Сергій Михайлович

Керівник

Галаган Роман Михайлович

**Анотація:** Дана дипломна робота присвячена розробці програмного забезпечення для аналізу ультразвукових зображень з метою діагностики полікістозу яєчників. У роботі проведено дослідження засобів комп'ютерного зору та вивчено особливості діагностики полікістозу яєчників. Була розроблена експериментальна система, яка використовує згорткову власну спроектовану нейронну мережу і метод візуалізації Grad-Cam для класифікації та детекції полікістозу яєчників на ультразвукових зображеннях. Результати дослідження показали ефективність розробленої системи, що дозволяє автоматизувати процес діагностики та сприяти швидкій та точній ідентифікації полікістозу яєчників з ультразвукових даних.

**Abstract:** This thesis is devoted to the development of software for analyzing ultrasound images for the purpose of diagnosing polycystic ovaries. In this work, the research of computer vision tools was carried out and the peculiarities of ovarian polycystic disease diagnosis were studied. An experimental system has been developed that uses a convolutional self-designed neural network and the Grad-Cam visualization method for classifying and detecting polycystic ovaries on ultrasound images. The results of the study showed the effectiveness of the developed system, which allows to automate the diagnostic process and facilitate the rapid and accurate identification of polycystic ovaries from ultrasound data.

<b>Вступ .....</b>	<b>7</b>
<b>1 Синдром полікістозних яєчників (СПКЯ) .....</b>	<b>9</b>
1.1 Структура та функція яєчників .....	9
1.2 Синдром полікістозних яєчників.....	10
1.3 Нормальний фолікулогенез.....	11
1.4 Змінений фолікулогенез при PCOS.....	11
1.5 Діагностика синдрому полікістозних яєчників .....	13
1.6 Ультразвукова діагностика синдрому полікістозних яєчників.....	15
<b>2 Ультразвукова діагностика .....</b>	<b>17</b>
2.1 Принципи ультразвукової діагностики у медицині .....	17
2.2 Проблеми ультразвукової діагностики .....	19
2.3 Сучасні методи автоматизації ультразвукової діагностики у медицині	21
<b>3 Використання машинного навчання в ультразвуковій діагностиці. 24</b>	<b>24</b>
3.1 Оптимізація та градієнтний спуск.....	24
3.2 Огляд використання методів машинного навчання в ультразвуковій діагностиці у медицині .....	25
3.3 Історичний огляд методів комп'ютерного зору. ....	26
3.4 Найбільш популярні задачі комп'ютерного зору в медицині. ....	32
<b>4 Огляд розповсюджених методів комп'ютерного зору .....</b>	<b>38</b>
4.1 Зіставлення шаблонів (Template Matching) .....	38
4.2 Метод опорних векторів (SVM) .....	40
4.3 Випадкові ліси .....	42
4.4 Згорткові нейронні мережі .....	44
4.5 Візуалізація та інтерпретація рішень згорткової нейронної мережі ..	46
<b>5 Імплементация алгоритму комп'ютерного зору.....</b>	<b>49</b>
5.1 Опис завдання.....	49
5.2 Опис даних.....	49
5.3 Опис обладнання .....	49
5.4 Опис імплементации.....	50
5.5 Правила аугментації та предобробки.....	50
5.6 Опис моделі .....	51

5.7 Візуалізація методом Grad-Cam .....	54
<b>6 Висновки .....</b>	<b>55</b>
<b>7 Список літератури .....</b>	<b>56</b>

## Вступ

Автоматизація аналізу ультразвукових зображень із використанням методів машинного навчання стала перспективним підходом у медичній діагностиці. Інтеграція ультразвукової візуалізації та алгоритмів машинного навчання відкриває потенціал для підвищення точності, ефективності та об'єктивності виявлення та діагностики захворювань. У цьому дипломі ми зосередимося на застосуванні машинного навчання для діагностики синдрому полікістозних яєчників (PCOS) із застосуванням ультразвукових зображень.

PCOS - поширене ендокринне захворювання, що вражає жінок репродуктивного віку та характеризується гормональним дисбалансом та наявністю кіст у яєчниках. Ультразвукова візуалізація широко використовується для діагностики PCOS завдяки своїй неінвазивності, економічності та здатності отримувати зображення в режимі реального часу. Однак інтерпретація ультразвукових зображень потребує великого досвіду і може бути суб'єктивною, що призводить до міжспостережної варіабельності.

Метою даного диплома є розробка автоматизованої системи діагностики PCOS з використанням можливостей машинного навчання, зокрема конволюційних нейронних мереж (CNN). CNN – це клас алгоритмів глибокого навчання, які продемонстрували значний успіх у завданнях розпізнавання та класифікації зображень. Навчаючи модель CNN на великому наборі даних анотованих ультразвукових зображень, ми прагнемо забезпечити точну та ефективну класифікацію випадків PCOS і PCOS.

Крім того, ми прагнемо забезпечити інтерпретованість прогнозів CNN-моделі, використовуючи сучасні методи візуалізації, такі як Gradient-weighted Class Activation Mapping (GRAD-CAM) та Local Interpretable Model-agnostic Explanations (LIME). GRAD-CAM дозволяє нам візуалізувати області на ультразвукових зображеннях, які роблять найбільший внесок у рішення



моделі, забезпечуючи цінне розуміння діагностичного процесу. LIME, з іншого боку, пропонує більш локалізоване пояснення, генеруючи інтерпретовані пояснення окремих прогнозів, роблячи процес прийняття рішення прозорішим і надійним.

Успішне виконання даного диплома сприятиме просуванню діагностики PCOS за рахунок автоматизації аналізу ультразвукових зображень, зниження суб'єктивності та підвищення точності діагностики. Поєднання алгоритмів машинного навчання та методів інтерпретації сприятиме розробці надійного та міцного діагностичного інструменту, який може допомогти медичним працівникам у прийнятті обґрунтованих рішень та забезпеченні кращого догляду за пацієнтами.

У наступних розділах ми обговоримо: діагноз та його постановку, методологію, отримання та попередню обробку набору даних, розробку та навчання моделі, метрики оцінки, а також реалізацію методів візуалізації для інтерпретації моделі. Ми очікуємо, що результати цього диплома матимуть значні наслідки для галузі медичної візуалізації та зроблять внесок у поточні зусилля щодо розробки систем охорони здоров'я з використанням ІІ.

# 1 Синдром полікістозних яєчників (СПКЯ)

Синдром полікістозних яєчників (PCOS) – це поширене ендокринне захворювання, яке вражає жінок репродуктивного віку. Воно характеризується поєднанням гормонального дисбалансу та наявністю кіст у яєчниках. PCOS може призвести до різних ускладнень і проблем зі здоров'ям у жінок, які страждають на це захворювання.

Поширеність PCOS є значною і зачіпає приблизно 5-10% жінок репродуктивного віку в усьому світі. За оцінками щорічно діагноз PCOS ставиться мільйонам жінок. Точна причина виникнення PCOS остаточно не з'ясована, але вважається, що у розвитку грають роль як генетичні, і екологічні чинники.

PCOS також створює проблеми для демографії через його вплив на фертильність. Жінки з ПКОС можуть відчувати труднощі з природним зачаттям і часто потребують застосування допоміжних репродуктивних технологій, таких як екстракорпоральне запліднення (ЕКЗ) для досягнення вагітності. Це може призвести до збільшення витрат на охорону здоров'я та емоційних потрясінь для окремих людей та пар, які прагнуть створити сім'ю. Крім того, довгострокові наслідки для здоров'я, пов'язані з PCOS, такі як діабет та серцево-судинні захворювання, можуть вплинути на загальний благополуччя та тривалість життя людей, які страждають на це захворювання.

## 1.1 Структура та функція яєчників

Яєчники людини - це два мигдалеподібні органи, розташовані по обидва боки від матки в жіночому тазі. Кожен яєчник має розміри приблизно 3-5 см завдовжки, 1,5-3 см завширшки і 0,5-1,5 см завтовшки, хоча їх розміри можуть коливатися протягом менструального циклу. Яєчники служать основними

жіночими репродуктивними органами та відповідають за дві ключові функції: виробництво яйцеклітин та секрецію гормонів, таких як естроген та прогестерон.

Тканини яєчників можна умовно розділити на дві області: кору та медулу. У зовнішній ділянці, корі, знаходяться фолікули яєчників. Ці фолікули складаються з незрілої яйцеклітини (ооциту), оточеної гранульозними клітинами, та розвиваються з різною швидкістю протягом менструального циклу. Внутрішня область, медулла, складається переважно з пухкої сполучної тканини, кровоносних, лімфатичних судин і нервів.

Менструальний цикл - це складний процес, який відбувається в середньому кожні 28 днів і включає дозрівання фолікула яєчника, овуляцію (вихід зрілої яйцеклітини) і підготовку ендометрію (внутрішньої оболонки матки) до потенційної імплантації заплідненої яйцеклітини. Цей процес регулюється такими гормонами, як фолікулостимулюючий гормон (ФСГ) та лютеїнізуючий гормон (ЛГ), які виробляються передньою часткою гіпофіза

## **1.2 Синдром полікістозних яєчників**

Синдром полікістозних яєчників (PCOS) – поширене ендокринне захворювання, що вражає приблизно 6-10% жінок репродуктивного віку. Це багатofакторне захворювання характеризується гормональним дисбалансом, що призводить до величезної кількості клінічних проявів, включаючи:

Полікістоз: Збільшені яєчники, що містять безліч дрібних фолікулів, зазвичай розташованих периферично навколо яєчника, що надає йому вигляду "нитки перлів" при ультразвуковому дослідженні. Ці незрілі фолікули що неспроможні випустити яйцеклітину, що зумовлює ановуляції.

Гіперандрогенізм: Надмірне виробництво чоловічих гормонів (андрогенів), таких як тестостерон, що може призвести до гірсутизму (надмірного росту волосся), акне та облісіння за чоловічим типом у жінок.

Оліго-овуляція або ановуляція: Нерегулярна, нечаста або відсутня овуляція, що призводить до порушень менструального циклу та безпліддя.

Щоб зрозуміти, як змінюється динаміка фолікулів при PCOS, необхідно зрозуміти нормальний процес фолікулогенезу і те, як він порушується при цьому захворюванні.

### **1.3 Нормальний фолікулогенез.**

У типовому менструальному циклі фолікулогенез починається з набору примордіальних фолікулів із оваріального резерву. Ці фолікули розвиваються в первинні, вторинні, а потім у преантральні та антральні фолікули. Під час цього зростання гранульозні клітини, що оточують ооцит, проліферують, і утворюється заповнена рідиною порожнину, яка називається антрумом. Когорта антральних фолікулів починає подальший розвиток на початку кожного менструального циклу і один домінуючий фолікул вибирається для остаточного дозрівання. Цей домінуючий фолікул виробляє зрілу яйцеклітину, яка вивільняється під час овуляції, тоді як інші підлеглі фолікули піддаються атрезії або дегенерації.

### **1.4 Змінений фолікулогенез при PCOS.**

У жінок з ПКОС нормальний процес фолікулогенезу порушується, що призводить до накопичення у яєчниках маленьких, незрілих антральних фолікулів. Наступні фактори сприяють зміні розвитку фолікулів:

а. Гормональний дисбаланс: PCOS характеризується підвищеним співвідношенням лютеїнізуючого гормону (ЛГ) та фолікулостимулюючого гормону (ФСГ), що призводить до надмірної стимуляції клітин папки для вироблення андрогенів. Високий рівень андрогенів перешкоджає нормальному відбору та дозріванню домінуючого фолікула, що призводить до зупинки розвитку фолікулів на стадії антрального періоду.

б. Порушення функції клітин гранулози: Гранулозні клітини відіграють життєво важливу роль у розвитку фолікулів та дозріванні яйцеклітин, реагуючи на ФСГ та виробляючи естроген. При ПКОС у цих клітинах знижено експресію рецепторів ФСГ та порушено активність ароматази, що перешкоджає їх здатності перетворювати андрогени на естрогени. Це сприяє зупинці зростання фолікулів та ановуляції.

с. Інсулінорезистентність: Загальна риса PCOS, резистентність до інсуліну та гіперінсулінемія погіршують гормональний дисбаланс, стимулюючи вироблення андрогенів яєчниками та знижуючи рівень глобуліну, що зв'язує статеві гормони (SHBG). Це призводить до підвищення рівня вільних андрогенів, які негативно впливають на зростання та дозрівання фолікулів.

Точна етіологія PCOS залишається незрозумілою, але вважається, що вона включає поєднання генетичних, метаболічних і екологічних факторів. Інсулінорезистентність, запалення та ожиріння є спільними рисами у жінок з PCOS, які ще більше посилюють гормональний дисбаланс та сприяють розвитку синдрому.

Враховуючи багатогранний характер захворювання, діагноз PCOS ґрунтується на Роттердамських критеріях, які вимагають наявності принаймні двох з наступних трьох ознак: полікістоз яєчників, гіперандрогенія та оліго-овуляція або ановуляція. Рання діагностика та лікування PCOS мають вирішальне значення для полегшення симптомів, запобігання таким ускладненням, як діабет 2 типу та серцево-судинні захворювання, а також для покращення якості життя жінок, які страждають на це захворювання.

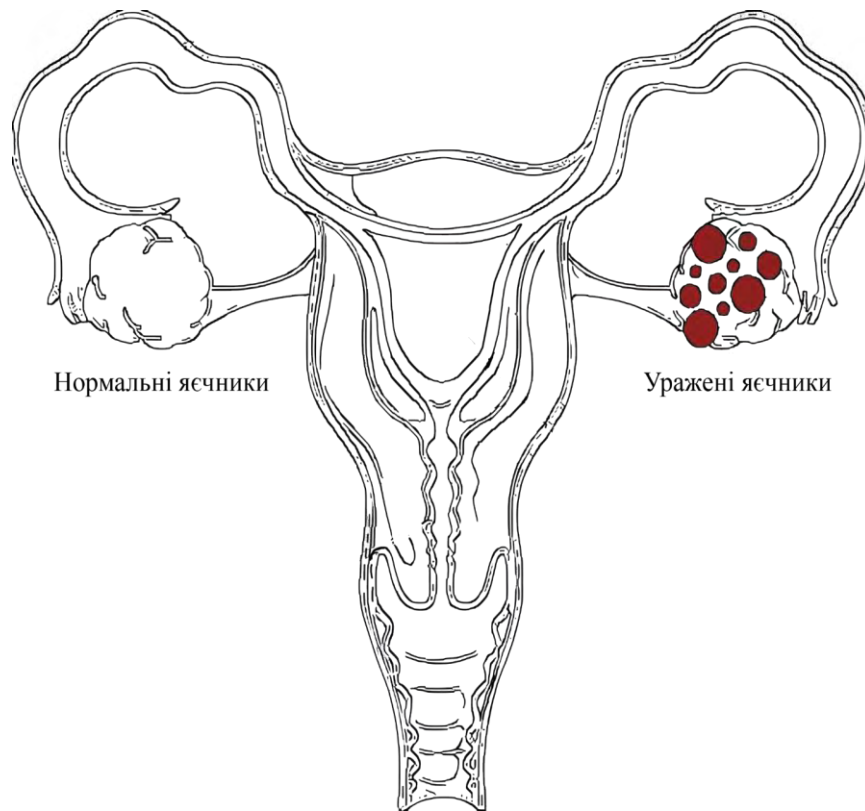


Рисунок 1.1 Ілюстрація анатомії матки

## 1.5 Діагностика синдрому полікістозних яєчників

Процес діагностики синдрому полікістозних яєчників (PCOS) включає поєднання клінічного огляду, лабораторних аналізів і візуалізуючих досліджень. Дуже важливо виключити інші захворювання зі схожими клінічними проявами, перш ніж встановити остаточний діагноз PCOS. Діагностичний підхід може бути викладений так:

Клінічне обстеження: Ретельний анамнез та фізикальне обстеження є життєво важливими для оцінки симптомів, пов'язаних з PCOS. Історія хвороби повинна включати інформацію про менструальний цикл, репродуктивний анамнез і сімейну історію PCOS або пов'язаних з ним захворювань. Під час фізичного обстеження медичний працівник шукатиме ознаки гіперандрогенії, такі як гірсутизм, акне та облісіння за чоловічим типом.

Лабораторні дослідження: Аналізи крові проводяться для оцінки гормонального рівня та виключення інших захворювань, які можуть імітувати ПКГЗ. Загальні аналізи включають:

- Рівень фолікулостимулюючого гормону (ФСГ) та лютеїнізуючого гормону (ЛГ): У жінок з ПКОС часто спостерігається підвищене співвідношення ЛГ/ФСГ, зазвичай більше 2:1.
- Рівні тестостерону, андростендіону та дегідроепіандростерону сульфату (DHEAS) у сироватці крові: Ці аналізи допомагають визначити ступінь гіперандрогенізму.
- Тести на функцію щитовидної залози (TSH та вільний T4): Ці аналізи необхідні для виключення дисфункції щитовидної залози як причини порушення менструального циклу.
- Підвищений рівень пролактину може викликати симптоми, схожі з симптомами PCOS, і тестування допомагає виключити гіперпролактинемію.
- Рівень інсуліну та глюкози у сироватці крові: Інсулінорезистентність є поширеною ознакою ПКГС, і такі тести, як швидка глюкоза, швидкий інсулін та пероральний тест на толерантність до глюкози (OGTT), можуть допомогти оцінити метаболічні порушення.

Візуалізаційні дослідження: Трансвагінальне ультразвукове дослідження є основним методом візуалізації, який використовується для оцінки морфології яєчників. Полікістозні яєчники виглядають збільшеними з безліччю дрібних периферичних фолікулів (2-9 мм у діаметрі), що часто описуються як "нитка перлів". Наявність 12 і більше фолікулів на яєчник або збільшений обсяг яєчників ( $>10 \text{ cm}^3$ ) підтверджує діагноз PCOS. Однак важливо відзначити, що полікістоз яєчників може спостерігатися у здорових жінок без PCOS.

Виняток альтернативних діагнозів: Перш ніж підтвердити діагноз PCOS, необхідно виключити стани, які можуть проявлятися подібними до PCOS клінічними ознаками, такі як вроджена гіперплазія надниркових залоз, синдром Кушинга, андрогенсекретуючі пухлини та гіпоталамічна аменорея.

Комплексний діагностичний підхід необхідний точного визначення PCOS, оскільки це забезпечує застосування відповідних стратегій лікування. Варіанти лікування можуть включати модифікацію способу життя, фармакологічні втручання та, в деяких випадках, хірургічні процедури для усунення гормонального дисбалансу, метаболічних порушень та клінічних симптомів, пов'язаних із цим захворюванням.

## **1.6 Ультразвукова діагностика синдрому полікістозних яєчників**

Ультразвукова візуалізація, зокрема трансвагінальне ультразвукове дослідження (TVUS), відіграє вирішальну роль у діагностичному процесі синдрому полікістозних яєчників (PCOS), забезпечуючи детальну візуалізацію яєчників та оцінюючи їхню морфологію. Трансвагінальне УЗД забезпечує більш високу роздільну здатність зображення порівняно з трансабдомінальним УЗД, що дозволяє більш точно ідентифікувати дрібні фолікули та покращити оцінку яєчників. Ключові аспекти ультразвукової діагностики при PCOS:

При трансвагінальному УЗД використовується високочастотний ендокавітальний зонд, який обережно вводиться в піхву для отримання зображень структур малого таза. Обстеження проводиться у літотомічному положенні пацієнтки, для підтримки гігієни використовується стерильний чохол. Для забезпечення комфорту пацієнтки під час процедури застосовується достатня кількість мастила.



Морфологія яєчників: Основною метою ультразвукового дослідження PCOS є оцінка морфології яєчників. Наступні ознаки вказують на полікістоз яєчників:

Кількість фолікулів: Наявність 12 або дрібніших (2-9 мм у діаметрі) фолікулів у кожному яєчнику є діагностичним критерієм полікістозу яєчників згідно з Роттердамським консенсусом. Фолікули зазвичай розташовуються периферично, надаючи яєчнику характерного вигляду "нитки перлів".

Об'єм яєчників: Збільшений об'єм яєчників ( $>10 \text{ см}^3$ ) є ще одним показником полікістозу яєчників. Об'єм яєчника розраховується за формулою для еліпсоїда: довжина  $\times$  ширина  $\times$  товщина  $\times 0,523$ .

Строма яєчника: При PCOS центральна стромальна область яєчника часто має вигляд гіперехогенної, щільної області при ультразвуковій візуалізації. Підвищена ехогенність строми пояснюється підвищеною продукцією андрогенів у клітинах теки та частіше зустрічається у жінок із клінічним гіперандрогенізмом.

Оцінка ендометрії: На додаток до оцінки яєчників, трансвагінальне УЗД може надати цінну інформацію про товщину та текстуру ендометрію, які можуть бути змінені при ПКГЗ через нерегулярні менструації та гормональний дисбаланс. Аномалії ендометрію, такі як потовщення або стоншення, можуть свідчити про потенційні ускладнення і можуть бути основою для подальшого дослідження.

Важливо наголосити, що хоча наявність полікістозних яєчників на УЗД є діагностичним критерієм ПКГС, цього недостатньо для самостійного діагнозу. Відповідно до Роттердамського консенсусу, для встановлення діагнозу PCOS необхідна наявність як мінімум двох з наступних трьох ознак: полікістоз яєчників, клінічний або біохімічний гіперандрогенізм, оліго-овуляція або ановуляція. Ультразвукове дослідження завжди має інтерпретуватися у контексті клінічної картини пацієнта та результатів лабораторних досліджень.

## 2 Ультразвукова діагностика

Ультразвукова діагностика, також відома як сонографія, - це неінвазивний метод візуалізації, який використовує високочастотні звукові хвилі для візуалізації внутрішніх органів, тканин та судин в організмі. Цей метод широко використовується в медицині в діагностичних цілях і надає цінну інформацію про різні захворювання та відхилення в організмі. Ультразвук є ефективним інструментом для медичних працівників у діагностиці широкого спектру медичних станів, від визначення потенційної вагітності до виявлення ракових пухлин та аномалій внутрішніх органів. У цьому дипломному розділі будуть розглянуті принципи ультразвукової діагностики, її переваги, обмеження та практичне застосування у сучасній медицині.

### 2.1 Принципи ультразвукової діагностики у медицині

Ультразвукова діагностика в медицині заснована на принципах поширення звукових хвиль та взаємодії звукових хвиль із різними тканинами всередині організму. Наступні фундаментальні принципи визначають формування та інтерпретацію ультразвукових зображень:

Генерація звукової хвилі: Ультразвукова візуалізація заснована на генерації високочастотних звукових хвиль за допомогою процесу, що називається п'єзоелектричним ефектом. П'єзоелектричні кристали, зазвичай виготовлені з таких матеріалів, як цирконат-титанат свинцю (PZT), перетворюють електричну енергію на механічні коливання під впливом прикладеної напруги. Ці коливання створюють звукові хвилі, які поширюються на тілі.

Розповсюдження звукових хвиль: Звукові хвилі, що генеруються, поширюються по тканинах тіла і зазнають декількох взаємодій. Швидкість

звуку у м'яких тканинах становить приблизно 1540 м/с. Коли звукові хвилі стикаються з різними тканинами з різними акустичними властивостями, такими як щільність і стисливість, вони піддаються відображенню, передачі, розсіювання, поглинання та заломлення.

**Відображення:** Коли звукові хвилі стикаються з межею розділу між тканинами з різним акустичним опором (твір щільності тканини та швидкості звуку), частина хвилі відбивається назад до датчика. Відбиті хвилі несуть інформацію про межі розділу тканин та використовуються для побудови ультразвукового зображення. Величина відображення залежить від кута падіння, характеру інтерфейсу та частоти ультразвукових хвиль.

**Передача:** Деякі звукові хвилі проходять через тканини і продовжують поширюватися всередину організму. Передача відбувається, коли кут падіння майже перпендикулярний до межі розділу тканин, що зводить до мінімуму відображення.

**Розсіювання:** Розсіювання відноситься до випадкового перенаправлення ультразвукових хвиль, коли вони стикаються з невеликими структурами або нерівностями тканин. Розсіяні хвилі роблять свій внесок у загальний сигнал зворотного розсіювання, одержуваний датчиком.

**Поглинання:** У міру поширення звукових хвиль по тілу вони поступово втрачають енергію через поглинання тканинами. На поглинання впливають властивості тканин, частота ультразвукових хвиль та пройдена відстань. Поглинена енергія перетворюється на тепло.

**Заломлення** Коли звукові хвилі проходять через тканини з різною швидкістю звуку, наприклад, при косому проходженні з однієї тканини в іншу, відбувається заломлення. Заломлення викликає зміни у напрямі звукових хвиль, потенційно впливаючи інтерпретацію ультразвукового зображення.

**Робота перетворювача:** В ультразвуковій візуалізації використовуються перетворювачі, які випромінюють, так і приймають звукові хвилі. Датчик складається з п'єзоелектричних елементів, які генерують звукові хвилі при подачі електричних імпульсів, і ці елементи приймають відбиті звукові хвилі.

Контролюючи час і послідовність випромінюваних імпульсів та аналізуючи отримані сигнали, можна відтворити детальні зображення внутрішніх структур.

Принципи ультразвукової діагностики математично описуються за допомогою хвильового рівняння та різних рівнянь, пов'язаних із поширенням, відображенням та згасанням звукових хвиль. Ці рівняння включають такі змінні, як частота ( $f$ ), довжина хвилі ( $\lambda$ ), швидкість звуку ( $c$ ), акустичний імпеданс ( $Z$ ) та коефіцієнт загасання ( $\alpha$ ). Конкретні математичні формулювання залежать від конкретного застосування та режиму візуалізації в ультразвуковій діагностиці.

Розуміння та застосування цих принципів дуже важливе для оптимізації отримання зображення, точної інтерпретації ультразвукових зображень та діагностики різних захворювань у медицині. Використовуючи фізику поширення та взаємодії звукових хвиль, ультразвукова діагностика стала безцінним інструментом у сучасній охороні здоров'я, полегшуючи неінвазивну візуалізацію та допомагаючи у діагностиці та лікуванні широкого спектру захворювань та розладів.

## **2.2 Проблеми ультразвукової діагностики**

Ультразвукова діагностика в медицині значно розширила можливості діагностичної візуалізації, але, як будь-яка інша медична технологія, вона не позбавлена своїх проблем і обмежень. Ось деякі проблеми, з якими стикається ультразвукова діагностика:

Обмежене проникнення в тканини: Ультразвукові хвилі мають обмежені можливості проникнення в тканини, особливо у пацієнтів з великою статурою або візуалізації структур, розташованих глибоко в тілі. Це може перешкоджати візуалізації структур, розташованих за кістками або газонаповненими

органами, що призводить до неповної візуалізації або низької якості зображення.

**Залежність від оператора:** Для отримання високоякісних ультразвукових зображень потрібні кваліфіковані оператори, які мають досвід правильного позиціонування датчика, оптимізації налаштувань візуалізації та точної інтерпретації отриманих зображень. Якість та діагностична точність ультразвукового дослідження значною мірою залежить від кваліфікації оператора, що призводить до потенційної варіабельності результатів у різних операторів.

**Артефакти зображення:** Ультразвукові зображення можуть бути схильні до різних артефактів, які можуть спотворювати або затушовувати справжню анатомічну інформацію. До поширених артефактів відносяться реверберація, затінення, акустичне посилення та артефакти заломлення. Ці артефакти можуть ускладнити інтерпретацію зображень та потенційно призвести до помилкового діагнозу чи неправильних оцінок.

**Обмежена візуалізація кісток:** Ультразвук насилу проникає в кістку, що обмежує його можливості отримання зображень структур, розташованих глибоко в кістці, або оцінці кісткової патології. Це може бути істотним недоліком в оцінці деяких захворювань опорно-рухового апарату або в оцінці переломів у певних анатомічних областях.

**Втома оператора:** Проведення ультразвукових досліджень може бути фізично важким, що вимагає від оператора твердої руки та утримання датчика у певних положеннях протягом тривалого часу. Це може призвести до втоми оператора, що може вплинути на якість зображення та потенційно збільшити ризик помилок або неправильної інтерпретації.

**Чинники пацієнта:** Певні фактори пацієнта можуть вплинути на якість та точність ультразвукової візуалізації. До таких факторів відносяться рух пацієнта, ожиріння, наявність повітря чи газу у шлунково-кишковому тракті та анатомічні відхилення. Такі фактори можуть створювати труднощі в

отриманні чітких та діагностично корисних зображень, вимагаючи додаткових зусиль або альтернативних методів візуалізації.

**Обмежена специфічність:** Хоча УЗД добре виявляє анатомічні структури, у деяких випадках воно може мати обмеження у розрізненні доброякісних та злоякісних тканин. Для підтвердження природи передбачуваної аномалії, виявленої на УЗД, можуть бути потрібні додаткові методи візуалізації або додаткові діагностичні дослідження.

**Обмеження обладнання:** Якість та можливості ультразвукового обладнання можуть бути різними, а застарілі або погано обслуговуються апарати можуть обмежувати діагностичну точність ультразвукових досліджень. Перехід на нові технології та забезпечення належного технічного обслуговування обладнання необхідні для оптимізації якості зображення та діагностичних характеристик.

Незважаючи на ці проблеми, ультразвукова діагностика залишається цінним і широко використовується методом візуалізації завдяки своїй неінвазивній природі, можливостям візуалізації в реальному часі та економічній ефективності. Постійний технологічний прогрес, покращені програми навчання та безперервні дослідження спрямовані на усунення цих обмежень та подальше підвищення корисності та діагностичної точності ультразвуку в медицині.

### **2.3 Сучасні методи автоматизації ультразвукової діагностики у медицині**

Автоматизація здатна підвищити ефективність, точність та відтворюваність ультразвукової діагностики в медицині. З розвитком технологій та алгоритмів машинного навчання було розроблено кілька методів автоматизації для допомоги медичним працівникам у різних аспектах ультразвукової візуалізації та інтерпретації. Ось деякі з сучасних методів автоматизації, що використовуються в ультразвуковій діагностиці:

Автоматизоване отримання зображення: Автоматизовані системи отримання зображень спрямовані на стандартизацію процесу ультразвукового сканування, спрямовуючи оператора при позиціонуванні датчика та оптимізуючи параметри візуалізації. Ці системи використовують заздалегідь визначені протоколи або алгоритми штучного інтелекту (ШІ) для забезпечення зворотного зв'язку в реальному часі розміщення датчика, кутового розташування та якості зображення. Завдяки зниженню варіабельності, яка залежить від оператора, автоматизоване отримання зображень може покращити послідовність отримання зображень та підвищити точність діагностики.

Комп'ютерна діагностика (CAD): CAD-системи аналізують ультразвукові зображення, щоб допомогти у виявленні та локалізації аномалій або пошкоджень. Ці системи використовують алгоритми машинного навчання для автоматичного визначення та виділення підозрілих областей на зображеннях. CAD може допомогти у ранньому виявленні патологій, таких як пухлини або кісти, забезпечуючи додатковий рівень аналізу та потенційно знижуючи помилки інтерпретації.

Сегментація та вимірювання зображень: Розробляються методи автоматизації сегментації анатомічних структур та вимірювання різних параметрів на ультразвукових зображеннях. Алгоритми машинного навчання можуть бути навчені розпізнавати конкретні анатомічні області, такі як серце чи печінку, і точно їх сегментувати. Автоматизовані вимірювання розмірів органів, швидкості кровотоку або характеристик тканин можуть надати кількісні дані для діагностики, моніторингу та планування лікування.

Глибоке навчання для інтерпретації зображень: Алгоритми глибокого навчання, зокрема згорткові нейронні мережі (CNN), застосовуються для автоматизації інтерпретації зображень в ультразвуковій діагностиці. Ці алгоритми навчаються великих масивах даних анованих ультразвукових зображень вивчення закономірностей і особливостей, що з певними станами чи аномаліями. Після навчання CNN можуть допомогти в ідентифікації та

класифікації різних патологій, таких як пухлини, кісти або судинні аномалії шляхом аналізу особливостей зображення. Оптимізація робочого процесу. Методи автоматизації також використовуються для оптимізації робочого процесу ультразвукової діагностики. Автоматизовані системи можуть допомогти в архівуванні зображень, створенні звітів та керуванні даними. Завдяки інтеграції із системами електронних медичних карток (EHR) засоби автоматизації можуть спростити процес документування, покращити комунікацію та підвищити загальну ефективність ультразвукових досліджень.

Забезпечення якості та стандартизація: Методи автоматизації використовуються для забезпечення якості та стандартизації в ультразвуковій діагностиці. Автоматизовані інструменти можуть контролювати та оцінювати якість ультразвукових зображень у режимі реального часу, забезпечуючи зворотний зв'язок з операторами та попереджаючи їх про потенційні проблеми, такі як неоптимальна якість зображення або неправильне позиціонування датчика. Крім того, автоматизація може допомогти у стандартизації інтерпретації зображень та звітності шляхом впровадження керівних принципів та протоколів.

Важливо відзначити, що хоча методи автоматизації показали багатообіцяючі результати, вони не покликані замінити медичних працівників, а швидше доповнити їх досвід. Роль автоматизації полягає в доповненні та допомозі у різних аспектах ультразвукової діагностики, що в кінцевому підсумку підвищує ефективність, точність та якість обслуговування пацієнтів.

Оскільки технології продовжують розвиватися, а алгоритми стають все більш складними, інтеграція методів автоматизації в рутинну клінічну практику здатна зробити революцію в ультразвуковій діагностиці, зробивши її доступнішою, стандартизованішою і ефективнішою, покращивши при цьому результати діагностики.



## 3 Використання машинного навчання в ультразвуковій діагностиці

### 3.1 Оптимізація та градієнтний спуск

Оптимізація в машинному навчанні – основне поняття, без якого саме машинне навчання би не існувало. Це процес налаштування параметрів(коефіцієнтів) моделі з метою максимізації її продуктивності або ефективності. Це може включати мінімізацію втрат (помилки) моделі, максимізацію точності передбачень, зменшення ризику або будь-якого іншого цільового критерію, залежно від контексту задачі.

Градiєнтний спуск є ітераційним алгоритмом оптимізації, який шукає мінімум функції втрат  $J(\theta)$  залежно від параметрів моделі  $\theta$ . Оновлення параметрів виконується відповідно до антиградієнтного напрямку функції втрат з кроком навчання (швидкістю навчання)  $\alpha$ .

Алгоритм градієнтного спуску використовує формулу:

$$\theta_{n+1} = \theta_n - \alpha \cdot \nabla J(\theta_n)$$

де:

- $\theta_n$  - значення параметрів на n-ому кроці
- $\alpha$  - крок навчання (швидкість навчання)
- $\nabla J(\theta_n)$  - градієнт функції втрат по вектору параметрів  $\theta_n$

Градiєнт  $\nabla J(\theta_n)$  обчислюється за допомогою похідних функції втрат по кожному параметру  $\theta_i$

$$\nabla J(\theta_n) = \left[ \frac{\partial J}{\partial \theta_1}(\theta_n), \frac{\partial J}{\partial \theta_2}(\theta_n), \dots, \frac{\partial J}{\partial \theta_m}(\theta_n) \right]$$

Цей процес повторюється до досягнення збіжності або заданої кількості ітерацій. Градієнтний спуск є одним з найпоширеніших алгоритмів оптимізації у машинному навчанні, що дозволяє навчити модель зі значень даних та параметрів, які мінімізують функцію втрат.

Градієнтний спуск може мати різні варіації, такі як стохастичний градієнтний спуск (SGD), міні-пакетний градієнтний спуск (mini-batch GD) та інші. Ці варіації використовуються для ефективнішого обчислення градієнту та оновлення параметрів моделі на основі підмножини тренувальних прикладів.

### **3.2 Огляд використання методів машинного навчання в ультразвуковій діагностиці у медицині**

Ультразвукова діагностика - широко використовується в медичній практиці як спосіб візуалізації, що забезпечує візуалізацію внутрішніх структур в режимі реального часу і допомагає в діагностиці та лікуванні різних захворювань. З швидким розвитком технологій та великою кількістю цифрових ультразвукових даних зростає інтерес до використання методів машинного навчання для розширення можливостей ультразвукової діагностики. Машинне навчання, підмножина штучного інтелекту продемонструвало великий потенціал у підвищенні точності, ефективності та інтерпретації ультразвукових зображень.

Цей розділ диплома спрямовано вивчення застосування методів машинного навчання в ультразвуковій діагностиці в медицині. У ньому будуть розглянуті різні способи, за допомогою яких алгоритми машинного навчання можуть допомогти медичним працівникам в аналізі зображень, виявленні відхилень, сегментації, вимірі та комп'ютерної діагностики. Крім того, буде розглянуто інтеграцію методів машинного навчання в оптимізацію робочого процесу, забезпечення якості та загальне поліпшення практики ультразвукової діагностики.

Інтеграція методів машинного навчання в ультразвукову діагностику дає безліч переваг. Використовуючи можливості цих алгоритмів, медичні працівники можуть скористатися перевагами автоматизованого аналізу зображень, стандартизованої інтерпретації та підвищення точності діагностики. Більш того, машинне навчання може допомогти в обробці постійно зростаючого обсягу ультразвукових даних, знижуючи робоче навантаження та забезпечуючи ефективніше обслуговування пацієнтів.

### **3.3 Історичний огляд методів комп'ютерного зору.**

Методи комп'ютерного зору не завжди ґрунтувалися на машинному навчанні, зокрема глибокому. Загалом можна виділити три епохи комп'ютерного зору, кожна з яких вирізняється основною характеристикою - ступенем участі машинного навчання і його власною якістю.

Спочатку всі методи автоматичного аналізу оптичної інформації програмувалися вручну, способом підбору параметрів під різні методи. Передобробка передбачала складні морфологічні перетворення зображень, колірну фільтрацію і так далі. Однак варто врахувати, що якість, а точніше розмір зображень відповідали технічним можливостям тієї епохи, а обчислювальна потужність просто б не дала змоги проводити такі операції, як проводимо ми в наші дні.

Разом ми маємо 3 етапи розвитку комп'ютерного зору:

1. "Класичний" комп'ютерний зір, який передбачав тривалий підбір методів оброблення самого зображення і різних фільтрів. До цього ж періоду можна зарахувати і перетворення Хоффа, детектори кутів, ліній.

2. Період, коли почало зароджуватися і саме машинне навчання. За допомогою методів опорних векторів або випадкових лісів, ми навчилися розрізняти окремі ознаки на зображенні, або їхні діапазони і на їхній основі проводити класифікацію.

3. "Глибокий" комп'ютерний зір. У цьому періоді ми зараз і перебуваємо. Використання операцій згорток, повнозв'язних "наскрізних" мереж для сегментації, трансформерів з механізмом "уваги".

Розглянемо список ключових для комп'ютерного зору методів:

### **Період "класичного" комп'ютерного зору:**

- Виявлення країв: Методи виявлення країв, як-от детектор країв Канні та оператор Собеля, спрямовані на визначення та локалізацію меж між різними областями зображення. Ці межі є важливими підказками для сегментації зображення і розпізнавання об'єктів.

- Виявлення кутів: Алгоритми виявлення кутів, як-от детектор кутів Гарріса та детектор кутів Ши-Томасі, визначають точки на зображенні, де градієнт інтенсивності значно змінюється в кількох напрямках. Кути є корисними характеристиками для вирівнювання зображення, відстеження і 3D-реконструкції.

- Фільтрація зображень: Методи фільтрації зображень, включно з гаусовим згладжуванням, медіанною фільтрацією та двосторонньою фільтрацією, змінюють значення пікселів на зображенні для досягнення різних цілей, як-от придушення шуму, поліпшення зображення та вилучення особливостей.

- Піраміди зображень: Піраміди зображень передбачають створення багатомасштабного представлення зображення шляхом побудови серії зменшених версій. Їх використовують для таких завдань, як виявлення інваріантних за масштабом ознак, змішування зображень і аналіз у декількох роздільних здатностях.

- Перетворення Хафа: Перетворення Хафа - це метод виявлення простих геометричних форм, таких як лінії та кола, на зображенні. Воно перетворює особливості зображення в простір параметрів, де наявність фігури позначається піками високої інтенсивності.

- Оптичний потік: алгоритми оптичного потоку оцінюють рух об'єктів між послідовними кадрами у відеопослідовності. Класичні методи, такі як Lucas-Kanade і Horn-Schunck, припускають сталість яскравості й обмеження гладкості для обчислення векторів руху на рівні пікселів.

- Сегментація зображень: Класичні алгоритми сегментації, як-от алгоритм вододілу, вирощування областей і розрізання графа, розділяють зображення на значущі області на основі таких критеріїв, як схожість кольору, текстури або інтенсивності. Сегментація має вирішальне значення для розпізнавання об'єктів і розуміння сцени.

- Виявлення масштабно-просторових плям: Масштабно-просторові методи виявлення плям, як-от лапласіан Гауса (LoG) і різниця гауссіанів (DoG), визначають цікавлять області зображення, які демонструють схожі на плями структури в різних масштабах.

- Зіставлення шаблонів: Зіставлення шаблонів включає в себе порівняння зображення шаблону з різними областями вхідного зображення для визначення найкращої відповідності. Для зіставлення шаблонів зазвичай використовують такі методи, як нормалізована крос-кореляція і сума квадратів різниць (SSD)

- (SSD).Активні контурні моделі (Snakes): Активні контурні моделі - це енергетичні методи, які прагнуть знайти оптимальну межу об'єкта на зображенні шляхом мінімізації енергетичного функціоналу. Змійки широко використовуються для таких завдань, як сегментація зображень, відстеження об'єктів і визначення меж.

### **Період використання не глибокого машинного навчання:**

- Машина опорних векторів (SVM): SVM - це популярний алгоритм машинного навчання, що використовується в комп'ютерному зорі. Його можна навчити для виконання таких завдань, як класифікація зображень, виявлення об'єктів і сегментація зображень шляхом навчання дискримінаційних меж між класами.

- Випадкові ліси: Випадкові ліси - це ансамблеві методи навчання, які об'єднують кілька дерев рішень для складання прогнозів. Вони можуть використовуватися для таких завдань, як класифікація зображень, вибір ознак і розпізнавання об'єктів.
- Bag of Visual Words (BoVW): BoVW - це метод, натхненний обробкою природної мови. Вона передбачає представлення зображень у вигляді гістограм частот візуальних слів. Вона використовується для вирішення таких завдань, як категоризація зображень, розпізнавання текстур і класифікація сцен.
- Умовні випадкові поля (CRF): CRF - це графічні моделі, які можуть відображати просторові залежності між пікселями або областями зображення. Їх використовують для вирішення таких завдань, як семантична сегментація, маркування зображень і знебарвлення зображень.
- Аналіз головних компонент (PCA): PCA - це метод зменшення розмірності, який може бути застосований до даних зображень. Він може використовуватися для таких завдань, як розпізнавання облич, стиснення зображень і реконструкція зображень.
- Приховані марковські моделі (HMM): HMM - це статистичні моделі, які широко використовуються для аналізу послідовностей. У комп'ютерному зорі HMM застосовуються для розв'язання таких завдань, як аналіз відео, розпізнавання жестів і дій.
- Каскадні класифікатори: Каскадні класифікатори, такі як система Віола-Джонса, використовують послідовність простих класифікаторів для ефективного виявлення об'єктів. Ці класифікатори можуть бути навчені за допомогою методів машинного навчання, таких як AdaBoost.
- Латентний розподіл Діріхле (LDA): LDA - це генеративна ймовірнісна модель, яку зазвичай використовують для моделювання тем в обробці природної мови. У комп'ютерному зорі її застосовують для вирішення таких завдань, як кластеризація зображень, розпізнавання об'єктів і візуальна категоризація.

- Моделі гауссових сумішей (GMM): GMM - це ймовірнісні моделі, які представляють розподіл даних у вигляді суміші гауссових розподілів. Вони використовуються для таких завдань, як сегментація зображень, моделювання текстури та виявлення аномалій.
- K-Nearest Neighbors (KNN): KNN - це непараметричний алгоритм, який класифікує дані на основі більшості голосів найближчих сусідів. Він використовується для таких завдань, як класифікація зображень, пошук зображень і розпізнавання об'єктів.

### **Період глибокого навчання:**

- Конволюційні нейронні мережі (CNN): CNN - це моделі глибокого навчання, спеціально розроблені для аналізу візуальних даних. Вони зробили революцію в завданнях комп'ютерного зору, таких як класифікація зображень, виявлення об'єктів, семантична сегментація і генерація зображень.
- Рекурентні нейронні мережі (RNN): RNN, а також їхні різновиди, як-от LSTM (Long Short-Term Memory) і GRU (Gated Recurrent Unit), використовують для послідовного аналізу даних у задачах комп'ютерного зору, як-от аналіз відео, створення титрів і розпізнавання дій.
- Генеративні адверсарні мережі (GAN): GAN складаються з мережі-генератора і мережі-дискримінатора, які конкурують одна з одною. Вони використовуються для таких завдань, як генерація зображення, переведення зображення в зображення і передача стилю.
- Сіамські мережі: Сіамські мережі вчать порівнювати і вимірювати схожість між двома вхідними даними. Їх використовують для вирішення таких завдань, як розпізнавання облич, відстеження об'єктів і пошук зображень.
- Автоенкодера: Автоенкодера - це моделі навчання без нагляду, метою яких є реконструкція вхідних даних. Їх використовують для розв'язання

таких завдань, як денуазинг зображень, інпайнтінг зображень і зменшення розмірності.

- **Механізми уваги:** Механізми уваги дають змогу моделям фокусуватися на певних регіонах або особливостях зображення, забезпечуючи тонкий аналіз. Вони використовуються в таких завданнях, як створення підписів до зображень, візуальні відповіді на запитання і виявлення об'єктів.

- **Трансформери:** Трансформери, спочатку розроблені для обробки природної мови, також успішно застосовуються в завданнях комп'ютерного зору. Вони досягли передових результатів у розпізнаванні зображень, виявленні об'єктів і генерації зображень.

- **Одномоментне навчання:** Методи одномоментного навчання спрямовані на розпізнавання або класифікацію об'єктів на основі одного або декількох прикладів. Для вирішення таких завдань використовують такі підходи глибокого навчання, як сіамські мережі, метричне навчання і мета-навчання.

- **Нейронне перенесення стилю:** Нейронні методи передачі стилю поєднують зміст одного зображення з художнім стилем іншого. Моделі глибокого навчання можуть навчитися переносити художні стилі на зображення, створюючи візуально привабливі результати.

- **Глибоке навчання з підкріпленням:** Глибоке навчання з підкріпленням об'єднує глибоке навчання і навчання з підкріпленням, даючи змогу машинам навчатися складної поведінки методом проб і помилок. Воно застосовується в завданнях комп'ютерного зору, таких як керування роботами, автономне водіння і відеоігри.



### 3.4 Найбільш популярні задачі комп'ютерного зору в медицині.

Методи комп'ютерного зору знайшли численні застосування в медицині, даючи змогу проводити автоматизований аналіз, інтерпретацію та розуміння медичних зображень і візуальних даних. Ці методи зробили революцію в різних завданнях медичної візуалізації, допомагаючи медичним працівникам у діагностиці, плануванні лікування та дослідженнях. У цьому розділі диплома розглядаються деякі з найпопулярніших завдань комп'ютерного зору в медицині.

**Класифікація зображень:** Класифікація зображень передбачає присвоєння мітки або категорії всьому зображенню на основі його візуального змісту. У медицині класифікація зображень відіграє важливу роль у діагностиці різних захворювань. Наприклад, алгоритми комп'ютерного зору можуть класифікувати ураження шкіри як доброякісні або злоякісні, визначати конкретні типи пухлин з гістопатологічних зрізів або розрізняти різні стадії захворювання на основі особливостей зображення. Моделі глибокого навчання, зокрема згорткові нейронні мережі (CNN), досягли значної точності у задачах класифікації медичних зображень.

**Виявлення об'єктів:** Виявлення об'єктів спрямоване на ідентифікацію і локалізацію декількох об'єктів, що становлять інтерес на зображенні. У медичній візуалізації виявлення об'єктів відіграє важливу роль у визначенні розташування анатомічних структур, аномалій або специфічних уражень. Наприклад, алгоритми комп'ютерного зору можуть автоматично виявити і локалізувати вузлики в легенях на рентгенівських знімках грудної клітки або визначити пухлини на знімках МРТ або КТ. Алгоритми виявлення об'єктів дають змогу підвищити ефективність скринінгу, ранньої діагностики та планування лікування завдяки автоматизації ідентифікації областей інтересу.

**Семантична сегментація:** Семантична сегментація спрямована на присвоєння конкретної мітки або класу кожному пікселю або вокселю зображення, забезпечуючи детальне розуміння структури зображення. У

медичних додатках семантична сегментація необхідна для виділення органів, пухлин та інших структур, що становлять інтерес. Завдяки точному визначенню меж алгоритми комп'ютерного зору допомагають в об'ємних вимірах, плануванні променевої терапії та хірургічному керівництві. Методи глибокого навчання, зокрема конволюційні нейронні мережі, значно підвищили точність і ефективність семантичної сегментації в медичній візуалізації. Реєстрація зображень: Реєстрація зображень – це вирівнювання та просторове відображення кількох зображень, отриманих за допомогою одного або різних методів. У медицині реєстрація зображень відіграє вирішальну роль в об'єднанні додаткової інформації, наприклад, при поєднанні МРТ та КТ або поєднанні передопераційних та інтраопераційних зображень. Алгоритми комп'ютерного зору забезпечують точну та надійну реєстрацію, допомагаючи в інтеграції мультимодальних даних для покращення діагностики, планування лікування та втручань за допомогою зображень.

Реконструкція зображень: Методи реконструкції зображень спрямовані на створення високоякісних зображень із необроблених або пошкоджених даних. У медичній візуалізації реконструкція зображень використовується для поліпшення якості зображення, зменшення шуму або відновлення інформації, що бракує. Алгоритми комп'ютерного зору у поєднанні з математичними моделями та передовими методами реконструкції дозволяють створювати чіткіші та інформативніші зображення. Це допомагає підвищити точність діагностики, візуалізувати анатомічні структури та виявити тонкі аномалії.

Застосування комп'ютерного зору у медицині продовжує розвиватися завдяки збільшенню доступності цифрових медичних зображень, підвищенню обчислювальної потужності та швидкому розвитку алгоритмів глибокого навчання. Ці завдання відіграють фундаментальну роль у використанні потенціалу даних медичної візуалізації, підвищенні точності діагностики, полегшенні планування лікування та покращенні результатів лікування пацієнтів.

Ось деякі потенційні проблеми, які можуть виникнути під час вирішення кожного зі згаданих завдань комп'ютерного зору.

### **Класифікація зображень:**

Обмежена кількість маркованих даних: Недостатня кількість точно помічених даних може ускладнити процес навчання моделей класифікації зображень, що призводить до зниження продуктивності та узагальненості.

Дисбаланс класів: Дисбаланс класів виникає, коли існує значна невідповідність кількості зображень, доступних для різних класів. Це може призвести до появи упереджених моделей, які не можуть класифікувати класи меншин.

Мінливість та неоднозначність: Медичні зображення можуть характеризуватись високою мінливістю та неоднозначністю, що ускладнює роботу алгоритмів класифікації зображень у складних випадках або для виявлення тонких відмінностей між класами.

### **Виявлення об'єктів:**

Маленькі або затемнені об'єкти: Виявлення невеликих або частково затемнених об'єктів на медичних зображеннях може бути складним завданням, оскільки їх риси можуть бути менш виразними або закритими іншими структурами, що може призвести до хибно-негативних результатів або неточної локалізації.

Непостійність зовнішнього вигляду об'єктів: Об'єкти, які мають інтерес, можуть мати різний зовнішній вигляд через варіації параметрів візуалізації, стану пацієнта або стадії захворювання. Ці відмінності можуть ускладнити алгоритми виявлення об'єктів узагальнення різних екземплярів одного і того ж об'єкта.

**Обчислювальна складність:** Алгоритми виявлення об'єктів часто включають складні моделі та інтенсивні обчислювальні вимоги, що може обмежити їх застосування в реальному часі або масштабованість у певних клінічних умовах.

### **Семантична сегментація:**

**Неоднозначні межі:** Визначення точних меж певних структур або пошкоджень може бути утруднено через нечіткі краї, що перекриваються областей або шуму на зображенні. Ці неоднозначності можуть призвести до неточностей результатів сегментації.

**Варіабельність між спостерігачами:** Різні експерти можуть по-різному визначати межі або сегментувати структури, що призводить до протиріч у мітках, отриманих у результаті. Така мінливість може вплинути на навчання та оцінку алгоритмів сегментації.

**Робота з різнорідними даними:** Медичні зображення, отримані за допомогою різних методів або сканерів, можуть мати різні характеристики та якість зображення, що створює труднощі розробки моделей сегментації, які можуть узагальнювати різні джерела даних.

### **Реєстрація зображень:**

**Помилки усунення:** Точна реєстрація зображень ґрунтується на точному вирівнюванні відповідних елементів або орієнтирів. Однак такі фактори, як рух пацієнта, артефакти зображення або розбіжності в протоколах візуалізації можуть вносити помилки реєстрації, що призводять до неправильного поєднання.

**Проблеми мультимодальної реєстрації:** Реєстрація зображень, отриманих за допомогою різних методів візуалізації з різними характеристиками зображення та просторовими спотвореннями, може бути особливо складним завданням. Відмінності у зовнішньому вигляді зображень,

роздільній здатності та анатомічних структурах між модальностями можуть призвести до неточної або ненадійної реєстрації.

Обчислювальна складність: алгоритми реєстрації зображень часто включають обчислювально інтенсивні методи оптимізації, які можуть вимагати значного часу обробки та обчислювальних ресурсів, що перешкоджає застосуванню в режимі реального часу або проведенню великомасштабних досліджень.

### **Реконструкція зображення.**

Шум та артефакти: Притаманні медичним зображенням шуми та артефакти можуть створювати проблеми для алгоритмів реконструкції зображень. Ці недоліки можуть негативно вплинути на якість та достовірність відновлених зображень.

Обмежена доступність даних: Алгоритмам реконструкції може знадобитися додаткова інформація або попередні знання, щоб компенсувати дані, що відсутні або пошкоджені. Однак отримання такої інформації не завжди можливе, що призводить до обмеження точності реконструкції зображень.

Компроміс між якістю реконструкції та обчислювальною ефективністю  
Для отримання високоякісних реконструкцій часто використовуються алгоритми, що вимагають великих обчислювальних витрат. Баланс між якістю реконструкції та обчислювальною ефективністю є важливим моментом, особливо для програм, що працюють у режимі реального часу.

Важливо вирішити ці потенційні проблеми, пов'язані із завданнями комп'ютерного зору медицини. Стратегії вирішення проблем можуть включати отримання великих і різноманітних наборів даних, розробку надійних алгоритмів, здатних впоратися з мінливістю і неоднозначністю, включення знань, специфічних для даної галузі, і строгих протоколів перевірки та оцінки. Співпраця між експертами з комп'ютерного зору,

медичними працівниками та фахівцями у цій галузі має вирішальне значення для вирішення цих проблем та розвитку комп'ютерного зору в медицині.

### **Етичні проблеми:**

Проблема інтерпретації рішення нейронних мереж у медицині породжує кілька важливих етичних проблем:

**Відповідальність:** Використання нейронних мереж у медицині вимагає встановлення відповідальності за прийняті рішення. Якщо медичний фахівець не може інтерпретувати рішення мережі та пояснити його пацієнту, виникає питання про те, хто несе відповідальність за можливі помилки або некоректні дії мережі.

**Довіра пацієнтів:** Інтерпретація рішення нейронної мережі має велике значення для пацієнтів. Якщо пацієнт не розуміє, як саме мережа прийшла до свого висновку, він може втратити довіру до медичної системи в цілому. Важливо забезпечити доступність та зрозумілість інформації про рішення мережі для пацієнтів.

**Рішення на основі особистих даних:** Використання нейронних мереж в медицині часто потребує збору та обробки особистих даних пацієнтів. Це викликає етичні питання щодо конфіденційності, приватності та безпеки цих даних. Пацієнти мають право знати, як саме їхні дані використовуються та як це впливає на прийняття рішень.

## 4 Огляд розповсюджених методів комп'ютерного зору

У цьому розділі ми розглянемо різні методи комп'ютерного аналізу з різних періодів його розвитку, які були детально описані в пункті 3.2 цієї роботи. Оскільки в практичній частині ми розв'язуватимемо задачу класифікації, то й описувані методи відповідатимуть цьому завданню.

### 4.1 Зіставлення шаблонів (Template Matching)

Зіставлення шаблонів - це фундаментальна техніка в комп'ютерному зорі, яка використовується для пошуку екземплярів шаблонного зображення у великому вхідному зображенні. Він включає в себе порівняння шаблону з різними областями вхідного зображення і пошук областей, які близько відповідають зовнішньому вигляду шаблону. Цей метод широко використовується для вирішення таких завдань, як виявлення об'єктів, розпізнавання об'єктів і вирівнювання зображень.

Процес зіставлення шаблонів можна коротко описати таким чином:

**Вибір шаблону:** Шаблонне зображення обирають на основі об'єкта або зразка, який має бути виявлений на вхідному зображенні. Шаблон повинен ідеально представляти зовнішній вигляд і характеристики об'єкта, що цікавить.

**Ковзне вікно:** Потім шаблон переміщається по вхідному зображенню, розглядаючи кожну позицію для визначення схожості між шаблоном і відповідною областю вхідного зображення. Ковзання може виконуватися послідовно або більш ефективно з використанням таких методів, як піраміди зображень.

**Метрика схожості:** У кожній позиції обчислюється метрика схожості, щоб виміряти, наскільки добре шаблон відповідає ділянці на вхідному зображенні. Зазвичай використовуються такі метрики подібності, як сума

квадратів різниці (SSD), нормалізована крос-кореляція (NCC) і коефіцієнт кореляції. Детальніше про них.

$$SSD = \sum_{i,j} [T(i,j) - I(x+i, y+j)]^2$$

SSD розраховує суму квадратів різниць пікселів між шаблоном і областю на вхідному зображенні. Вона мінімізується, коли шаблон і регіон мають високий ступінь подібності.

$$NCC = \frac{\sum_{i,j} [T(i,j) \cdot I(x+i, y+j)]}{\sqrt{\sum_{i,j} T(i,j)^2 \cdot \sum_{i,j} I(x+i, y+j)^2}}$$

NCC вимірює нормалізований точковий добуток між нормалізованим шаблоном і нормалізованим регіоном. Він відображає схожість у термінах інтенсивності пікселів, незалежно від відмінностей у яскравості або контрастності між шаблоном і регіоном.

Обчислення бала збігу: Метрика схожості використовується для розрахунку бала збігу для кожної позиції, де шаблон порівнюється з відповідним регіоном. Що вищий бал збігу, то краща відповідність між шаблоном і регіоном.

Порогова оцінка: Оцінки збігу зазвичай порівнюють із заздалегідь визначеним порогом, щоб визначити, чи було знайдено збіг. Позиції з оцінками, вищими за поріг, вважаються збігами, що вказує на присутність шаблону у вхідному зображенні.

Післяобробка: Після того як збіги визначено, можна застосувати кроки післяобробки, як-от не максимальне придушення або кластеризація, для уточнення результатів і усунення дублюючих виявлень.

Поєднання шаблонів може бути ефективним, якщо шаблон і об'єкт на вхідному зображенні мають схожий зовнішній вигляд, умови освітлення і масштаб. Однак при цьому можуть виникати такі проблеми, як оклюзія, часткова оклюзія, зміни точки зору, а також зміни масштабу або повороту. Незважаючи на ці обмеження, зіставлення шаблонів залишається цінним методом у комп'ютерному зорі, особливо для простих завдань виявлення або



вирівнювання об'єктів. Він служить основою для більш досконалих методів і знаходить застосування в різних галузях, включно з робототехнікою, спостереженням і медичною візуалізацією. Дослідники продовжують вивчати і вдосконалювати метод зіставлення шаблонів, щоб усунути його обмеження і поліпшити його ефективність у різних сценаріях.

## 4.2 Метод опорних векторів (SVM)

Support Vector Machines (SVM) - це потужний алгоритм машинного навчання, який може бути застосований до завдань класифікації зображень. Він працює шляхом навчання межі прийняття рішення, яка максимально розділяє різні класи зображень на основі їхніх витягнутих ознак. Нижче наводиться опис того, як SVM може бути використана для класифікації зображень:

Витяг ознак: Першим кроком у застосуванні SVM для класифікації зображень є витяг значущих ознак із зображень. Зазвичай для цього використовують такі методи, як гістограма орієнтованих градієнтів (HOG), масштабно-інваріантне перетворення ознак (SIFT) або конволюційні нейронні мережі (CNN). Ці методи витягують дискримінаційні ознаки, які відображають важливі візуальні характеристики зображень.

Підготовка навчальних даних: Для навчання моделі SVM необхідний набір даних із мітками. Цей набір даних має складатися з набору зображень і відповідних їм міток класів. Кожне зображення представлено вектором ознак, отриманим на попередньому етапі, а мітки класів вказують на категорію або клас кожного зображення.

Навчання моделі SVM: Наступним кроком є навчання моделі SVM з використанням набору мічених даних. SVM прагне знайти оптимальну гіперплощину, яка найкращим чином розділяє різні класи в просторі ознак. Гіперплощину навчають шляхом максимізації маржі, яка являє собою відстань між гіперплощиною і найближчими навчальними вибірками кожного класу.

SVM використовує функцію втрат і різні алгоритми оптимізації, такі як послідовна мінімальна оптимізація (SMO), для пошуку оптимальної гіперплощини.

Математично метод опорних векторів можна описати наступним чином:

1. Нехай маємо навчальну вибірку з елементами  $((x_i, y_i))$ , де  $(x_i)$  - вектор ознак,  $(y_i)$  - мітка класу (+1 або -1).

2. Знаходимо гіперплощину, яка максимально розділяє класи. Ця гіперплощина може бути описана рівнянням:

$$[w^T x + b = 0]$$

де  $(w)$  - вектор нормалі до гіперплощини,  $(b)$  - зсув (bias) гіперплощини.

3. Мета полягає у знаходженні оптимальних параметрів  $(w)$  і  $(b)$ , які мінімізують помилки класифікації та максимізують розмір проміжку між гіперплощиною і найближчими до неї точками обох класів. Це еквівалентно розв'язанню оптимізаційної задачі:

$$[\min w, b \frac{1}{2} \|w\|^2] \text{ за умови: } [y_i(w^T x_i + b) \geq 1 \quad \forall i]$$

4. Розв'язуючи цю оптимізаційну задачу, отримуємо оптимальні параметри  $(w)$  і  $(b)$ . Після цього нові зразки можуть бути класифіковані шляхом оцінки  $(w^T x + b)$ .

Класифікація: Після навчання моделі SVM вона може бути використана для класифікації нових, зображень. Для класифікації витягуються характеристики тестового зображення, і SVM-модель застосовує функцію прийняття рішення для віднесення зображення до одного з навчених класів. Функція прийняття рішення обчислює відстань вектора ознак тестового зображення від вивченої гіперплощини. Тестовому зображенню присвоюється мітка класу з найбільшою довірою або найменшою відстанню.

SVM мають низку переваг для завдань класифікації зображень. Вони можуть працювати з високорозмірними просторами ознак і ефективно

справляються з невеликими наборами навчальних даних. SVM також менш схильні до перенавчання порівняно з іншими алгоритмами машинного навчання. Крім того, завдяки використанню функції ядра, SVM можуть вловлювати нелінійні зв'язки між ознаками зображення і мітками класів, що дає їм змогу навчатися складним межам ухвалення рішень.

Однак SVM мають певні обмеження при застосуванні до класифікації зображень. Вони можуть погано масштабуватися під час роботи з великими наборами даних через складність обчислень. Крім того, SVM не дають прямих імовірнісних результатів, що ускладнює оцінку достовірності прогнозів.

### **4.3 Випадкові ліси**

Random Forest - це популярний алгоритм машинного навчання, який можна використовувати для задач класифікації зображень. Він належить до сімейства ансамблевого навчання, в якому кілька дерев рішень об'єднуються для складання прогнозів. Ось опис того, як Random Forest може бути застосований для класифікації зображень:

Підготовка навчальних даних: Першим кроком є підготовка маркованого набору даних для навчання моделі Random Forest. Набір даних має складатися з набору зображень і відповідних їм міток класів. Кожне зображення представлено набором ознак, витягнутих із зображення, таких як гістограма орієнтованих градієнтів (HOG) або дескриптори на основі кольору.

Навчання моделі Random Forest: Процес навчання включає в себе побудову ансамблю дерев рішень. Кожне дерево рішень навчається на окремій підмножині навчальних даних і використовує випадкову підмножину ознак для прийняття рішень. Ця випадковість допомагає декорелювати дерева рішень і запобігти надмірній підгонці. Кількість дерев рішень у лісі та глибина кожного дерева є гіперпараметрами, які необхідно визначити.

Побудова дерева рішень: У кожному вузлі дерева рішень вибирають ознаку на основі її здатності якнайкраще розділити дані на однорідні

підмножини. Зазвичай для цього використовують такі метрики, як домішка Джині або інформаційний виграш. Дерево рекурсивно розбиває дані на основі різних ознак доти, доки не буде досягнуто критерію зупинки, наприклад, максимальної глибини дерева або мінімальної кількості зразків у кожному вузлі листа.

Класифікація: Після того як модель Random Forest навчена, її можна використовувати для класифікації нових, невидимих зображень. Для класифікації ознаки зображення пропускаються через кожне дерево ухвалення рішень у лісі, і кожне дерево присвоює мітку класу на основі вивчених правил. Остаточний прогноз визначається за допомогою механізму голосування, де зображенню присвоюється клас, який отримав більшість голосів від дерев прийняття рішень.

Випадкові ліси мають низку переваг для задач класифікації зображень. Вони стійкі до шумів і викидів, ефективно обробляють високорозмірні простори ознак і можуть працювати з незбалансованими наборами даних. Крім того, випадкові ліси надають міри важливості ознак, що дає змогу краще зрозуміти основні характеристики зображення, які сприяють класифікації.

Однак у випадкових лісів можуть бути обмеження при застосуванні до класифікації зображень. Порівняно з деякими іншими алгоритмами вони працюють повільніше, особливо при використанні великої кількості дерев рішень або складних наборів даних. Крім того, випадкові ліси можуть не справлятися з уловлюванням дрібних деталей або складних закономірностей у зображеннях, які вимагають більш складних моделей, таких як архітектури глибокого навчання.

## 4.4 Згорткові нейронні мережі

Згорткові нейронні мережі (CNN/ЗНМ) є одним з найпотужніших інструментів для аналізу зображень та відео. Вони складаються з багатьох шарів, які виконують різні операції на вхідних даних, такі як згортка, пулінг, активація та повнозв'язання. Згортка є основною операцією ЗНМ, яка полягає в застосуванні невеликого фільтра (ядро) до кожної ділянки вхідного зображення або попереднього шару. Результатом є нове зображення (карта ознак), яке містить інформацію про наявність певних ознак на вхідному зображенні. Пулінг є операцією, яка зменшує розмір карти ознак за допомогою певної функції, наприклад максимуму або середнього. Активація є операцією, яка застосовує нелінійну функцію до кожного елемента карти ознак, наприклад сигмоїду або ReLU. Повнозв'язання є операцією, яка перетворює карти ознак у вектори та застосовує лінійне перетворення до них.

ЗНМ мають багато переваг перед традиційними методами обробки зображень. По-перше, вони можуть автоматично вивчати оптимальні фільтри для розпізнавання різних ознак на зображеннях, тоді як традиційним методам потребується ручне визначення функцій. По-друге, вони можуть обробляти зображення будь-якого розміру та формату, тоді як традиційним методам потребується попередня обробка та нормалізація даних. По-третє, вони можуть ефективно використовувати ієрархичну структуру даних, тобто використовувати інформацію з попереднього шару для аналізу наступного шару.

Одна з основних проблем ЗНМ є їх складність та велика кількість параметрів, які потребують тренування. Це призводить до високих вимог до обчислювальних ресурсів та пам'яті. Крім того, ЗНМ схильні до перенавчання на обмежених наборах даних. Для подолання цих проблем існують різні

методи оптимізації та регуляризації ЗНМ, такі як стохастичний градієнтний спуск (SGD), dropout, batch normalization та data augmentation.

Основними компонентами згорткової мережі є:

- Згорткові шари. Використовують невеликі фільтри (або ядра), які пересуваються по вхідному зображенню або попередньому шару і обчислюють скалярний добуток між фільтром і вхідними даними. Таким чином, вони виділяють основні ознаки зображення, такі як ребра, кути, текстури тощо. Математично можна описати процес згортки наступним чином. Нехай  $X$  - це вхідне зображення розміру  $H \times W$ , а  $K$  - це фільтр розміру  $h \times w$ . Тоді вихідне зображення  $Y$  розміру  $(H - h + 1) \times (W - w + 1)$  обчислюється за формулою:

$$Y_{ij} = \sum_{m=0}^{h-1} \sum_{n=0}^{w-1} X_{i+m, j+n} K_{mn},$$

де  $i = 0, 1, \dots, H - h$  і  $j = 0, 1, \dots, W - w$ . Ця операція може бути представлена як матричний добуток між матрицею  $X$  і матрицею  $K$ , якщо перетворити їх у вектори.

- Агрегувальні шари. Зменшують розмірність вихідних даних згорткового шару, застосовуючи операції пулінгу (або субдискретизації), такі як максимальний пулінг, середній пулінг або сумарний пулінг. Таким чином, вони забезпечують інваріантність до невеликих зміщень та зменшують обчислювальне навантаження. Математично можна описати процес пулінгу наступним чином. Нехай  $Y$  - це вихідне зображення згорткового шару розміру  $H \times W$ , а  $P$  - це операція пулінгу з розміром ядра  $h \times w$  і кроком  $s$ .

Тоді вихідне зображення  $Z$  розміру  $\left\lfloor \frac{H-h}{s} + 1 \right\rfloor \times \left\lfloor \frac{W-w}{s} + 1 \right\rfloor$  обчислюється за формулою:

$$Z_{ij} = P(Y_{is:is+h-1, js:js+w-1}),$$

де  $i = 0, 1, \dots, \left\lfloor \frac{H-h}{s} \right\rfloor$  і  $j = 0, 1, \dots, \left\lfloor \frac{W-w}{s} \right\rfloor$ . Операція  $P$  може бути максимальною, середньою або сумарною функцією на своїх аргументах.

- Повноз'єднані шари. Приймають на вхід векторизоване представлення даних попереднього шару і виконують логістичну регресію для класифікації або регресії. Вони є останньою частиною згорткової мережі і визначають категорію або значення вихідного результату.

#### 4.5 Візуалізація та інтерпретація рішень згорткової нейронної мережі

У сфері комп'ютерного зору згорткові нейронні мережі (CNN) досягли значних успіхів у вирішенні різних завдань, таких як класифікація зображень і виявлення об'єктів. Однак внутрішня робота цих моделей глибокого навчання часто може здаватися "чорною скринькою", що ускладнює розуміння того, як і чому вони роблять певні прогнози. Щоб вирішити цю проблему, дослідники розробили методи візуалізації, які проливають світло на процес прийняття рішень ШНМ. У цьому контексті два відомі методи, а саме Grad-CAM і LIME, стали цінними інструментами для інтерпретації та пояснення прогнозів CNN. Ці методи дають уявлення про важливі області вхідних зображень і локальні внески специфічних особливостей у рішення моделі. У цьому розділі ми заглибимося в роботу Grad-CAM і LIME, демонструючи їхню важливість для глибшого розуміння моделей CNN і зміцнення довіри до їхніх прогнозів.

## 1. Grad-CAM (Gradient-weighted Class Activation Mapping):

Grad-CAM - це метод візуалізації, який виділяє важливі області на вхідному зображенні, які роблять найбільший внесок у прогноз CNN. Вона надає теплову карту, яка показує області зображення, що є найбільш релевантними для рішення моделі. Ось як працює Grad-CAM:

- Спочатку модель CNN навчається на конкретному завданні, наприклад, класифікації зображень.

- Під час виведення обчислюються градієнти моделі по відношенню до прогнозованої оцінки класу.

- Потім ці градієнти поєднуються з картами ознак останнього шару згортки.

- Отримані зважені карти ознак агрегуються для створення теплової карти, де більш високі значення представляють області, які є більш важливими для прийняття рішення моделлю.



*Рисунок 4.1 Візуалізація методом GRAD-CAM*

Grad-CAM допомагає візуалізувати, на яких частинах зображення модель фокусується, щоб зробити певний прогноз. Це забезпечує інтерпретованість і розуміння процесу прийняття рішень моделями CNN.



## 2. LIME (Local Interpretable Model-Agnostic Explanations):

LIME - це ще один метод інтерпретації та пояснення прогнозів моделей CNN. Він спрямований на пояснення окремих прогнозів на локальному рівні, тобто для конкретних вхідних даних. Ось огляд того, як працює LIME:

- LIME апроксимує межу прийняття рішень складної моделі CNN, створюючи простішу модель, яку можна інтерпретувати, наприклад, лінійну модель або модель дерева рішень.

- Це відбувається шляхом генерування набору збурених зразків навколо вихідних даних і спостереження за відповідними прогнозами CNN.

- Модель, яку можна інтерпретувати, навчається на цих збурених зразках разом з відповідними прогнозами.

- Потім LIME досліджує навчену інтерпретовану модель, щоб визначити особливості або частини вхідних даних, які найбільше впливають на рішення моделі.

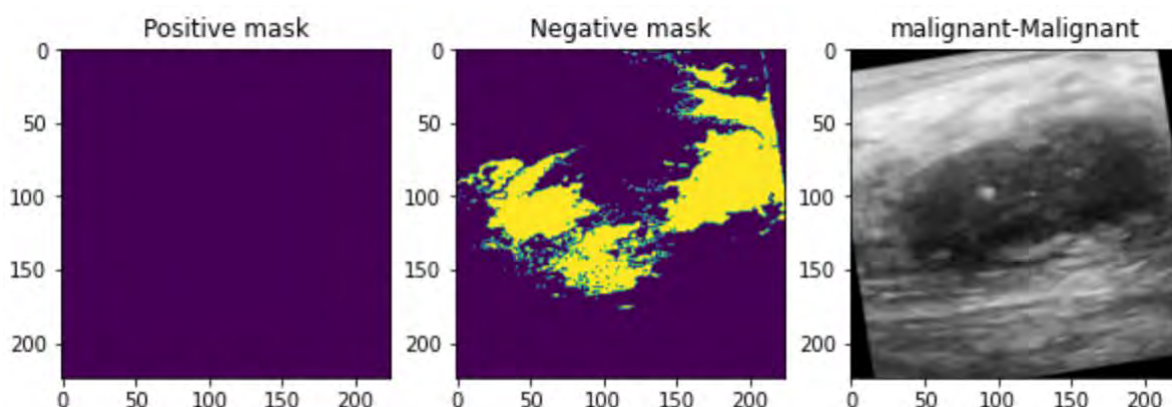


Рисунок 4.2. Візуалізація за допомогою методу LIME

Використовуючи LIME, дослідники можуть отримати уявлення про те, як конкретні вхідні дані впливають на прогнози, зроблені моделями CNN. Це допомагає зрозуміти процес прийняття рішень моделлю на локальному рівні і забезпечує інтерпретацію складних моделей.

Отже, як Grad-CAM, так і LIME є цінними інструментами в галузі комп'ютерного зору, що дозволяють дослідникам і практикам краще розуміти та інтерпретувати моделі CNN, надаючи уявлення про процес прийняття рішень і допомагаючи зміцнити довіру до їхніх прогнозів.

## 5 Імплементация алгоритму комп'ютерного зору

### 5.1 Опис завдання

Завдання цього проєкту можна сформулювати так - реалізувати систему класифікації ультразвукового знімка для визначення на ньому полікістозу яєчників. Додатково - реалізувати систему підсвічування критичних зон, за якими модель глибокого навчання зробила висновки про своє передбачення.

### 5.2 Опис даних

Дані для розв'язання задач у сфері ультразвукової діагностики наразі дуже складно знайти, оскільки потребують оцифрування і згоди пацієнтів. Тому ми знайшли відкриті дані для завдання класифікації та застосуємо GRAD-CAM для псевдо-сегментації зображень.

<https://www.kaggle.com/datasets/anaghachoudhari/pcos-detection-using-ultrasound-images>

На жаль, на платформі Kaggle не вказана додаткова інформація про походження цих даних.

Розподіл за класами та вибірками відображено в таблиці нижче

	Інфіковано	Не інфіковано
Тренувальна вибірка	781	1 143
Валідаційна вибірка	787	1 145
Усього спостережень	1568	2288

### 5.3 Опис обладнання

Усю попередню обробку даних і обчислення проводилися на такому обладнанні.

Центральний процесор	AMD Ryzen 7 5800X 8-Core Processor
Оперативна пам'ять	96 ГБ DDR-4 2666 MHz
Диск	CT1000P2SSD8 – 1000 Gb
Графічний процесор	NVIDIA GeForce RTX 3060 12 Gb VRAM

## 5.4 Опис імплементації

Для реалізації моделі класифікації ми спроектували архітектуру згорткової нейронної мережі.

Спочатку було зроблено спробу тренування претренерованої мережі з архітектурою ResNet50/101/152. Однак моделі показували невтішні результати з фіксацією перетренування або низької точності на обох вибірках (тренувальній і валідаційній).

Також було зроблено спробу використання сучасного фреймворка autogluon для автоматичного навчання моделі. Цей підхід дав би змогу продемонструвати донавчання моделі при отриманні нових вхідних даних. Однак, ця бібліотека не дає змоги використовувати GRAD-CAM та інші методи візуалізації та інтерпретації висновків.

## 5.5 Правила аугментації та предобробки

Для збільшення кількості даних у вибірці ми провели аугментацію даних. Дані генеруються в тому разі, якщо наявних у початковій вибірці даних недостатньо на новій епосі.

Аугментація даних - це метод розширення навчального набору шляхом створення нових варіацій існуючих даних. Цей підхід широко використовується в машинному навчанні, особливо в області комп'ютерного зору. Він полягає у застосуванні різних трансформацій до зображень, таких як обертання, зсув, масштабування, зміна контрастності, відображення та інші. Це дозволяє отримати нові зображення, які відрізняються від оригінальних, але все ще містять корисну інформацію. Аугментація даних допомагає збільшити різноманітність навчального набору і поліпшити здатність моделі до узагальнення на нові дані.

Агументацію та предобробку проводили за такими правилами.

Нормалізація	Ділення на максимум
Діапазон зсуву	+/- 40%
Діапазон наближення	+/- 40%
Горизонтальне віддзеркалення	Дозволено
Вертикальне віддзеркалення	Довзолоно
Діапазон повороту	60 градусів від вертикальної осі
Тренувальна вибірка	60%
Валідаційна вибірка	40%
Метод заповнення	Заповнення найближчим

## 5.6 Опис моделі

Моделі приймає на вхід зображення розміром 224 на 224 за всіма трьома каналами RGB. Після цього проходить операція двовимірної згортки тричі з почерговою операцією пулінгу 4x4 для зниження розмірності за максимальним значенням. У всіх операціях згорток функції активації ReLU. Після чого ознаки розгортаються назад у вектор із розміром 72 і двічі проходять через тренувані повнозв'язні шари. Перший шар має активацію ReLU, а другий, кінцевий - сигмоїду.

Функція активації ReLU (Rectified Linear Unit) використовується в нейронних мережах для нелінійного перетворення вхідних даних. Вона визначається як  $f(x) = \max(0, x)$ , де  $x$  - вхідне значення. Ця функція повертає 0 для всіх негативних значень і саме вхідне значення для всіх не-негативних значень.

Функція активації сигмоїда, також відома як логістична функція, використовується для зведення значень на проміжок між 0 та 1. Вона визначається як  $f(x) = 1/(1 + \exp(-x))$ , де  $x$  - вхідне значення. Великі від'ємні значення наближаються до 0, а великі позитивні значення наближаються до 1.

Загалом кількість тренованих параметрів у моделі 9452, що порівняно з ResNet або VGG16/19 вкрай мало. Однак цим же забезпечується і стійкість моделі під час застосування її з новими даними, оскільки на останніх епохах практично відсутнє перетренування

Layer (type)	Output Shape	Params #
conv2d (Conv2D)	(None, 220, 220, 12)	912
max_pooling2d(MaxPooling2D)	(None, 55, 55, 12)	0
conv2d_1 (Conv2D)	(None, 51, 51, 10)	3010
max_pooling2d_1(MaxPooling 2D)	(None, 12, 12, 10)	0
conv2d_2 (Conv2D)	(None, 10, 10, 8)	728
max_pooling2d_2 (MaxPooling 2D)	(None, 3, 3, 8)	0
flatten (Flatten)	(None, 72)	0
dense (Dense)	(None, 64)	4672
dense_1 (Dense)	(None, 2)	130

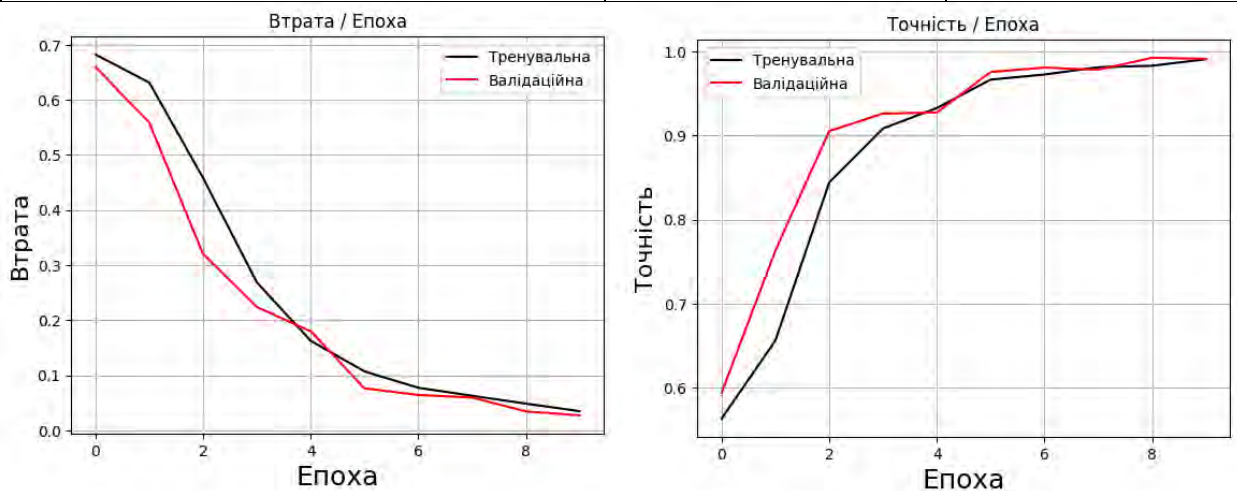


Рисунок 5.1 Результати моделі на кожній епосі

Як можна побачити на графіках вище, після 4-ї епохи модель практично ідеально підлаштувалася під валідаційні дані, тому ми з упевненістю можемо сказати про відсутність факту перетренування або недонавчання моделі.

Confusion Matrix	Інфіковано	Не інфіковано
Інфіковано	778	0
Не інфіковано	3	1141

Також ми можемо в цьому переконатися, подивившись на матрицю зіставлення відповідей. Комірки відображають кількість клас/прогнозування.

На останній 10-й епісі ми домоглися точності 0.9905 на тренувальній і 0.9909 на валідаційній вибірці відповідно.

## 5.7 Візуалізація методом Grad-Cam

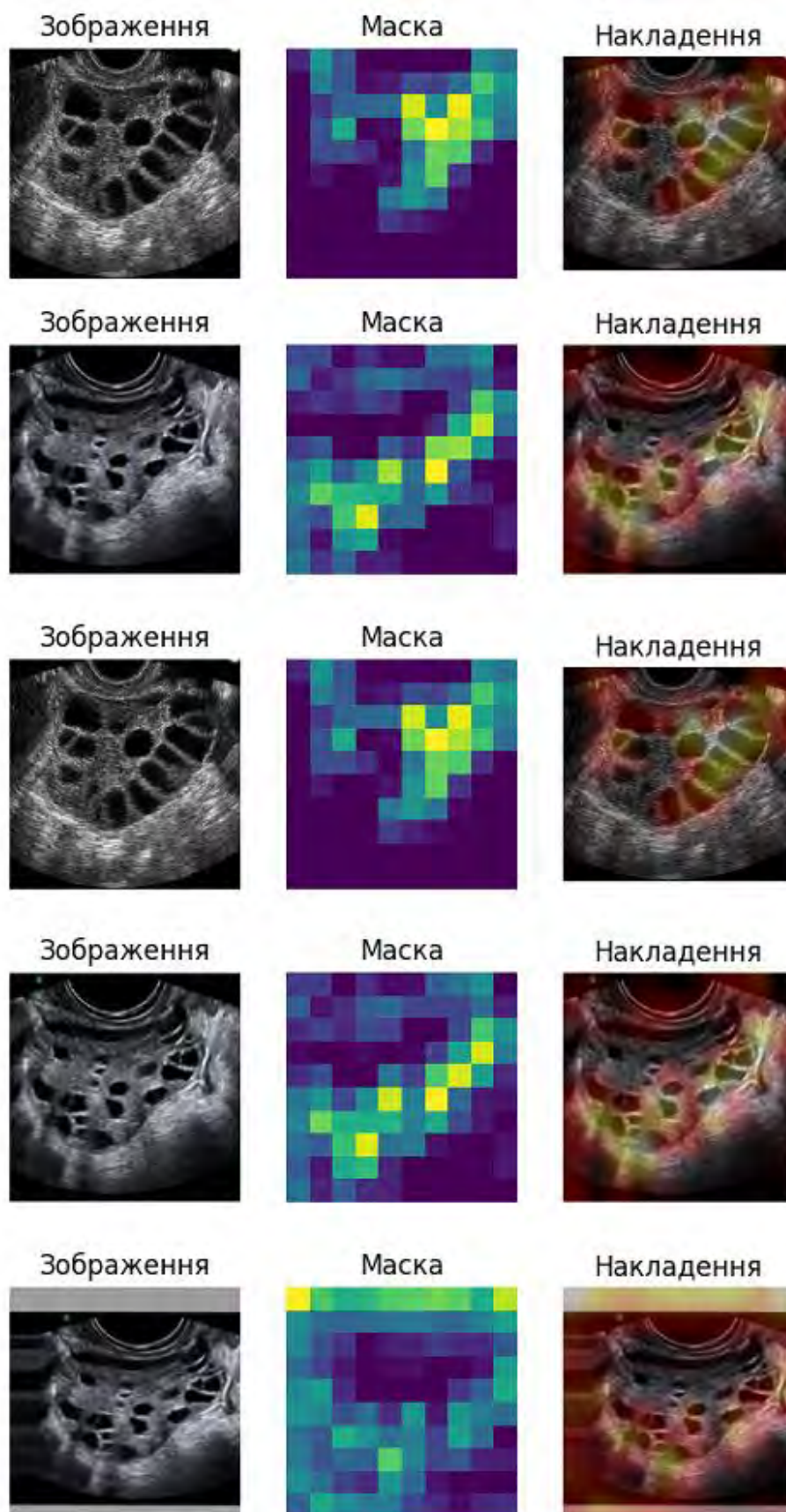


Рисунок 5.2 Візуалізація активаційних шарів

За допомогою методу GRAD-CAM ми реалізували модуль інтерператації передбачення моделі, що дасть змогу медичним співробітникам швидше виявляти дефекти тканин, що потенційно спростить їхній робочий або навчальний процес.

## **6 Висновки**

У дипломній роботі на тему "Програмне забезпечення для аналізу ультразвукових зображень" з використанням згорткової нейронної мережі та GRAD-CAM для інтерпретації, була досягнута мета автоматизації ультразвукової діагностики полікістозу яєчників. Для цього було спроектовано та навчено власну згорткову нейронну мережу, яка здатна розпізнавати ознаки полікістозу на ультразвукових зображеннях жіночих яєчників.

Використання мережі дозволило ефективно виявляти ознаки полікістозу на ультразвукових зображеннях. Модель була навчена на великому обсязі даних, що дозволило їй розпізнавати широкий спектр ознак, пов'язаних з полікістозом. Застосування GRAD-CAM підсилило інтерпретованість результатів, надаючи важливу інформацію про області, на які модель звертала увагу при класифікації. Отримані результати вказують на потенційну можливість використання розробленого програмного забезпечення для автоматичного виявлення полікістозу на ультразвукових зображеннях.

Дані отримані з дослідження підтверджують перспективи використання штучного інтелекту та глибокого навчання в області медичної діагностики. Програмне забезпечення, розроблене в рамках дипломної роботи, може стати цінним інструментом для лікарів та медичних спеціалістів, які займаються діагностикою полікістозу яєчників. Подальші дослідження і вдосконалення системи можуть привести до покращення її точності та розширення застосування на інші області медицини, де важливим є аналіз ультразвукових зображень.



## 7 Список літератури

1. A review: Brief insight into Polycystic Ovarian syndrome by Jeshica Bulsara, Priyanshi Patel, Arun Soni and Sanjeev Acharya. -  
<https://www.sciencedirect.com/science/article/pii/S266639612100008X>
2. Polycystic ovary syndrome: improving policies, awareness, and clinical care by Rhonda M. Garad and Helena J. Teede. -  
<https://www.sciencedirect.com/science/article/pii/S2451965020300417>
3. The Prevalence of Polycystic Ovary Syndrome: A Brief Review.-  
<https://pubmed.ncbi.nlm.nih.gov/33627974/>
4. The Prevalence of Polycystic Ovary Syndrome: A Brief Systematic Review by Ritu Deswal, Vinay Narwal, Amita Dang and Chandra S. Pundir. -  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7879843/>
5. Epidemiology, diagnosis, and management of polycystic ovary syndrome by Susan M Sirmans and Kristen A Pate. -  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3872139/>
6. Polycystic Ovarian Disease.-  
<https://www.ncbi.nlm.nih.gov/books/NBK459251/>
7. Basics of Ultrasound Imaging by Vincent Chan and Anahi Perlas. -  
[https://www.researchgate.net/publication/225262445\\_Basics\\_of\\_Ultrasound\\_Imaging](https://www.researchgate.net/publication/225262445_Basics_of_Ultrasound_Imaging)
8. Preclinical Ultrasound Imaging—A Review of Techniques and Imaging Applications by Carmel M. Moran and Adrian J. W. Thomson. -  
<https://www.frontiersin.org/articles/10.3389/fphy.2020.00124/full>
9. Deep learning-enabled medical computer vision by Andre Esteva, Katherine Chou, Serena Yeung, Nikhil Naik, Ali Madani, Ali Mottaghi, Yun Liu, Eric Topol, Jeff Dean and Richard Socher. -  
<https://www.nature.com/articles/s41746-020-00376-2>

10. Deep learning in computer vision: A critical review of emerging techniques and application scenarios -  
<https://www.sciencedirect.com/science/article/pii/S2666827021000670>
11. Support Vector Machines Applications -  
<https://link.springer.com/book/10.1007/978-3-319-02300-7>
12. Convolutional Neural Network in Medical Image Analysis: A Review by Sapna Singh Kshatri and Deepak Singh.-  
<https://link.springer.com/article/10.1007/s11831-023-09898-w>
13. Convolutional neural networks: an overview and application in radiology.-  
<https://insightsimaging.springeropen.com/articles/10.1007/s13244-018-0639-9>
14. Random Forests -  
<https://link.springer.com/article/10.1023/A:1010933404324>
15. Computer Vision - <https://link.springer.com/book/10.1007/978-1-84882-935-0>
16. Deep Learning in Computer Vision: Principles and Applications -  
[https://www.researchgate.net/publication/339569977\\_Deep\\_Learning\\_in\\_Computer\\_Vision\\_Principles\\_and\\_Applications](https://www.researchgate.net/publication/339569977_Deep_Learning_in_Computer_Vision_Principles_and_Applications)
17. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization - <https://link.springer.com/article/10.1007/s11263-019-01228-7>
18. R. M. Galagan and A. S. Momot, "The use of ART-2 neural network for processing information signals of non-destructive testing", 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kiev, 2017, pp. 981-985 (doi: 10.1109/UKRCON.2017.8100395)
19. Momot A.S., Galagan R.M., Gluhovskii V.Yu. Deep Learning Automated System for Thermal Defectometry of Multilayer Materials. Devices and Methods of Measurements. 2021; Vol. 12(2): pp. 98-107.
20. Momot, A., Galagan, R., Zabolueva M. Automation of ultrasound breast cancer images classification using deep neural networks. Sciences of Europe. 2022. No96. pp. 38–41.

21. Галаган Р.М., Андреев С.М., Петрик В.Ф., Баженов В.Г., Лисенко Ю.Ю. Виявлення дефектів бетонних конструкцій на основі аналізу зображень за допомогою згорткових нейронних мереж. Вчені записки ТНУ імені В.І. Вернадського. Серія: Технічні науки. 2023. Том 34 (73), № 2. с. 138-144.