

3.4 Методичний підхід з інтелектуальної оцінки та прогнозування радіоелектронної обстановки

Вступ

Системи підтримки прийняття рішень (СППР) активно використовуються в усіх сферах життєдіяльності людей. Створення інтелектуальних СППР стало природним продовженням широкого застосування СППР класичного типу. Інтелектуальні СППР знайшли широке використання для вирішення специфічних завдань військового призначення, а саме [161, 162]: планування розгортання, експлуатації систем зв'язку та передачі даних; автоматизація управління військами та зброєю; збір, обробка та узагальнення розвідувальних відомостей про стан об'єктів розвідки та ін.

Одним з таких завдань, де використовуються інтелектуальні СППР є аналіз та прогнозування радіоелектронної обстановки, що також зводиться до багатокритерійного аналізу альтернатив, враховуючи велику чисельність та різнотипність угруповань військ (сил). З практичної точки зору від ефективності аналізу та прогнозування радіоелектронної обстановки залежить ефективність використання радіочастотного ресурсу та ефективність функціонування систем зв'язку (радіоелектронного подавлення). Разом з тим, при аналізі завдань радіоелектронної обстановки (РЕО) виникає ряд проблемних питань, а саме:

1. Сигнали аналізуються в складній радіоелектронній обстановці на фоні різних за походженням завад.
2. Вхідні сигнали не збігаються з еталонами за рахунок впливу різних типів завад.
3. Інтерпретація сигналів залежить від досвіду оператора (експерта), повноти додаткової інформації по конкретному завданні (умови невизначеності).

Проведений аналіз робіт [1–16] свідчить про те, що найбільш ефективними засобами обробки радіоелектронної обстановки є згорточні нейронні мережі (ЗНМ). Згорточна нейронна мережа будується на основі операції згортки, що дозволяє навчати ЗНМ на окремих частинах даних, ітераційно збільшуючи

локальну область навчання окремого ядра згортки.

Встановлено, що на даний момент не існує методології побудови гібридних згорткових нейронних мереж, які забезпечують покращення точності розв'язання задачі при зниженні об'єму обчислювальних та часових витрат. Тому, виникає потреба у визначенні властивостей унікальних блоків та їх використання для створення нових топологій, а саме, гібридних згорткових нейронних мереж.

Матеріали та методи дослідження

При розробці методичного підходу з інтелектуальної оцінки та прогнозування радіоелектронної обстановки використовувалися методи:

гібридних згорткових нейронних мереж – для отримання узагальненої оцінки стану радіоелектронної обстановки для вирішення задачі аналізу та прогнозування стану РЕО;

штучних нейронних мереж, що еволюціонують – для навчання та корегування баз знань радіоелектронної обстановки. Для навчання штучних нейронних мереж, що еволюціонують використаний розроблений в попередніх роботах метод навчання штучних нейронних мереж, який дозволяє проводити глибоке навчання штучних нейронних мереж, який полягає в навчанні архітектури, виду та параметрів функції належності.

Моделювання методики оцінки та прогнозування РЕО в інтелектуальних системах підтримки прийняття рішень проводилося при наступних параметрах:

– засоби радіозв'язку з псевдовипадковою перестройкою робочої частоти (ППРЧ). Діапазон частот – 30–512 МГц; потужність передавача – 10 Вт; ширина смуги частот, що випромінюється – 12,5 кГц, чутливість приймача –120 дБ; кількість засобів радіозв'язку (ЗРЗ) в мережі – 5. Кількість частотних каналів для переналаштування засобів радіозв'язку (ЗРЗ) – 10000; кількість перестроювання – від 333,5 до 1000 стрибків/сек;

– комплекси радіоелектронного подавлення (РЕП)–2. Діапазон частот комплексів РЕП – 30–2000 МГц; потужність передавача – 2000 Вт; максимальна

смуга частот, що може бути подавлена одночасно – 80 МГц. Тип завади комплексів РЕП – шумова загороджувальна завада з частотною маніпуляцією, як одна з найбільш поширених та вплив якої добре відомий; стратегія комплексу РЕП-динамічна.

Моделювання проводилося з використанням програмного забезпечення MathCad 2014 та ПЕОМ Intel Core i3. До ПЕОМ було підключено 5 програмованих прийомопередавачів LimeSDR (США) з програмним забезпеченням GNU Radio (Німеччина) та підключено до генератора шуму RIGOL DG5252 (Німеччина), що імітував роботу комплексу РЕП.

Для проведення експерименту використовувалася навчальна вибірка, що містить дані про стан радіоелектронної обстановки угруповання. Для експериментів використовувалося 5000 спостережень із цієї вибірки. Навчальна вибірка містила 3000 спостережень, тестова – 2000 спостережень.

Під час розв'язання прикладних задач з метою підвищення точності та зниження складності виникають кілька завдань. Перше завдання – пошук оптимальної топології мережі. Друге – структурна (визначення кількості прихованих шарів та нейронів в них, міжнейронних зв'язків окремих НМ) і параметрична (налаштування вагових коефіцієнтів) оптимізація. Це дозволяє розв'язувати складні задачі, в першу чергу, збір, обробка та узагальнення відомостей про радіоелектронну обстановку, аналіз якої неможливо розв'язати на основі окремих методів або технологій.

Результатом впровадження даної роботи є інтелектуальні системи аналізу радіоелектронної обстановки.

Залежно від типу архітектурних модифікацій ЗНМ можна, в цілому, розділити на сім різних класів, а саме: ЗНМ на основі просторового використання; ЗНМ на основі використання глибини; розгалужені ЗНМ; ЗНМ з безліччю підключень на основі ширини; ЗНМ на основі використання карти ознак; ЗНМ на основі бустінгу каналів; ЗНМ на основі використання механізму уваги. Класифікацію архітектур ЗНМ наочно представлено на рис. 1 [171].

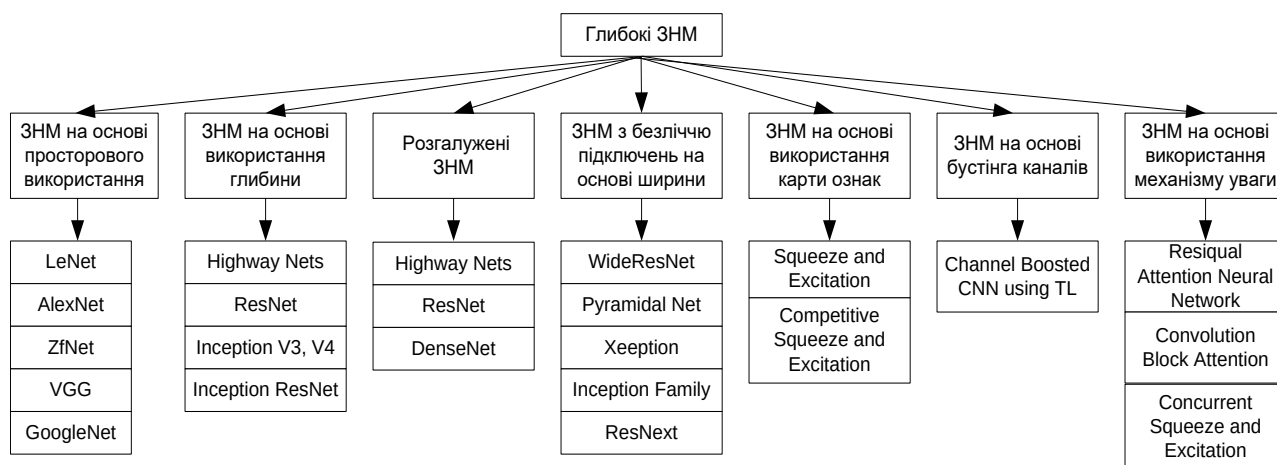


Рис. 1. Класифікація глибоких архітектур ЗНМ

Характерною особливістю сучасних ЗНМ є наявність унікальних блоків, які визначають їх суттєві особливості. Наприклад, блок стиснення та збудження, згортковий блок модуля уваги, модуль уваги каналу, модуль просторової уваги, залишковий блок, початковий модуль, блок ResNeXt [172–178]. Топології вищеназваних блоків (модулів) представлено на рис. 1. До унікальних блоків (шарів) слід віднести пакетну нормалізацію (англ. Batch Normalization) та блок спрощення (англ. Block Reduction).

Пакетна нормалізація

Пакетну нормалізацію можна розглядати як ще один шар, який можна вставити в архітектуру моделі, як і повнозв'язний або згортковий шари [178]. На практиці шари пакетної нормалізації вставляються після згорткового або повнозв'язного шару, але до того, як вихідні дані передаються в функцію активації. Пакетна нормалізація використовується для нормалізації вхідного шару шляхом повторного центрування і масштабування. Кожен шар нейронної мережі має входи з відповідним розподілом, на яке в процесі навчання впливає випадковість в ініціалізації параметрів і у вхідних даних. Вплив цих джерел випадковості на розподіл входів у внутрішні шари при навчанні визначається як внутрішній коваріантний зсув. Коротко зупинимося на функціональних характеристиках окремих унікальних блоках.

Блок спрощення

Проблема глибоких згорткових нейронних мереж полягає в тому, що кількість карт ознак часто збільшується із глибиною мережі. Ця проблема може призвести до різкого збільшення кількості параметрів та обчислень, необхідних при використанні більших розмірів фільтрів (ядер згортки), таких як 5×5 та 7×7 .

Для вирішення цієї проблеми може бути використаний згортковий шар 1×1 , який пропонує об'єднання каналів, яке часто називають субдискретизацією (агрегацією) карт ознак або проєкційним шаром [170]. Цей простий прийом можна використовувати для зменшення розмірності, зменшуючи кількість карт ознак, зберігаючи при цьому свої помітні риси. Це також може бути використано безпосередньо для створення індивідуальної проєкції карт ознак для об'єднання ознак по каналах або для збільшення кількості карт ознак, наприклад, після традиційних шарів субдискретизації.

Блок стиснення і збудження

Структуру блока стиснення і збудження (БСЗ) (англ. Squeeze and excitation block (SEB)) представлено на рис. 2.

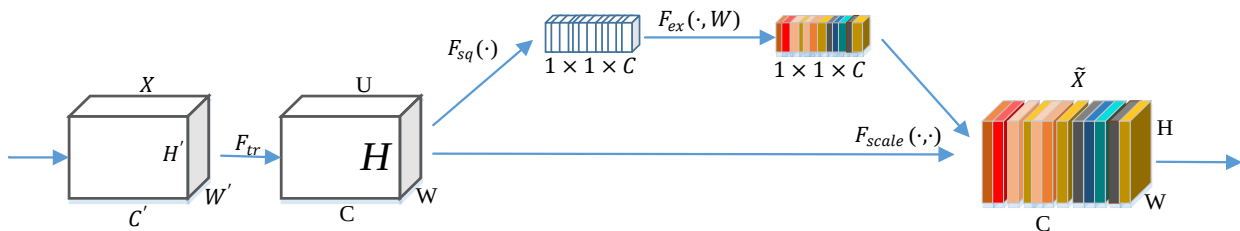


Рис. 2. Структура блоку стиснення і збудження

У блоці стиснення і збудження виконуються наступні операції: перетворення ознак або просту операцію згортки до вхідних даних X , щоб отримати ознаки U ; операція стиснення для отримання одного значення виходу U для кожного каналу; операція збудження, яка застосовується до виходів стислих даних, щоб отримати вагові коефіцієнти для кожного каналу; зміна масштабу карти ознак U з цими активаціями, щоб отримати результуючі вихідні дані блоку БСЗ. Роль, яку ця операція виконує на різній глибині, розрізняється по мережі.

На більш ранніх рівнях БСЗ збуджує інформативні ознаки незалежно від класу, посилюючи загальні уявлення нижнього рівня. На більш пізніх рівнях блоки БСЗ стають все більш спеціалізованими і реагують на різні входні дані, які строго залежать від специфіки класу.

Блок магістральної мережі

Магістральні мережі дозволяють безперешкодний потік інформації через багатоманітні шари інформації за рахунок міжшарового підключення. Структуру блока магістральної мережі (БММ) (англ. Highway network block (HNB)) представлено на рис. 3.

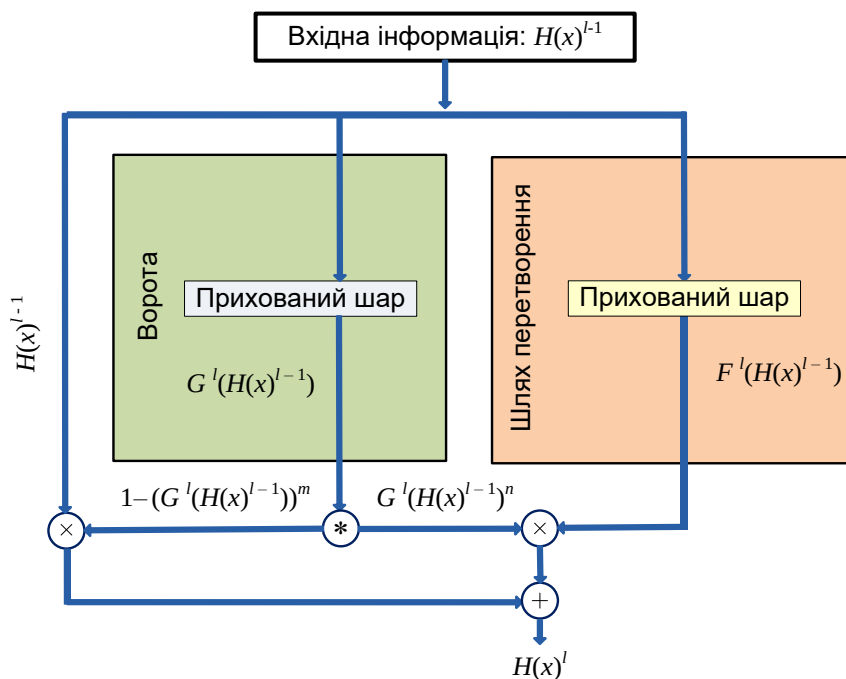


Рис. 3. Структура блоку магістральної мережі: $H(x)^{l-1}$ – інформація на шарі $(l-1)$ магістральної мережі, $G^l(H(x)^{l-1})$ – вихід модуля селекції, $F^l(H(x)^{l-1})$ – перетворення навчене на прихованому шарі l для входу $H(x)^{l-1}$

В роботі [177] навчання блоку магістральної нейронної мережі переформульовано для досягнення наступного: віддавати перевагу оптимізації на ранніх етапах навчання, коли на селекцію в основному спрямовують неперетворені ознаки; зосередитися на перетворенні навчальних ознак на більш пізніх етапах навчання, коли через селекції в основному проходять перетворені ознаки; покладатися на набагато меншу кількість селекцій для навчання

декількох шарів перетворення ознак, оскільки перехід селекцій більш ефективний для оптимізації і узагальнення моделі. Для досягнення вищезазначених характеристик навчання пропонується новий магістральний блок вигляду (рис. 4).

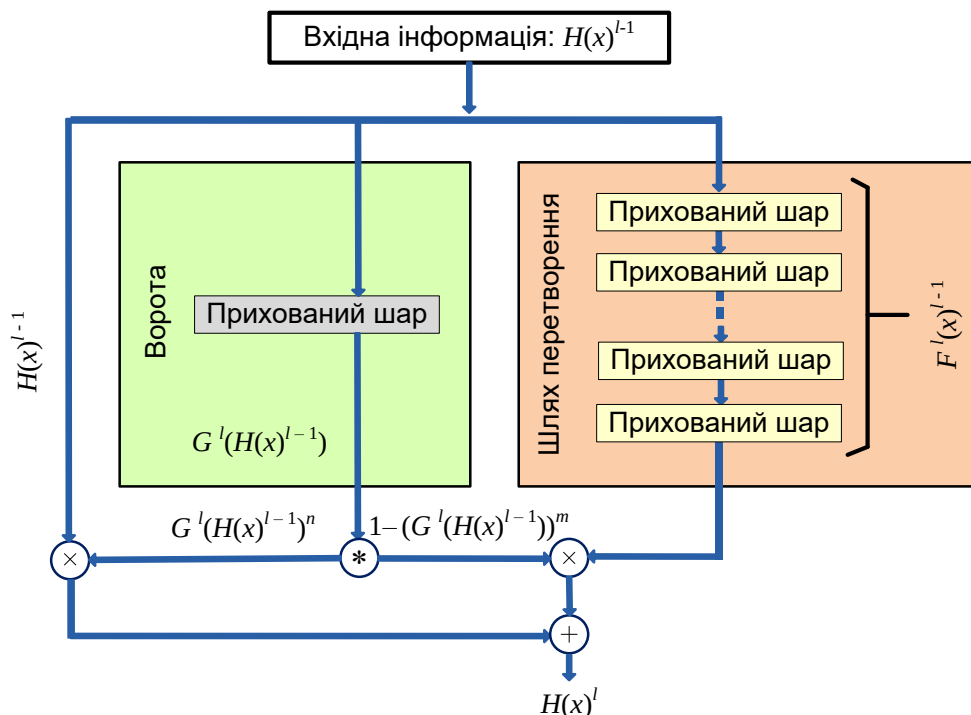


Рис. 4. Запропонований блок магістральної мережі: $H(x)^{l-1}$ – інформація на шарі $(l-1)$ магістральної мережі, $G^l(H(x)^{l-1})$ – вихід модуля селекції, $F^l(H(x)^{l-1})$ – перетворення навчене на прихованому шарі l для входу $H(x)^{l-1}$

Блоки магістральної мережі доцільно використовувати в ГЗНМ з метою покращення їх ефективності в задачах обробки відеоінформації.

Залишковий блок

Основна ідея розробки залишкового блоку полягає у наступному. Глибокі мережі витягають низько-, середньо-і високорівневі ознаки наскрізним багатошаровим способом, а збільшення кількості шарів або блоків, що складаються з декількох шарів може збагатити рівні ознак. Однак зі збільшенням глибини мережі навчання стає нестабільним, а досягається точність починає зменшуватися (деградувати). Це пов'язано із затуханням градієнта при зворотному проході в методі зворотного поширення помилки і, як наслідок,

погіршення роботи нейронної мережі. Структурну схему залишкового блока (англ. residual block), зображено на рис. 5.

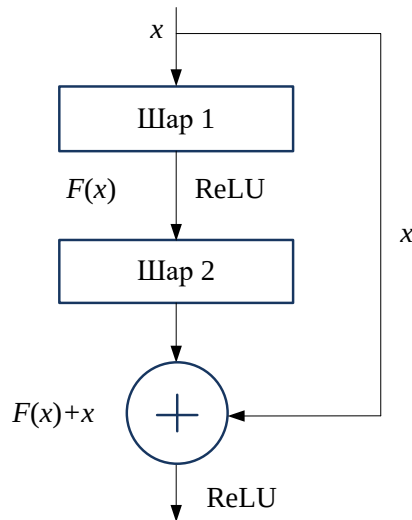


Рис. 5. Структурна схема залишкового блока

У даному блоці використовуються так звані швидкі з'єднання (англ. shortcut connections), тобто явно додається тотожне відображення. В результаті при зворотному проході в методі зворотного поширення помилки отримуємо: $dF(x)/dx + 1$. Таким чином, загасання градієнта не відбудеться, тому що завжди буде виконаний зворотній прохід. Такий дизайн вимагає, щоб вихідні дані двох згорткових шарів мали ту ж форму, що і вхідні, щоб їх можна було скласти докупи. Якщо потрібно змінити кількість каналів, необхідно ввести блок спрощення (додатковий згортковий шар 1×1), щоб перетворити вхідну інформацію x в бажану форму для виконання операції додавання. Згортка 1×1 просто відображає вхідний піксель з усіма його каналами на вихідний піксель, не дивлячись ні на що навколо себе. Така згортка часто використовується для зменшення кількості каналів глибини. З урахуванням необхідності використання пакетної нормалізації та додаткового згорткового шару 1×1 структура залишкового блоку має вигляд, представлений на рис. 6. Таким чином, наступний шар не витрачає свої ступені свободи на зсув і масштабування вхідних даних, а займається тільки оцінюванням їх структурних властивостей, як наслідок, отримується прискорення збіжності процесу навчання.

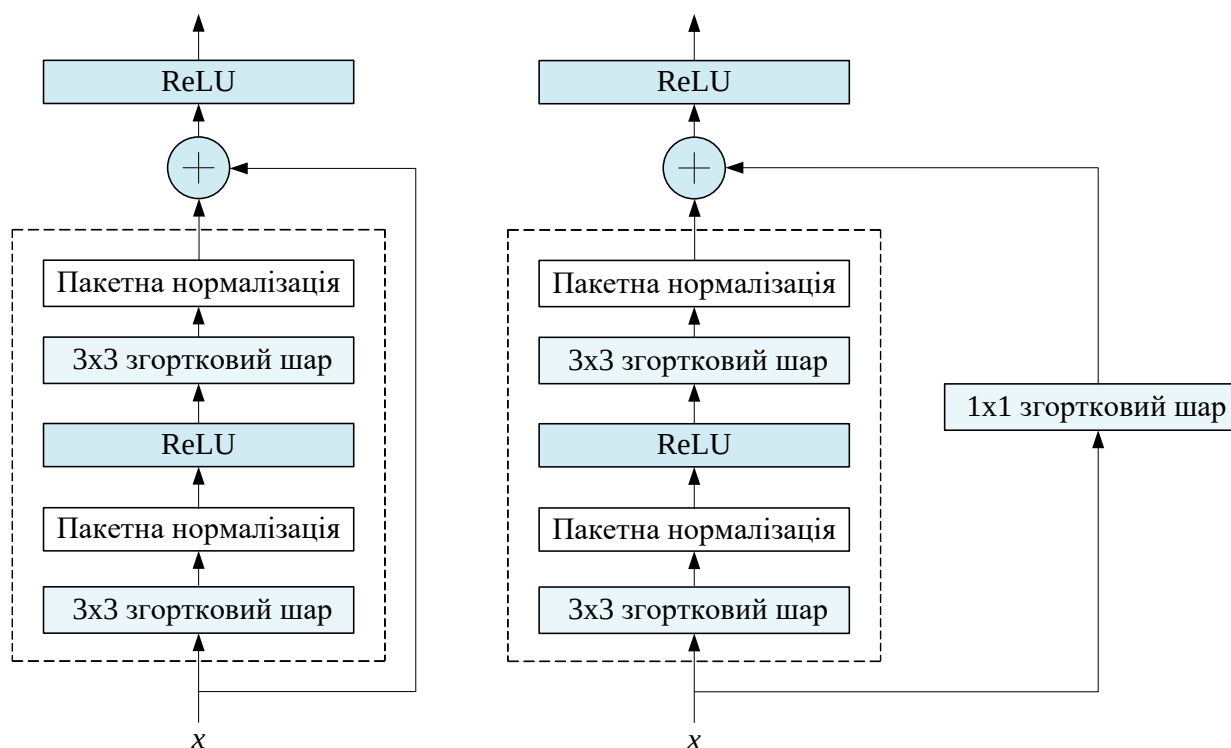


Рис. 6. Залишковий блок при наявності та відсутності згорткового шару 1×1

Використання залишкового блоку в складі ГЗНМ дозволяє знизити обчислювальні витрати та покращити результати обробки за рахунок збільшення кількості шарів НМ.

Блок Inception

Існує два способи збільшення якості мережі – збільшення глибини і збільшення ширини, але це пов’язано з підвищенням ймовірності перенавчання та неефективним використанням обчислювальних ресурсів – розріджена структура згорток неефективна в обчислювальному сенсі. Для усунення цих недоліків використовується кореляційна структура активацій попередніх шарів. Структуру блоку Inception наведено на рис. 7.

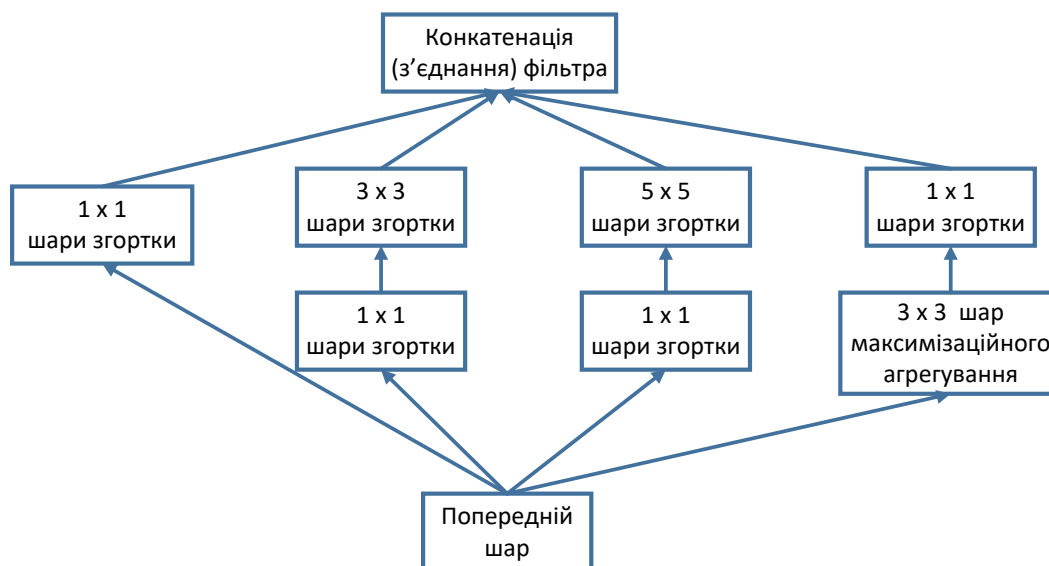


Рис. 7. Структура блока Inception

У кожному блоці є шари згортки з ядром різного розміру, щоб розпізнавати ознаки різного масштабу. Також в даній моделі активно використовуються згортки 1×1 , щоб зменшувати розмірність тензорів, які будуть подаватися на вхід наступного шару. Для того щоб не втратити інформацію, отриману в попередньому шарі також застосовується субдискритизуючий шар. Після нього також застосовується згортковий шар з ядром згортки 1×1 , в даному випадку для того, щоб вирівняти розмірність тензорів на виході після кожного паралельного шару. Потім відбувається конкатенація карт ознак, отриманих на кожному паралельному шарі.

Блок уваги

Механізми уваги – це підхід в машинному навчанні, що полягає у виділенні частини вхідних даних. Проте, згорткові мережі, які є найбільш популярним методом вирішення такого завдання, витрачають однакову кількість обчислювальних ресурсів на всі частини даних. Блок (модуль) уваги реалізується в двох варіантах: модуль уваги каналу та модуль просторової уваги. Модуль уваги каналу. Створюється карта уваги каналу, використовуючи взаємозв'язок ознак між каналами. Оскільки кожен канал карти ознак розглядається як детектор об'єктів, увага каналу зосереджується на тому, що має значення з урахуванням вхідних даних. Щоб ефективно обчислити увагу каналу, стискаємо

просторовий розмір карти вхідних характеристик. Для узагальнення просторової інформації дотепер загальноприйнятим є усереднювальне агрегування (рис. 8).

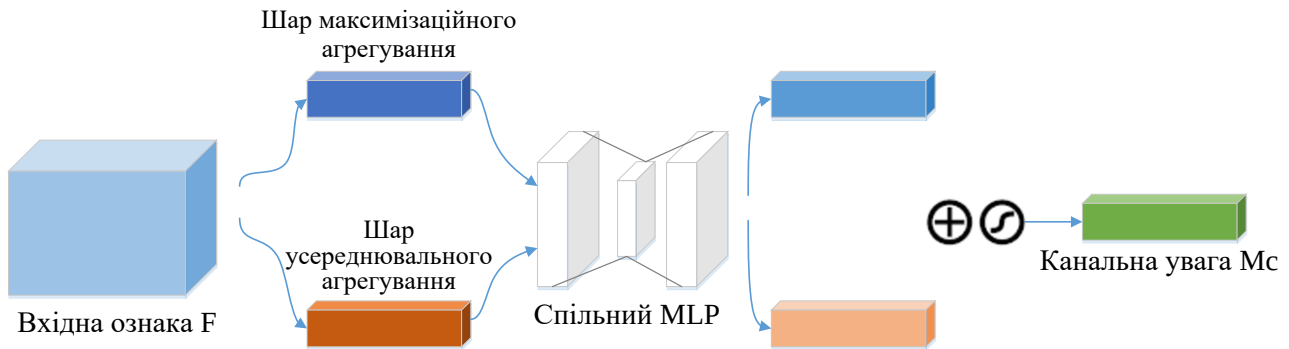


Рис. 8. Структурна схема модуля каналної уваги (англ. Channel Attention Module)

Модуль просторової уваги. Генерується просторова карта уваги, використовуючи міжпросторовий взаємозв'язок ознак. На відміну від уваги каналу, просторова увага зосереджується на тому, де є інформативна частина, яка доповнює увагу каналу. Для обчислення просторової уваги спочатку застосовуємо операції середнього та максимального об'єднання уздовж осі каналу та об'єднуємо їх для створення ефективного дескриптора ознак. Показано, що застосування операцій об'єднання уздовж осі каналу ефективно впливає на виділення інформативних областей. На об'єднаному дескрипторі ознак застосовано шар згортки, щоб сформувати просторову карту $M_s(F) \in \mathbb{R}^{H \times W}$ уваги, яка кодує, де слід підкреслити або придушити. Детальну структурну схему наведено на рис. 9.

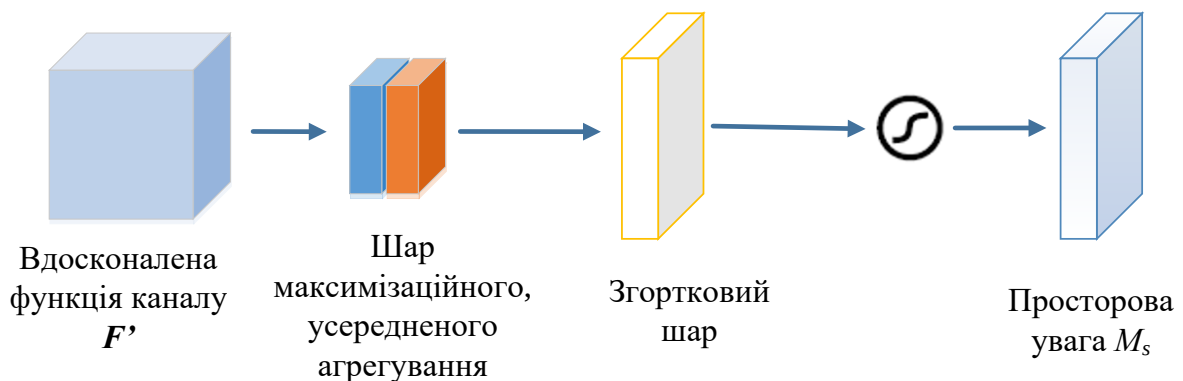


Рис. 9. Структурна схема модуля просторової уваги (англ. Spatial Attention Module)

Однією з основних проблем, що стримує подальший прогрес використання ЗНМ, є великий архітектурний простір параметрів, включаючи тип унікального блоку, місце розміщення в структурі ЗНМ, його зв'язки з іншими блоками та шарами. В результаті виникає задача структурно-параметричного синтезу ЗНМ.

Синтез гібридних згорткових нейронних мереж, які використовуються головним чином для обробки зображень значно складніше ніж синтез згорткових нейронних мереж. Такий синтез вимагає визначення типу унікальних блоків, які будуть використовуватися, узгодженні з сусідніми блоками, місць їх розміщення.

Ці складнощі вимагають спочатку розв'язання структурно-параметричного синтезу класичної ЗНМ на основі заданої навчальної вибірки. Класична ЗНМ складається із згорткових шарів, кожен нейрон яких виконує згортку деякої області попереднього шару, шарів агрегації (субдискретизації), що виконують функції зменшення розмірності карти ознак, і повнозв'язних шарів (класифікатор, який знаходиться на виході мережі). Згорткові шари та шари агрегації можуть чергуватися. Найчастіше шари агрегації розміщують за шарами згортки [176, 177]. Такий метод структурно-параметричного синтезу звичайних ЗНМ наведено в розділі 5. Далі ця мережа буде розглядатися як базова. Утворення гібридної структури за рахунок включення до складу базової ЗНМ унікальних блоків або ансамблю гібридних згорткових нейронних мереж дає нові можливості для підвищення ефективності розв'язання поставленої задачі. Задача структурно-параметричного синтезу гібридних ЗНМ розв'язується на основі застосування багатоетапної процедури визначення структури та параметрів гібридної ЗНМ з формуванням її двійкового представлення.

На основі вище наведеного методологію синтезу гібридних нейронних мереж можна представити у вигляді такої послідовності дій:

1. Структурно-параметричний синтез базової ЗНМ.
2. Структурно-параметричний синтез гібридної ЗНМ з визначенням типів та послідовності унікальних блоків, які вводяться до складу гібридної ЗНМ.
3. Структурно-параметричний синтез ансамблю, що складається з окремих

гібридних ЗНМ.

Структурно-параметричний синтез базової ЗНМ

Основою побудови ЗНМ є використання операції згортки з метою можливості навчання ЗНМ на окремих частинах вхідних даних. Розмір цих даних визначається розмірністю відповідного ядра згортки. При цьому слід зазначити, що нейрони, які відповідають одному і тому ж ядру згортки мають спільні ваги, що забезпечує зниження обчислювальних витрат ЗНМ в порівнянні, наприклад з багат шаровим персептроном. Шари НМ, які побудовані таким чином, називаються згортковими шарами.

З метою зменшення обчислювальних витрат до складу ЗНМ входять шари агрегації(субдискретизації), що виконують функції зменшення розмірності карти ознак. На основі визначених ознак формується аномальна область, для класифікації якої використовуються повнозв'язні шари (класифікатор, який знаходиться на виході мережі). Згорткові мережі будуються за правилом, а саме: спочатку розміщуються згорткові шари, їх кількість та визначається в результаті розв'язання задачі структурно-параметричного синтезу, а далі розміщується шар агрегації [172, 173]. Кількість таких повторів в мережі залежить від складності задачі.

Основними параметрами згорткової нейронної мережі [172] є: розмір ядра згортки (фільтра); кількість ядер згортки (залежить від кількості згорткових шарів); величина зміщення при пересуванні ядра згортки по матриці вхідних даних (крок ядра згортки); параметри (по вертикалі та горизонталі) урахування крайових ефектів (початкове положення ядра згортки на матриці або карті ознак перед початком переміщення з метою побудови карт ознак); початкове заповнення ядер згортки.

Як можна побачити, кількість параметрів, оптимальне значення яких треба визначити при структурно-параметричному синтезі ЗНМ може бути дуже великим. Тому необхідно скоротити кількість параметрів. Загальні правила не сильно впливають на цей процес – все залежить від навчальної вибірки. Тому пропонується визначати найбільш значущі з точки зору ефективності параметри

ЗНМ в результаті експерименту над згортковою нейронною мережею. Експеримент проводиться наступним чином: крок за кроком змінюється один з параметрів ЗНМ при фіксованих інших параметрах і визначається наскільки сильно змінився вихід.

Враховуючи велику кількість параметрів, за якими треба провести оптимізацію вищенаведених критеріїв для оптимізації структури та параметрів ЗНМ використано багатокритеріальний генетичний алгоритм [163].

Оптимізація структури та параметрів гібридної ЗНМ

Як було сказано вище, відповідно до запропонованої методології у результаті виконання першого етапу було отримано базову згорткову НМ, яка не має унікальних блоків. Назвемо гібридною згортковою мережею мережу, до складу якої входять різні унікальні блоки, вибір і розміщення яких визначимо на підставі використання генетичного алгоритму. Можливо використання більш складних блоків, а саме: Attention-Inception Block; Inception-Dense Block; Inception-Residual Block; Inception-Recurrent Residual Block; Deep Pyramid-Highway Block; Deep Pyramid-Self Attention, Squeeze and Excitation-Attention Block; Squeeze and Excitation-Dense Block; Squeeze and Excitation-Inception Block; Squeeze and Excitation-Residual Block; Residual-Squeeze and Excitation-Attention Block; Residual-Attention Block.

Багато сучасних мережних структур [174, 175] можуть ділитися на кілька угруповань. У кожному угрупованні геометричні розміри (ширина, висота й глибина) куба угруповання залишаються без змін. Сусідні угруповання зв'язані просторовою операцією субдискретизації. Структура базової НМ у загальному випадку представляє чергування двох згорткових шарів, за яким йде шар субдискретизації, тому гібридна ЗНМ будується на основі базової ЗНМ шляхом заміни згорткових шарів на угруповання. Всі згорткові операції в одному угрупованні мають одну й ту саму кількість фільтрів або каналів.

Надається двійкове строкове представлення для структури мережі в обмеженому випадку. Насамперед відзначимо, що багато сучасних мережних структур можуть бути поділені на кілька блоків. У кожному блоці геометричні

розміри (ширина, висота й глибина) куба шару залишаються без змін. Сусідні блоки зв'язані просторовою операцією субдискретизації (англ. pooling), що може змінити просторову роздільність. Всі згорткові операції на одному блоці мають одну й ту саму кількість фільтрів або каналів (рис. 10).

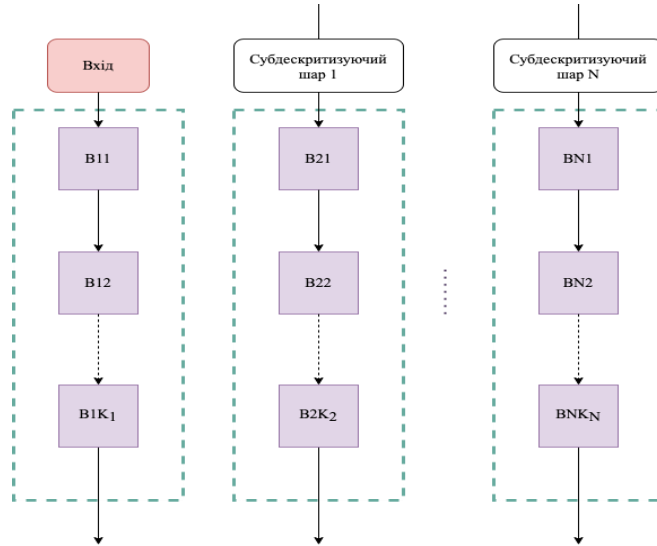


Рис. 10. Багатоблокова гібридна згорткова нейронна мережа

Кожна гібридна ЗНМ складається із S угруповань, а s -е угруповання, $s=1, 2, \dots, S$, містить K_s блоків, позначуваних v_{s,k_s} , $k_s=1, 2, \dots, K_s$. Вузли у кожному блоці впорядковані, і дозволяється з'єднання від вузла з меншим номером до вузла з більш високим номером. Кожному вузлу відповідає унікальний блок. Повнозв'язна частина мережі не кодується. У кожному угрупованні використаємо $\frac{1}{2} K_s (K_s - 1)$ бітів для кодування міжвузлових зв'язків. Перший біт представляє зв'язок між $(v_{s,1}, v_{s,2})$, тоді наступні два біти представляють з'єднання, між $(v_{s,1}, v_{s,3})$ і $(v_{s,2}, v_{s,3})$, тощо. Цей процес триває доти, поки останні біти $K_s - 1$ не будуть використовуватися для подання з'єднання між $v_{s,1}, v_{s,2}, \dots, v_{s,K_s-1}$ і v_{s,K_s} . Для $1 \leq i < j \leq K_s$, якщо код, що відповідає $(v_{s,i}, v_{s,j})$, дорівнює 1, існує ребро, що з'єднує $v_{s,i}$ і $v_{s,j}$, тобто $v_{s,j}$ приймає вихід $v_{s,i}$ у складі по-елементного підсумовування, і навпаки. Для навчання згорткової нейронної мережі в даній роботі використовується нормалізована ініціалізація, яка носить назву ініціалізація Glorot [172].

Процес селекції виконується на початку кожного покоління. До t -го

покоління n -му індивідові $M_{t-1,n}$ призначається фітнес-функція, що визначається як швидкість розпізнавання $r_{t-1,n}$, отримана в попередньому поколінні або ініціалізації $r_{t-1,n}$ прямо впливає на ймовірність, що $M_{t-1,n}$ зберігається в процесі вибору [164].

Процес мутації окремого $M_{t,n}$ включає незалежну заміну кожного біта з імовірністю q . На практиці q часто мале, наприклад 0,05, так що мутація навряд чи занадто сильно змінить один індивід. Це робиться для того, щоб зберегти гарні якості індивіда, що вижив, надаючи можливість випробувати нові можливості. Процес кросовера включає одночасну зміну двох індивідів. Замість того, щоб розглядати кожний біт окремо, основною ланкою кросовера є етап, що мотивується необхідністю зберігати локальні структури усередині кожної частки. Подібно мутації, кожна пара відповідних часток обмінюється з невеликою ймовірністю q [164]. Для налаштування вагових коефіцієнтів НМ використовується двоетапний алгоритм: на першому етапі застосовується генетичний алгоритм, а на другому, з метою покращення результатів налаштування, градієнтний алгоритм AdaM.

Агрегація результатів відбувається одним із трьох методів:

- Ensemble-ADD (використовується на рис. 11): об'єднує результати Mask R-CNN, DeepLabV3 і Deep Pyramid Attention Module для створення остаточної маски сегментації;

- Ensemble-Comparison-Large: вибирає велику сегментовану область, порівнюючи кількість даних у вихідних даних всіх сегментатор.

- Ensemble-Comparison-Small: навпаки, вибирає меншу сегментовану область на виході всіх сегментатор.

В роботі [170] для побудови ансамблю використано метод Ensemble-ADD. Ensemble-ADD(CША) об'єднує результати як Mask R-CNNADD (CША), DeeplabV3+ADD (CША), Deep Pyramid-Attention ModuleADD (CША) щоб створити остаточною маску сегментації.

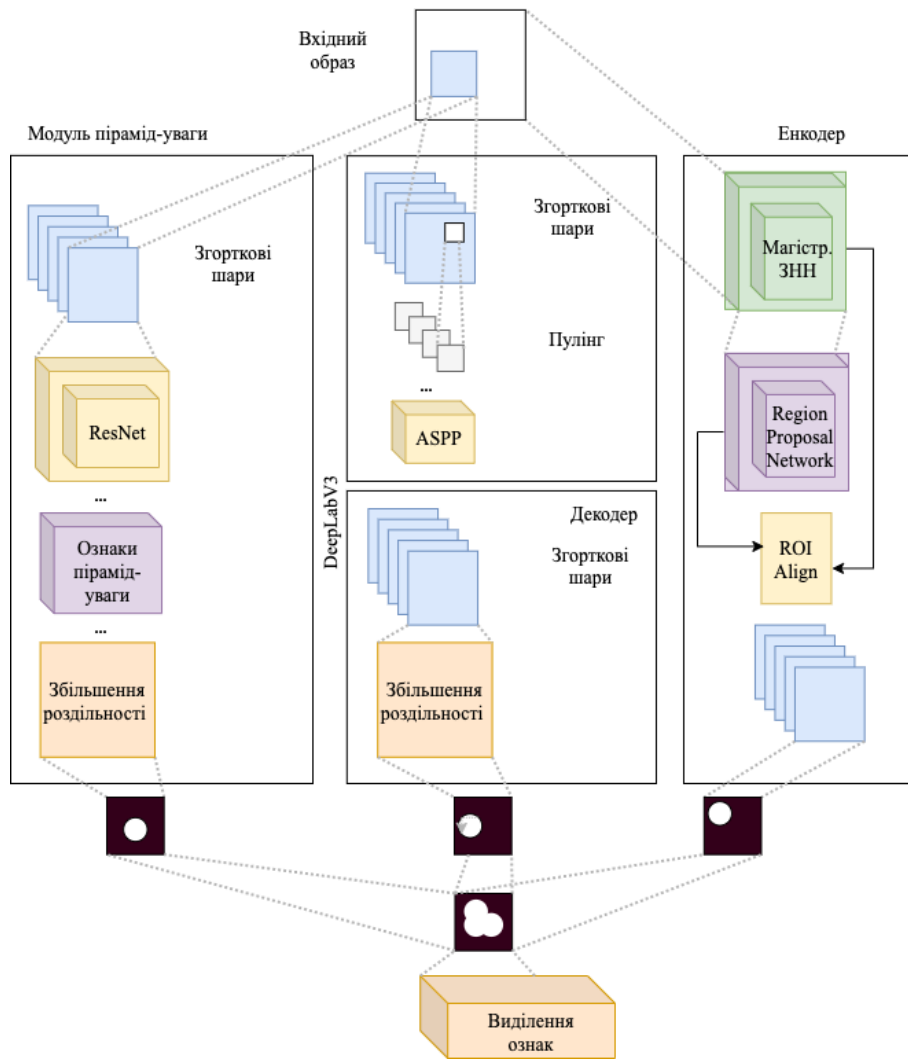


Рис. 11. Архітектура ансамблю сегментації

Оцінимо ефективність алгоритмів за допомогою індексу подібності Жаккара (JSI), чутливості, специфіки, точності, коефіцієнта подібності гральної кістки (Dice) і коефіцієнту кореляції Метью (MCC) [174]:

$$\text{Чутливість} = \frac{TP}{TP + FN}, \quad (3)$$

$$\text{Специфіка} = \frac{TN}{FP + TN}, \quad (4)$$

$$\text{Точність} = \frac{TP + TN}{TP + FP + TN + FN}, \quad (5)$$

$$\text{Коефіцієнт подібності Жаккара} = \frac{TP}{(TP + FP + FN)}, \quad (6)$$

$$\text{Dice} = \frac{2 + \text{TP}}{(2 + \text{TP} + \text{FP} + \text{FN})}, \quad (7)$$

$$\text{MCC} = \frac{\text{TP} + \text{TN} - \text{FP} + \text{FN}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}}. \quad (8)$$

Чутливість визначається в рівнянні (3), де TP – справжні позитивні результати і FN – хибнонегативні. Висока чутливість (близька до 1,0) вказує на хорошу продуктивність під час сегментації, всі поразки були успішно сегментовані. З іншого боку, Конкретність (як у рівнянні (4)) показує частку істинно негативних (TN) серед неушкоджень. Висока специфічність вказує на здатність методу не сегментувати осередок ураження. Точність методів сегментації визначає відсоток джерел радіовипромінювання, які були правильно класифіковані, як у рівнянні (5). JSI і Dice – це міра того, наскільки схожі передбачення і достовірність, шляхом вимірювання кількості виявлених TP і штрафів за FP, знайдені методом, як в рівнянні (6) і (7) відповідно. MCC має діапазон від –1 (абсолютно неправильний двійковий класифікатор) до 1 (повністю правильний двійковий класифікатор). Використовується для оцінювання ефективності алгоритмів сегментації, заснованих на бінарній класифікації (ураження у порівнянні з неураженням), як у рівнянні (8).

На основі ансамблевої сегментації виділяються ознаки, які є вхідною інформацією для розв'язання задачі класифікації. Вхідні дані можуть розбиватись на певні групи для обробки різними НМ чи подаватись на всі мережі одночасно. Основна складність об'єднання мереж в ансамблі – це навчання всіх складових для розв'язання задачі. З метою підвищення ефективності навчання НМ навчаються окремо (якщо це можливо), а потім об'єднуються в єдину структуру. Однак, у випадку, якщо алгоритми налаштування вибраних топологій відносяться до різних класів навчання, вимагається синхронне навчання всіх ШНМ, що входять в ансамбль, а отже необхідно розробити єдиний алгоритм налаштування всіх ЗНМ ансамблю.

На основі вищенаведеного аналізу в роботі використовується паралельна структура ансамблю (бегінг) з шаром об'єднання. Необхідною й достатньою

умовою для побудови ансамблю ЗНМ, який має більшу точність розв'язання задачі класифікації ніж кожна окрема ЗНМ є включення до його складу ЗНМ, які відповідають критеріям точності та різноманітності. Оскільки різноманітність ансамблю зменшується зі збільшенням точності членів ансамблю, то розв'язання задачі створення ефективного ансамблю зводиться до пошуку компромісу.

Однією із важливих вимог для побудови ансамблів є забезпечення різноманітності між моделями класифікації, що може бути виконано шляхом оперування процесом моделювання або вхідними даними. Оперування процесом моделювання полягає в побудові моделей класифікації з використанням або різних алгоритмів навчання, або одного алгоритму навчання, але з різними параметрами. Більш популярні підходи до оперування вхідними даними – це оперування навчальними вибірками і вхідними ознаками. Алгоритми, що використовуються для управління навчальними вибірками, можуть генерувати кілька гіпотез, в яких алгоритм навчання застосовується до різних підмножин навчальних вибірок. Алгоритм збору даних складається з випадкового поділу вихідного набору даних на кілька навчальних підмножин однакового розміру на основі вибірки з заміною, яка може застосовуватися до будь-якого алгоритму навчання. Аналогічним чином алгоритм підвищення об'єднує результати класифікації з використанням одного і того ж алгоритму навчання; однак цей тип алгоритму є ітеративним, коли кожна нова модель заснована на результаті раніше побудованої моделі. Алгоритми оперування вхідними характеристиками створюють ансамблі на основі різних підмножин ознак, доступних алгоритму навчання. Цим процесом може бути, наприклад, випадкове поділ набору ознак на підмножини [170, 172] або використання алгоритму вибору ознак в поєднанні з оперуванням навчальними вибірками. Одна з проблем, що впливають на продуктивність класифікаторів, полягає в тому, як визначити, які функції мають значення для опису образів (зразків), які цікавлять. Отже, алгоритми вибору ознак [174] можуть використовуватися у разі побудови ансамблю для досягнення високої ефективності. Основним внеском цього дослідження є моделі вибору підмножини ознак, засновані на конкретних групах ознак, і алгоритми вибору

ознак для оперування вхідними ознаками. Відбір ознак для кожного класифікатора в подальшому ансамблі відбувається різними методами: коефіцієнт кореляції Пірсона; коефіцієнт підсилення на основі вибору ознак; Relief-F; аналіз основних компонентів; вибір функцій на основі кореляції.

Використання різнорідних класифікаторів (дерево прийняття рішень, багатошаровий персептрон, нейро-нечіткі мережі, Баєсівський класифікатор) на різних вибірках і ознаках для отримання остаточного класу оброблюваних даних. Агрегація відбувається одним із методів: зважування; мажоритарне голосування; мета-модель.

Точність розв'язків окремих згорткових класифікаторів наведено у табл. 1–5.

Таблиця 1

Точність розв'язків окремих згорткових класифікаторів без пропонованої архітектури сегментації

Класифікатор	Тренувальна точність, %	Тестова точність, %
Residual-Squeeze Module	92.11	90.39
Squeeze and Excitation-Attention Module	97.98	91.93
Inception-Attention Block	97.11	91.74
Squeeze and Excitation-Residual Module	90.66	90.02

Таблиця 2

Точність розв'язку повного згорткового мажоритарного ансамблю

Тренувальна точність	Тестова точність
97.65%	93.36%

Таблиця 3

Точність розв'язку окремих класифікаторів з запропонованою архітектурою сегментації

Класифікатор	Тренувальна точність, %	Тестова точність, %
Multilayer perceptron	94.23	91.82
Decision Tree	95.91	92.56
NefClassM	92.67	90.41

Таблиця 4

Точність повного ансамблю класифікаторів з ансамблем сегментації

Тренувальна точність, %	Тестова точність, %
96.6	97.14

Таблиця 5

Точність інших різноманітних архітектур ансамблів

Ансамбль	Тренувальна точність, %	Тестова точність, %
Random Forest	97.32	87.15
AdaBoost	54.96	54.25
Bagging Decision Tree	94.41	83.85
CNN bagging	94.96	91.39

Використання наведеної структури дозволить підвищити якість розв'язання задачі класифікації, особливо в складних умовах при наявності великої кількості ознак різної природи.

Висновки

1. Обґрунтовано вибір унікальних блоків (модулів) на основі аналізу і функціональних можливостей з метою їх використання в ГЗНМ. Це дає можливість підвищити якість виявлення ознак при зборі, обробці та узагальненні інформації.

2. Запропоновано методичний підхід з інтелектуальної оцінки та прогнозування радіоелектронної обстановки, який полягає в розв'язанні послідовності задач структурно-параметричного синтезу базової ЗНМ, гібридної ЗНМ з визначенням типів та послідовності унікальних блоків, які вводяться до складу гібридної ЗНМ, ансамблю, що складається з окремих гібридних ЗНМ, що дозволить підвищити точність розв'язання задач класифікації та скоротити обчислювальні витрати.

Методичний підхід базується на поступовому включенні до складу базової ЗНМ унікальних блоків або їх послідовності, що дозволить підвищити точність розв'язання задач класифікації та скоротити обчислювальні витрати.

Запропоновано підходи побудови гібридних згорткових нейронних мереж ансамлевої структури як для розв'язання задач сегментації, так і класифікації, що дозволить підвищити точність аналізу радіоелектронної обстановки (табл. 1–5).