

РОЗПІЗНАВАННЯ ХАРАКТЕРИСТИК ВКЛЮЧЕННЯ МЕТОДАМИ НЕЙРОННИХ МЕРЕЖ ПРИ МОДЕЛЮВАННІ СТАЦІОНАРНИХ ПРОЦЕСІВ

Роман Гавдутьський¹, Любов Журавчак²

^{1,2} Національний університет “Львівська політехніка”,
кафедра програмного забезпечення, Львів, Україна

¹ E-mail: roman.i.havdulskyi@lpnu.ua, ORCID 0009-0004-6737-6018

² E-mail: liubov.m.zhuravchak@lpnu.ua, ORCID 0000-0002-1444-5882

© Гавдутьський Р., Журавчак Л., 2025

Анотація. Виявлення та ідентифікація включень при моделюванні стаціонарних процесів є важливою задачею в багатьох технічних галузях, зокрема, матеріалознавстві, електроніці та неруйнівному контролю. Наявність включень може впливати на механічні, теплові та електричні властивості матеріалу, а тому точність визначення їхніх геометричних та фізичних характеристик є важливою. Використання сучасних чисельних методів та методів глибинного навчання відкриває нові можливості для покращення ефективності та точності прогнозування результатів.

У статті розглянуто застосування непрямого методу приграничних елементів для розв’язування прямих задач теорії потенціалу та різних типів нейронних мереж для розв’язування обернених, тобто для розпізнавання вказаних характеристик включення (їхнього розміщення, розмірів та коефіцієнта провідності). Проведено низку обчислювальних експериментів для знаходження розв’язків прямої задачі, які дозволили отримати набори даних для навчання та тестування моделей з використанням згорткових, рекурентних та глибинних нейронних мереж.

Для підвищення ефективності аналізу даних протестовано використання методу нормалізації даних. Запропоновано новий підхід виділення ознак для прогнозування фізичної характеристики включення (коефіцієнта провідності), що суттєво покращило точність її обчислення. Після проведеного тестування вибрано низку архітектур нейронних мереж для знаходження геометричних і фізичних характеристик та побудовано ансамблі нейронних мереж, які об’єднують найкращі аспекти кожної з них. Результати тестування ансамблевої моделі показали вищу точність визначення коефіцієнта провідності та геометричних характеристик включення порівняно з окремими моделями.

Отримані результати можуть бути корисними для вдосконалення методів неруйнівного контролю та виявлення дефектів чи включень. Подальші дослідження будуть спрямовані на розроблення підходу до визначення характеристик декількох включень та створення варіативніших наборів даних.

Ключові слова – нейронні мережі, розпізнавання включень, стаціонарні процеси, згорткові нейронні мережі, машинне навчання, неруйнівний контроль, теорія потенціалу, обчислювальний експеримент.

Постановка проблеми

Виявлення та ідентифікація включень у фізичних об’єктах є важливим завданням у багатьох галузях техніки, оскільки їхня наявність може суттєво змінювати фізичні властивості об’єкта або середовища, погіршуючи його надійність. Можливість виявлення цих дефектів так само важлива, як і подальше попередження виникнення нових. З математичної точки зору такі задачі належать до

обернених, зокрема, під час моделювання стаціонарних процесів, до задач теорії потенціалу у неоднорідних об'єктах. У цій статті нами запропоновано підхід до попереднього опрацювання та видобування ознак для покращення точності визначення фізичної характеристики.

Класичні методи розв'язування оберненої задачі не завжди є ефективними. Використання нейронних мереж може покращити точність або ж зменшити необхідний час для отримання результату. Враховуючи важливість ефективного вирішення завдання виявлення та ідентифікації включення, гостро постає проблема вибору оптимального методу попереднього опрацювання даних та побудови архітектури нейронної мережі або ансамблів нейронних мереж.

Враховуючи вищезгадану проблематику, основним завданням дослідження є побудова підходу до розпізнавання характеристик включення з урахуванням ефективності нейронних мереж. Крім того, при розробленні алгоритму розв'язування оберненої задачі теорії потенціалу важливо дослідити методи попередньої обробки даних з метою виділення ознак для подальшого навчання.

Аналіз останніх досліджень та публікацій

Наявні дослідження фокусуються на виявленні або лише геометричних характеристик (глибина розміщення, радіус (Travassos, Vieira, Ida, Vollaie, & Nicolas, 2008) та інші (Al-Athel, Alhasan, Alomari, & Arif, 2022)), або ширшого спектру параметрів, включаючи фізичні характеристики (Khodadad & Ardakani, 2008). Деякі дослідження акцентують свою увагу на знаходженні включень складної форми (Yin, Yang, & Meng, 2023) або мікротріщин у мікрочипах (Weitz, Clausner, & Zschech, 2024).

Сучасні чисельні методи, зокрема, скінченних (МСЕ) та граничних елементів (МГЕ) широко використовують для моделювання та аналізу різноманітних стаціонарних та динамічних процесів у неоднорідних об'єктах (Jaddi, Coulombier, Raskin, & Pardoen, 2019) та прогнозування терміну їхньої служби, враховуючи появу дефектів (Alshoaibi & Fageehi, 2024). Широке застосування мають спеціалізовані програмні забезпечення STRAND7 (Han, Gan, & Setiawan, 2014) та COMSOL (Kurnyta-Mazurek, Wrabel, & Kurnyta, 2025), розроблені на основі МСЕ.

Виявлення геометричних та фізичних характеристик відбувається завдяки використанню як методів нейронних мереж [1, 2], так і класичних методів (генетичний алгоритм (ГА) та метод спряженого градієнта (МГГ) [2]). Для отримання даних використовують сканування об'єктів багатьма різноманітними методами (Müller, Netzelmann, & Valeske, 2022; Xisto L. Travassos, Avila, & Ida, 2021).

Серед методів нейронних мереж, які активно використовують для виявлення тріщин та включень, варто виділити CNN (згорткові нейронні мережі, ЗНМ) (Yin et al., 2023), DNN (глибинні нейронні мережі) (Yang et al., 2019) та їх модифікації. Інколи використовують мережі RNN (рекурентні нейронні мережі) та LSTM (довга короткочасна пам'ять) (Yuan & Lee, 2020), LVQ (навчальне векторне квантування) (Lema, Pedrayes, Usamentiaga, Venegas, & Garcia, 2022; L. Zhao, He, Wang, Su, & Lu, 2020) та YOLO (You only look once) v5 (Hussain, 2024). Використання останньої мережі передбачає передавальне навчання (Hosna et al., 2022), що дозволяє значно економити час шляхом донавчання моделі під подібну, але раніше не відому їй, задачу. Вибір мережі першочергово залежить від характеру даних, зокрема, RNN та LSTM використовують для опрацювання послідовностей.

Не менш важливим аспектом є методи попереднього опрацювання даних. Метод нормалізації даних (Jo, 2019) вважають класичним і часто необхідним для коректного опрацювання даних нейронними мережами.

У статті (Travassos et al., 2008) описано метод для виявлення та визначення характеристик включень у бетонних конструкціях за даними, отриманими з допомогою георадару. Як метод попередньої обробки, для зменшення обсягу даних, використовують метод головних компонент (МГК). У першу чергу це необхідно для прискорення навчання нейронної мережі. Як архітектуру використовують Parallel Layer Perceptron (Tavares, Saldanha, & Vieira, 2015). На вхід мережі надходять оброблені дані про сканування, на виході отримують радіус та глибину включення. Протестовано роботу методу для включення радіусом від 0.02 м до 0.1 м, розташованого на глибині від 0.05 м до

0.25 м. Дані для проведення дослідження отримано в результаті обчислювального експерименту з використанням методу Finite-Difference Time-Domain. Результатом дослідження стала модель, похибка якої була меншою за 1 % для параметрів глибини та радіусу включення.

У статті (Al-Athel et al., 2022) досліджено метод виявлення дефектів у композитних матеріалах за допомогою гібридного інфрачервоного сканування. Використано нейронну мережу з 10-ма прихованими шарами. Результатом є дані про глибину розміщення включення та його розміри. Залежно від форми досліджуваного об'єкта одержано такі похибки: при виявленні глибини дефекту від 9.5 % до 18.7 %, для діаметру – 4.4 %, визначення ширини й довжини – 25 %. Підкреслено ефективність використання методів нейронних мереж щодо економії ресурсів.

У статті (Khodadad & Ardakani, 2008) описано ідентифікацію включень з використанням МГЕ, ГА та МСГ, досліджено можливість ідентифікації характеристик внутрішніх структур у неоднорідному тілі за вимірними зміщеннями його поверхні. Розв'язано задачу визначення місцезнаходження, форми та фізичних характеристик. Запропонований підхід базується на розв'язуванні прямої задачі еластичності МГЕ. Для оберненої задачі цей метод допомагає оцінити невідомі характеристики включення. Алгоритм пошуку характеристик засновано на такому оптимізаційному підході:

- ГА оцінює геометрію включення та параметри матеріалу завдяки операторам кросінговеру та мутації для забезпечення різноманітності розв'язку та конвергенції;
- МСГ уточнює початкове припущення ГА;
- точність оцінки визначають як квадрат різниці між вимірними та обчисленими значеннями; для стабілізації оптимізації використовують регуляризацию.

Запропонований метод під час тестування продемонстрував похибку біля 10 % для задачі виявлення матеріалу та геометричних характеристик. Він залишався надійним і за наявності шуму інтенсивністю до 3 %, а при її збільшенні значно втрачав свою точність. Важливою відмінністю статті варто виділити гібридний підхід для виявлення характеристик.

У статті (Yin et al., 2023) описано розв'язування оберненої задачі для процесу розсіювання з метою виявлення тріщин у неоднорідних включеннях з використанням методу ЗНМ. Скановано невелику область за умов відсутності та наявності шуму та враховано неоднорідність матеріалів. Запропонований розв'язок використовує модель ЗНМ, яка видобуває ознаки, використовуючи згорткові та пулінг шари, для обробки послідовності даних. На основі видобутих даних ця модель будує передбачення про наявність тріщини. Розв'язок використовує оптимізатор Adam для мінімізації помилки реконструкції тріщини. Дані для тренування видобуто на основі хвильового сканування. Архітектура згорткової мережі включала в себе два згорткові шари, один шар пулінгу та два повністю з'єднані шари. Проведений експеримент продемонстрував позитивні результати на достатньо великих наборах даних (10000 елементів) та ефективність методу за умови наявності шуму інтенсивністю 10-20 %. Відзначено, що при звужуванні зони сканування точність зменшувалася, як і при наявності декількох тріщин.

У статті (Travassos et al., 2021) досліджено інтеграцію машинного навчання (МН) та нейронних мереж для неруйнівного контролю з використанням георадару як джерела даних. Основними цілями є виявлення та ідентифікація закопаних об'єктів або геологічних аномалій, оцінка фізичних та геометричних характеристик, зменшення часу, необхідного для інтерпретації даних. Підхід, розроблений у статті, базується на використанні нейронних мереж для сегментації та виявлення регіону, який буде піддаватися подальшому дослідженню. Алгоритми МН використовують для розпізнавання патернів та класифікації об'єктів у межах виділених регіонів. Результатом статті є запропонований гібридний алгоритм, який демонструє можливість використання нейронних мереж та методів МН для задач інтерпретації даних георадару з метою виявлення та ідентифікації включень.

У статті (Müller et al., 2022) досліджено метод виявлення дефектів у об'єктах за допомогою аналізу термографічних послідовностей. Для виявлення дефектів використано модель U-Net та комбінацію згорткового шару з LSTM. Результати виявлення дефектів оцінено кількісно за допомогою метрики перетину над об'єднанням (Intersection over Union, IoU). Результатом дослідження є те, що запропоновані підходи демонструють підвищення точності порівняно з традиційними методами неруйнівного контролю з використанням інфрачервоної термографії.

У статті (Yang et al., 2019) досліджено використання глибинного навчання для визначення тріщин у матеріалах. Як вхідні дані використовують інформацію про розподіл тепла, отриманий методом активного інфрачервоного сканування. Вхідні дані поділено на три класи проникальні тріщини, непроникальні тріщини та пошкодження поверхні. Зібрано набір даних з 3000 елементів. Важливою особливістю дослідження є модифікація моделі нейронної мережі, відомої як Faster R-CNN (Ren, He, Girshick, & Sun, 2015), для детальнішого виявлення малих об'єктів, запропонована модель досягнула точності 95.54 %. Реалізовано багаторівневе злиття ознак, яке покращило ефективність розпізнавання малих тріщин та просторове відображення. Окрім цього, CNN продемонструвала здатність точно відрізнати проникальні та непроникальні тріщини.

У статті (Yuan & Lee, 2020) досліджено виявлення дефектів місць спаювання за допомогою MCE та нейронних мереж типу RNN і LSTM. MCE використано для проведення обчислювального експерименту й генерації набору даних. Згенеровано 81 пара даних, з яких 27 використано для послідовного навчання нейронних мереж. Ключовими результатами статті є такі: 27 із 81 пари даних достатньо для навчання, додавання додаткових прихованих шарів (від 2 до 3) покращило ефективність на один порядок, LSTM незначно переважає RNN, отримана модель достатньо добре узагальнює нові вхідні дані. Середня похибка для моделей RNN та LSTM становить 1.213×10^{-4} та 1.190×10^{-4} відповідно. Зазначено, що використання МН значно зменшує часові витрати порівняно з використанням лише MCE.

У статті (Zhao et al., 2020) досліджено використання засобів нейронних мереж для виявлення дефектів у місцях спаювання в дрібних електронних компонентах, для яких характерний монтаж методом перевернутого чипа. Дані для навчання отримано з використанням MCE, симулюючи їхній збір шляхом активного інфрачервоного сканування. Для фільтрації шуму використано просторову адаптивну фільтрацію. Тепліші ділянки, які відображають місця спаювання, виділяють із загального набору даних методом сегментації, а репрезентативні ознаки – з отриманих сегментів. Використовуючи нейронну мережу LVQ, проведено класифікацію отриманих даних. Результатом дослідження є доведення ефективності використання описаного методу для виявлення дефектів у електронних пристроях високої щільності.

У статті (Lema et al., 2022) досліджено виявлення приповерхневих дефектів шляхом використання активного інфрачервоного сканування та методів глибинного навчання. Серед запропонованої новизни варто виділити зміщення фокусу з сегментації чи класифікації включень на їх виявлення. Вхідні дані представлено як триканальне кольорове зображення, сформоване з використанням РСТ (principal component thermography). Обраний формат представлення даних допомагає зменшити кількість шуму в них. Результатом дослідження є перевірка ефективності різноманітних підходів глибинного навчання. Найкращу ефективність показало використання моделі YOLO v5, яка продемонструвала ефективне та точне визначення приповерхневих дефектів.

У статті (Patra & Dutta, 2013) описано метод розпізнавання характеристик новоутворень з використанням нейронних мереж на основі даних дифузно-оптичної томографії, який використовує нейронні мережі для виявлення та локалізації включень. Як вхідні параметри мережа отримує інформацію про пікову амплітуду, пропускну здатність та кут детектора. Для перевірки ефективності запропоновано використати метрики: MSE, Object Centroid Error та Misclassification Ratio. Результатом дослідження стала модель, середньоквадратична точність якої близька до 0.006.

У статті (X. Zhao et al., 2023) досліджено використання методів МН та штучного інтелекту (ШІ) для виявлення дефектів у електронних компонентах. За вхідні дані беруть інформацію про розподіл температури, отриманий шляхом активного інфрачервоного сканування. Для виявлення дефектів у сонячних панелях використовують методи ШІ та нейронних мереж, щоб знайти розв'язки задач видобування ознак, класифікації дефектів та обчислення ймовірності виникнення проблем під час експлуатації. Використано модель CNN для видобування ознак та класифікації типу дефектів та алгоритм CSDM для аналізу хаотичних термальних сигналів. ГА використано для оптимізації вибору ознак з метою підвищення точності виявлення дефектів. Для аналізу та виявлення аномалій в інших електронних компонентах запропоновано глибинне навчання. Результатами дослідження є підвищення ефективності локалізації дефектів, автоматичне визначення їхніх типів, допомога в ранньому виявленні дефектів.

Автори статті (Oswald-Tranta, Lopez De Uralde Olavera, Gorostegui-Colinas, & Westphal, 2023) наводять результати використання ЗНМ для виявлення тріщин методом сегментації за такими категоріями: гарячі місця, тріщини, оточення (фон). ЗНМ складається з п'ятьох згорткових шарів. Важливою особливістю дослідження є перевірка ефективності як і на штучно згенерованих даних, так і на невеликій кількості реальних.

Узагальнені результати аналізу актуальних досліджень наведено у табл. 1. Як можна помітити, більшу частину з них сфокусовано на знаходженні тріщин різного роду. Деякі дослідження зосереджені на пошуку дефектів у об'єктах, без уточнення його виду, інші акцентують увагу на виявленні самого включення без можливості обчислення його фізичних чи геометричних характеристик. Варто зазначити, що переважно для отримання даних використовують складніший, з точки зору обсягу пам'яті та методики обчислень, МСЕ, рідше МГЕ.

Таблиця 1

Порівняння особливостей підходів до виявлення чужорідних включень

Рік	Автори	Методи	Підхід	Спосіб отримання набору даних
1	2	3	4	5
2009	Xisto Lucas Travassos, Douglas A. G. Vieira, Nathan Ida, Christian Vollaire, Alain Nicolas	Parallel Layer Perceptron (Tavares et al., 2015)	Визначення радіусу та глибини залягання включення	GPR (Ground Penetrating Radar) сканування
2022	Khaled S. Al-Athel, Motaz M. Alhasan, Ahmed S. Alomari, Abul Fazal M. Arif	Нейронна мережа з 10-ма прихованими шарами	Виявлення глибини та геометричних характеристик включення	Обчислювальний експеримент за МСЕ
2008	Mahmud Khodadad, Mohsen Dashti Ardakani	ГА, МСТ	Прогнозування форми та фізичних характеристик включення	Обчислювальний експеримент за МГЕ
2023	Weishi Yin, Ziyuan Yang, Pinchao Meng	ЗНМ	Прогнозування параметрів форми асиметричної тріщини	Обчислювальний експеримент за МСЕ
2020	Xisto L. Travassos, Sérgio L. Avila, Nathan Ida	Поєднання нейронних мереж та алгоритмів МН	Виявлення включень, визначення їх фізичних та геометричних характеристик	GPR сканування
2020	David Müller, Udo Netzelmann, Bernd Valeske	Поєднання CNN та LSTM	Виявлення дефектів типу тріщин у об'єктах шляхом термального сканування	Обчислювальний експеримент, COMSOL
2019	Jun Yang, Wei Wang, Guang Lin, Qing Li, Yeqing sun, Yixuan Sun, Lei Su,	Faster R-CNN	Виявлення тріщин в об'єктах шляхом активного термального сканування	Експеримент
2020	Cadmus C.A. Yuan, Chang-chi Lee	Нейронні мережі типу RNN та LSTM	Виявлення дефектів у місцях пайки	Обчислювальний експеримент за МСЕ
2019	Libo Zhao, Zhenzhi He, Zengxiang Wang, Xiangning Lu	LVQ	Виявлення дефектів у місцях спаювання, фокусування на виявленні незапаяних місць	Обчислювальний експеримент, COMSOL

Продовження табл. 1

1	2	3	4	5
2022	Darío G. Lema, Oscar D. Pedrayes, Rubén Usamentiaga, Pablo Venegas, Daniel F. García	YOLO v5	Виявлення приповерхневих тріщин в об'єктах шляхом активного термального сканування	Експеримент
2013	Rusha Patra, Pranab K. Dutta	Backpropagation ANN (Patra & Dutta, 2013)	Виявлення геометричних характеристик включення	Обчислювальний експеримент
2023	Xinfeng Zhao, Yangjing Zhao, Shunchang Hu, Hongyan Wang, Yuyan Zhang, and Wuyi Ming	CNN, CSDM, DNN	Виявлення дефектів з використанням активного інфрачервоного сканування	Експеримент
2023	B. Oswald-Tranta, P. Lopez de Uralde Olavera, E. Gorostegui-Colinas, Ph. Westphal	CNN з 5-ма шарами	Виявлення тріщин шляхом сегментації	Обчислювальний експеримент для основного набору, частина реальних даних для перевірки ефективності

Таким чином, розробка підходів, що застосовують нейронні мережі для розпізнавання геометричних і фізичних характеристик включень при моделюванні стаціонарних процесів у неоднорідних об'єктах, є актуальним напрямом досліджень.

Мета дослідження – аналіз методів та засобів для розпізнавання фізичних та геометричних характеристик чужорідного включення.

Об'єкт дослідження – розпізнавання фізичних та геометричних характеристик чужорідного включення.

Предмет дослідження – методи та засоби нейронних мереж для розпізнавання характеристик включення.

Формулювання мети статті

Основною метою статті є перевірка ефективності використання нейронних мереж для розпізнавання характеристик одного включення при моделюванні стаціонарних процесів. Результати виконаного дослідження слугуватимуть теоретичним та практичним підґрунтям для подальшого удосконалення алгоритму виявлення декількох включень. З метою підготовки теоретичної бази перед початком практичної частини дослідження проведено аналіз наукових статей, які описують розв'язання оберненої задачі для різних фізичних процесів з використанням варіації методів ІІІ.

Однією із ключових цілей практичної частини є збір та формування достатнього набору даних, доповнення його варіативності та забезпечення репрезентативності. Для проведення обчислювального експерименту використано непрямий метод приграничних елементів (НМПГЕ) (Zhuravchak & Zabrodskaya, 2024). Не менш важливим аспектом є вибір методів попереднього опрацювання даних, зокрема класичних підходів (нормалізація даних, розрахунок градієнту, що дозволяє оцінити швидкість зміни величин у вхідних даних). У межах цього дослідження нейронні мережі спершу тренуватимуться на невеликому наборі даних для оцінки ефективності кожної з них. Опісля, оцінивши ефективність, буде внесено зміни у моделі нейронної мережі, проведено тренування й тестування на більшому наборі даних.

Виклад основного матеріалу

Розв'язування прямої задачі теорії потенціалу для кусково-однорідного тіла

Перед розробленням підходу до розв'язування оберненої задачі, необхідно отримати низку наборів даних, які відповідають розв'язкам прямої задачі з різними геометричними та фізичними характеристиками включення. Математичні моделі стаціонарних процесів, побудовані для методів, які базуються на використанні природних та штучних потенціальних полів (гравітаційних, магнітних, електричних, теплових, фільтраційних) для виявлення неоднорідності об'єкта, складаються з рівнянь Лапласа чи Пуассона:

$$\mathbf{P}_0^{(m)}(u^{(m)}(x)) = \mathbf{D}u^{(m)}(x) = -y^{(m)}(x), x \in W_m, W_m \subset \mathbf{R}^2, m=1, \dots, M, \quad (1)$$

доповнених заданими крайовими умовами:

$$u^{(m)}(x) = f_G^{(1)}(x), x \in \mathbb{T}W^{(1)}, \frac{\mathbb{T}u^{(m)}(x)}{\mathbb{T}n^{(m)}(x)} = f_G^{(2)}(x), x \in \mathbb{T}W^{(2)}, \quad (2)$$

та умовами ідеального контакту на межі поділу середовищ:

$$u^{(m)}(x) = u^{(s)}(x), l_m \frac{\mathbb{T}u^{(m)}(x)}{\mathbb{T}n^{(m)}(x)} = -l_s \frac{\mathbb{T}u^{(s)}(x)}{\mathbb{T}n^{(s)}(x)}, x \in \mathbb{T}W^{ms}. \quad (3)$$

Тут W_m – однорідна область з постійною фізичною характеристикою l_m (провідністю); $y^{(m)}(x)$ – інтенсивність заданих внутрішніх джерел в W_m ; $\mathbf{n}^{(m)}(x) = (\mathbf{n}_1^{(m)}(x), \mathbf{n}_2^{(m)}(x))$ – вектор зовнішньої одиничної однозначно визначеної нормалі до межі $\mathbb{T}W_m$; $\mathbf{D} = \frac{\mathbb{T}^2}{\mathbb{T}x_1^2} + \frac{\mathbb{T}^2}{\mathbb{T}x_2^2}$ – оператор Лапласа, $x = (x_1, x_2)$,

$\mathbb{T} \in \mathbb{R}^2_{q=1} \mathbb{T}W^{(q)} = \mathbb{T}W$; $\mathbb{T}W^{(1)} \cap \mathbb{T}W^{(2)} = \mathbb{A}$, $W = \bigcup_{m=1}^M (W_m \cap \mathbb{T}W_m)$, $\mathbb{T}W^{ms} = \mathbb{T}W_m \cap \mathbb{T}W_s$ – межа поділу

середовищ W_m та W_s , $x = (x_1, x_2)$ – система декартових координат.

Для знаходження розв'язків крайової задачі (1)–(3) використано НМПГЕ, який продемонстрував вищу порівняно з МГЕ точність та менші вимоги до математичної кваліфікації програміста.

Зовнішні приграничні області дискретизуємо на приграничні елементи (ПЕ) G_v^m ($v=1, \dots, V_m$) та уведемо на них сімейства точок G_v^{mk} ($k=1, \dots, K_{mv}$). Для кожного G_v^m введемо невідомі функції $\varphi_v^{(m)}(x)$, які описують колективну поведінку фіктивних джерел, поміщених у точках G_v^{mk} .

Оскільки для оператора Лапласа $\mathbf{P}_0^{(m)}(u^{(m)}(x))$ існує й відомий фундаментальний розв'язок (ФР) $\mathbf{U}^{(m)}(x, x)$, то інтегральне представлення розв'язків рівнянь (1) та похідних від них за нормаллю мають вигляд:

$$\begin{aligned} u^{(m)}(x) &= \mathbf{F}^{(m)}(x, \mathbf{U}^{(m)}) + b^{(m)}(x, \mathbf{U}^{(m)}) + C_m, \\ \mathbb{T}u^{(m)}(x) / \mathbb{T}n^{(m)}(x) &= \mathbf{F}^{(m)}(x, \mathbf{Q}^{(m)}) + b^{(m)}(x, \mathbf{Q}^{(m)}), \end{aligned} \quad (4)$$

$$\text{де } \mathbf{F}^{(m)}(x, \mathbf{F}^{(m)}) = \sum_{v=1}^{V_m} \sum_{k=1}^{K_{mv}} \Phi^{(m)}(x, \xi^{v0k}) \varphi_v^{(m)}(\xi^{v0k}), \quad b^{(m)}(x, \mathbf{F}^{(m)}) = \oint_{W_m} \mathbf{Q} \mathbf{F}^{(m)}(x, x) y^{(m)}(x) dW_m(x)$$

$$\mathcal{U}^{(m)}(x, \xi) = -\frac{1}{2\pi\lambda_m} \ln |r / r_0|; \quad \mathcal{Q}^{(m)}(x, x) = \mathring{\mathbf{a}} \sum_{l=1}^2 \mathcal{Q}_l^{(m)}(x, x) \mathbf{n}_l^{(m)}(x); \quad \mathcal{Q}_l^{(m)}(x, x) = -\frac{y_l}{2\pi r^2};$$

$|x|$ - модуль x ; r_0 - константа, яка використовується для покращення обчислювального процесу,
 $x = (x_1, x_2) \in \mathbf{R}^2$; $r^2 = \mathring{\mathbf{a}} \sum_{i=1}^2 y_i^2$, $y_i = x_i - x_i$.

Спрямувавши в (4) x з середини області W до межі $\mathbb{W}$ для задоволення умов (2), (3) і додавши рівняння, що виражає рівність нулю сумарної дії всіх джерел, що діють в $\mathbf{R}_m^2$, одержимо граничні інтегральні рівняння (ГІР):

$$\begin{aligned} \mathbf{F}^{(m)}(x, \mathcal{U}^{(m)}) + C_m &= f_G^{(1)}(x) - b^{(m)}(x, \mathcal{U}^{(m)}), \quad x \in \mathbb{W}^{(1)}, \\ \mathbf{F}^{(m)}(x, \mathcal{Q}^{(m)}) &= f_G^{(2)}(x) - b^{(m)}(x, \mathcal{Q}^{(m)}), \quad x \in \mathbb{W}^{(2)}, \end{aligned} \quad (5)$$

$$\begin{aligned} \mathbf{F}^{(m)}(x, \mathcal{U}^{(m)}) + C_m - \mathbf{F}^{(s)}(x, \mathcal{U}^{(s)}) - C_s &= -b^{(m)}(x, \mathcal{U}^{(m)}) + b^{(s)}(x, \mathcal{U}^{(s)}), \quad x \in \mathbb{W}^{ms}, \\ \mathbf{I}_m \mathbf{F}^{(m)}(x, \mathcal{Q}^{(m)}) + \mathbf{I}_s \mathbf{F}^{(s)}(x, \mathcal{Q}^{(s)}) &= -\mathbf{I}_m b^{(m)}(x, \mathcal{Q}^{(m)}) - \mathbf{I}_s b^{(s)}(x, \mathcal{Q}^{(s)}), \end{aligned} \quad (6)$$

$$\mathbf{F}^{(m)}(x, 1) = -b^{(m)}(x, 1), \quad x \in \mathbf{R}_m^2. \quad (7)$$

ГІР (5)-(7) після апроксимації функцій $\varphi_v^{(m)}(x)$ невідомими постійними d_v^m запишемо у вигляді системи лінійних алгебраїчних рівнянь (СЛАР):

$$\begin{aligned} \mathring{\mathbf{a}} \sum_{v=1}^{V_m} \mathcal{A}_v^m(x^{mw}, \mathcal{U}^{(m)}) d_v^m + C_m &= f_G^{(1)}(x^{mw}) - b^m(x^{mw}, \mathcal{U}^{(m)}), \quad x^{mw} \in \mathbb{W}^{(1)}, \quad w=1, \dots, V_m^{(1)}, \\ \mathring{\mathbf{a}} \sum_{v=1}^{V_m} \mathcal{A}_v^m(x^{mw}, \mathcal{Q}^{(m)}) d_v^m &= f_G^{(2)}(x^{mw}) - b^m(x^{mw}, \mathcal{Q}^{(m)}), \quad x^{mw} \in \mathbb{W}^{(2)}, \quad w=V_m^{(1)}+1, \dots, V_m^{(2)}, \\ \mathring{\mathbf{a}} \sum_{v=1}^{V_m} \mathcal{A}_v^m(x^{mw}, \mathcal{U}^{(m)}) d_v^m + C_m - \mathring{\mathbf{a}} \sum_{v=1}^{V_s} \mathcal{A}_v^s(x^{mw}, \mathcal{U}^{(s)}) d_v^s - C_s &= \\ = b^s(x^{mw}, \mathcal{U}^{(s)}) - b^m(x^{mw}, \mathcal{U}^{(m)}), \quad x^{mw} \in \mathbb{W}^{ms}, \quad w=V_m^{(2)}+1, \dots, V_m, \\ \mathbf{I}_m \mathring{\mathbf{a}} \sum_{v=1}^{V_m} \mathcal{A}_v^m(x^{mw}, \mathcal{Q}^{(m)}) d_v^m + \mathbf{I}_s \mathring{\mathbf{a}} \sum_{v=1}^{V_s} \mathcal{A}_v^s(x^{mw}, \mathcal{Q}^{(s)}) d_v^s &= -\mathbf{I}_s b^s(x^{mw}, \mathcal{Q}^{(s)}) - \mathbf{I}_m b^m(x^{mw}, \mathcal{Q}^{(m)}), \\ \mathring{\mathbf{a}} \sum_{v=1}^{V_m} \mathcal{A}_v^m(x, 1) d_v^m &= -b^{(m)}(x, 1), \quad x \in \mathbf{R}_m^2, \end{aligned} \quad (8)$$

$$\text{де } \mathcal{A}_v^{(m)}(x, \mathbf{F}^{(m)}) = \sum_{k=1}^{K_{mv}} \Phi^{(m)}(x, \xi^{v0k}).$$

Знайшовши з (8) невідомі d_v^m та C_m , використаємо формули, аналогічні (4), для обчислення шуканих функцій $u^{(m)}(x)$ та похідних від неї за нормаллю як у внутрішніх точках кожної W_m , так і на межах поділу середовищ, оскільки тепер всі зони W_m розглядають як цілком незалежні області:

$$\begin{aligned} u^{(m)}(x^{mz}) &= \mathring{\mathbf{a}} \sum_{v=1}^{V_m} \mathcal{A}_v^m(x^{mz}, \mathcal{U}^{(m)}) d_v^m + b^m(x^{mz}, \mathcal{U}^{(m)}) + C_m, \quad x^{mz} \in Z_m, \\ \frac{\mathbb{W} u^{(m)}(x^{mz})}{\mathbb{W} \mathbf{n}^{(m)}(x)} &= \mathring{\mathbf{a}} \sum_{v=1}^{V_m} \mathcal{A}_v^m(x^{mz}, \mathcal{Q}^{(m)}) d_v^m + b^m(x^{mz}, \mathcal{Q}^{(m)}), \end{aligned}$$

де Z_m - множини точок спостереження $x^{mz} \in \mathbb{W} W_m \in W_m$.

Зазначимо, що використання сімейств точок у ПЕ дозволило, на відміну від МГЕ, уникнути попереднього аналітичного виділення особливості (головного значення) під час обчислення похідних за нормаллю від шуканих функцій. Сімейства точок значно спрощують алгоритм розв'язування задачі, оскільки замінюють інтегрування по ПЕ сумуванням добутків ФР на значення невідомої інтенсивності джерел. Тому їх можна рекомендувати при знаходженні початкових наближень як експрес-метод для розв'язування обернених задачах математичної фізики, коли більшу роль відіграє оптимізація часу, ніж точності.

Формування набору даних

Розглянуто об'єкт, складений з вміщувального середовища та одного включення (рис. 1, на якому декартові координати x_1 , x_2 позначено через X , Y), та розв'язано низку прямих двовимірних задач теплопровідності, використовуючи як приграничні елементи сімейства точок.

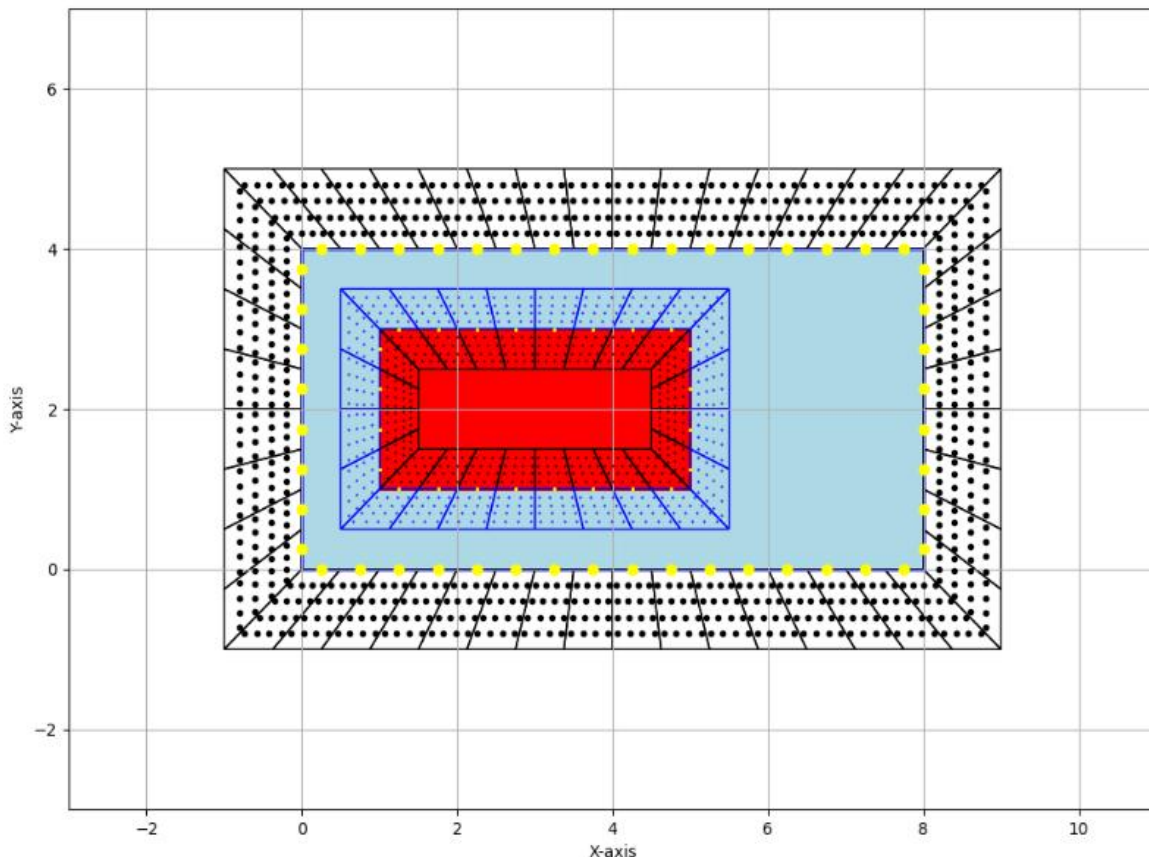


Рис. 1. Візуальне представлення об'єкта з включенням з урахуванням розбиття приграничної області на приграничні елементи

Верхня та нижня межі прямокутника теплоізолювані, на лівій та правій межах задано крайову умову першого роду: температуру 0 та 10 градусів Цельсія відповідно. Коефіцієнт теплопровідності вміщувального середовища рівний $1 \frac{\text{Вт}}{\text{м}\cdot\text{К}}$. Розподіл температури в об'єкті наведено на рис. 2. Крім інформації про температуру, кожен елемент набору містить дані про координати нижнього кута включення, його висоту, ширину та коефіцієнт теплопровідності. Дані збережені у вигляді JSON файлів, кожен з яких представляє інформацію про обчислювальний експеримент з певними варіативними змінними.

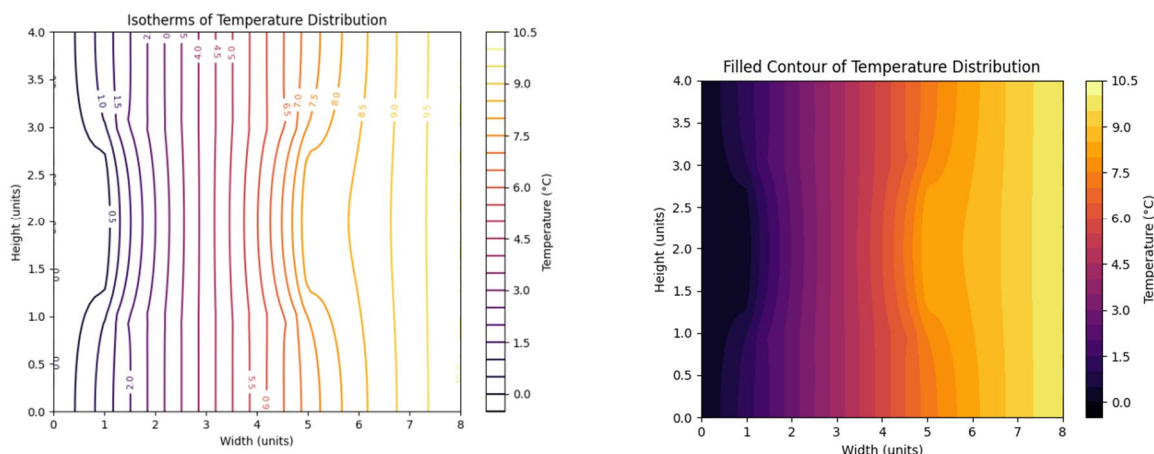


Рис. 2. Візуальне відображення розподілу ізотермами

Для швидкої апробації потенційних архітектур нейронної мережі та методів попередньої обробки сформовано менший набір даних з такими параметрами: кількість елементів, що описують об'єкт – 8546; висота об'єкта – 4; ширина об'єкта – 8; варіативність висоти включення – 1, 2; варіативність ширини включення – 1, 2, 3, 4, 5; варіативність координати X включення – 1, 2, 3, 4, 5, 6; варіативність координати Y включення – 1, 2; варіативність коефіцієнта теплопровідності – від 0.1 до 15 (не включаючи 1) з кроком 0.1; мінімальна відстань від включення до краю об'єкта – 1; кількість фіктивних точкових джерел у кожному ПЕ – 9 рядків, 9 стовпців, разом 81; кількість ПЕ на довгих сторонах прямокутника – по 16, на коротших – по 8, разом 48; кількість ПЕ на межі контакту – по 8 на довгих сторонах, по 4 на коротших, разом 24; висота ПЕ на межі об'єкта – 1; висота ПЕ на межі контакту – 0.5.

Методика розв'язування оберненої задачі

Після розв'язання прямої задачі теплопровідності та отримання за допомогою обчислювального експерименту наборів даних, запропоновано таку методику розв'язування оберненої задачі.

1. Описати архітектури нейронних мереж так, щоб кожна мережа апроксимувала лише одну характеристику, тобто або висоту чи ширину, або координату X чи Y, або коефіцієнт теплопровідності включення; так нейронна мережа розв'язуватиме задачу регресії.
2. Для оцінки ефективності використати такі метрики:
 - середня абсолютна похибка $MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$;
 - середня квадратична похибка $MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$.
3. Використати такі мережі для тренування на меншому наборі даних: ЗНМ з кількістю згорткових шарів від 3 до 9, глибинні нейронні мережі: UNet (Ronneberger, Fischer, & Brox, 2015), Transformer (Vaswani et al., 2017), ResNet (He, Zhang, Ren, & Sun, 2015).
4. З метою оптимізації часу тренування використати техніку ранньої зупинки (Montesinos López, Montesinos López, & Crossa, 2022).
5. Розбити набір на три піднабори: тренувальний, валідаційний та тестовий у пропорції 70:15:15.
6. Забезпечити відтворюваність шляхом використання статично заданих ключів для генератора псевдовипадкових чисел, який бере участь у поділі наборів та інших операціях.
7. Зберігати найкращу версію моделі, отриману протягом навчання, та інформацію про навчання: значення функції втрати для тренувального та валідаційного наборів наприкінці кожної епохи; значення метрик на тестовому наборі; результати апроксимації, отримані на тестовому наборі, для наочної оцінки точності.

8. Проаналізувавши ефективність моделей та підходів, протестованих на малому наборі даних, вибрати оптимальні, внести необхідні корективи та провести тренування й тестування на більшому наборі даних з такими параметрами: кількість елементів, що описують об'єкт – 114730; варіативність висоти включення – 1, 1.5, 2, 2.5; варіативність ширини включення – 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0; варіативність координати X включення – 0.6, 1.1, 1.6, 2.1, 2.6, 3.1, 3.6, 4.1, 4.6, 5.1, 5.6, 6.1; варіативність координати Y включення – 0.6, 1.1, 1.6, 2.1; мінімальна відстань від включення до краю об'єкта – 0.6; інші параметри такі ж, як в малому наборі.
9. Сформулювати висновки на основі отриманих даних для продовження дослідження з метою покращення ефективності моделі та розширення сценарію її застосування – пошуку декількох включень.

Класичний метод попереднього опрацювання

Для підвищення точності визначення характеристик попередньо опрацюємо дані. Спочатку нормалізуємо їхній розподіл. Щоб підтвердити позитивний вплив класичного методу попереднього опрацювання наведемо результати навчання деяких моделей без опрацювання та з нею для визначення координати X (табл. 2) та апроксимації коефіцієнта теплопровідності (табл. 3).

Таблиця 2

Результати попереднього навчання для координати X

Модель	Без попередньої обробки		З попередньою обробкою	
	MAE	MSE	MAE	MSE
DNN (2 шари з Dropout)	0.86	1.19	0.14	0.06
DNN (3 шари)	0.23	0.20	0.009	0.0015

Таблиця 3

Результати попереднього навчання для коефіцієнта теплопровідності

Модель	Без попередньої обробки		З попередньою обробкою	
	MAE	MSE	MAE	MSE
DNN (3 шари)	2.11	7.54	0.67	1.17
DNN (5 шарів)	2.35	9.06	0.81	1.82

Результати навчання на меншому наборі даних

Похибки моделей на малому наборі даних із використанням класичного методу попереднього опрацювання, описаного у попередньому підрозділі, наведено у табл. 4 та 5.

Таблиця 4

Метрики похибок визначення геометричних характеристик для різних моделей (менший набір даних)

Модель	X		Y		Ширина		Висота	
	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE
1	2	3	4	5	6	7	8	9
CNN (3 шари)	0.0077	0.0005	0.0035	9.84e-5	0.0116	0.0029	0.0019	0.0004
CNN (5 шарів)	0.0056	0.0001	0.0006	5.51e-6	0.0132	0.0030	0.0003	6.21e-7
CNN (7 шарів)	0.0030	0.0001	0.0001	2.43e-7	0.0092	0.0007	0.0003	3.15e-7

Продовження табл. 4

1	2	3	4	5	6	7	8	9
DNN (2 шари з Dropout)	0.1416	0.0639	0.0328	0.0045	0.1017	0.0370	0.0220	0.0016
DNN (3 шари)	0.0093	0.0015	0.0024	0.0004	0.0471	0.0222	0.0027	0.0004
DNN (5 шарів)	0.0068	0.0011	0.0015	0.0004	0.0230	0.0153	0.0005	0.0009
ResNet	0.0693	0.0224	0.0362	0.0057	0.0704	0.0290	0.0135	0.0022
Transformer	0.0105	0.0098	0.0022	3.17e-5	0.0141	0.0032	0.0072	0.0008
UNet	0.0167	0.0018	0.0047	0.0001	0.0251	0.0085	0.0062	0.0002

Таблиця 5

**Метрики похибок визначення коефіцієнта теплопровідності
для різних моделей(менший набір даних)**

Модель	MAE	MSE
DNN (3 шари)	0.6713	1.1670
DNN (5 шарів)	0.8102	1.8165
Transformer	1.2083	3.1892
CNN1D (3 шари)	0.6579	1.0213

Аналізуючи отримані результати метрик похибок, бачимо, що переважно усі моделі достатньо вдало узагальнюють дані й передбачають такі характеристики: значення координат, висоту та ширину. Натомість точність визначення коефіцієнта теплопровідності, попри використання класичного методу попереднього опрацювання, все ще залишається достатньо низькою. Зауважимо, що використання складніших моделей не дало значного покращення результату. Тому для покращення точності варто розглянути додаткові підходи до попереднього опрацювання даних.

Проаналізувавши метрики похибок, запропоновано таке.

1. Не використовувати складніші моделі (Transformer, UNet, ResNet) для визначення геометричних характеристик (висота, ширина, координати нижнього лівого кута включення) через більшу витрату часу на подальше навчання й збільшений ризик перенавчання (Guan & Burton, 2022).
2. Для визначення геометричних характеристик включення дослідити ефективність використання моделі CNN з 7-ма та 9-ма згортковими шарами, а як альтернативну модель – DNN з трьома шарами.
3. Покращити класичний метод попереднього опрацювання даних для визначення коефіцієнта теплопровідності.
4. Для визначення коефіцієнта теплопровідності використати такі моделі нейронних мереж: DNN з 3-ма та 9-ма шарами, CNN для одновимірних задач.
5. Реалізувати й вдосконалити альтернативну модель нейронної мережі ThermalNet для визначення коефіцієнта теплопровідності.
6. Побудувати ансамбль нейронних мереж (на основі згаданих вище) для підвищення точності й виділення сильних сторін кожної з моделей. Використати метод зважених коефіцієнтів, визначивши їх для кожної з моделей шляхом тренування додаткового шару, відповідно до “уміння” моделі узагальнювати дані.

Запропонований підхід до попереднього опрацювання

Оскільки класичний метод попереднього опрацювання даних для визначення коефіцієнта теплопровідності не був достатньо ефективним, запропоновано підхід на основі градієнта температури на межі контакту.

1. Знайти горизонтальний відрізок, який проходить вздовж усього об'єкта чітко посередині включення.
2. Визначити градієнт для чотирьох точок перед межею контакту, на межі контакту й для чотирьох наступних за межею контакту.
3. Це зменшить розмірність даних до 1D та дозволить чіткіше виділити більшу кількість інформативних ознак.
4. Візуальне представлення підходу наведено на рис. 3.

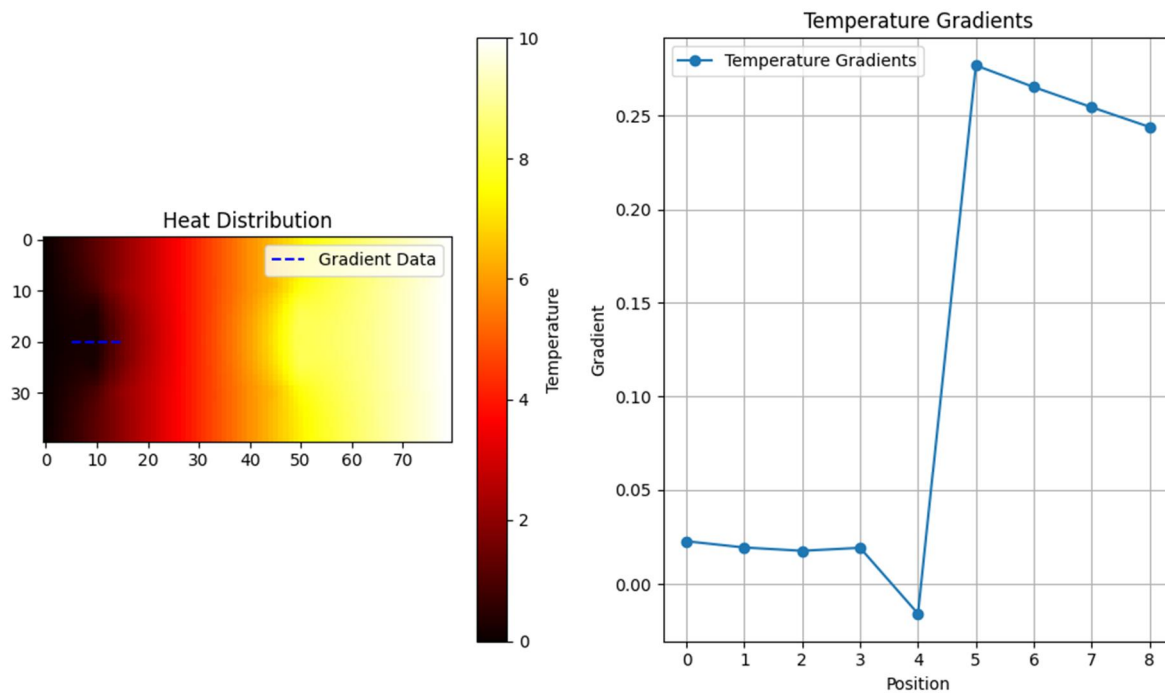


Рис. 3. Візуалізація запропонованого підходу до попереднього опрацювання на основі градієнта на межі контакту

Проектування альтернативної моделі ThermalNet для визначення коефіцієнта теплопровідності

Першочерговою мотивацією для реалізації альтернативної моделі є такі фактори:

- фокусування на ознаках, наявних в попередньо опрацьованих даних, для їхнього ефективного узагальнення;
- спрощення моделі, зменшення кількості внутрішніх параметрів – зменшення ймовірності перенавчання;
- підвищення ефективності щодо точності й часу тренування та виконання прогнозів.

Характерною ознакою моделі є зменшення кількості входів й виходів на лінійних шарах мережі. Опишемо її архітектуру:

- дані (градієнт на межі контакту) діляться навпіл: одна частина належить об'єкту та межі контакту, інша – включенню. Ці два масиви опрацьовуємо паралельно, видобуваючи ознаки, двома лінійними шарами, сполученими функцією активації ReLU (Dubey & Jain, 2019);

- ознаки об'єднуються і опрацьовуються трьома лінійними шарами, з'єднаними попарно ReLU шарами;
- на виході модель надає передбачення коефіцієнта теплопровідності.

Для пошуку оптимальної архітектури спроектуюмо другу версію моделі, яка має такі зміни:

- береться до уваги градієнт, який належить об'єкту та включенню, межа контакту ігнорується з метою отримання симетрії на шарах видобування ознак;
- збільшено кількість виходів на шарах видобування та опрацювання ознак удвічі та учетверо відповідно.

Третя версія архітектури має такі відмінності:

- функція активації замінена на LeakyReLU (Dubey & Jain, 2019);
- збільшено кількість виходів на шарах опрацювання ознак удвоє.

Результати навчання на більшому наборі даних

Метрики похибок моделей, обраних для тестування на більшому наборі даних, наведено в табл. 6. Останньою моделлю в ній є ансамблева модель, створена з використанням вагових коефіцієнтів, отриманих в процесі її навчання. Тому ансамблева модель намагається максимально використовувати сильні сторони кожної з моделей для досягнення вищої точності, що підтверджено діаграмами (рис. 4).

Таблиця 6

**Метрики похибок визначення геометричних характеристик
для різних моделей (більший набір даних)**

Модель	X		Y		Ширина		Висота	
	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE
CNN (7 шарів)	0.0033	6.09e-5	0.0002	2.24e-6	0.0053	0.0003	0.0002	1.44e-5
CNN (9 шарів)	0.0067	0.0002	6.29e-5	1.374e-7	0.0077	0.0002	0.0003	3.82e-5
DNN (3 шари)	0.0150	0.0001	0.0087	0.0006	0.0173	0.0001	0.0048	0.0006
Ансамбль мереж	0.0020	4.4e-5	0.0002	1.2e-7	0.0054	0.0001	0.0002	1.1e-5

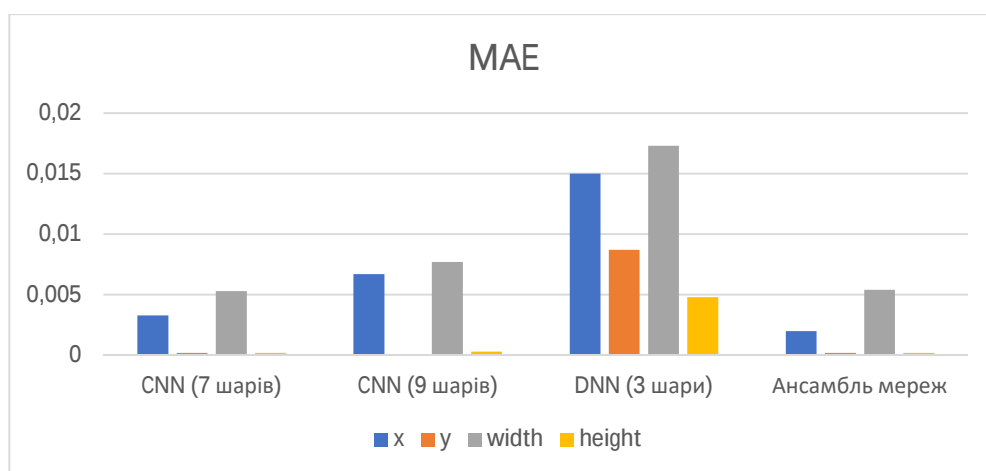


Рис. 4. Похибка прогнозування моделей (для різних геометричних характеристик), виражена в метриці MAE

Аналізуючи отримані результати, бачимо, що задача визначення позиції включення, його ширини й висоти достатньо ефективно розв'язується кожною з зазначених моделей. Наявну похибку, в межах варіативності набору й завдання дослідження, можна вважати несуттєвою.

На великому наборі даних моделі, які показали себе добре на малому, продемонстрували ідентичний або кращий результат, інколи на один порядок. Це показує, що моделі з великим ступенем ймовірності узагальнюють дані, а не просто перенавчаються. Також варто виділити те, що ансамблі нейронних мереж переважно показують найвищу, або близьку до неї, точність.

Аналізуючи метрики похибок визначення коефіцієнта теплопровідності, наведені в табл. 7, бачимо багаторазове покращення точності передбачення. Візуалізацію суттєвого покращення точності запропонованого підходу до попереднього опрацювання даних наведено на діаграмах для моделі CNN1D з трьома шарами (рис. 5).

Таблиця 7

**Метрики похибок визначення коефіцієнта теплопровідності
для різних моделей (більший набір даних)**

Модель	MAE	MSE
CNN1D (3 шари)	0.0208	0.0007
DNN (3 шари)	0.0121	0.0003
DNN (9 шарів)	0.0165	0.0006
ThermalNet #0	0.0306	0.0018
ThermalNet #1	0.0195	0.0008
ThermalNet #2	0.0146	0.0005
Ансамбль мереж	0.0096	0.0002

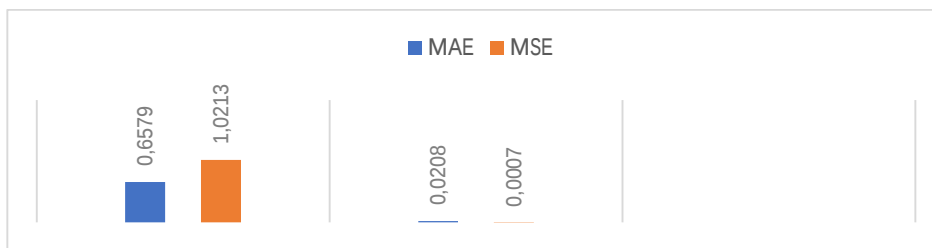


Рис. 5. Порівняння похибок визначення коефіцієнта теплопровідності класичним методом попереднього опрацювання та запропонованим підходом для двох метрик

Аналізуючи три версії запропонованої альтернативної моделі, бачимо зростання її точності від версії до версії. Отже, запропоновані зміни мали позитивний вплив. Точність третьої версії цієї моделі перевищує точність моделей CNN1D та DNN з 9-ма шарами. Це свідчить про доцільність подальших експериментів зі змінами її архітектури. Візуалізацію порівняння похибок прогнозування моделей наведено на рис. 6.

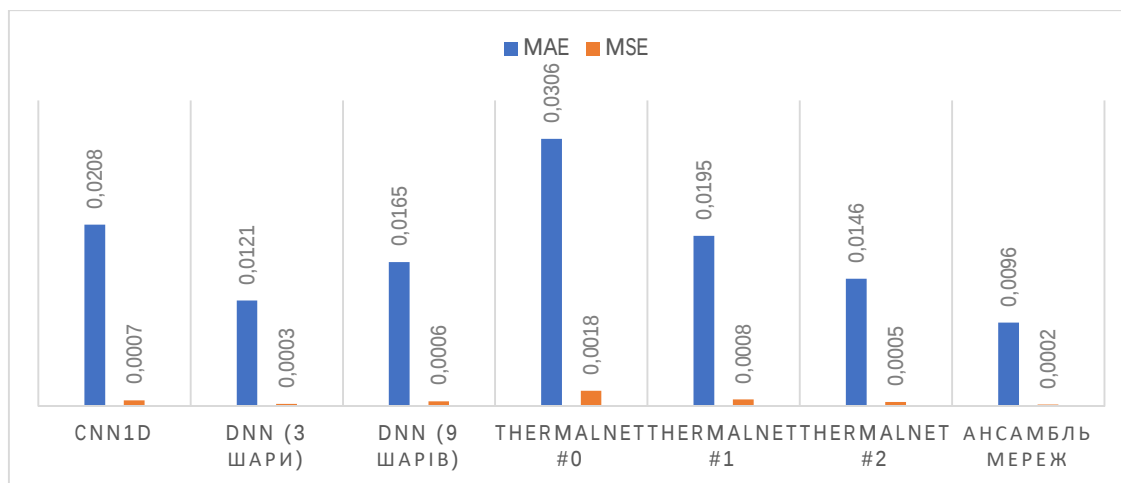


Рис. 6. Похибки прогнозування фізичної характеристики різними моделями

Точність ансамблевої мережі, як і раніше, була вищою, ніж усіх інших розглянутих моделей, що вказує на доцільність побудови ансамблів нейронних мереж для розв'язування цієї задачі.

Подальші дослідження

Незважаючи на високу точність моделей, наведених у статті, набори даних для тренування мають все ще недостатньо високу варіативність геометричних характеристик і включення, і об'єкта. Тому одним з кроків подальших досліджень є створення більш варіативного набору даних. Враховуючи, що для тренування моделей визначення геометричних характеристик необхідна інформація про розподіл температури, а для визначення фізичної характеристики – попередньо опрацьовані дані меншої розмірності, є сенс створення двох окремих наборів. Фокусування в одному з них має бути на варіативності геометричних характеристик, в іншому – на фізичній характеристиці. Також подальші дослідження мають включати можливості виявлення двох та більшої кількості включень.

Висновки

Аналіз предметної області показав, що попри значні досягнення у використанні неруйнівних методів контролю структур різного роду, покращення їхньої ефективності щодо точності визначення геометричних та фізичних характеристик включень має значну актуальність. Зменшення при цьому часу, необхідного для збору та опрацювання даних, та адаптація до нових сфер застосування й умов використання спричинює розвиток нових та удосконалення наявних підходів, які ґрунтуються на поєднанні ефективних чисельних методів розв'язування задач математичної фізики з новітніми архітектурами нейронних мереж.

Проаналізовано наукові праці, спрямовані на виявлення тріщин, включень та різноманітних дефектів у об'єктах. Після дослідження запропонованих у цих працях методів нейронних мереж (для виявлення включень і для опрацювання даних різного роду) вибрано декілька для їхньої апробації та подальшого вдосконалення згідно вимог предметної області.

Досліджено можливість визначення геометричних та фізичних характеристик однорідного включення у двовимірному об'єкті. Показано ефективність поєднання НМПГЕ з методами нейронних мереж для такого визначення. Використовуючи НМПГЕ, побудовано два набори даних, які відрізняються за розміром та варіативністю. Завдяки меншій (порівняно з МГЕ) обчислювальній складності цього методу вдалося зібрати великий набір даних зі значною варіативністю, щоб зменшити ймовірність перенавчання нейронної мережі. Запропоновано підхід до попереднього опрацювання даних, який покращує точність прогнозування фізичної характеристики включення у декілька разів. Досліджено методи нейронних мереж для ефективного розпізнавання геометричних характеристик включення. Розглянуто широке різноманіття архітектур нейронних мереж, виділено архітектури, оптимальні за складністю та точністю передбачення.

REFERENCES

1. Al-Athel, K. S., Alhasan, M. M., Alomari, A. S., & Arif, A. F. M. (2022). Damage characterization of embedded defects in composites using a hybrid thermography, computational, and artificial neural networks approach. *Heliyon*, 8(8), e10063. <https://doi.org/10.1016/j.heliyon.2022.e10063>
2. Alshoaibi, A. M., & Fageehi, Y. A. (2024). Advances in Finite Element Modeling of Fatigue Crack Propagation. *Applied Sciences*, 14(20), 9297. <https://doi.org/10.3390/app14209297>
3. Dubey, A. K., & Jain, V. (2019). Comparative Study of Convolution Neural Network's Relu and Leaky-Relu Activation Functions. B. S. Mishra, Y. R. Sood, & A. Tomar (Ред.), *Applications of Computing, Automation and Wireless Systems in Electrical Engineering* (с. 873–880). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-13-6772-4_76
4. Guan, X., & Burton, H. (2022). Bias-variance tradeoff in machine learning: Theoretical formulation and implications to structural engineering applications. *Structures*, 46, 17–30. <https://doi.org/10.1016/j.istruc.2022.10.004>

5. Han, A. L., Gan, B. S., & Setiawan, Y. (2014). The Influence of Single Inclusions to the Crack Initiation, Propagation and Compression Strength of Mortar. *Procedia Engineering*, 95, 376–385. <https://doi.org/10.1016/j.proeng.2014.12.196>
6. He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition* (Версія 1). Версія 1. arXiv. <https://doi.org/10.48550/ARXIV.1512.03385>
7. Hosna, A., Merry, E., Gyalmo, J., Alom, Z., Aung, Z., & Azim, M. A. (2022). Transfer learning: A friendly introduction. *Journal of Big Data*, 9(1), 102. <https://doi.org/10.1186/s40537-022-00652-w>
8. Hussain, M. (2024). *YOLOv5, YOLOv8 and YOLOv10: The Go-To Detectors for Real-time Vision* (Версія 1). Версія 1. arXiv. <https://doi.org/10.48550/ARXIV.2407.02988>
9. Jaddi, S., Coulombier, M., Raskin, J.-P., & Pardoën, T. (2019). Crack on a chip test method for thin freestanding films. *Journal of the Mechanics and Physics of Solids*, 123, 267–291. <https://doi.org/10.1016/j.jmps.2018.10.005>
10. Jo, J.-M. (2019). Effectiveness of Normalization Pre-Processing of Big Data to the Machine Learning Performance. *The Journal of the Korea institute of electronic communication sciences*, 14(3), 547–552. <https://doi.org/10.13067/JKIECS.2019.14.3.547>
11. Khodadad, M., & Ardakani, M. D. (2008). Inclusion Identification by Inverse Application of Boundary Element Method, Genetic Algorithm and Conjugate Gradient Method. *American Journal of Applied Sciences*, 5(9), 1158–1166. <https://doi.org/10.3844/ajassp.2008.1158.1166>
12. Kurnyta-Mazurek, P., Wrąbel, R., & Kurnyta, A. (2025). Computer-Aided Supporting Models of Customized Crack Propagation Sensors for Analysis and Prototyping. *Sensors*, 25(2), 566. <https://doi.org/10.3390/s25020566>
13. Lema, D. G., Pedrayes, O. D., Usamentiaga, R., Venegas, P., & Garcia, D. F. (2022). Automated Detection of Subsurface Defects Using Active Thermography and Deep Learning Object Detectors. *IEEE Transactions on Instrumentation and Measurement*, 71, 1–13. <https://doi.org/10.1109/TIM.2022.3169484>
14. Montesinos López, O. A., Montesinos López, A., & Crossa, J. (2022). Fundamentals of Artificial Neural Networks and Deep Learning. B O. A. Montesinos López, A. Montesinos López, & J. Crossa, *Multivariate Statistical Machine Learning Methods for Genomic Prediction* (c. 379–425). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-89010-0_10
15. Müller, D., Netzelmann, U., & Valeske, B. (2022). Defect shape detection and defect reconstruction in active thermography by means of two-dimensional convolutional neural network as well as spatiotemporal convolutional LSTM network. *Quantitative InfraRed Thermography Journal*, 19(2), 126–144. <https://doi.org/10.1080/17686733.2020.1810883>
16. Oswald-Tranta, B., Lopez De Uralde Olavera, P., Gorostegui-Colinas, E., & Westphal, P. (2023). Convolutional neural network for automated surface crack detection in inductive thermography. B N. P. Avdelidis (Ред.), *Thermosense: Thermal Infrared Applications XLV* (c. 16). Orlando, United States: SPIE. <https://doi.org/10.1117/12.2663485>
17. Patra, R., & Dutta, P. K. (2013, Березень 6). *Improved DOT reconstruction by estimating the inclusion location using artificial neural network* (R. M. Nishikawa & B. R. Whiting, Ред.). Lake Buena Vista (Orlando Area), Florida, USA. <https://doi.org/10.1117/12.2007905>
18. Ren, S., He, K., Girshick, R., & Sun, J. (2015). *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks* (Версія 3). Версія 3. arXiv. <https://doi.org/10.48550/ARXIV.1506.01497>
19. Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation* (Версія 1). Версія 1. arXiv. <https://doi.org/10.48550/ARXIV.1505.04597>
20. Tavares, L. D., Saldanha, R. R., & Vieira, D. A. G. (2015). Extreme learning machine with parallel layer perceptrons. *Neurocomputing*, 166, 164–171. <https://doi.org/10.1016/j.neucom.2015.04.018>
21. Travassos, Xisto L., Avila, S. L., & Ida, N. (2021). Artificial Neural Networks and Machine Learning techniques applied to Ground Penetrating Radar: A review. *Applied Computing and Informatics*, 17(2), 296–308. <https://doi.org/10.1016/j.aci.2018.10.001>
22. Travassos, X.L., Vieira, D. A. G., Ida, N., Vollaie, C., & Nicolas, A. (2008). Characterization of Inclusions in a Nonhomogeneous GPR Problem by Artificial Neural Networks. *IEEE Transactions on Magnetics*, 44(6), 1630–1633. <https://doi.org/10.1109/TMAG.2007.915332>
23. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2017). *Attention Is All You Need*. <https://doi.org/10.48550/ARXIV.1706.03762>
24. Weitz, S., Clausner, A., & Zschech, E. (2024). Microcracking in On-Chip Interconnect Stacks: FEM Simulation and Concept for Fatigue Test. *Journal of Electronic Materials*, 53(8), 4401–4409. <https://doi.org/10.1007/s11664-024-11091-z>
25. Yang, J., Wang, W., Lin, G., Li, Q., Sun, Y., & Sun, Y. (2019). Infrared Thermal Imaging-Based Crack Detection Using Deep Learning. *IEEE Access*, 7, 182060–182077. <https://doi.org/10.1109/ACCESS.2019.2958264>

26. Yin, W., Yang, Z., & Meng, P. (2023). Solving Inverse Scattering Problem with a Crack in Inhomogeneous Medium Based on a Convolutional Neural Network. *Symmetry*, 15(1), 119. <https://doi.org/10.3390/sym15010119>
27. Yuan, C. C. A., & Lee, C.-C. (2020). Solder Joint Reliability Modeling by Sequential Artificial Neural Network for Glass Wafer Level Chip Scale Package. *IEEE Access*, 8, 143494–143501. <https://doi.org/10.1109/ACCESS.2020.3014156>
28. Zhao, L., He, Z., Wang, Z., Su, L., & Lu, X. (2020). Simulation and Experimental Investigation on Active Thermography Test of the Solder Balls. *IEEE Transactions on Industrial Informatics*, 16(3), 1617–1624. <https://doi.org/10.1109/TII.2019.2945583>
29. Zhao, X., Zhao, Y., Hu, S., Wang, H., Zhang, Y., & Ming, W. (2023). Progress in Active Infrared Imaging for Defect Detection in the Renewable and Electronic Industries. *Sensors*, 23(21), 8780. <https://doi.org/10.3390/s23218780>
30. Zhuravchak, L., & Zabrodska, N. (2024). Algorithm for determining inclusion parameters in solving inverse problems of geoelectrical exploration using the profiling method. *Geodynamics*, 1(36)2024(1(36)), 98–107. <https://doi.org/10.23939/jgd2024.01.098>

RECOGNITION OF INCLUSION CHARACTERISTICS USING NEURAL NETWORK METHODS IN STATIONARY PROCESS MODELING

Roman Havdul'skyi¹, Liubov Zhuravchak²

^{1,2} Lviv Polytechnic National University,
Software Development Department, Lviv, Ukraine

¹ E-mail: roman.i.havdul'skyi@lpnu.ua, ORCID 0009-0004-6737-6018

² E-mail: liubov.m.zhuravchak@lpnu.ua, ORCID 0000-0002-1444-5882

© Havdul'skyi R., Zhuravchak L., 2025

Summary. Detection and identification of inclusions in the modeling of stationary processes is a crucial task in many technical fields, including materials science, electronics, and non-destructive testing. The presence of inclusions can affect the mechanical, thermal, and electrical properties of a material, making the accurate determination of their geometric and physical characteristics essential. The use of modern numerical methods and deep learning techniques opens new opportunities for improving the efficiency and accuracy of prediction results.

This paper explores the application of the indirect near-boundary element method for solving direct problems in potential theory and various types of neural networks for solving inverse problems, i.e., recognizing the specified characteristics of inclusions (their location, size, and conductivity coefficient). A series of computational experiments were conducted to solve the direct problem, allowing the generation of datasets for training and testing models utilizing convolutional, recurrent, and deep neural networks.

To enhance data analysis efficiency, the application of data normalization techniques was tested. A novel approach to feature extraction for predicting the physical characteristic of inclusions (conductivity coefficient) was proposed, significantly improving the accuracy of its computation. After extensive testing, several neural network architectures were selected for determining the geometric and physical characteristics, and neural network ensembles were constructed to integrate the best aspects of each model. The results of ensemble model testing demonstrated higher accuracy in determining the conductivity coefficient and geometric characteristics of inclusions compared to individual models.

The obtained results may contribute to the advancement of non-destructive testing methods and the detection of defects or inclusions. Future research will focus on developing an approach for identifying the characteristics of multiple inclusions and creating more diverse datasets.

Keywords - neural networks, inclusion recognition, stationary processes, convolutional neural networks, machine learning, non-destructive testing, potential theory, computational experiment.