

ЗАСТОСУВАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ В ОБРОБЦІ 3D СЕЙСМІЧНИХ ЗОБРАЖЕНЬ

Анотація. У статті представлено комплексний технічний огляд застосування методів машинного навчання (МН) та глибокого навчання (ГН) для аналізу 3D сейсмічних зображень у геофізичній розвідці. Розглядається проблема обробки терабайтних обсягів сейсмічних даних, які традиційно вимагають місяців ручної інтерпретації, та обґрунтовується необхідність переходу до автоматизованих методів аналізу.

Детально описуються основні архітектури нейронних мереж (CNN, U-Net, TransUNet) та їх застосування для ключових завдань: виявлення розломів, делімітації соляних тіл та класифікації літофацій. Особлива увага приділяється практичним аспектам впровадження МН, включаючи роботу з форматом SEG-Y через Python-бібліотеки, методи подолання дефіциту розмічених даних через синтетичну генерацію та аугментацію, а також використання публічних наборів даних (F3, FORCE) для навчання моделей.

Висвітлюються сучасні напрямки досліджень, зокрема самоконтрольоване навчання (SSL) для роботи з нерозміченими даними, фізико-інформовані нейронні мережі (PINNs) для інтеграції геофізичних законів, та методи квантифікації невизначеності результатів.

Ключові слова: машинне навчання, глибоке навчання, 3D сейсмічні дані, згорткові нейронні мережі, геофізична інтерпретація, виявлення розломів, Python, SEG-Y.

Вступ. Дослідження та характеристика підземної частини Землі є фундаментальними для енергетичного сектору, екологічного моніторингу та оцінки геологічних ризиків. Протягом десятиліть 3D сейсмічна візуалізація була основним інструментом для візуалізації підземних структур, надаючи неоціненні дані для ідентифікації вуглеводневих резервуарів, потенційних місць для секвестрації вуглецю та джерел геотермальної енергії. Ця технологія використовує звукові хвилі для генерації тривимірних об'ємних зображень, або “кубів”, геологічних формацій глибоко під поверхнею. Процес включає збір величезних обсягів сейсмічних даних, які потім піддаються інтенсивній обчислювальній обробці для створення інтерпретованого зображення підземної частини.

Однак, величезний обсяг і складність сучасних сейсмічних наборів даних представляють значні виклики. Одне 3D дослідження може генерувати терабайти даних, роблячи ручну інтерпретацію трудомістким, часозатратним і суб'єктивним процесом. Ця проблема “великих даних” довела традиційні методи інтерпретації до їх меж, створюючи попит на більш ефективні, об'єктивні та потужні аналітичні інструменти.

У відповідь, галузь геофізики зазнає цифрової трансформації, все більше інтегруючи методи машинного навчання (МН) та глибокого навчання (ГН) у робочі процеси сейсмічного аналізу. Ці технології пропонують нову парадигму, переходячи від ручної, гіпотезо-орієнтованої інтерпретації до автоматизованого, даних-орієнтованого розпізнавання шаблонів. Алгоритми МН можуть аналізувати величезні, багатоатрибутні набори даних для ідентифікації тонких геологічних особливостей, автоматизувати повторювані завдання, такі як виділення розломів, і забезпечувати більш повторювані та об'єктивні результати. Цей звіт надає комплексний огляд застосування машинного навчання до 3D сейсмічних даних, детально описуючи методології, експерименти, результати та майбутні напрямки цієї швидко розвиваючої галузі.

Пов'язані роботи. Застосування машинного навчання в сейсмічній розвідці зросло експоненціально, з алгоритмами, що впроваджуються практично на кожному етапі робочого процесу обробки та інтерпретації. Літературний аналіз публікацій з основних геонаукових репозиторіїв, таких як SEG та EAGE, показує, що впровадження МН в основному зосереджені на автоматизації окремих завдань для скорочення часу обробки та покращення ефективності та результатів.

Для сейсмічної інтерпретації основними цілями для застосування МН є виявлення розломів, класифікація літофацій та ідентифікація геотіл, таких як соляні куполи.

Виявлення розломів: Це завдання є критичним для розуміння структурних пасток і міграції флюїдів, але традиційно є одним з найбільш часозатратних аспектів ручної інтерпретації. Глибоке навчання, особливо з використанням згорткових нейронних мереж (CNN) та архітектур U-Net, виявилось дуже ефективним і результативним. Ці методи формулюють ідентифікацію розломів як проблему сегментації зображення, класифікуючи кожен воксель як розлом або не-розлом. Дослідження показують, що підходи на основі CNN можуть виробляти чистіші, більш послідовні результати за години, порівняно з місяцями, необхідними для ручного виділення. Останні роботи досліджували передові архітектури, такі як TransUnet та Dual U-Net, щоб краще захоплювати як локальні, так і глобальні особливості розломів.

Делімітація геотіл і солі: Ідентифікація соляних тіл є критичною в багатьох нафтових басейнах, але складною через погане сейсмічне зображення на їх межах. CNN та моделі U-Net були успішно застосовані для сегментації цих складних структур, часто перевершуючи традиційні методи, базовані на атрибутах, шляхом навчання особливостей безпосередньо з сирих сейсмічних даних амплітуди.

Класифікація літофацій і стратиграфії: Це включає ідентифікацію різних типів порід і осадових середовищ з сейсмічних шаблонів. Використовуються як контрольовані, так і неконтрольовані методи навчання. Неконтрольовані робочі процеси, часто поєднуючи аналіз головних компонентів (PCA) для зменшення розмірності та самоорганізуючі карти (SOM) для кластеризації, є потужними для дослідницького аналізу в областях з обмеженими даними. Контрольовані методи, включаючи різні архітектури CNN, використовуються, коли доступні дані свердловин для маркування, дозволяючи безпосереднє прогнозування розподілу фацій по всьому сейсмічному об'єму.

Незважаючи на ці успіхи, залишаються значні виклики. Дефіцит розмічених даних істинних значень є основним вузьким місцем для контрольованого навчання. Крім того, моделі МН часто борються з узагальнюваністю, де модель, навчена на одному наборі даних, погано працює на іншому з різного геологічного середовища. Це стимулювало дослідження в нових областях, таких як самоконтрольоване навчання (SSL) для навчання з нерозмічених даних, фізико-інформовані нейронні мережі (PINNs) для інтеграції фізичних законів у моделі, та квантифікація невизначеності (UQ) для оцінки довіри моделі.

Цілі дослідження. Основною метою цього звіту є надання структурованого та комплексного аналізу застосування та розробки алгоритмів машинного навчання для роботи з 3D сейсмічними зображеннями. Ключові цілі дослідження:

Огляд сучасного стану: Систематично дослідити та синтезувати поточний ландшафт застосувань МН в сейсмічній інтерпретації, спираючись на академічну літературу та промислові практики для ідентифікації ключових тенденцій, домінуючих методологій і встановлених еталонів.

Деталізація основних методологій: Надати поглиблене пояснення основних архітектур МН і робочих процесів, використовуваних для фундаментальних завдань інтерпретації, включаючи делімітацію систем розломів, ідентифікацію соляних тіл і стратиграфічну класифікацію фацій.

Окреслити практичний робочий процес МН: Описати наскрізний процес застосування МН до сейсмічних даних, покриваючи критичні кроки, такі як надходження та обробка даних, стратегії створення навчальних даних для подолання вузького місця анотації, та використання публічних наборів даних для експериментування та валідації моделей.

Ідентифікувати та обговорити ключові виклики та майбутні напрямки: Проаналізувати основні перешкоди, що заважають широкому впровадженню та надійності МН в цій області, такі як узагальнюваність моделей та потреба в квантифікації невизначеності. Дослідити межі досліджень, включаючи фізико-інформоване та самоконтрольоване навчання, що має на меті вирішити ці виклики та визначити майбутнє інтелектуальної сейсмічної інтерпретації.

Методологія. Методології застосування машинного навчання до 3D сейсмічних даних охоплюють обробку даних, створення навчальних даних та впровадження специфічних моделей навчання, адаптованих для завдань геологічної інтерпретації.

Обробка та підготовка даних. Стандартним форматом для сейсмічних даних є SEG-Y, який може бути незручним для фреймворків МН. Екосистема Python пропонує бібліотеки з відкритим кодом, такі як `segyio` для ефективного, низькорівневого читання та запису, та `segysak`, який використовує `hdf5` та `dask` для високорівневої, масштабованої маніпуляції даними на наборах даних, які занадто великі для розміщення в пам'яті.

Типовий робочий процес починається з підготовки даних, яка включає очищення, нормалізацію та інженерію ознак. Для багатоатрибутного аналізу загальним першим

кроком є генерація десятків сейсмічних атрибутів (наприклад, амплітуда, частота, когерентність), призначених для підкреслення специфічних геологічних особливостей. Для управління цими високорозмірними даними часто використовується аналіз головних компонентів (РСА) для зменшення розмірності та вибору найбільш інформативних, некорельованих атрибутів для входу моделі.

Анотація навчальних даних. Контрольоване навчання, найпотужніша парадигма для багатьох завдань, сильно обмежене дефіцитом розмічених даних. Використовується кілька стратегій для подолання цього “вузького місця анотації”:

- Ручне та AI-асистоване маркування: Експерт вручну маркує особливості, процес, який є повільним, але забезпечує високоякісні дані. Для прискорення цього AI-асистовані інструменти можуть пропонувати початкові позначки для уточнення експертом-людиною, створюючи робочий процес “людина в циклі”.

- Генерація синтетичних даних: Потужною альтернативою є створення 3D геологічних моделей з ідеально відомими особливостями (наприклад, розломи) і потім використання прямого моделювання для симуляції відповідних сейсмічних даних. Це генерує необмежені, ідеально розмічені синтетичні дані для навчання моделей.

- Збільшення даних: Це включає створення нових навчальних зразків шляхом застосування трансформацій до існуючих даних, таких як додавання шуму або зсув даних, що змушує модель вивчати більш стійкі особливості та покращує узагальнення. Генеративні змагальні мережі (GANs) є передовою технікою, використовуваною для генерації нових, реалістичних синтетичних сейсмічних хвильових форм на основі статистичних характеристик реальних даних, додатково збагачуючи навчальний набір.

Моделі машинного навчання. Специфічні архітектури МН виявилися особливо ефективними для завдань сейсмічної інтерпретації.

U-Net для сегментації розломів і солі: Архітектура U-Net, первинно з біомедичної візуалізації, є сучасним стандартом для завдань семантичної сегментації, таких як делімітація розломів і солі. Її структура кодер-декодер з “пропускними з’єднаннями” дозволяє їй захоплювати особливості на кількох масштабах, поєднуючи широку контекстну інформацію з точними просторовими деталями для створення гострих, точних меж сегментації. Варіанти, такі як TransUnet, інтегрують блоки Transformer для кращого захоплення далекосяжних залежностей та покращення континуальності прогнозованих розломів.

CNN для класифікації фацій і солі: Згорткові нейронні мережі (CNN) є дуже ефективними для завдань класифікації. У загальному підході CNN навчається на невеликих 3D патчах сейсмічних даних для класифікації центрального вокселя (наприклад, як ‘сіль’ або специфічний тип фацій). Цей наскрізний метод дозволяє мережі автоматично вивчати оптимальні дискримінативні особливості безпосередньо з сирих сейсмічних даних амплітуди, уникаючи упередженого ручного вибору атрибутів. Більш передові моделі кодер-декодер можуть класифікувати цілі сейсмічні секції відразу.

Неконтрольована кластеризація (PCA та SOM): Для аналізу сейсмічних фацій в областях з обмеженими даними свердловин поширений неконтрольований робочий процес. Після використання PCA для вибору ключових атрибутів використовується са-

моорганізуюча карта (SOM) для кластеризації. SOM є неконтрольованою нейронною мережею, яка відображає високорозмірні дані атрибутів на 2D сітку, де різні кластери представляють окремі сейсмічні фації, геологічне значення яких потім інтерпретується геовченим.

Експерименти. Розробка та валідація згаданих методологій спираються на стандартизовані експерименти з використанням публічно доступних еталонних наборів даних. Ці набори даних є критичними для забезпечення об'єктивного порівняння між різними моделями та сприяння відтворюваних досліджень. Ключові портали для цих ресурсів включають репозиторій TerraNubis та урядові джерела, такі як Геологічна служба США (USGS).

Набір даних Нідерландського блоку F3. Одним з найбільш широко використовуваних публічних наборів даних для досліджень МН є Нідерландський блок F3, розташований на морі в Північному морі. Це 3D сейсмічне дослідження відоме своїм високоякісним зображенням різноманітних геологічних особливостей, включаючи:

- Великомасштабні сигмоїдальні кліноформи з великої дельтової системи.
- Окремі сейсмічні фації, канали та піщані хвилі.
- Аномалії, такі як газові димарі та яскраві плями від біогенного газу.
- Глибокий соляний купол Цехштейн та різні системи розломів.

Набір даних, наданий як проект OpendTest, включає 3D сейсмічний об'єм, численні похідні об'єми атрибутів, інтерпретовані горизонти та дані каротажу з чотирьох свердловин. Ця багата комбінація робить його ідеальним випробувальним стендом для широкого спектру експериментів МН, від неконтрольованої класифікації фацій до контрольованої сегментації соляних тіл.

Набір даних конкурсу FORCE 2020 ML. Для специфічного тестування стійкості та узагальнюваності алгоритмів виявлення розломів норвезький консорціум FORCE організував конкурс машинного навчання в 2020 році. Наданий набір даних складається з двох 3D сейсмічних досліджень з Північно-Західного шельфу Австралії:

- Навчальні дані: Дослідження поля Іхтіос разом з деякими синтетичними позначками розломів.
- Сліпі тестові дані: Сусіднє дослідження Адель.

Геологія характеризується добре вираженою полігональною системою розломів у більш поверхневих секціях та більш складною, дифузною системою розломів у більш глибокій юрській секції. Конкурс мав на меті оцінити, наскільки добре моделі МН, часто навчені на чистих синтетичних даних, можуть працювати на “не дуже ідеальних” реальних даних зі складними патернами розломовування, роблячи його критичним еталоном для узагальнюваності.

Результати та обговорення. Застосування методологій машинного навчання до сейсмічних наборів даних дало трансформативні результати, водночас підкресливши значні виклики та області для майбутніх досліджень.

Продуктивність та досягнення автоматизації. Моделі МН послідовно демонстрували здатність автоматизувати завдання інтерпретації зі значними здобутками в

швидкості та об'єктивності порівняно з ручними методами. Для виявлення розломів модель глибокого навчання може делімітувати цілу мережу розломів у 3D об'ємі за години, завдання, яке зайняло б у експерта-людини місяці. Результати є не лише швидшими, але й більш повторюваними та менш схильними до упередженості окремого інтерпретатора. У класифікації фацій неконтрольовані методи, такі як SOM, можуть швидко аналізувати десятки атрибутів одночасно, виявляючи тонкі шаблони та геологічні тенденції, які є непомітними для людського ока.³¹ Контрольовані CNN, коли навчені на даних свердловин, можуть точно прогнозувати 3D розподіл літофацій, забезпечуючи детальні моделі резервуарів. Ці автоматизовані робочі процеси звільняють геовчених від нудних ручних завдань, дозволяючи їм зосередитися на аналізі та прийнятті рішень вищого рівня.

Ключові виклики та майбутні напрямки. Незважаючи на ці успіхи, залишається кілька фундаментальних викликів, що визначають межі поточних досліджень.

Узагальнюваність: Основною слабкістю багатьох моделей глибокого навчання є їх неспроможність узагальнити з одного сейсмічного дослідження на інше. Модель може вивчити особливості, специфічні для шуму або параметрів збору її навчальних даних, що призводить до поганої продуктивності на нових даних. Вирішення цього потребує розробки більш стійких моделей через техніки, такі як трансферне навчання та агресивне збільшення даних.

Фізико-інформоване машинне навчання: Для покращення узагальнюваності та вирішення природи “чорної скриньки” чисто даних-орієнтованих моделей дослідники розробляють фізико-інформовані нейронні мережі (PINNs). PINNs інтегрують керуючі фізичні закони, такі як рівняння акустичної хвилі, безпосередньо в функцію втрат моделі. Це обмежує модель до виробництва фізично правдоподібних рішень, покращуючи стійкість, особливо там, де дані розріджені або зашумлені.

Навчання без міток: Дефіцит розмічених даних залишається найбільшою перешкодою для розгортання контрольованого навчання в масштабі.⁴⁰ Самоконтрольоване навчання (SSL) пропонує переконливе рішення, дозволяючи моделям вивчати значущі представлення з величезних обсягів нерозмічених даних. Навчаючись на “претекстовому завданні”, такому як прогнозування замаскованих частин сейсмічного зображення, модель вивчає базову структуру даних, яка потім може бути налаштована для специфічного завдання з дуже невеликою кількістю міток.

Квантифікація невизначеності (UQ): Для рішень розвідки з високими ставками недостатньо однієї детерміністичної передбачення. UQ має на меті квантифікувати довіру моделі до її передбачень. Техніки, такі як байєсівські нейронні мережі або ансамблі моделей, використовуються для оцінки діапазону можливих результатів, розрізняючи невизначеність від зашумлених даних (алеаторну) та невизначеність від самої моделі (епістемічну). Результуючі карти ймовірності є набагато більш цінними для прийняття рішень на основі ризику, ніж одна інтерпретація.

Висновок. Інтеграція машинного навчання в аналіз 3D сейсмічних зображень означає ключову еволюцію в дослідженні земної підповерхні. Цей технологічний зсув вирішує властиві обмеження ручної інтерпретації в епоху експоненціально зростаючих

обсягів даних, пропонуючи драматичні покращення в ефективності, об'єктивності та глибині геологічного розуміння. Основні завдання інтерпретації, такі як делімітація розломів, соляних тіл та стратиграфічних фацій, тепер успішно автоматизуються спеціалізованими архітектурами глибокого навчання, такими як U-Nets та CNNs, які послідовно перевершують традиційні методи як за швидкістю, так і за точністю.

Практичне застосування цих моделей підтримується розширюючою екосистемою інструментів з відкритим кодом для обробки даних та публічно доступними еталонними наборами даних, які є необхідними для валідації нових алгоритмів. Однак шлях до широкого впровадження не позбавлений викликів. Критичне вузьке місце дефіциту розмічених даних продовжує стимулювати інновації в областях, таких як генерація синтетичних даних та само контрольоване навчання.

Дивлячись вперед, межа досліджень зосереджена на побудові наступного покоління інтелектуальних інструментів інтерпретації, які є не лише точними, але й стійкими, надійними та заслуговуючи довіри. Ключові напрямки досліджень — покращення узагальнюваності моделей, інтеграція фізичних принципів через PINNs, дозвіл навчання без міток через SSL, та квантифікація довіри моделі з UQ — глибоко взаємопов'язані. Прогрес у цих областях приведе до систем ШІ, які можуть бути розгорнуті більш широко та з більшою впевненістю. Ця триваюча революція продовжуватиме трансформувати роль геовченого, надаючи їм потужні аналітичні інструменти для прийняття більш обґрунтованих, даних-орієнтованих рішень про ресурси та навколишнє середовище Землі.

ЛІТЕРАТУРА

1. 3-D Seismic Interpretation. M. Bacon, R. Simm, T. Redshaw, Cambridge University Press, 2003
2. Fault detection seismic data using graph convolutional network. Patitapaban Palo, Aurobinda Routray, Rahul Mahadik, Sanjai Kumar Singh, The Journal of Supercomputing, 03.2023, DOI:10.1007/s11227-023-05173-8
3. Introduction to Seismic Imaging. Prof. Gerald Gardner, Society of Exploration Geophysicists
4. Discovering hidden hydrocarbons: using seismic-imaging technology to map formations far below the earth's surface. ExxonMobil, 09.2018
5. Introduction to seismic data and processing. Cambridge University Press, Hua-Wei Zhou
6. Solving the challenge of seismic data management. SLB, Victor Aarre
7. Sebai, D., Zouaoui, M. & Ghorbel, F. Seismic data compression: an overview. Multimedia Systems 30, 38 (2024). <https://doi.org/10.1007/s00530-023-01233-4>
8. Pitfalls and limitations in seismic attribute interpretation of tectonic features. SEG Library, Kurt J. Marfurt, Tiago M. Alves, doi.org/10.1190/INT-2014-0122.1
9. Lappin, M. (2004). 3D Seismic Technology: Are We Realising Its Full Potential?. Geological Society, London, Memoirs. Geological Society of London. doi: 10.1144/GSL.MEM.2004.029.01.01.

10. Machine learning for seismic exploration: Where are we and how far are we from the holy grail?. Farbod Khosro Anjom, Francesco Vaccarino, Laura Valentina Socco, GEOPHYSICS - SEG Library, <https://doi.org/10.1190/geo2023-0129.1>
11. Seismic facies analysis using machine learning. Thilo Wrona, Indranil Pan, Robert L. Gawthorpe, and Haakon Fossen, GEOPHYSICS - SEG Library, <https://doi.org/10.1190/geo2017-0595.1>
12. Unravel Complex Strike-Slip System in Frontier Collision Margin of Banggai-Sula Basin, Eastern Indonesia: A Machine-Learning Augmentation for 3D Seismic Interpretation. Atha Khawarizmy, Fakhriar Naufaldi, Krishna Pratama Laya, Wahyudin Suwarlan, Iswani Waryono, Ghufon Fauzi, OB Pertamina Medco Tomori Ltd, 2 Pi Energy, Badley Geoscience
13. Machine learning for seismic exploration: Where are we and how far are we from the holy grail? Farbod Khosro Anjom, Francesco Vaccarino, Laura Valentina Socco, GeoScience-World, Geophysics, vol. 89, no. 1 (january-february 2024); p. wa157–wa178, 17 figs., 2 tables. [10.1190/geo2023-0129.1](https://doi.org/10.1190/geo2023-0129.1)
14. Seismic Fault Detection using Neural Networks. GeoConvention, Ayda Azad Khorasani, J.Kim Welford, Alison Malcolm, Department of Earth Sciences, Memorial University of Newfoundland, 05.2023
15. Comparing Modern Deep Learning with Traditional Methods for Seismic Fault Interpretation. N. Ngcobo, M. Manzi, G. Nwaila, J. Bourdeau, S. Zhang, European Association of Geoscientists & Engineers, NSG 2024 5th Conference on Geophysics for Mineral Exploration and Mining, Sep 2024, Volume 2024, p.1 – 5, <https://doi.org/10.3997/2214-4609.202420194>
16. Artificial intelligence methods for predicting hydrocarbon deposits from three-dimensional seismic images. I. Dmytriieva; A. Dmytrenko, System technologies, 2025, <https://doi.org/10.34185/1562-9945-4-159-2025-16>

REFERENCES

1. 3-D Seismic Interpretation. M. Bacon, R. Simm, T. Redshaw, Cambridge University Press, 2003
2. Fault detection seismic data using graph convolutional network. Patitapaban Palo, Aurobinda Routray, Rahul Mahadik, Sanjai Kumar Singh, The Journal of Supercomputing, 03.2023, DOI:10.1007/s11227-023-05173-8
3. Introduction to Seismic Imaging. Prof. Gerald Gardner, Society of Exploration Geophysicists
4. Discovering hidden hydrocarbons: using seismic-imaging technology to map formations far below the earth's surface. ExxonMobil, 09.2018
5. Introduction to seismic data and processing. Cambridge University Press, Hua-Wei Zhou
6. Solving the challenge of seismic data management. SLB, Victor Aarre
7. Sebai, D., Zouaoui, M. & Ghorbel, F. Seismic data compression: an overview. Multimedia Systems 30, 38 (2024). <https://doi.org/10.1007/s00530-023-01233-4>
8. Pitfalls and limitations in seismic attribute interpretation of tectonic features. SEG Library, Kurt J. Marfurt, Tiago M. Alves, doi.org/10.1190/INT-2014-0122.1

9. Lappin, M. (2004). 3D Seismic Technology: Are We Realising Its Full Potential?. Geological Society, London, Memoirs. Geological Society of London. doi: 10.1144/GSL.MEM.2004.029.01.01.
10. Machine learning for seismic exploration: Where are we and how far are we from the holy grail? Farbod Khosro Anjom, Francesco Vaccarino, Laura Valentina Socco, GEOPHYSICS - SEG Library, <https://doi.org/10.1190/geo2023-0129.1>
11. Seismic facies analysis using machine learning. Thilo Wrona, Indranil Pan, Robert L. Gawthorpe, and Haakon Fossen, GEOPHYSICS - SEG Library, <https://doi.org/10.1190/geo2017-0595.1>
12. Unravel Complex Strike-Slip System in Frontier Collision Margin of Banggai-Sula Basin, Eastern Indonesia: A Machine-Learning Augmentation for 3D Seismic Interpretation. Atha Khawarizmy, Fakhriar Naufaldi, Krishna Pratama Laya, Wahyudin Suwarlan, Iswani Waryono, Ghufron Fauzi, OB Pertamina Medco Tomori Ltd, 2 Pi Energy, Badley Geoscience
13. Machine learning for seismic exploration: Where are we and how far are we from the holy grail? Farbod Khosro Anjom, Francesco Vaccarino, Laura Valentina Socco, GeoScience-World, Geophysics, vol. 89, no. 1 (january-february 2024); p. wa157–wa178, 17 figs., 2 tables. [10.1190/geo2023-0129.1](https://doi.org/10.1190/geo2023-0129.1)
14. Seismic Fault Detection using Neural Networks. GeoConvention, Ayda Azad Khorasani, J.Kim Welford, Alison Malcolm, Department of Earth Sciences, Memorial University of Newfoundland, 05.2023
15. Comparing Modern Deep Learning with Traditional Methods for Seismic Fault Interpretation. N. Ngcobo, M. Manzi, G. Nwaila, J. Bourdeau, S. Zhang, European Association of Geoscientists & Engineers, NSG 2024 5th Conference on Geophysics for Mineral Exploration and Mining, Sep 2024, Volume 2024, p.1 – 5, <https://doi.org/10.3997/2214-4609.202420194>
16. Artificial intelligence methods for predicting hydrocarbon deposits from three-dimensional seismic images. I. Dmytriieva, A. Dmytrenko, System technologies, 2025, <https://doi.org/10.34185/1562-9945-4-159-2025-16>

Received 25.01.2026.
Accepted 29.01.2026.

Application of machine learning methods in the processing of 3D seismic images

This paper presents a comprehensive technical review of machine learning (ML) and deep learning (DL) methods applied to 3D seismic image analysis in geophysical exploration. The exponential growth in seismic data acquisition has created a "big data" challenge where single 3D surveys generate terabytes of complex volumetric data, making traditional manual interpretation increasingly impractical and subjective. This work systematically examines the paradigm shift from hypothesis-driven manual interpretation to data-driven automated pattern recognition, addressing the critical need for more efficient, objective, and repeatable analytical tools in the energy sector, environmental monitoring, and geological risk assessment.

The paper provides an in-depth analysis of state-of-the-art neural network architectures and their specific applications in seismic interpretation. Convolutional Neural Networks (CNNs) and U-Net architectures are examined for their effectiveness in fault detection,

achieving clean, consistent results in hours compared to months of manual work. Advanced architectures like TransUnet and Dual U-Net are explored for their ability to capture both local and global geological features. The study covers supervised learning approaches for lithofacies classification and stratigraphic interpretation, as well as unsupervised methods combining Principal Component Analysis (PCA) with Self-Organizing Maps (SOM) for exploratory analysis in data-limited regions.

Significant emphasis is placed on practical implementation challenges and solutions. The paper details workflows for handling SEGY format data using Python ecosystems (segvio, segysak with xarray and dask), strategies for multi-attribute analysis, and approaches to overcome the critical bottleneck of labeled training data scarcity through synthetic data generation, augmentation techniques, and utilization of public datasets (F3, FORCE, TerraNubis). Emerging research frontiers are thoroughly discussed, including self-supervised learning (SSL) for leveraging unlabeled data, physics-informed neural networks (PINNs) for incorporating domain knowledge, and uncertainty quantification (UQ) methods for assessing model confidence.

Keywords: machine learning, deep learning, 3D seismic data, convolutional neural networks, geophysical interpretation, fault detection, Python, SEGY.

Дмитрієва Ірина Сергіївна - к.т.н., доцент, доцент кафедри Інформаційних технологій і систем, Український державний університет науки і технологій.

ORCID: <https://orcid.org/0009-0008-3298-7563>

Дмитренко Андрій Миколайович - аспірант кафедри Інформаційних технологій і систем, Український державний університет науки і технологій.

ORCID: <https://orcid.org/0009-0009-4939-987X>

Dmytriieva Iryna - Ph.D., Ass.Professor, Associate Professor of the Department of Information Technologies and Systems, Ukrainian State University of Science and Technology.

ORCID: <https://orcid.org/0009-0008-3298-7563>

Dmytrenko Andrii - Postgraduate Student of the Department of Information Technologies and Systems, Ukrainian State University of Science and Technology.

ORCID: <https://orcid.org/0009-0009-4939-987X>