

АНАЛІЗ LArTPC-ДАНИХ З ВИКОРИСТАННЯМ МЕТОДІВ МАШИННОГО НАВЧАННЯ

А. Фалько^{ORCID}, О. Гогота^{ORCID}, Р. Єрмоленко^{ORCID}, І. Каденко^{ORCID}

Київський національний університет імені Тараса Шевченка,

e-mails: falko@knu.ua, olga.gogota@knu.ua, ruslan.yermolenko@knu.ua, imkadenko@knu.ua

(Отримано 12 серпня 2023 р.; в остаточному вигляді — 18 січня 2024 р.; прийнято до друку — 31 січня 2024 р.;
опубліковано онлайн — 05 березня 2024 р.)

Глибокі згорткові нейронні мережі (CNN, ЗНН) продемонстрували високу ефективність для аналізу даних у багатьох галузях фізичних досліджень. У фізиці елементарних частинок та високих енергій однією із задач для розв'язання є аналіз треків частинок, отриманих у трекових детекторах, як-от часопроеційні камери на зрідженому аргоні (англ. LArTPC).

Оскільки на детектори нейтринних експериментів накладено вимогу масштабованості, вона висувається й щодо реалізацій моделей CNN. Масштабованість для обробки даних LArTPC досягається внаслідок розрідженості даних, які мають вигляд тонких траєкторій. Для обробки розріджених даних запропоновано просторово розріджені згорткові нейронні мережі.

У статті подано результати семантичної сегментації змодельованих LArTPC-даних PИLArNET з використанням просторово розріджених архітектур та протестовані різні модифікації класичної архітектури U-Net: Attention U-ResNet, U-ResNet3+, U-ResNet із додатковим деер-блоком, а також функцій утрат: focal loss, balanced cross-entropy для підвищення точності ідентифікації мішель-електронів. Найкращі результати були отримані для balanced cross-entropy. Подальше поліпшення можливе завдяки комбінуванню декількох методів.

Ключові слова: нейтрино, класифікація, нейронні мережі, машинне навчання.

DOI: <https://doi.org/10.30970/jps.28.1802>

1. ВСТУП

У цій статті наведено результати адаптації та застосування моделей нейронних мереж, що продемонстрували свою ефективність у задачах комп'ютерного бачення, до даних, отриманих з відомих у фізиці нейтрино, детекторів LArTPC (Liquid Argon Time Projection Chambers) [1].

LArTPC — тип трекових детекторів, що дає змогу отримати 2D- або 3D-зображення треків частинок із надзвичайно високою роздільною здатністю (\approx мм/піксель). Оскільки нейтрино — частинка з найбільшою проникною здатністю в стандартній моделі, нейтринні детектори мають надзвичайно великий активний об'єм. Відомі сучасні нейтринні експерименти, як-от: MicroBooNE [2], SBND [3], ICARUS [4], DUNE [5] — мають активні об'єми масою 90, 112, 600, та 40 000 тонн відповідно.

Дослідження треків частинок (протонів, мюонів, електронів, мезонів), що є продуктами реакцій нейтрино з атомами аргону, дають змогу отримати докладну й повну картину взаємодії та визначити аромат й енергію нейтрино. Автори статті [6] розробили послідовність алгоритмів реконструювання фізичних подій усередині LArTPC методами машинного навчання. Першою ланкою в реконструюванні подій нейтрино є семантична сегментація — класифікація пікселів стосовно належності до траєкторії певного класу частинок.

Зображення великих LArTPC-детекторів мають розмір 10^6 – 10^9 пікселів (зображення з детектора MicroBooNE — 80 мегапікселів). Траєкторії, що утворені внаслідок йонізації і мають поперечні розміри в де-

кілька пікселів, є тонкими та неперервними. За великих розмірів самого зображення траєкторії тонкі, а тому займають малий об'єм. У підсумку LArTPC-зображення є розрідженими. Водночас самі треки неперервні, тобто ці зображення характеризуються локально високою щільністю активних пікселів.

Ці характеристики призводять до того, що матрична алгебра, яку використовують традиційні згорткові нейронні мережі (CNN), є обчислювально неефективною для LArTPC-зображень, що переважно заповнені нульовими значеннями.

Постає питання: як ефективно, з погляду розміщення даних у пам'яті, подати розріджені дані та як побудувати для них ефективну матричну алгебру згортки? Для представлення даних LArTPC доцільно використати особливий тип даних — розріджені тензори (англ. sparse tensors) [7], а для операцій згортки, замість класичної операції згортки (*convolution*), — розріджені згортки (англ. *sparse convolution*) [8]. Ми називатимемо такі нейронні мережі зі згортковими шарами, що приймають як вхідні дані розріджені тензори, просторо розрідженими згортковими мережами.

Автори публікації [10] використали розріджену архітектуру U-ResNet для 2D- та 3D-семантичної сегментації змодельованих даних PИLArNET [11]. У роботах [12–14] її застосовують під час реконструювання взаємодії нейтрино.

У цій праці ми поставили за мету реалізувати задачу 3D-семантичної сегментації даних PИLArNET та протестували різні модифікації попередньої архітектури [10] для поліпшення точності сегментації треків заряджених частинок, зокрема треків мішель-електронів [25] (мішель-електрон — електрон, що утворюється під час розпаду мюона в стані спокою).



II. НАБІР ДАНИХ P1LArNET

Набір даних P1LArNET [11] отриманий моделюванням методом Монте Карло з допомогою Geant4 [15] та LArSoft [16] та містить по 100 000 зображень розмірами 192, 512, 768 пікселів. Лінійні розміри пікселя становлять 3 мм. Усього 300 000 зображень представилих 2D та 3D-екземплярами.

Набір даних P1LArNET [11], отриманий моделюванням методом Монте Карло за допомогою Geant4 [15] та LArSoft [16], містить по 100 000 зображень розмірами 192, 512, 768 пікселів. Лінійні розміри пікселя становлять 3 мм. Усього 300 000 зображень, репрезентованих 2D- та 3D-екземплярами.

Моделювали дані в три етапи:

1. Моделювання генератора події.

Для генерації кожної події використовували два типи генераторів первинних частинок:

(i) *MultiPartVertex*: від 1 до 6 первинних частинок, включно з електронами, гамма-променями, мюонами, антимюонами, зарядженими піонами та протонами, що виходять з однієї вершини.

(ii) *MultiPartRain*: декілька вершин з ізольованими первинними частинками: електрон, мюон, антимюон або протон. Десять таких частинок генеруються в більшому об'ємі, і записується обрізаний 3D-об'єм.

2. Моделювання взаємодії частинок із середовищем детектора.

Подальша взаємодія частинок із середовищем передбачає як розсіювання енергії в середовищі, так і вторинні взаємодії внаслідок генерації вторинних частинок. Ці процеси змодельовані за допомогою Geant4. Водночас середовище вокселізувалося в кубічні 3D-комірки однакових лінійних розмірів 3 мм, а поглинена середовищем енергія розподілена по окремих вокселях.

3. Просторове розмиття енергії для імітації дифузії електронів.

Процеси дрейфу електронів та формування сигналу зчитувальною електронікою не моделювали. Натомість, для імітації дифузії поглинута середовищем енергія була перерозподілена між сусідніми пікселями вздовж траєкторії за нормальним розподілом ($\sigma = 1$ мм) (див. деталі у [11]). Це дає змогу наближено відтворювати роздільну здатність справжнього детектора.

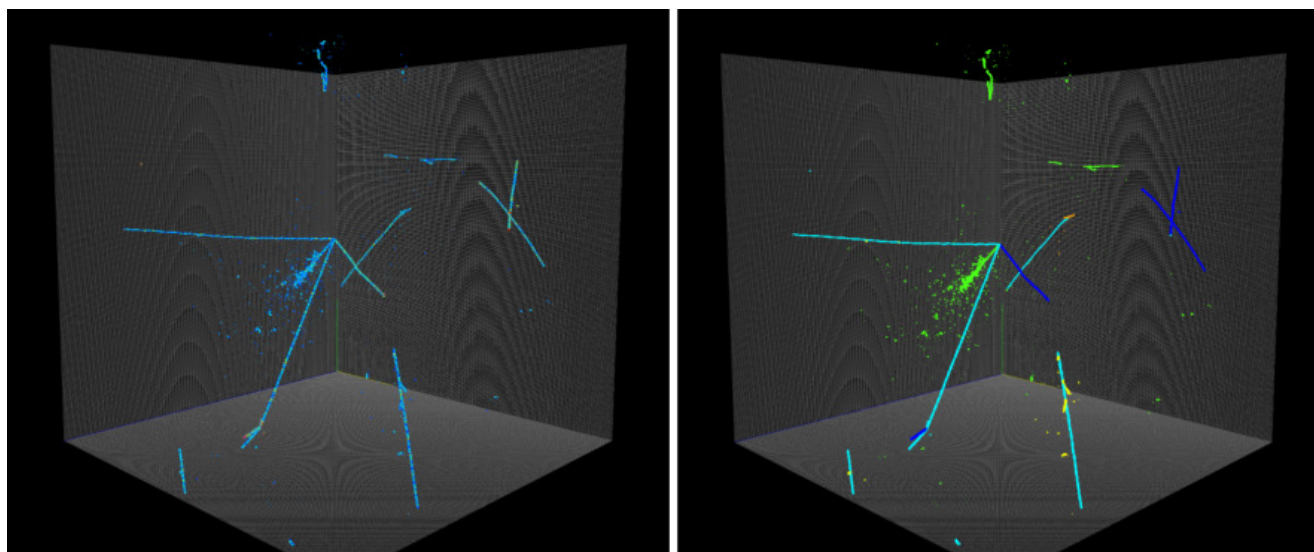


Рис. 1. Змодельована подія LArTPC. **Зліва** — розбита по кубічних комітках поглинута зрідженим аргонем енергія заряджених частинок. Кольором позначена енергетична шкала. **Справа** — розмічене зображення, на якому окремому пікселю відповідає 1 із 5 кольорів для кожного з 5 класів [10]

Fig. 1. Simulated LArTPC event. **Left**: energy deposition recorded per cubical cells. The color represents the energy scale. **Right**: pixels are categorized into 5 colors which represent 5 particle types [10]

Згенеровані дані призначені для задачі машинного навчання з “учителем” (англ. Supervised machine learning [17]). Зокрема, для навчання задачі 3D-семантичної сегментації LArTPC-зображень у P1LArNET усім ненульовим пікселям кожного зображення поставлено у відповідність мітку, що відповідає одному з п'яти класів:

1. Протони: переважно короткі траєкторії — клас високо йонізуювальних частинок (ВЙЧ).
2. Мінімально йонізуювальні частинки (МЙЧ), як-

от мюони та піони, з переважно довгими треками.

3. Електромагнітні зливи (ЕМ-зливи), спричинені електронами, позитронами або фотонами, з кінетичною енергією, більшою за критичну (близько 45 MeV в аргоні [19]).
4. Дельта (δ)-електрони (e_{δ}^{-}).
5. Мішель-електрони (e_{M}^{-}), продукти розпаду мюонів.

У таблиці 1 подано загальну статистику кількості пікселів, що належать кожному класу. Загальна кількість ненульових пікселів в одному зображенні має близько 0.01%. У таблиці наведено класи, до яких ці ненульові пікселі належать. Приклад зображення даних P(L)ArNET подано на рис. 1.

	ВЙЧ	МЙЧ	ЕМ зливи	e_{δ}^-	e_M^-
Частка, %	17	34	47	1	1

Табл. 1. Відсоток кількості пікселів, що належать кожному класу від загальної кількості ненульових пікселів
Table 1. The ratio of the number of pixels in percentages belonging to each class to the total number of nonzero pixels

III. АРХІТЕКТУРА МОДЕЛІ МАШИННОГО НАВЧАННЯ SPARSE U-RESNET

A. Загальний вигляд архітектури U-ResNet

Архітектура U-Net [18] була розроблена для задачі сегментації медичних зображень і стала класичною архітектурою семантичної сегментації. Вона складає-

ться з двох частин: енкодер та декодер (рис. 2). Енкодер, приймаючи на вхід тензор даних, застосовує операції зменшення просторових розмірів, повертаючи в підсумку тензор менших просторових розмірів та більшої кількості каналів, у яких закодовано семантичну інформацію. Кількість операцій зменшення просторових розмірів є гіперпараметром нейронної мережі, її прийнято називати *глибиною*. Декодер застосовує до отриманого тензора операції збільшення просторових розмірів, повертаючи тензор початкових просторових розмірів та з кількістю каналів, рівною кількості класів.

Кількість операцій зменшення просторових розмірів дорівнює кількості операцій збільшення просторових розмірів, тому кожному шару певної просторової розмірності на шляху енкодера відповідає шар цієї ж просторової розмірності на шляху декодера. Уздовж шляху енкодера втрачається багато просторової інформації. Це спричиняє операція зменшення просторових розмірів, що об'єднує декілька сусідніх пікселів, формуючи внаслідок тензор меншої просторової роздільної здатності. Для відновлення просторової інформації кожен шар енкодера подається до відповідного шару декодера через скорочувальні зв'язки (skip connections), а два тензори об'єднуються вздовж осі каналів.

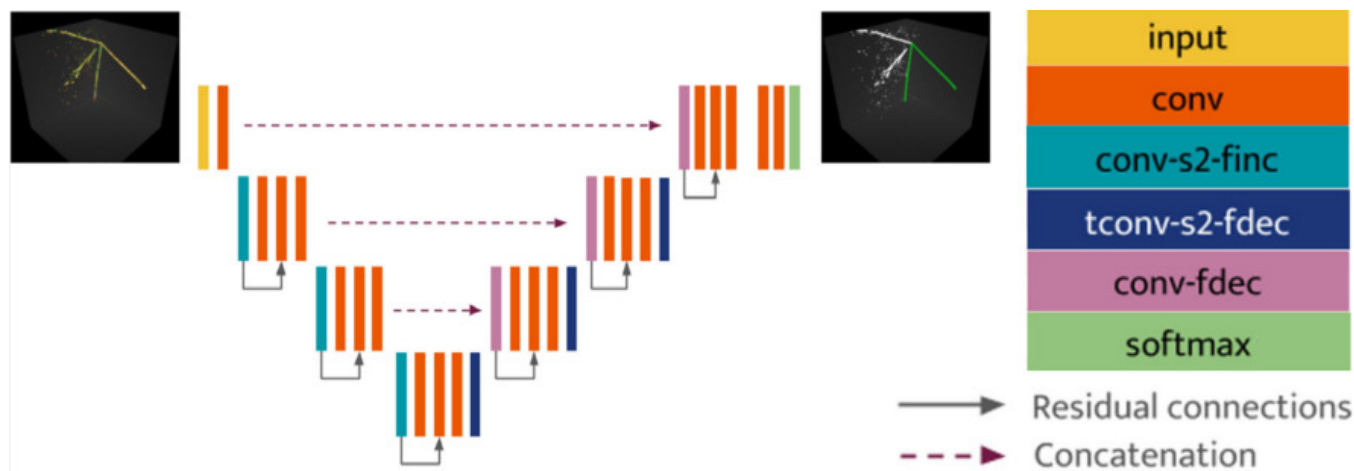


Рис. 2. Архітектура U-ResNet для семантичної сегментації. У цьому прикладі U-ResNet має глибину 3, оскільки виконуються 3 операції зменшення просторових розмірів. Бірюзові поля репрезентують згортки з кроком 2 зі збільшенням кількості фільтрів. Темно-сині поля — це транспоновані згортки з кроком 2 та зменшенням кількості фільтрів. Фіолетові — згортки з кроком 1, які зменшують кількість фільтрів. Червоний — 3×3 згортка, що не змінює кількості фільтрів та просторових розмірів [10]

Fig. 2. U-ResNet architecture for semantic segmentation. In this example, the U-ResNet has a depth of 3, since three spatial dimensionality reduction operations are performed. Turquoise boxes represent convolutions with stride 2 and increasing the number of filters. Dark blue boxes are transposed convolutions with stride 2 and decreasing the number of filters. Purple boxes are convolutions with stride 1 that decrease the number of filters. Red box is a 3×3 convolution, which does not change the number of filters or spatial dimensions [10]

Схемою на рис. 2 зображено архітектуру U-ResNet, що відрізняється від описаної вище U-Net лише наявністю додаткових залишкових зв'язків (в англ. resi-

dual connections, додаткові зв'язки між різними шарами, що пропускають деякі проміжні шари архітектури), які дають змогу архітектурі навчатися швидше

та мати більшу глибину.

Ця архітектура була використана для 3D-семантичної сегментації даних PILArNET [10].

Б. Просторово розріджені згорткові мережі

Для особливого випадку розрідженості даних, ефективна розмірність яких менша за простір, де містяться ці дані (приклад: 1D-криві, або 2D-поверхні у 3D-просторі), розроблено Submanifold Sparse Convolutional Networks¹ (SSCN, далі SSC-мережі) [8, 9]. Ключовим їхнім елементом є операція розрідженої згортки на підмноговиді (submanifold sparse convolution).



Рис. 3. Приклад дилатації кривої після застосування згортки. **Зліва** — початкова крива. **Посередині** — результат застосування 3×3 згортки з вагами, рівними $1/9$. **Справа** — результат застосування 3×3 згортки з вагами, рівними $1/9$ до шару посередині [12]

Fig. 3. Example of curve dilation after applying convolutions. **Left:** Original curve. **Middle:** Result of applying a regular 3×3 convolution with weights $1/9$. **Right:** Result of applying the same convolution again [12]



Рис. 4. Ділянка, до якої застосовано ядро SSC ($\bullet, \bullet, 3$), зображена 3×3 квадратом із червоними та зеленими пікселями. Активні елементи в ньому позначені зеленим кольором. Червоні ігноруються SSC, унаслідок чого крива зберігає свою структуру [9]

Fig. 4. SSC($\bullet, \bullet, 3$) receptive field represented as a box with red and green pixels. Active locations in the box are shown in green. Red locations are ignored by SSC so the pattern of active locations remains unchanged [9]

Традиційні операції згортки не доцільні для даних із указаною структурною особливістю з таких причин:

- Надмірна кількість арифметичних операцій, залучених під час виконання матричних або тензорних операцій, операнди яких містять переважно нульові елементи, призводить до непомірного споживання обчислювальних ресурсів.

- Швидке руйнування розрідженої структури даних у наступних шарах мережі. Один активний елемент у попередньому шарі активує f^d елементів у наступному (рис. 3), де f^d — кількість елементів у ядрі згортки, f — розмір ядра, d — розмірність. Цю проблему автори роботи [9] називають проблемою дилатації (dilation — розширення).

Ідея SSC-мереж полягає в тому, щоб зберігати рівень розрідженості даних у наступних шарах згорткової мережі (рис. 4). У публікації [9] розглянуто дві операції. Перша — розріджена згортка (sparse convolution) $SC(n, m, f, s)$, де n — кількість вхідних каналів, m — кількість вихідних каналів, f — розміри ядра та s — крок (stride). Головна відмінність SC від стандартної операції згортки в тому, що обчислювальна складність пропорційна до кількості активних елементів у вхідному шарі, а не до загальної кількості елементів. Усі операції виконуються лише з ненульовими елементами, потребуючи меншого споживання обчислювальних ресурсів. Друга — розріджена згортка на підмноговиді (submanifold sparse convolution) $SSC(n, m, f)$, яка, окрім обчислювальної переваги, що закладена в $SC(n, m, f, s)$, зберігає ще й розрідженість у вихідному тензорі (рис. 4).

Сучасною архітектурою sparse-мережі є Sparse Tensor Networks [7] (далі STN або ST-мережі), у якій узагальнено всі операції над розрідженими тензорами, що використовують під час побудови CNN і можуть бути виконані над даними довільної структури. Реалізацією Sparse Tensor Networks є pytorch-бібліотека Minkowski Engine².

Ця архітектура є більш гнучкою та документованою, ніж SSC-мережі, завдяки оптимізації вимагає менше обчислювальних ресурсів. Так, автори роботи з реконструювання подій нейтрино [6] та робіт [12–14] замінили SSCN на Minkowski Engine. Також Minkowski Engine використано для аналізу даних з ProtoDUNE [21].

Визначена операція згортки над розрідженими тензорами Minkowski Engine математично збігається з розрідженою згорткою на підмноговиді, властивості якої обговорені вище [7].

IV. ПРОВЕДЕННЯ ВИМІРЮВАНЬ

А. Деталі експерименту

Задача семантичної сегментації була виконана на 3D-даних PILArNET розміром 192 пікселів та 768 пікселів. Менші зображення були використані для порівняння різних архітектур семантичної сегментації (розділ IV B), а більші — для тестування однієї архітектури, обраної за результатами попереднього аналізу, на

¹ pytorch реалізація SparseConvNet: <https://github.com/facebookresearch/SparseConvNet/tree/main?tab=readme-ov-file>.

² <https://nvidia.github.io/MinkowskiEngine/overview.html>.

аугментованих даних (розділ IV Г). Файли були розділені в такому співвідношенні, щоб об'єм тренувальної вибірки становив 80 000 подій, а валідаційної — 20 000. Спосіб розподілення вибірки зі 100 000 подій по 80 000 і 20 000 значення не має, оскільки дані рандомізовані на етапі моделювання. Тренувальна вибірка складається з перших 8 файлів 192-піксельних даних і перших 16 файлів 768-піксельних даних. Валідаційна — останні 2 та 4 файли відповідно 192- та 768-піксельних даних.

Загальна будова архітектури sparse U-ResNet зображена на рис. 2. Робота виконана на основі реалізацій sparse U-ResNet із бібліотекою SparseConvNet³ та бібліотекою MinkowskiEngine⁴.

Відмінність реалізації розрідженої архітектури від класичної [18] у тому, що останній шар, до якого застосовується активація softmax, не згортковий, а лінійний⁵, і, що є суттєвішим, кількість фільтрів зростає після кожного застосування операції зменшення просторових розмірів за лінійним законом $n_i = n_0 \cdot i$, а не в геометричній прогресії.

Операція зменшення просторових розмірів реалізована згорткою з ядром $2 \times 2 \times 2$ та кроком $2 \times 2 \times 2$. У межах одного блоку згорткових шарів розміри ядра становлять $3 \times 3 \times 3$, а кроку — $1 \times 1 \times 1$.

Тренування виконувалося за оптимізаційним алгоритмом ADAM з темпом навчання 0.001 та розміром міні-пакета (mini-batch size) 64 у випадку тренування на 192-піксельних зображеннях та 32 у випадку 762-ох.

Б. Метрики

Критеріями точності нейронної мережі, аналогічно до наведених у публікації [10], є такі дві метрики:

- Nonzero assuarcy (далі просто точність, або *загальна* точність): загальна частка ненульових пікселів, клас яких був визначений правильно;
- Classwise nonzero assuarcy (далі точність класифікації пікселів *певного класу*, або *покласова* точність): частка ненульових правильно класифікованих пікселів, що належать до певного класу.

В. Тестування архітектур семантичної сегментації

Ми виконали семантичну сегментацію 192-піксельних зображень P1LArNET з використанням

³ Реалізація зі SparseConvNet: https://github.com/Temigo/ur_esnet_pytorch.

⁴ Реалізація із MinkowskiEngine: https://github.com/DeepLearnPhysics/lartpc_mlreco3d.

⁵ <https://pytorch.org/docs/stable/generated/torch.nn.Linear.html>.

розріджених U-Net-подібних архітектур із глибиною $depth=6$ та кількістю фільтрів у першому та останньому шарах 16 та 96, відповідно.

На рис. 5 зображено залежність функції втрат, загальної точності класифікації та точності визначення мішель-електронів на валідаційній вибірці від тривалості навчання в епохах.

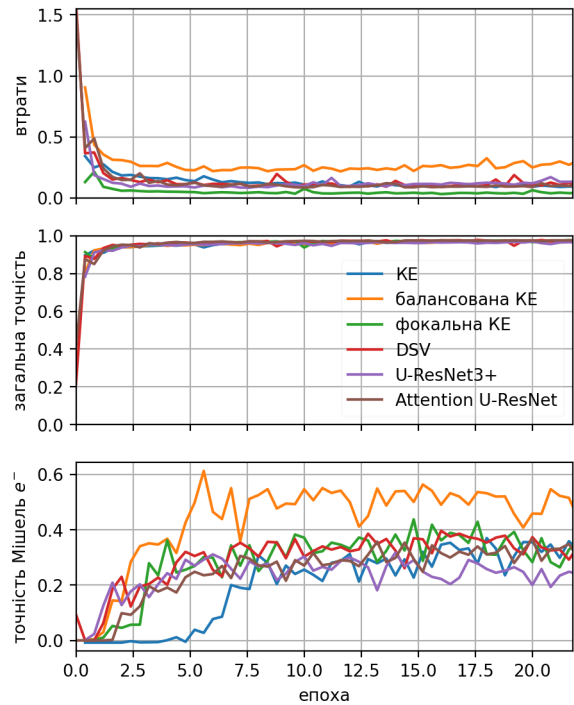


Рис. 5. **Перший** графік — функція втрат на валідаційній вибірці залежно від кількості епох тренування. **Другий** — загальна точність. **Третій** — точність визначення мішель-електронів. KE — U-ResNet із крос-ентропією. DSV — U-ResNet із повним глибоким навчанням. Фокальна KE — U-ResNet із фокальною функцією втрат, $\gamma = 2$. Fig. 5. **First** plot: loss on the validation set depending on the number of training epochs. **Second**: overall accuracy. **Third**: classification accuracy of Michel-electrons. KE — U-ResNet with cross-entropy. DSV — UResNet with deep supervision. Focal KE — U ResNet with focal loss, $\gamma = 2$

Референтну архітектуру U-ResNet із функцією втрат, відомою в літературі з машинного навчання як cross entropy [22], позначено на графіку синьою лінією. Функція втрат крос-ентропія має вигляд $loss_{CE} = -\sum_{ti} \log(p_{ti})$, де $p_{ti} \in [0, 1]$ — ступінь упевненості нейронної мережі, або ймовірність, що i -й піксель належить класу t . Підсумовування проходить по всіх пікселях та класах.

Кривою помаранчевого кольору (*балансована KE*) на графіку позначено архітектуру U-ResNet із функцією втрат крос-ентропія з ваговими коефіцієнтами, в англ. balanced cross-entropy loss [22], яка має такий вигляд: $loss_{balanced CE} = -\sum_{ti} w_t \log(p_{ti})$, де w_t — статистична вага цього класу t . Функція втрат із ваговими коефіцієнтами частково усуває статистичний дисбаланс між мішель-електронами та іншими кла-

сами. У цій роботі прийнято $w_t = \frac{1}{n_t}$, де n_t — кількість пікселів, що належать класу t . Як видно з графіка, статистичні вагові коефіцієнти продемонстрували найкращу точність ідентифікації мішель-електронів.

Зеленою кривою позначено залежність, що відповідає тесту U-ResNet із фокальною крос-ентропією (англ. focal loss [22]), яка має вигляд $\text{loss}_{\text{focal}} = -\sum_{ti}(1 - p_t)^\gamma \log(p_{ti})$, де γ — фокусувальний параметр.

Фокальна функція втрат під час тренування дає змогу змінювати значення вагових коефіцієнтів, так що клас t із високим ступенем упевненості передбачення $p_t \rightarrow 1$ має ваговий коефіцієнт, що наближається до нуля $w_t \rightarrow 0$, де $w_t = (1 - p_t)^\gamma$. Зміст динамічних

статистичних вагових коефіцієнтів полягає в тому, що під час тренування модель фокусується саме на важких для точного передбачення прикладах, тобто на тих класах, що класифікуються з меншим ступенем упевненості, для яких p_t має менше значення та, відповідно, більший ваговий коефіцієнт.

Також у процесі роботи реалізовано та протестовано такі архітектури:

U-ResNet3+. Архітектуру U-ResNet3+ реалізовано через заміну в архітектурі U-ResNet декодера (рис. 6). Ця архітектура побудована на основі архітектури UNet 3+ [23] додаванням додаткових скорочувальних зв'язків (res. — residual).

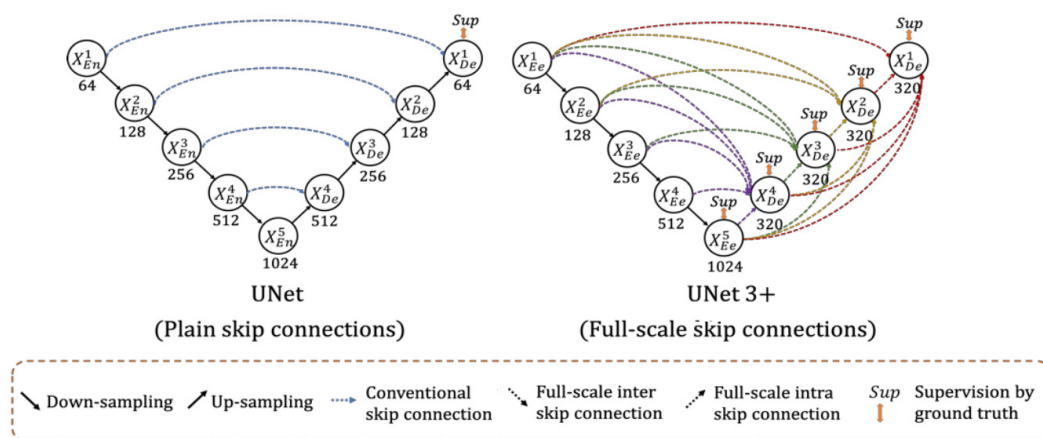


Рис. 6. Схема UNet 3+ [23]. З рисунка видалено частини, що не стосуються UNet або UNet 3+
Fig. 6. U-Net3+ architecture [23]. The part not related to the U-Net or U-Net3+ figure has been removed

Модель	Функція втрат	Загальна, %	ВЙЧ, %	МЙЧ, %	ЕМ зливи, %	e_{δ}^{-} , %	e_{M}^{-} , %
UResNet	Крос-ентропія	97.3	93	95.9	96.4	70	39.9
UResNet	Балансована KE	95	94.4	89.9	95.1	78.3	61.2
UResNet	Фокальна KE	96.5	88.3	96.3	95.4	69.1	43.7
UResNet	DSV	97.5	94.1	95.2	96.8	67.1	39.6
UResNet3+	Крос-ентропія	96.2	91	94.4	96.1	64.7	31.7
Attention UResNet	Крос-ентропія	97.6	96.4	96	96.2	69.6	44

Табл. 2. Загальна та покласова точності класифікації на валідаційній вибірці всіх архітектур,%. KE — крос-ентропія
Table 2. Overall and class-wise classification accuracies of all architectures on the validation sample, %. KE — cross-entropy

Кожен блок декодера в стандартній архітектурі U-Net приймає на вхід два тензори: перший із попереднього блоку декодера, другий із відповідного за розмірами блоку енодера. Натомість в архітектурі UNet3+ кожен блок декодера приймає на вхід кількість тензорів, що дорівнює глибині (кількості блоків енодера) архітектури. Результат роботи цієї архітектури зображено на рис. 5 фіолетовою кривою.

Full-scale Deep Supervision. Автори U-Net3+ [23] продемонстрували ефективність навчання на промі-

жних шарах декодера вздовж усього його тракту, тобто навчання відбувається на всіх згорткових блоках декодера, починаючи від найнижчої просторової розмірності до розмірності вхідного зображення, а не лише на вихідному шарі декодера. У статті навчання таким способом називається повномасштабним глибоким навчанням (Full-scale Deep Supervision).

Full-scale Deep Supervision реалізовано з архітектурою U-ResNet. Кожен блок декодера додатково подається на операцію збільшення просторових розмі-

рів, результатом якої є тензор початкової просторової розмірності, тобто $192 \times 192 \times 192$. Отриманий тензор подається на вхід до операції згортки з 5-фільтрами розмірами $3 \times 3 \times 3$. Кількість фільтрів відповідає кількості класів. Для отриманого тензора з розмірами $192 \times 192 \times 192 \times 5$ визначаються втрати за допомогою крос-ентропії. У підсумку крос-ентропію рахуємо по кожному з шарів декодера $\text{loss}_{DSV} = \sum_{i=1}^{\text{depth}} \text{CE}_i$, де i – номер шару декодера, CE – функція втрат крос-ентропії. Результати застосування повномасштабного глибокого навчання до класичної архітектури U-ResNet позначено червоною кривою на рис. 5, де можна бачити незначене поліпшення результатів порівняно з референтною архітектурою.

Attention U-ResNet. У публікації [24] репрезентовано модифікацію архітектури U-Net, відому як Attention U-Net. В оригінальній архітектурі кожен блок декодера приймає два тензори: перший — тензор даних з відповідного блоку енкодера, а другий — тензор з попереднього блоку декодера. Ці два тензори з'єднуються вздовж осі каналів. Однак, на відміну від U-Net, в Attention U-Net перший тензор множиться на вагові коефіцієнти, що дає змогу надавати різним просторовим координатам (або пікселям) першого тензора різні вагові коефіцієнти. Вони є тренуваними параметрами й обчислюються в блоці, відомому як attention gate. Attention gate приймає тензор енкодера та тензор декодера на вході й повертає тензор енкодера, помножений на вагові коефіцієнти. Результати застосування архітектури attention U-ResNet позначені кривою коричневого кольору на рис. 5.

Порівняння метрик різних моделей. У таблиці 2 наведено загальні та покласові точності для всіх протестованих моделей. Для кожної з них ми подали значення метрик, що відповідають найвищій точності класифікації мішель-електронів. Згідно з наведеними значеннями, найвищої точності класифікації досягнуто за використання балансованої крос-ентропії (61.2%).

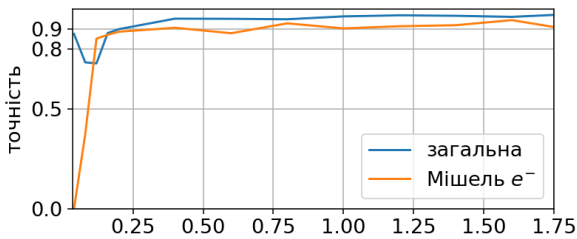


Рис. 7. Точність класифікації $768 \times 768 \times 768$ розмірних даних

Fig. 7. Classification accuracy of $768 \times 768 \times 768$ pixel data

Г. Тестування архітектури U-ResNet на аугментованих даних

Ми застосували архітектуру U-ResNet із балансованою крос-ентропією до семантичної сегментації 768-піксельних зображень. Залежності загальної точності

та точності класифікації мішель-електронів від епохи тренування показані на рис. 7. Використання 768-піксельних зображень замість 192-х є еквівалентним до 64-кратної аугментації даних.

Д. Порівняння отриманих результатів із результатами попередніх авторів

Автори публікації [10] провели низку схожих експериментів, а саме: семантичну сегментацію даних PILLArNET за допомогою розрідженої U-ResNet. У цій публікації подано отримані метрики для стандартної крос-ентропії, проте не наведено для балансованої. У таблиці 3 порівняно результати балансованої крос-ентропії, отримані в нашій роботі, із результатами попередньої праці, де наведені метрики для незбалансованої крос-ентропії.

Результати	$192 \times 192 \times 192$	$768 \times 768 \times 768$
Ця робота	61.2%*	94.5%
Публікація [10]	36.5%	70.4%

*Після тренування архітектури на 768^3 -піксельних даних ми одержали точність 76% (див. розділ V Г).

Табл. 3. Порівняння точності класифікації мішель-електронів

Table 3. Comparison of Michel-electron classification accuracy

V. АНАЛІЗ РЕЗУЛЬТАТІВ

А. Порівняння архітектур семантичної сегментації та функцій втрат

Унаслідок експерименту на даних з $192 \times 192 \times 192$ просторовою розмірністю всі модифікації функції втрат, а саме: фокальна крос-ентропія, крос-ентропія з статистичними ваговими коефіцієнтами і full-scale deep supervision — поліпшували точність ідентифікації мішель-електронів.

Окрім того, виявилось, що всі модифікації прискорюють процес навчання, тобто вимагають меншої кількості ітерацій для досягнення точності класифікації мішель-електронів у 30%. Передусім це стосується функції втрат із ваговими коефіцієнтами, максимальна точність якої досягнула за 7000 ітерацій, або 5.6 епох.

Архітектура U-ResNet3+ не дала бажаного ефекту, проте вона навчалася за меншу кількість ітерацій. За 7500 ітерацій (6 епох) досягнуто 30-відсоткової точності з виходом на плато. Водночас вона мала на 33% меншу кількість параметрів⁶.

⁶ Кількість параметрів U-ResNet — 4 738 405. Кількість параметрів UResNet3+ — 3 177 105.

Використання архітектури Attention U-ResNet дало змогу помітно поліпшити результат (44%), хоча дані PILDaNET не мають виділених просторових ділянок, яким потрібно було би надати вагових коефіцієнтів. На нашу думку, поліпшення результатів пов'язане з тим, що attention gate спричиняє появу додаткової кількості нових параметрів, які покращили результати валідаційної вибірки. Окрім того, наведені в таблиці 2 значення досягнуті за 76 500 (61.2 епохи) ітерацій, що відповідає досить тривалому навчанню.

Порівняймо ефективність роботи архітектури U-ResNet із статистичними ваговими коефіцієнтами та без них за допомогою матриць похибок, поданих на рис. 8. З обох матриць видно, що мішель-електрони найчастіше сплутуються з електромагнітними зливами. Водночас зазначимо, що більшу частину пікселів, що належать до треків мішель-електронів, нейронна мережа з небалансованою функцією втрат класифікує як електромагнітні зливи, правильно ідентифікуючи лише 40% пікселів. За балансованої крос-ентропії статистика правильних прогнозів краща — до 60% відсотків “угадувань”.

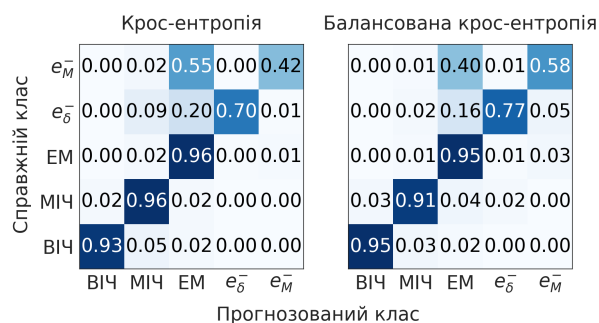


Рис. 8. Матриці похибок U-ResNet, отриманих на 192-піксельних зображеннях. **Зліва** — небалансована крос-ентропія. **Справа** — крос-ентропія із статистичними ваговими коефіцієнтами

Fig. 8. U-ResNet confusion matrices inferred on 192-pixel images. **Left** — unbalanced cross-entropy. **Right** — cross-entropy with statistical weights

Б. Оптимальність вибору архітектури та ступеня її навченості

Оскільки значення точності класифікації на даних з $192 \times 192 \times 192$ просторовою розмірністю в межах 55–60% (рис. 6), що близько до 50% випадкового прогнозу з імовірністю 1/2, архітектура U-ResNet в експериментах на $192 \times 192 \times 192$ даних із балансованою крос-ентропією майже не відрізняє мішель-електронів від електромагнітних злив.

Можливими поясненнями щодо недостатньої точності класифікації є, наприклад, недостатня тривалість навчання, недостатній об'єм тренувальної вибірки або ж невдалий вибір архітектури.

Щодо відповідності до цієї задачі архітектури U-ResNet, то в розділі IVГ ми навели істотно ліпші метрики (точність визначення мішель-електронів перевищувала 90%), отримані на даних більшого розміру із цією самою архітектурою, тому вона є принципово здатною до розрізнення EM-злив та мішель-електронів.

Також ми вважаємо, що протестували модель з оптимальним рівнем навченості. На рис. 9 на *першому* графіку *пурпуровою лінією* показано зростання точності класифікації мішель-електронів під час тренування. Зауважимо, що осциляції точності на тренувальній вибірці (*пурпурова крива*) наявні через навчання міні-пакетним (нагадаємо, що розмір пакета — 64 зображення). Точність класифікації на тренувальній та валідаційній вибірках у середньому збільшуються до 7000 ітерацій. Проте після 7000 точність на валідаційній поступово зменшується. Щоб визначити, чи досягнуто достатнього рівня навченості моделі, потрібно також розглянути поведінку функції втрат (*другий* графік) на тренувальній (*пурпурова крива*) та валідаційній (*сині відмітки*) вибірках. Мінімуму функція втрат досягає за 5000 ітерацій, та не збільшується до 20000 ітерацій.

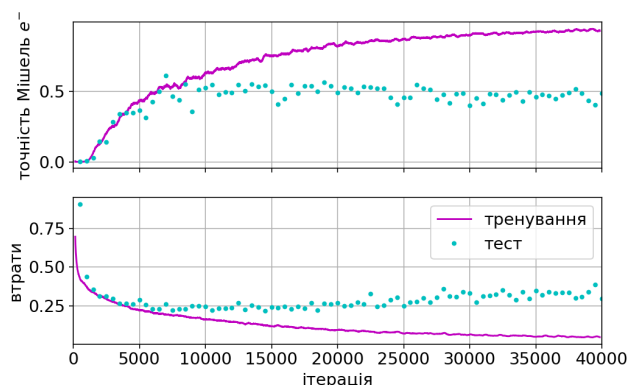


Рис. 9. **Перший графік** — точність класифікації мішель-електронів на тренувальній вибірці в процесі навчання (пурпурна лінія) та на валідаційній вибірці (сині точкові позначки). **Другий** — функція втрат на тренувальній (пурпурна лінія) та валідаційній (сині точкові позначки) вибірках

Fig. 9. **First plot**: accuracy of Michel-electrons classification during training on the training (magenta line) and on the validation (blue dots marks) sets. **Second**: loss function on the training (magenta line) and validation (blue dot marks) sets

В. Доцільність аугментації даних

Тож найліпше пояснення низької точності на валідаційній вибірці $192 \times 192 \times 192$ даних серед висунутих у попередньому розділі — недостатній об'єм тренувальної вибірки. Ми вважаємо недоцільним збільшення кількості параметрів архітектури (як збільшення

глибини архітектури, так і кількості фільтрів), оскільки ускладнення архітектури приведе лише до пере-запам'ятовування даних. Типова поведінка точності на валідаційній та тренувальній вибірках продемон-стрована на рис. 9. Точність на тренувальній вибірці продовжує збільшуватися, тоді як на валідаційній пі-сля досягнення максимуму зменшується або виходить на плато. За балансової крос-ентропії функція втрат на валідаційній вибірці зростає після досягнення гло-бального мінімуму (рис. 9).

Ми не виключаємо можливостей незначних покращень унаслідок застосування методів регуляризації, проте згідно з результатами, наведеними в розділі IV Г і наступному, справді ефективною виявилася аугментация даних завдяки використанню даних у 64 рази більшої розмірності.

Зауважимо також, що подібна аугментация була би неможливою в разі застосування конвенціональних тензорів та згорток. Для традиційних ЗНН n -кратне збільшення просторових розмірів вхідного тензора даних повинне супроводжуватися n^D -кратним збільшенням використання обчислювальних ресурсів, де D — розмірність простору, у якому дані розташовані. Якщо збільшення з 192-піксельних 3D-зображень до 768, то маємо 64-кратне збільшення об'єму. Натомість, оскільки LAtTPC-дані — переважно одновимірні треки, збільшення простого розміру вхідного зображення повинно результувати не в степеневу зростанні обчислювальних ресурсів ($\propto n^D$), а в лінійному ($\propto n$). Ми спостерігали до 6-кратного збільшення в споживанні GPU-пам'яті (час операції forward).

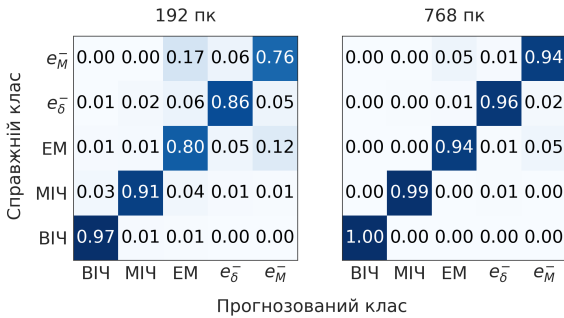


Рис. 10. Матриці похибок U-ResNet із балансованою KE, тренуваною на 768-піксельних даних. **Зліва** — тест на 192-піксельних даних. **Справа** — 768

Fig. 10. U-ResNet architecture with balanced CE confusion matrices, trained on 768-pixel data. **Left** — inferred on 192-pixel data. **Right** — 768

Г. Результати тренування на аугментованих даних

Ми провели тренування архітектури, тренування розрідженої архітектури U-ResNet із балансованою

крос-ентропією (деталі тренування див. у розд. IV А) на 768³-вибірці та протестували її на двох валідаційних вибірках: з 192³-піксельних та 768³-піксельних даних (рис. 10).

Унаслідок тестування архітектури на 768³-розмірних даних (матриця на рис. 10), по-перше, усі покласові точності, одержані в аналізі валідаційної вибірки, мають високі значення (від 94%) і, по-друге, наведена матриця похибок, узята лише для двох класів: електромагнітних злив та мішель-електронів, — виявляється симетричною: мішель-електрони та електромагнітні зливи сплутуються між собою з однаковою частотою. Унаслідок тестування на 192³-розмірних даних одержано найбільшу точність класифікації мішель-електронів на 192³-вибірці (76%).

Д. Можливі причини втрати точності

1. Пов'язані з фізикою мішель-електронів

Із наведених на рис. 10 матриць похибок видно, що архітектура продовжує сплутувати ЕМ-зливи та мішель-електрони. Ми пояснюємо таку поведінку нейронної мережі фізичними властивостями мішель-електронів.

Мішель-електрони є продуктами розпаду мюонів, що майже втратили свою енергію на йонізацію середовища: $\mu \rightarrow e + \nu_e + \nu_{\mu}$.

Кінематика розпаду приводить до характерного спектра [26] з енергетичним порогом у 53 MeV, який дорівнює пів маси мюона. Однак особливості взаємодії мюонів з атомами спричиняють спотворення спектра так, що максимум зміщується в ділянку нижчих енергій, а хвіст розподілу розтягується до маси мюона. Можна наближено вважати, що енергія мішель-електронів є в межах в 0–60 MeV.

Електрони, рухаючись у середовищі, втрачають енергію, або йонізуючи та збуджуючи атоми, з яких складається матеріал, або випромінюючи радіаційні фотони. Електрони, що мають енергію вищу за критичну, переважно втрачають її через радіаційні втрати, а не через йонізаційні. Критична енергія електрона в зрідженому аргоні дорівнює 45 MeV⁷. За цієї енергії радіаційні втрати дорівнюють йонізаційним. Утрати енергії внаслідок йонізації в часо-проекційній камері відбуваються безперервно, утворюючи неперервні трекові сегменти. Фотони, випромінені електроном, не будуть одразу віддавати енергію в детектувальний об'єм, а, ймовірніше, рухатимуться без взаємодії, доки не відбудеться комптонівське розсіювання або утвориться електрон-позитронна пара. Значні відстані, що можуть пройти фотони без взаємодії, спричиняють йонізацію атомів аргону за межами початкової траєкторії мішель-електрона (рис. 11).

⁷ Дані взято з бази даних NIST ESTAR <http://physics.nist.gov/Star>.

2. Пов'язані із особливістю моделювання даних

Отже, мішель-електрони утворюють сліди, із характерною для електромагнітних злив топологією. Власне це ми вважаємо поясненням матриці хиб, зображеної на рис. 10 справа. Однак залишається питання, чому одна й та сама архітектура дає високі значення метрик на 768³-вибірці та нижчі на 192³-вибірці (рис. 10).

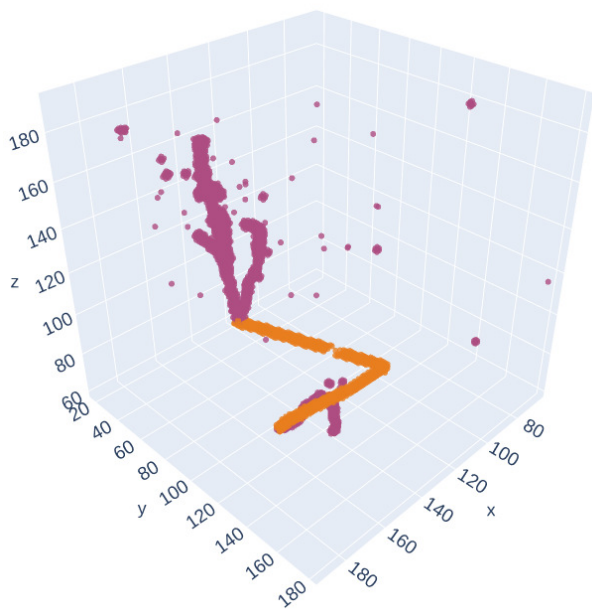


Рис. 11. Приклад події з тестової вибірки 192³-вимірних даних, де мішель-електрон утворює електромагнітну зливу. Оранжевим кольором позначені пікселі, що належать мюонам, пурпуровим — мішель-електронам

Fig. 11. Example of an event from the test sample of 192³-pixel data. Michel-electron forms an electromagnetic shower. The pixels belonging to the muons track are marked in orange, and the pixels belonging to the Michel-electrons track are marked in magenta

Ми вважаємо, що недостатня точність на 192-піксельній валідаційній вибірці пов'язана з особливостями моделювання даних. Як згадувалося в розділі II, під час моделювання ініціалізуються первинні частинки та їхня початкова кінематика. Особливість моделювання полягає в тому, що вершини частинок можуть виходити за межі ділянки простору, що вокселізується. Наприклад, для ідентифікації мішель-електрона необхідно визначити точку, де зупинився мюон та з якої почав рух дочірній мішель-електрон. Якщо ж у деякому зображенні події поряд із слідом, що міг бути утворений мішель-електроном, мюона не буде, відрізнити клас мішель-електрона від електромагнітної зливи, імовірно, не вдасться. А саме такі зображення є у валідаційній вибірці 192³-піксельних даних. На рис. 12 подано дві події, де вершини частинок не увійшли в об'єм, і тому відрізнити мішель-електрон від ЕМ-зливи неможливо.

Отже, метрики, переходячи з аналізу 768³-піксель-

них зображень до аналізу 192³-піксельних даних, повинні погіршуватися, оскільки в останніх має бути більше “обрізаних” траєкторій частинок. Водночас “обрізані” треки мішель-електронів очікувано ідентифікуються, як електромагнітні зливи (лівий графік на рис. 10).

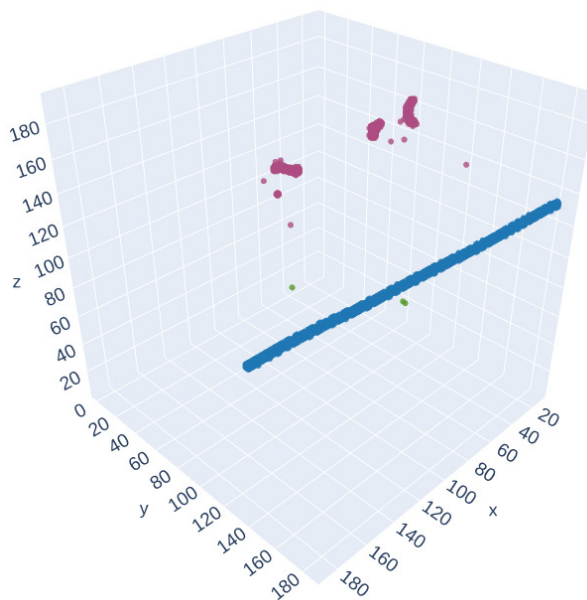


Рис. 12. Приклад події з тестової вибірки 192³-вимірних даних, де немає мюона, проте є мішель-електрон. Синім кольором позначені пікселі, що належать треку протона, пурпуровим — мішель-електронам. Оскільки в об'ємі немає мюона, мішель-електрон складно відрізнити від ЕМ-зливи

Fig. 12. Example of an event from the test sample of 192³-pixel data. There is no muons track, but there is a Michel-electron. The blue color marks pixels belonging to the proton track, and the purple marks pixels belonging to Michel electrons. Since there is no muon in the volume, it is difficult to distinguish a Michel-electron from an EM shower

VI. ВИСНОВКИ

Унаслідок проведеної роботи перевірено нові архітектури семантичної сегментації на даних PILLArNET, які не згадувалися в попередніх публікаціях [6, 10, 12–14], та нові функції втрат. Визначено, що через статистичний дисбаланс класів, архітектура U-ResNet із крос-ентропією не здатна відрізнити мішель-електронів від електромагнітних злив, тож було застосовано балансовану крос-ентропію для усунення статистичного дисбалансу.

Автори попередньої публікації [10] підкреслювали вплив статистичного дисбалансу класів та неоднозначності між способами визначення класів електромагнітних злив, дельта-електронів та мішель-електронів. Ми в розділі V звернули увагу, що відрізнити мішель-електрони від електромагнітних злив важко саме тоді, коли траєкторії частинок значно виходять за межі об'єму запису подій. Ця гіпотеза не-

прямо підтверджується тим, що нам вдалося отримати значно точніші результати на 768^3 -піксельних даних. Ї можна дослідити докладніше, модифікувавши дані P1LArNET так, щоб виключити з аналізу ті “обрізані” треки мішель-електронів, що не надають достатньої інформації для ідентифікації.

У підсумку ми отримали кращі метрики, що дають змогу точніше класифікувати треки. Проаналізувавши можливі причини втрат точності в 192-піксельних зображеннях у розділі V, ми дійшли висновку, що архітектура U-ResNet із балансованою крос-ентропією дозволяє ефективно виконувати складні задачі з розрізнення мішель-електронів та електромагнітних злив. Ми вважаємо, що в цій задачі потрібно приділити особливу увагу

1. наявності частини треку мюона, де він зупиняється, та вершини мішель-електрона в зображенні,
2. достатній кількості даних для тренування.

Точне вимірювання електронів від нейтринних взаємодій необхідне для майбутніх дослідницьких програм, що вивчають нестандартні нейтринні осциляції, порушення CP лептонного сектора та ієрархію мас нейтрино [27]. Крім того, очікується, що нейтрино, які утворюються в спалахах галактичних наднових, можуть породжувати електрони в діапазоні енергій 5-50 MeV, що робить вивчення електронів низьких енергій критично важливим для розуміння цих астрофізичних подій [27].

VII. ПОДЯКА

Робота виконана за підтримки Міністерства освіти і науки України: Грант Міністерства освіти і науки України для перспективного розвитку наукового напрямку: “Математичні та природничі науки” в Київському національному університеті імені Тараса Шевченка.

-
- [1] P. Benetti *et al.*, Nucl. Instrum. Methods Phys. Res. A **332**, 395 (1993); [https://doi.org/10.1016/0168-9002\(93\)90297-U](https://doi.org/10.1016/0168-9002(93)90297-U).
 - [2] R. Acciarri *et al.* (MicroBooNE Collaboration), J. Instrum. **12**, P02017 (2017); <https://doi.org/10.1088/1748-0221/12/02/P02017>.
 - [3] M. Antonello *et al.*, preprint arXiv:1503.01520 (2015); <https://doi.org/10.48550/arXiv.1503.01520>.
 - [4] S. Amerio *et al.* (ICARUS Collaboration), Nucl. Instrum. Methods Phys. Res. A **527**, 329 (2004); <https://doi.org/10.1016/j.nima.2004.02.044>.
 - [5] R. Acciarri *et al.*, preprint arXiv:1601.02984 (2016); <https://doi.org/10.48550/arXiv.1601.02984>.
 - [6] F. Drielsma, K. Terao, L. Dominé, D. H. Koh, preprint arXiv:2102.01033 (2021); <https://doi.org/10.48550/arXiv.2102.01033>.
 - [7] C. Choy, JY. Gwak, S. Savarese, in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (Long Beach, CA, USA, 2019), p. 3070; <https://doi.org/10.1109/CVPR.2019.00319>.
 - [8] B. Graham, L. van der Maaten, preprint arXiv:1706.01307 (2017); <https://doi.org/10.48550/arXiv.1706.01307>.
 - [9] B. Graham, M. Engelcke, L. van der Maaten, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (IEEE, 2018)*, p. 9224; http://openaccess.thecvf.com/content_cvpr_2018/html/Graham_3D_Semantic_Segmentation_CVPR_2018_paper.html.
 - [10] L. Dominé, K. Terao (DeepLearnPhysics Collaboration), Phys. Rev. D **102**, 012005 (2020); <https://doi.org/10.1103/PhysRevD.102.012005>.
 - [11] C. Adams, K. Terao, T. Wongjirad, preprint arXiv:2006.01993 (2020); <https://doi.org/10.48550/arXiv.2006.01993>.
 - [12] D. H. Koh *et al.*, preprint arXiv:2007.03083 (2020); <https://doi.org/10.48550/arXiv.2007.03083>.
 - [13] F. Drielsma *et al.*, preprint arXiv:2007.01335 (2020); <https://doi.org/10.48550/arXiv.2007.01335>.
 - [14] L. Dominé *et al.*, Phys. Rev. D **104**, 032004 (2021); <https://doi.org/10.1103/PhysRevD.104.032004>.
 - [15] S. Agostinelli *et al.* (GEANT4), Nucl. Instrum. Meth. A **506**, 250 (2003); [https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
 - [16] E. L. Snider, G. Petrillo, J. Phys. Conf. Ser. **898**, 042057 (2017); <https://doi.org/10.1088/1742-6596/898/4/042057>.
 - [17] S. J. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. (Prentice Hall, 2010).
 - [18] O. Ronneberger, P. Fischer, T. Brox, preprint arXiv:1505.04597 (2015); <https://doi.org/10.48550/arXiv.1505.04597>.
 - [19] M. J. Berger, J. S. Coursey, M. A. Zucker, J. Chang, *ESTAR, PSTAR, and ASTAR: Computer Programs for Calculating Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions (version 1.2.3)* (National Institute of Standards and Technology, 2005); <https://doi.org/10.18434/T4NC7P>.
 - [20] R. Acciarri *et al.*, J. Instrum. **12**, P09014 (2017); <https://doi.org/10.1088/1748-0221/12/09/P09014>.
 - [21] Adam Abed Abud on behalf of the DUNE collab., J. Phys. Conf. Ser. **2438**, 012125 (2023); <https://doi.org/10.1088/1742-6596/2438/1/012125>.
 - [22] T.-Y. Lin, P. Goyal, R. Girshick, K. M. He, P. Dollár, preprint arXiv:1708.02002 (2017); <https://doi.org/10.48550/arXiv.1708.02002>.
 - [23] H. Huang *et al.*, preprint arXiv:1708.02002 (2020); <https://doi.org/10.48550/arXiv.1708.02002>.
 - [24] O. Oktay *et al.*, preprint arXiv:1804.03999 (2018); <https://doi.org/10.48550/arXiv.1804.03999>.
 - [25] L. Michel, Proc. Phys. Soc. A **63** 514 (1950); <https://doi.org/10.1088/0370-1298/63/5/311>; **63**, 1371 (1950); <https://doi.org/10.1088/0370-1298/63/12/512>.
 - [26] M. Bardon *et al.*, Phys. Rev. Lett. **14**, 449 (1965); <https://doi.org/10.1103/PhysRevLett.14.449>.
 - [27] R. Acciarri *et al.*, J. Instrum. **12**, P09014 (2017); <https://doi.org/10.1088/1748-0221/12/09/P09014>.

ANALYSIS OF LArTPC DATA USING MACHINE LEARNING METHODS

A. Falko, O. Gogota, R. Yermolenko, I. Kadenko
Taras Shevchenko National University of Kyiv, Ukraine
e-mails: falko@knu.ua, olga.gogota@knu.ua, ruslan.yermolenko@knu.ua, imkadenko@knu.ua

Deep Convolutional Neural Networks (CNNs) have exhibited remarkable efficacy in data analysis across various domains of physical research. Within the realm of particle and high-energy physics, an imperative challenge revolves around the scrutiny of particle tracks acquired through track detectors, exemplified by liquefied argon time projection chambers (LArTPCs). These cutting-edge neural networks have showcased unparalleled performance in tackling this intricate task, demonstrating their profound potential in enhancing particle track analysis.

Since the scalability requirement is imposed on neutrino detectors, it is also imposed on CNN model implementations. The scalability for LArTPC data processing is achieved due to the sparsity of the data, which have the form of thin trajectories. To process sparse data, subspecies sparse convolutional networks (SSCNs) and sparse tensor networks (STNs) have been proposed.

In this paper, we present the results of semantic segmentation of the LArTPC-simulated PILArNET data using STNs and test various modifications of the classical U-Net architecture: Attention U-ResNet, U-ResNet3+, U-ResNet with an additional deep supervision block, as well as loss functions: focal loss, balanced cross-entropy to improve the accuracy of Michel electron identification. The best results were obtained for balanced cross-entropy. Further improvement is possible by combining several methods.

Key words: neutrino, classification, neural network, machine learning.