

ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ
ФГАОУ ВО НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Факультет компьютерных наук
Образовательная программа «Прикладная математика и информатика»

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

Исследовательский проект на тему:

Оптимизация размера сенсоров для физики элементарных частиц

Выполнил студент:

группы #БПМИ203, 4 курса Зиманов Алихан

Принял руководитель ВКР:

Болдырев Алексей Сергеевич
Научный сотрудник
Факультет компьютерных наук НИУ ВШЭ

Москва 2024

Содержание

Аннотация	3
Abstract	4
1 Введение	5
2 Обзор литературы	5
2.1 Предшествующие работы	5
2.2 Используемые методы	9
3 Метод	9
3.1 Постановка задачи	9
3.2 Модели	9
3.3 Идеи обучения	9
4 Эксперименты	9
4.1 Данные	9
4.2 Метрики	9
4.3 Процесс обучения и замеров	9
4.4 Аугментации	9
5 Результаты	9
6 Обсуждение	9
7 Заключение	9
8 Благодарности	9
Список литературы	10
A Пример секции аппендикса	12

Аннотация

В исследованиях фундаментальных частиц, точность и эффективность измерения метрик имеет первостепенное значение. Данное исследование представляет новый подход к оптимизации размеров датчиков в электромагнитных калориметрах (ЭМК) Большого адронного коллайдера (БАК) для точной реконструкции энергии и позиции фотонов. Используя методы глубинного обучения, особенно адаптированные из области компьютерного зрения, мы стремимся идентифицировать оптимальную конфигурацию матрицы датчиков, которая максимизирует точность реконструкции энергии и позиции, учитывая ограничения по стоимости. Введя архитектуру модели, способную адаптироваться к различным размерам датчиков, данное исследование рассматривает баланс между высоким разрешением измерений и экономической целесообразностью. Предварительные результаты показывают, что специально адаптированная модель глубокого обучения может значительно улучшить дизайн датчиков, предлагая многообещающий путь для будущих экспериментальных физических установок. Эта работа не только способствует текущему дискурсу по оптимизации датчиков, но и демонстрирует потенциал глубинного обучения в продвижении исследований в области физики частиц.

Ключевые слова

Физика элементарных частиц, оптимизация размера сенсоров, глубинное обучение, компьютерное зрение, реконструкция энергии фотонов, БАК, электромагнитные калориметры

Abstract

In particle physics research, the accuracy and effectiveness of measurement metrics have a primary importance. This study presents a new approach to optimizing detector sizes in electromagnetic calorimeters (EMC) of the Large Hadron Collider (LHC) for precise photon energy and position reconstruction. Utilizing deep learning methods, particularly those adapted from computer vision, we aim to identify the optimal detector matrix configuration that maximizes reconstruction accuracy while considering cost constraints. By introducing a model architecture capable of adapting to various detector sizes, this research explores the trade-off between high measurement resolution and economic feasibility. Preliminary results indicate that a specially adapted deep learning model can significantly improve detector design, offering a promising path for future experimental physics setups. This work not only contributes to the current discourse on detector optimization but also demonstrates the potential of deep learning in advancing research in particle physics.

данное исследование рассматривает баланс между высоким разрешением измерений и экономической целесообразностью. Предварительные результаты показывают, что специально адаптированная модель глубокого обучения может значительно улучшить дизайн датчиков, предлагая многообещающий путь для будущих экспериментальных физических установок. Эта работа не только способствует текущему дискурсу по оптимизации датчиков, но и демонстрирует потенциал глубинного обучения в продвижении исследований в области физики частиц.

Keywords

Физика элементарных частиц, оптимизация размера сенсоров, глубинное обучение, компьютерное зрение, реконструкция энергии фотонов, БАК, электромагнитные калориметры

1 Введение

2 Обзор литературы

Машинное обучение достаточно широко используется в задачах физики высоких энергий. Основными примерами таких задач являются задачи реконструкции свойств элементарных частиц, проходящих через датчики по типу электромагнитных калориметров в Большом адронном коллайдере. Так как структура данных, собираемых калориметром является матрицей, состоящей из вещественных чисел и имеет некоторые сходства с фотографиями (см. далее), то многие работы вполне естественно адаптируют модели из области компьютерного зрения для решения задач физики высоких энергий.

2.1 Предшествующие работы

Одна из работ [2] из данной области исследует использование методов глубинного обучения для генерации и реконструкции при столкновении частиц в физике высоких энергий. Авторы этой статьи решают две задачи:

- Восстановление энергии и идентификация элементарных частиц;
- Моделирование ливня элементарных частиц на калориметр.

Обе задачи решаются с помощью методов глубинного обучения. Для первой задачи в качестве подходящих архитектур были выбраны полносвязная нейронная сеть, трехмерная сверточная сеть и модель основанная на GoogLeNet. Для задачи моделирования авторы статьи используют архитектуру 3DGAN, основанную на трехмерных сверточных слоях.

Всего рассматриваются четыре типа частиц: электроны e , фотоны γ , заряженные пионы π и нейтральные пионы π^0 с зарядами равномерно распределенными между 10 и 510 GeV. (смотрите на Fig. 15 photon variable energy regression).

В другой статье [3] предлагается использовать методы компьютерного зрения и глубинного обучения для улучшения реконструкции потока частиц (PFlow) в экспериментах по физике высоких энергий. Алгоритмы PFlow предназначены для идентификации и измерения энергии частиц, образующихся в результате столкновений, что является критически важным для понимания фундаментальных физических процессов.

Ключевые моменты статьи:

- **Задача:** Различение нейтральных частиц (таких как фотоны и нейтральные пионы) от заряженных частиц (например, заряженных пионов) в данных калориметра, особенно когда их энергетические отложения перекрываются. Предлагаемое решение: Использование моделей компьютерного зрения и глубокого обучения для анализа изображений калориметра и прямого предсказания распределения нейтральной энергии.
- **Изученные модели:**
 - Сверточные нейронные сети (CNN): Включая UNet с блоками ResNet, подходящие для обычных геометрий калориметров.
 - Графовые нейронные сети (GNN): Предлагают гибкость для нерегулярных геометрий и превосходят CNN по энергетическому и пространственному разрешению.
 - Глубокие множества (DS): Еще один подход для нерегулярных геометрий с хорошей производительностью.
- **Супер-разрешение:** В статье также исследуется использование глубокого обучения для супер-разрешения, что позволяет восстанавливать изображения с более высокой детализацией, чем собственное разрешение детектора. Это может быть полезно для идентификации частиц и проектирования будущих детекторов.
- **Результаты:** Предложенные модели глубокого обучения значительно улучшают реконструкцию энергетических отложений нейтральных частиц по сравнению с традиционными параметрическими методами PFlow. Они достигают лучшего энергетического разрешения, пространственного разрешения и уменьшают смещение.

Дополнительные соображения: В статье рассматривается упрощенный сценарий, включающий только заряженные и нейтральные пионы. Оценка производительности в более сложных событиях с несколькими типами частиц является критически важной. Исследование других архитектур глубокого обучения и стратегий обучения может привести к дальнейшим улучшениям. Важно исследовать вычислительные затраты и возможность реализации этих методов в реальных экспериментальных условиях.

В целом, эта статья представляет собой перспективный подход к реконструкции PFlow с использованием компьютерного зрения и глубокого обучения, предлагая потенциальные достижения в исследованиях по физике высоких энергий.

В данной статье [1] исследуется потенциал глубинных нейронных сетей, в частности сверточных нейронных сетей (CNN) и графовых нейронных сетей (GNN), для реконструкции

энергии частиц в высокогранулированных калориметрах. Авторы сравнивают эти методы с традиционными методами реконструкции энергии и обнаруживают, что CNN и GNN предлагают значительные преимущества, особенно в улучшении энергетического разрешения для адронных ливней.

Ключевые выводы:

- CNN превосходят традиционные методы: CNN, обученная на смоделированных пионных ливнях, продемонстрировала существенное улучшение энергетического разрешения как для одиночных пионов, так и для струй по сравнению с традиционными методами, такими как простое суммирование энергии и коррекция f_{em} (аналогично двухканальным калориметрам).
- CNN эффективны для различных типов частиц: CNN также показала хорошую производительность для реконструкции электронов и фотонов, даже несмотря на то, что она была обучена только на пионах.
- Понимание преимуществ CNN: Исследование предполагает, что CNN достигают превосходного разрешения, используя скрытые особенности пространственного распределения отложения энергии, которое коррелирует с физикой адронных ливней (например, множественность и угол рождения вторичных частиц).
- GNN и информация о времени: Включение информации о времени с помощью GNN еще больше улучшает энергетическое разрешение, особенно при точных измерениях времени ниже 100 пикосекунд.
- Потенциал для одноканальных калориметров: Способность CNN оценивать f_{em} и f_{had} предполагает возможность достижения высокоточной реконструкции энергии с помощью одноканальных калориметров, устраняя необходимость в сложных многоканальных конструкциях.

В целом, это исследование подчеркивает перспективный потенциал методов глубокого обучения для продвижения технологии калориметров и достижения более точной реконструкции энергии в будущих экспериментах по физике высоких энергий.

В следующей статье [4] исследуется применение графовых нейронных сетей (GNN) для реконструкции фотонов в электромагнитном калориметре (ECL) эксперимента Belle II. Цель состоит в улучшении энергетического разрешения и снижении фоновых шумов, особенно при высоких уровнях светимости.

Основные проблемы:

- Высокий уровень фонового излучения пучка: Повышенная светимость приводит к большему фоновому шуму, что влияет на энергетическое разрешение.
- Перекрывающиеся фотоны: Важно различать энергетические отложения от нескольких фотонов.
- Сложная геометрия детектора: ECL имеет неоднородную структуру с кристаллами разных размеров и форм.

Предложенное решение:

- Нечеткая кластеризация с использованием GNN: Этот подход назначает кристаллы разным классам (например, фотоны, фон) с различной степенью принадлежности.
- Архитектура GNN: Сеть использует блоки GravNet для изучения пространственных и характеристических представлений, а затем передает сообщения между связанными узлами.
- Входные данные: Используются свойства кристаллов (положение, масса), измерения (энергия, время, дискриминация формы импульса) и усредненные характеристики области интереса.

Результаты:

- Значительное улучшение энергетического разрешения: Более чем на 30% улучшение для фотонов низкой энергии и высокого фонового излучения пучка по сравнению с базовым алгоритмом.
- Сокращение хвостов в распределении энергии: GNN эффективно уменьшает фоновый шум, что приводит к более точной реконструкции энергии.
- Улучшенная производительность для перекрывающихся фотонов: Сеть успешно разделяет энергетические отложения от нескольких фотонов, особенно для асимметричных пар энергий.
- Многообещающая устойчивость к изменяющимся фонам пучка: GNN демонстрирует способность к обобщению на разные уровни фонового излучения.

Это исследование демонстрирует потенциал GNN для реконструкции калориметров в экспериментах по физике высоких энергий. Значительные улучшения в энергетическом разрешении и подавлении фона имеют прямые последствия для физических возможностей Belle II, особенно для анализов, связанных с фотонами и недостающей энергией.

2.2 Используемые методы

ResNet, Vision Transformer, Image augmentations

3 Метод

3.1 Постановка задачи

3.2 Модели

3.3 Идеи обучения

4 Эксперименты

4.1 Данные

4.2 Метрики

4.3 Процесс обучения и замеров

4.4 Аугментации

5 Результаты

6 Обсуждение

7 Заключение

8 Благодарности

Модель	Метрика	Размер матрицы					
		10×10	15×15	20×20	25×25	30×30	40×40
SumModel	\mathcal{L}_{total}	0.1549	0.1555	0.1563	0.1575	0.1560	0.1561
	\mathcal{L}_{eng}	0.0213	0.0221	0.0226	0.0229	0.0220	0.0222
	\mathcal{L}_{pos}	0.2887	0.2891	0.2901	0.2924	0.2901	0.2901
LinearModel	\mathcal{L}_{total}	0.0728	0.0738	0.0741	0.0745	0.0748	0.0751
	\mathcal{L}_{eng}	0.0213	0.0221	0.0226	0.0231	0.0221	0.0223
	\mathcal{L}_{pos}	0.1244	0.1258	0.1257	0.1264	0.1278	0.1282
MyResnet18	\mathcal{L}_{total}	0.0926	0.0823	0.0248	0.0254	0.0707	0.0271
	\mathcal{L}_{eng}	0.0814	0.0705	0.0215	0.0215	0.0565	0.0207
	\mathcal{L}_{pos}	0.1042	0.0941	0.0284	0.0297	0.0851	0.0336
MyCNN	\mathcal{L}_{total}	0.0229	0.0240	0.0682	0.5922	0.0260	0.1326
	\mathcal{L}_{eng}	0.0197	0.0200	0.0201	0.8934	0.0197	0.0196
	\mathcal{L}_{pos}	0.0262	0.0281	0.1166	0.2912	0.0324	0.2462
MyViT	\mathcal{L}_{total}	0.0221	0.0232	0.0243	0.0242	0.0253	0.0264
	\mathcal{L}_{eng}	0.0194	0.0198	0.0203	0.0194	0.0196	0.0193
	\mathcal{L}_{pos}	0.0250	0.0270	0.0285	0.0295	0.0311	0.0336

Таблица 5.1

Функция потерь	Метрика	Размер матрицы					
		10×10	15×15	20×20	25×25	30×30	40×40
RMSE/E	\mathcal{L}_{eng}	0.0194	0.0198	0.0201	0.0192	0.0193	0.0195
	\mathcal{L}_{pos}	0.0250	0.0269	0.0284	0.0294	0.0310	0.0338
MAE/E	\mathcal{L}_{eng}	0.0198	0.0200	0.0204	0.0199	0.0199	0.0200
	\mathcal{L}_{pos}	0.0251	0.0269	0.0284	0.0294	0.0312	0.0336
RMSLE	\mathcal{L}_{eng}	0.0193	0.0196	0.0204	0.0191	0.0194	0.0194
	\mathcal{L}_{pos}	0.0249	0.0269	0.0285	0.0294	0.0310	0.0338
RMSE	\mathcal{L}_{eng}	0.0222	0.0222	0.0238	0.0214	0.0216	0.0219
	\mathcal{L}_{pos}	0.0290	0.0294	0.0332	0.0363	0.0356	0.0385

Таблица 5.2

Список литературы

- [1] N. Akchurin, C. Cowden, J. Damgov, A. Hussain и S. Kunori. “On the use of neural networks for energy reconstruction in high-granularity calorimeters”. в: *Journal of Instrumentation* 16.12 (дек. 2021), P12036. ISSN: 1748-0221. DOI: [10.1088/1748-0221/16/12/p12036](https://doi.org/10.1088/1748-0221/16/12/p12036). URL: <http://dx.doi.org/10.1088/1748-0221/16/12/P12036>.
- [2] Dawit Belayneh, Federico Carminati, Amir Farbin, Benjamin Hooberman, Gulrukh Khattak, Miaoyuan Liu, Junze Liu, Dominick Olivito, Vitória Barin Pacela, Maurizio Pierini, Alexander Schwing, Maria Spiropulu, Sofia Vallecorsa, Jean-Roch Vlimant, Wei Wei и Matt Zhang. “Calorimetry with deep learning: particle simulation and reconstruction for collider physics”. в: *The European Physical Journal C* 80.7 (июль 2020). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-020-8251-9](https://doi.org/10.1140/epjc/s10052-020-8251-9). URL: <http://dx.doi.org/10.1140/epjc/s10052-020-8251-9>.

- [3] Francesco Armando Di Bello, Sanmay Ganguly, Eilam Gross, Marumi Kado, Michael Pitt, Lorenzo Santi и Jonathan Shlomi. “Towards a computer vision particle flow”. в: *The European Physical Journal C* 81.2 (февр. 2021). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-021-08897-0](https://doi.org/10.1140/epjc/s10052-021-08897-0). URL: <http://dx.doi.org/10.1140/epjc/s10052-021-08897-0>.
- [4] F. Wemmer, I. Haide, J. Eppelt, T. Ferber, A. Beaubien, P. Branchini, M. Campajola, C. Cecchi, P. Cheema, G. De Nardo, C. Hearty, A. Kuzmin, S. Longo, E. Manoni, F. Meier, M. Merola, K. Miyabayashi, S. Moneta, M. Remnev, J. M. Roney, J.-G. Shiu, B. Shwartz, Y. Unno, R. van Tonder и R. Volpe. “Photon Reconstruction in the Belle II Calorimeter Using Graph Neural Networks”. в: *Computing and Software for Big Science* 7.1 (дек. 2023). ISSN: 2510-2044. DOI: [10.1007/s41781-023-00105-w](https://doi.org/10.1007/s41781-023-00105-w). URL: <http://dx.doi.org/10.1007/s41781-023-00105-w>.

А Пример секции аппендикса