

Зиманов Алихан

Исследовательский проект
ВКР



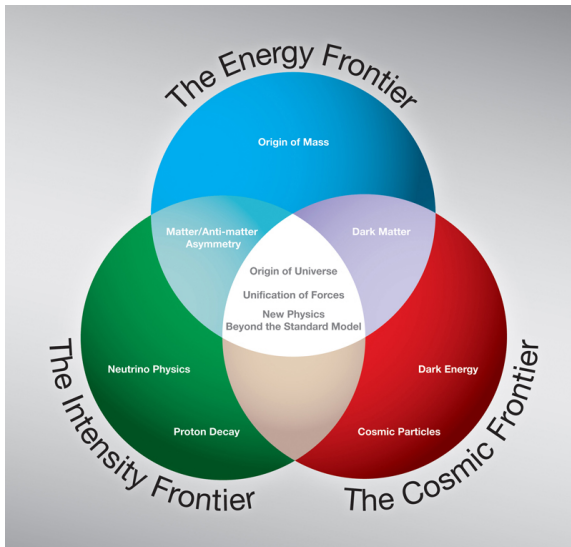
Оптимизация размера сенсоров для физики элементарных частиц

Научный руководитель Болдырев А.С.

Москва, 2024

- 1 Введение
- 2 Обзор литературы
- 3 Методология
- 4 Результаты
- 5 Заключение
- 6 Приложения

Постановка задачи



Актуальность работы

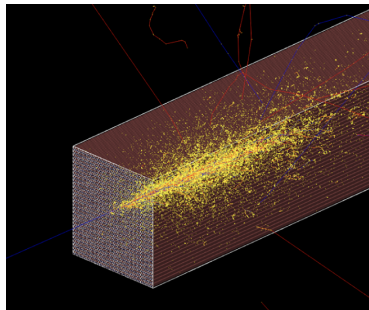
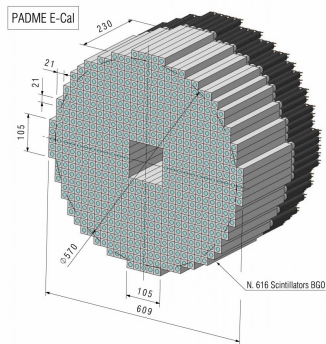


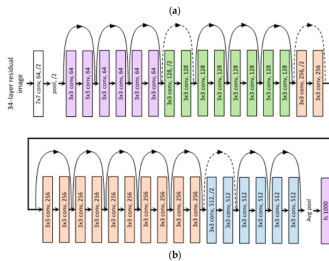
Рис. 1: Электромагнитный калориметр PADME и симуляция электромагнитного ливня с помощью GEANT4.

Смежные работы

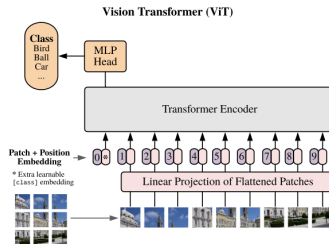
Применение глубинного обучения в физике элементарных частиц

- ❖ Идентификация частиц и реконструкция энергии, моделирование калориметрических ливней [1].
- ❖ Разделение остаточной энергии заряженных и нейтральных частиц, а также сравнение с PFlow алгоритмами [2].
- ❖ Реконструкция энергий фотонов с помощью графовых нейронных сетей [3].

Используемые модели



(a) ResNet



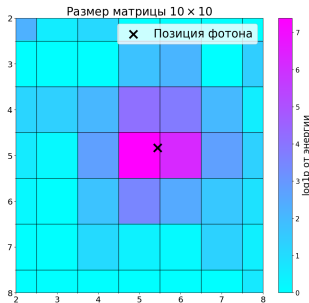
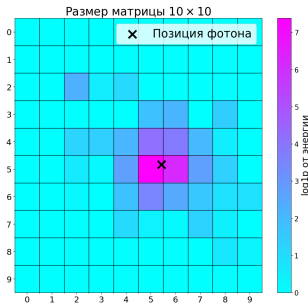
(b) Vision Transformer

ResNet и Vision Transformer являются самыми показательными моделями из области компьютерного зрения.

Данные

Данные сгенерированы с помощью GEANT4. Каждый элемент состоит из:

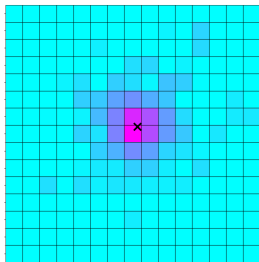
- ❖ Матрица неотрицательных чисел (размерность матрицы может быть от 10×10 до 40×40)
- ❖ Исходная энергия фотона (от 1 до 100 ГэВ)
- ❖ Положение входной точки фотона (в центральной ячейке калориметра)



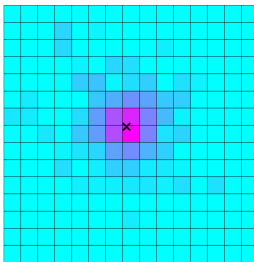
Аугментации

Применяемые аугментации

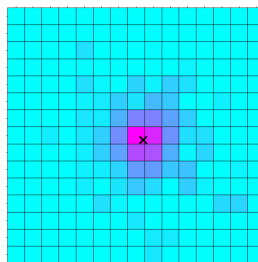
- ❖ Случайное горизонтальное или вертикальное отражение
- ❖ Случайный поворот на угол вида 0° , 90° , 180° и 270° .



(a) Оригинал



(b) Отражение



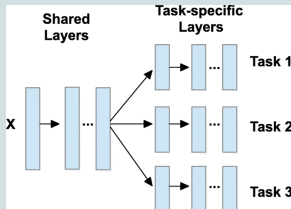
(c) Поворот

Модели

Использованные модели

- ❖ Аналитическая модель (AnaModel)
- ❖ Линейная регрессия (LinReg)
- ❖ ResNet18
- ❖ Сверточные сети (CNN)
- ❖ Vision Transformer (ViT)

Модели обучались одновременно на две задачи: реконструкция энергии и восстановление позиции фотона.



Метрики

$\{(X_i, E_i, P_i), (\hat{E}_i, \hat{P}_i)\}_{i=1}^n$ — выборка данных, где

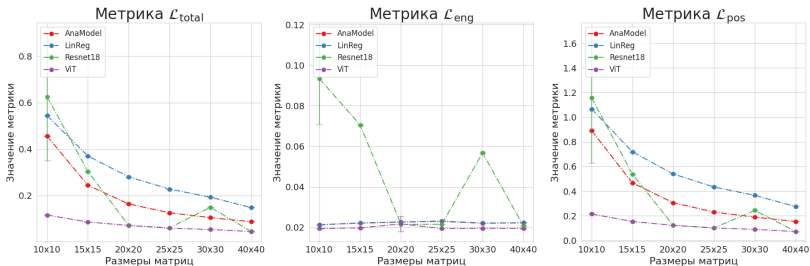
- ❖ X_i — считанные калориметром значения ($\mathbb{R}_+^{D \times D}$)
- ❖ E_i — исходная энергия фотона (\mathbb{R}_+)
- ❖ $P_i = (P_i^x, P_i^y)$ — позиция входа фотона (\mathbb{R}^2)
- ❖ \hat{E}_i — предсказанная энергия (\mathbb{R}_+)
- ❖ $\hat{P}_i = (\hat{P}_i^x, \hat{P}_i^y)$ — предсказанная позиция (\mathbb{R}^2)

$$\mathcal{L}_{\text{eng}} = \text{RMSE}/E(\hat{E}, E) = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{\hat{E}_i - E_i}{E_i} \right)^2}$$

$$\mathcal{L}_{\text{pos}} = \text{RMSE}(\hat{P}, P) = \sqrt{\frac{1}{2n} \sum_{i=1}^n \left((\hat{P}_i^x - P_i^x)^2 + (\hat{P}_i^y - P_i^y)^2 \right)}$$

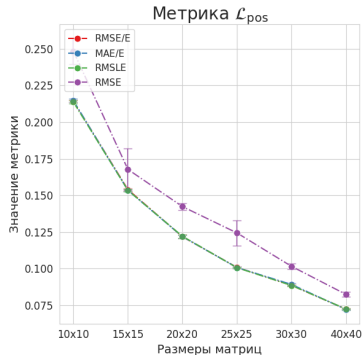
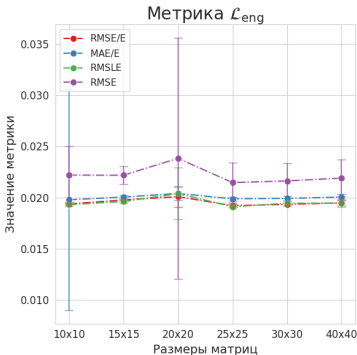
$$\mathcal{L}_{\text{total}} = \alpha \cdot \mathcal{L}_{\text{eng}} + (1 - \alpha) \cdot \mathcal{L}_{\text{pos}}, \quad \alpha \in [0, 1]$$

Сравние моделей



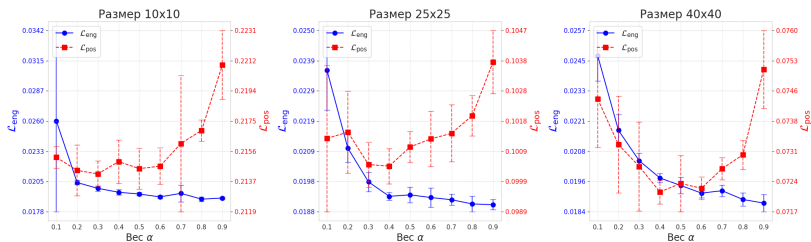
Модель ViT показывает лучшие и стабильные результаты.

Сравние функций потерь



Обучение на нормализованную (относительную) ошибку приводит к лучшему качеству.

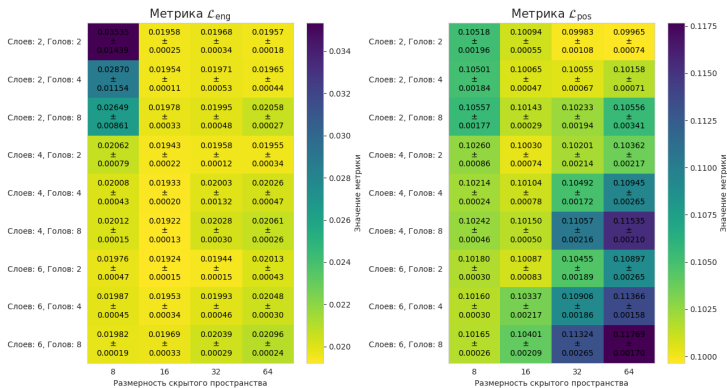
Отношение важности задач



Влияние гиперпараметра α на качество модели^а.

$$^a \mathcal{L}_{total} = \alpha \cdot \mathcal{L}_{eng} + (1 - \alpha) \cdot \mathcal{L}_{pos}.$$

Размер модели



Оптимальные параметры модели ViT это 4 слоя, 2 головы и размерность скрытого слоя 16.

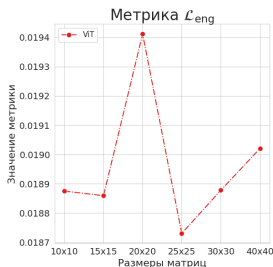
Эффективность аугментаций

		Размер матрицы			
		15 × 15			
Отражения	Повороты	$\mathcal{L}_{\text{eng}}^{\text{train}}$	$\mathcal{L}_{\text{pos}}^{\text{train}}$	$\mathcal{L}_{\text{eng}}^{\text{val}}$	$\mathcal{L}_{\text{pos}}^{\text{val}}$
✗	✗	0.0182	0.1456	0.0217	0.1535
✓	✗	0.0186	0.1481	0.0194	0.1537
✗	✓	0.0185	0.1476	0.0193	0.1528
✓	✓	0.0185	0.1475	0.0192	0.1525

Применение аугментаций

- ✦ Улучшение качества на валидационной выборке
- ✦ Сокращение разрыва между обучающей и валидационной выборкой

Итоговое качество



Модель ViT одновременно

- ❖ решает задачу восстановления энергии с относительной ошибкой в 1.3%
- ❖ решает задачу реконструкции позиции с точностью, в 20 раз меньшую размера ячейки калориметра

Заклучение

Результаты работы

- ❖ Исследование и сравнение моделей глубинного обучения
- ❖ Оптимизация метрик и функций потерь
- ❖ Анализ влияния размера модели и аугментации данных
- ❖ Достижение высокой точности реконструкции
- ❖ Практическое применение и перспективы

Список литературы I



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. URL: <http://dx.doi.org/10.1140/epjc/s10052-020-8251-9>.



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. URL: <http://dx.doi.org/10.1140/epjc/s10052-021-08897-0>.

Список литературы II



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. Schwartz, 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. URL: <http://dx.doi.org/10.1007/s41781-023-00105-w>.

Метрики

- ❖ Корень из среднеквадратичной ошибки (RMSE)
- ❖ Средняя абсолютная ошибка (MAE)
- ❖ Корень из среднеквадратичной логарифмической ошибки (RMSLE):

$$\text{RMSLE}(a, y) = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log(a_i + 1) - \log(y_i + 1))^2}.$$

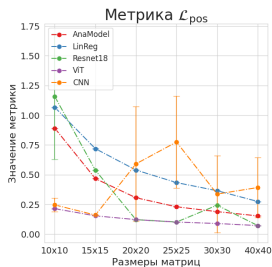
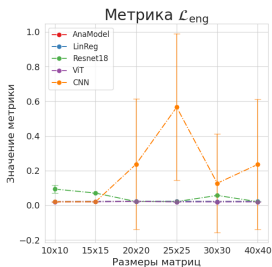
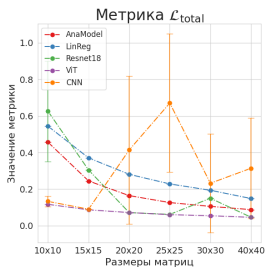
- ❖ Взвешенный корень из среднеквадратичной ошибки (RMSE/E):

$$\text{RMSE/E}(a, y) = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{a_i - y_i}{y_i} \right)^2}.$$

- ❖ Взвешенная средняя абсолютная ошибка (MAE/E):

$$\text{MAE/E}(a, y) = \frac{1}{n} \sum_{i=1}^n \frac{|a_i - y_i|}{y_i}.$$

Полное сравнение моделей



Модель CNN показывает слабые результаты, поэтому не была включена в основные слайды.

Таблица результатов лучшей модели

Метрика	Размер матрицы					
	10×10	15×15	20×20	25×25	30×30	40×40
$\mathcal{L}_{\text{total}}$	0.1153	0.0852	0.0702	0.0588	0.0535	0.0453
\mathcal{L}_{eng}	0.0189	0.0189	0.0194	0.0187	0.0189	0.0190
\mathcal{L}_{pos}	0.2117	0.1515	0.1211	0.0989	0.0881	0.0715
Размер одной ячейки	6.0600	4.0400	3.0300	2.4240	2.0200	1.5150
$\text{MAE}/E_{\text{eng}}$	0.0131	0.0127	0.0132	0.0124	0.0130	0.0129

Модель ViT способна решать задачу реконструкции позиции с точностью, в 20 раз меньшую чем длина стороны центральной ячейки. Более того, данная модель решает задачу восстановления энергии с относительной ошибкой в 1.3%.