

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЯДЕРНЫЙ УНИВЕРСИТЕТ «МИФИ»

УДК 531.3, 539.1.05

ОТЧЕТ
О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ
**РЕКОНСТРУКЦИЯ СИГНАЛОВ В ДЕТЕКТОРЕ FARICH
УСТАНОВКИ SPD НА КОЛЛАЙДЕРЕ NICA**

Научный руководитель

к. ф-м. н.

Студент

_____ А. С. Жемчугов

_____ К. Ю. Массалов

Москва 2024

СОДЕРЖАНИЕ

Введение	4
1 Экспериментальная установка Spin Physics Detector на строящемся в ОИЯИ коллайдере NICA	5
1.1 Конструкция детектора	5
1.2 FARICH	6
2 Алгоритмы распознавания колец	7
2.1 Разделение частиц по вторым моментам	7
2.2 Итерационный метод	7
2.3 Отношение функций максимального правдоподобия	7
2.4 Метод масок	7
2.5 Нечеткая кластеризация	7
2.6 Преобразование Хафа	8
2.7 COP	8
2.8 Фитирование эллипса	8
3 Моделирование детектора	9
3.1 GEANT4	9
3.2 Моделирование детектора и событий	9
4 Применение нейросетевых методов	11
4.1 Сверточные нейронные сети	11
4.2 Постановка задачи	12
4.3 Облако точек	12
4.3.1 Архитектура нейросети	12
4.3.2 Обучение и результаты	14
4.4 Изображение	15
4.4.1 Архитектура нейросети	15
4.4.2 Обучение и результаты	15
4.5 Разбиение входного изображения	17
Заключение	18
Список использованных источников	19

ВВЕДЕНИЕ

Относительный вклад видимой барионной материи во Вселенную составляет около 5% от общей массы. Хотя барионная материя представляется достаточно изученной, вопрос о том, почему нуклоны имеют именно такую структуру и свойства, остается открытым. Одним из главных нерешенных вопросов в квантовой хромодинамике (КХД) остается понимание структуры и основных свойств нуклонов исходя из динамики кварков и глюонов внутри них.

Нуклон обладает спином, который отвечает за его магнитный момент и другие свойства. Изучение спиновой структуры нуклонов является важной задачей. Центральной задачей является понимание того, как спин нуклона формируется из спинов и орбитальных моментов кварков и глюонов.

За последние 25 лет эксперименты по глубоко неупругому рассеянию (CERN, DESY, JLab, SLAC) и высокоэнергетические столкновения протонов (RHIC) были основными источниками информации о спиновых структурных функциях нуклонов. Наши знания о внутренней структуре нуклонов все еще ограничены, особенно в отношении глюонного вклада.

Spin Physics Detector — универсальная установка для исследования спиновой структуры нуклонов и других спиновых явлений. Установка будет использовать поляризованные пучки протонов и дейтронов. Проект основан на большом опыте работы с поляризованными пучками в ОИЯИ. Основная цель эксперимента - комплексное исследование неполяризованного и поляризованного глюонной составляющей нуклона. Измерения спина в SPD на коллайдере NICA откроют новые возможности для понимания спиновой структуры нуклона.

На установке планируется частота соударений до 4 МГц и число выходных каналов с детекторов ~ 700000 . Таким образом, ожидается поток данных примерно 20 Гбайт/с или 200 Пбайт/год. Такой объем информации необходимо уменьшать, избавляясь от неинформативных событий и шумов. Для этого планируется использовать онлайн-фильтр, представляющий собой высокопроизводительную систему, которая будет включать в себя гетерогенные вычислительные платформы, аналогичные многим высокопроизводительным вычислительным кластерам. Для реконструкции и отбрасывания неинтересных событий можно использовать методы машинного обучения, поскольку проблема классических алгоритмов в том, что они имеют невысокую скорость и плохо параллелятся, в отличие от нейронных сетей. Проблема же нейронных сетей в сложности интерпретации их решений: никогда нельзя знать наверняка, почему сеть выдала именно такой результат, и насколько она права.

Таким образом, планируется использовать нейронные сети и классические алгоритмы одновременно: первые обеспечат скорость работы, вторые позволят проверять правильность результатов нейросетей.

Цель работы состоит в создании нейросетевых методов реконструкции сигналов с детектора FARICH, изучении их характеристик и сравнении их с классическими подходами.

1 ЭКСПЕРИМЕНТАЛЬНАЯ УСТАНОВКА SPIN PHYSICS ДЕТЕКТОР НА СТРОЯЩЕМСЯ В ОИЯИ КОЛЛАЙДЕРЕ NICA

Коллаборация Spin Physics Detector предлагает установить универсальный детектор во второй точке взаимодействия строящегося ускорителя NICA в Дубне. Этот детектор будет использоваться для изучения спиновой структуры протона и дейтрона и других спиновых эффектов с помощью поляризованных пучков протонов и дейтронов с энергией столкновения до 27 ГэВ и светимостью до $10^{32} \text{ см}^{-2} \text{ с}^{-1}$. Эксперимент SPD в NICA позволит заполнить кинематический пробел между измерениями при низких энергиях на установках ANKE-COSY и SATURNE, измерениями при высоких энергиях на ускорителе тяжелых ионов и будущими экспериментами на Большом адронном коллайдере (БАК). Уникальность NICA заключается в возможности работы с пучками поляризованных дейтронов при этих энергиях.

1.1 КОНСТРУКЦИЯ ДЕТЕКТОРА

Экспериментальная установка SPD (Рис. 1) проектируется как универсальный детектор 4π с расширенными возможностями для отслеживания и идентификации частиц на основе современных технологий. Вершинный детектор (vertex detector) на основе кремния обеспечит разрешение по позиции вершин на уровне менее 100 микрон, необходимое для восстановления вторичных вершин от распадов D- мезонов. Система трекинга (straw tracker), помещенных в соленоидное магнитное поле с напряженностью до 1 Тесла на оси детектора, обеспечивает разрешение по поперечной импульсу $\sigma_{p_T}/p_T \approx 2\%$ для частиц с импульсом 1 ГэВ/с. Время-пролетная система (time-of-flight system) с временным разрешением около 60 пикосекунд обеспечивает разделение 3σ для пионов/каонов и для каонов/протонов до энергий 1.2 и 2.2 ГэВ соответственно. Использование черенковского детектора на базе аэрогеля — FARICH — поможет расширить этот диапазон вплоть до 6 ГэВ. Обнаружение фотонов обеспечивается электромагнитным калориметром (electromagnetic calorimeter) с энергетическим разрешением $5\%/\sqrt{E}$. Для идентификации мюонов планируется использовать мюонную систему (range system). Она также может использоваться как грубый адронный калориметр. За локальную поляриметрию и контроль светимости будет отвечать пара счетчиков пучков (beam-beam counter) и калориметров малого угла, или малоапертурных калориметров (zero-degree calorimeter).

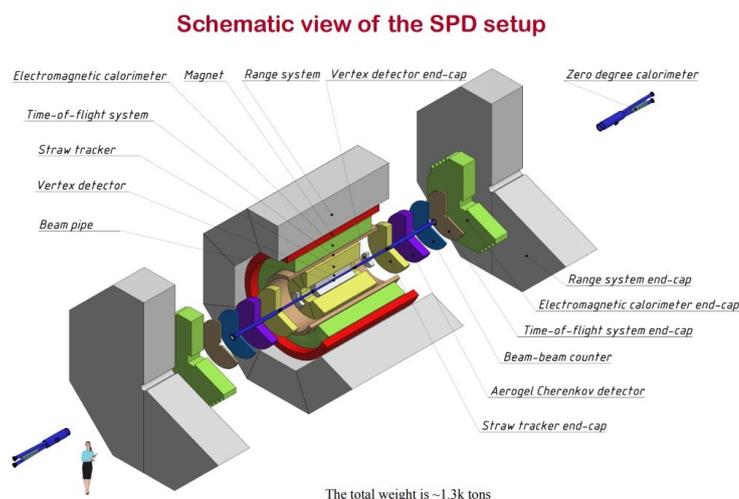


Рисунок 1 — Конструкция детектора

1.2 FARICH

FARICH (Focusing Aerogel Ring-Imaging CHerenkov detector) — черенковский детектор с аэрогелем, используемым в качестве рабочего тела. Черенковский детектор работает на основе эффекта Черенкова, который заключается в излучении света заряженными частицами, движущимися в среде со скоростью, превышающей фазовую скорость света в этой среде: $\beta n > c$. Это излучение имеет вид конуса, ось которого совпадает с направлением движения частицы, а угол раствора зависит от скорости частицы и показателя преломления среды. Излучение регистрируется массивом фотодетекторов, в результате чего на выходе получается срез конуса плоскостью. Если частица летела перпендикулярно детектору, получается набор точек, формирующих окружность, которую в дальнейшем на онлайн-фильтре надо будет реконструировать при помощи методов машинного обучения.

В качестве среды планируется использовать четыре слоя аэрогеля каждый со своим показателем преломления, что позволит улучшить разрешающую способность детектора [1].

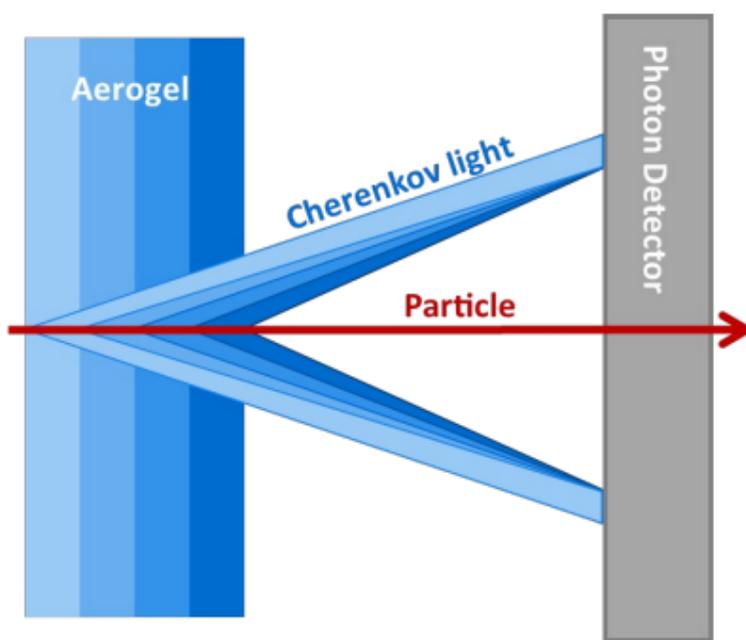


Рисунок 2 — Принципиальная схема детектора FARICH

2 АЛГОРИТМЫ РАСПОЗНАВАНИЯ КОЛЕЦ

2.1 РАЗДЕЛЕНИЕ ЧАСТИЦ ПО ВТОРЫМ МОМЕНТАМ

Простейшим методом идентификации является разделение по вторым моментам. Для каждого значения импульса второй центральный момент для кольца будет меняться в зависимости от типа частицы. Зная зависимость вторых моментов от импульса для разных частиц, можно определить их тип.

2.2 ИТЕРАЦИОННЫЙ МЕТОД

Для данного метода необходима дополнительная информация о предположительных центрах колец, которую можно получить независимо, например, в результате экстраполяции трека из трекового детектора.

Принцип работы следующий: при инициализации рассчитывается расстояние от каждого хита до каждого трека. Полученными значениями с весом 1 заполняют гистограммы (по числу треков). Затем, на каждой итерации обновляют веса хитов по формуле:

$$w_k = \frac{y_k}{\sum_i y_i}, \quad (1)$$

где w_k — вес хита в k -той гистограмме, y_k — содержание ячейки гистограммы, в которую попал данный хит. В результате получаются пики, которые соответствуют центрам колец [2].

2.3 ОТНОШЕНИЕ ФУНКЦИЙ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

Данный метод также требует дополнительной информации о центрах колец. На основании информации о конструкции детектора и параметрах трека делается предсказание положения черенковского фотона при гипотезе о типе частицы. На основе предсказания и наблюдения строится функция правдоподобия, а затем выбирается функция с максимальным правдоподобием [2].

2.4 МЕТОД МАСОК

На предварительно центрированное изображение накладывается трафарет с известными параметрами. Наложение заключается в сложении пикселей маски и изображения по модулю два. Чем больше нулей в результате, тем больше данная маска соответствует изображению [2].

2.5 НЕЧЕТКАЯ КЛАСТЕРИЗАЦИЯ

Метод заключается в минимизации следующего функционала:

$$J = \sum_{i=1}^C \sum_{j=1}^N u_{ij} d_{ij}^2, \quad (2)$$

где u_{ij} — весовой коэффициент, показывающий вероятность принадлежности хита i к кластеру j ($\sum_j u_{ij} = 1$), d_{ij} — расстояние между i хитом и j кластером, C — количество колец, N — количество хитов.

В данном методе необходимо также знать центры колец и их количество [2].

2.6 ПРЕОБРАЗОВАНИЕ ХАФА

Изначально преобразование Хафа предназначено для распознавания прямых и заключается в идее перейти из пространства координат в пространство параметров и нахождения там точки с наибольшей «интенсивностью», т.е. точки с наибольшим числом пересечений прямых. Полученная точка в пространстве параметров и есть параметры, описывающие прямую, проходящую через исходные точки.

Затем данное преобразование можно обобщить и на случай окружностей, перейдя в трехмерное пространство параметров [3].

2.7 СОР

Алгоритм СОР (Chernov-Ososkov-Pratt) позволяет по точкам оценить параметры окружности [4].

Для подгонки окружности методом наименьших квадратов необходимо минимизировать функционал:

$$L(a, b, R) = \sum_{i=1}^N (\sqrt{(x_i - a)^2 + (y_i - b)^2} - R)^2, \quad (3)$$

где a, b — x и y координаты центра окружности, R — ее радиус, x_i и y_i — координаты i -й точки, N — число точек.

Однако, такой метод требует достаточно много времени для вычисления. Для ускорения работы алгоритма можно использовать следующий функционал:

$$M(a, b, c) = \sum_{i=1}^N ((x_i - a)^2 + (y_i - b)^2 - R^2)^2 \quad (4)$$

Минус такого подхода в том, что при измерениях, не покрывающих всю окружность, результат может оказаться довольно далеко от истины.

В алгоритме СОР предлагается использовать функционал:

$$M(a, b, c) = \sum_{i=1}^N [((x_i - a)^2 + (y_i - b)^2 - R^2)^2 / R^2] \quad (5)$$

Нахождение минимума такого функционала эквивалентно минимизации (3), однако вычислительно менее затратно.

2.8 ФИТИРОВАНИЕ ЭЛЛИПСА

Для фитирования эллипса предлагается использовать функционал вида [5]:

$$L(x_{F_1}, y_{F_1}, x_{F_2}, y_{F_2}, a) = \sum_{i=1}^N (\sqrt{(x_i - x_{F_1})^2 + (y_i - y_{F_1})^2} + \sqrt{(x_i - x_{F_2})^2 + (y_i - y_{F_2})^2} - 2a)^2, \quad (6)$$

где $x_{F_1}, y_{F_1}, x_{F_2}, y_{F_2}$ — координаты x и y первого и второго фокуса эллипса соответственно, a — длина большой полуоси.

Минимизация осуществляется при помощи градиентного спуска.

3 МОДЕЛИРОВАНИЕ ДЕТЕКТОРА

3.1 GEANT4

GEANT4 — это программный инструмент на C++ для моделирования экспериментов в области физики элементарных частиц. Он разрабатывается и поддерживается Европейской организацией по ядерным исследованиям (CERN) [6–8].

GEANT4 используется для моделирования прохождения частиц через материю, включая взаимодействия, которые происходят при столкновении частиц с атомами в материале. Эта информация может быть использована для проектирования экспериментов, понимания результатов этих экспериментов и создания симуляций будущих экспериментов.

В данной работе пакет GEANT4 использовался для моделирования детектора FARICH.

3.2 МОДЕЛИРОВАНИЕ ДЕТЕКТОРА И СОБЫТИЙ

Модель детектора (Рис. 3) — два диска радиусом 847 мм. Первый диск состоит из четырех слоев аэрогеля на основе диоксида кремния (94% по массе) с небольшой примесью диоксида циркония (6% по массе) с изменяющимся показателем преломления: 1.037, 1.041, 1.043, 1.047. Толщины слоев 7 мм, 10, 9 мм, 10 мм соответственно. Второй — чувствительный детектор, который регистрирует черенковские фотоны в диапазоне примерно от 400 нм до 450 нм — в области чувствительности кремниевых фотодетекторов. Расстояние между аэрогелем и фотоприемником — 164 мм.

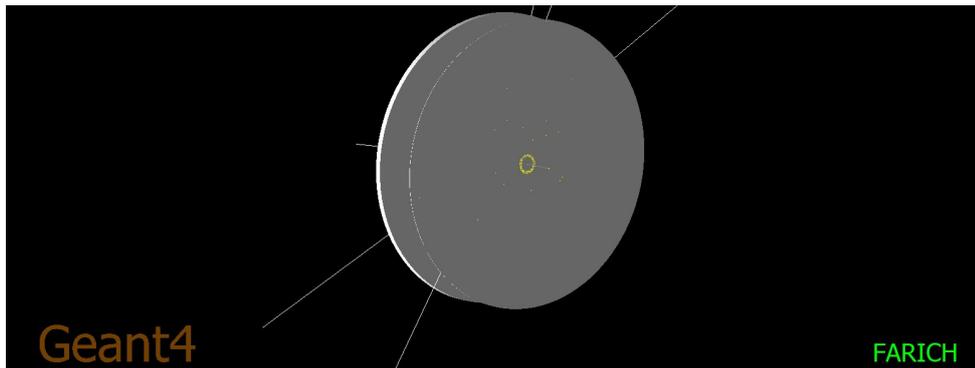


Рисунок 3 — Моделирование детектора FARICH

Во время моделирования для каждого из четырех типов частиц при пяти значениях импульсов (1-5 ГэВ) запускалось 100 событий. В каждом событии частица падала перпендикулярно детектору в точке (0,0). Координаты и энергия черенковских фотонов из диапазона 400-450 нм сохранялись. Затем строились гистограммы для числа фотонов в каждом событии (Рис. 4), а также вторые центральные моменты для колец по формуле:

$$\mu_2 = \frac{1}{n} \sum_{i=1}^n (\sqrt{x_i^2 + y_i^2}) - \frac{1}{n} \sum_{j=1}^n (\sqrt{x_j^2 + y_j^2}) \quad (7)$$

где x_i и y_i — координаты i -той точки, n — число точек в кольце.

В результате моделирования были построены спектры фотонов, гистограммы для радиусов колец и вторых моментов для e^- , π^+ , K^+ , p^+ для импульсов от 1 до 5 ГэВ.

Из приведенных графиков видно, что по моментам возможно разделить электроны и пионы до импульса примерно в 1-2 ГэВ. Также неплохо разделяются каоны и протоны до импульсов в 5 ГэВ. Однако, требуется дальнейшее совершенствование метода для пионов и каонов.

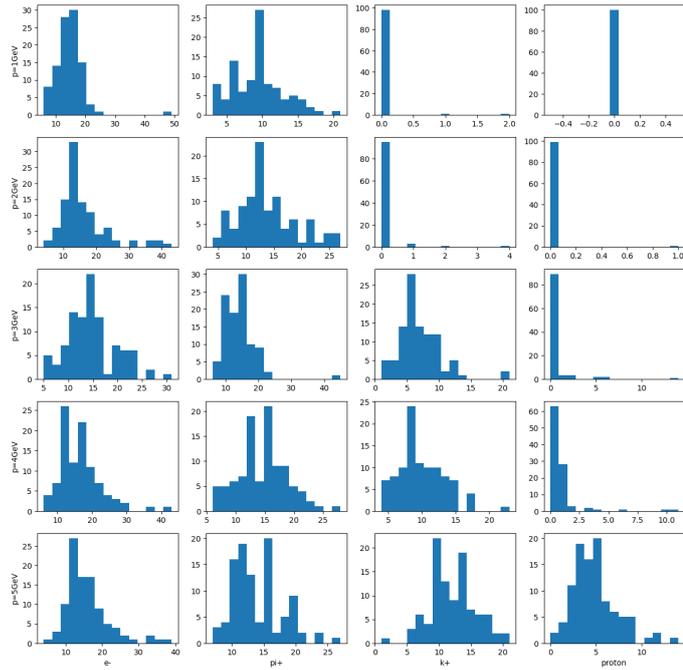


Рисунок 4 — Энергетические спектры для разных импульсов и частиц

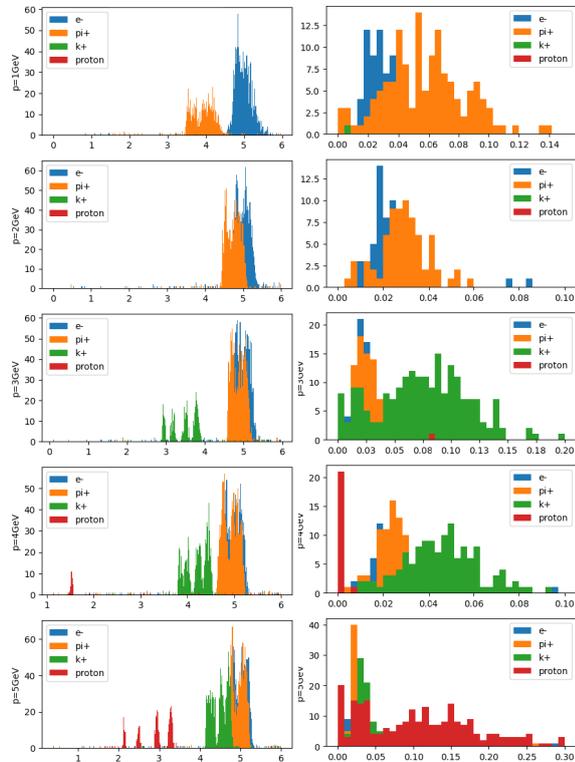


Рисунок 5 — Радиусы(слева) и вторые моменты(справа) для разных импульсов и частиц. Радиусы взяты из истинной информации Монте-Карло генерирования.

4 ПРИМЕНЕНИЕ НЕЙРОСЕТЕВЫХ МЕТОДОВ

В связи с быстрым развитием методов машинного обучения и в частности нейросетевых подходов, возникает возможность применения этих методов для анализа данных, полученных с FARICH детектора. Нейросетевые методы обладают уникальной способностью выявлять сложные зависимости в данных, что делает их идеальным инструментом для распознавания колец и анализа параметров частиц, регистрируемых в эксперименте. Такой подход не только повышает точность распознавания, но и позволяет создавать более эффективные модели, способствуя глубокому пониманию физических процессов.

В рамках поставленной задачи было решено использовать сверточные нейронные сети.

4.1 СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ

Сверточные нейронные сети (Convolutional Neural Networks, CNN) - это класс нейронных сетей, специально разработанный для обработки структурированных данных, таких как изображения. Они состоят из нескольких слоев, включая сверточные слои, слои пулинга и полносвязные слои.

Сверточные слои (Рис. 6): основной компонент CNN. Они используют фильтры для извлечения различных признаков из входных данных. Например, первый сверточный слой может выявлять базовые черты, такие как грани или углы, тогда как последующие слои могут выявлять более сложные особенности.

Input		Kernel		Output		
0	1	2	*	=	19	25
3	4	5			0	1
6	7	8			2	3

Рисунок 6 — Пример свертки с ядром 2×2

Слои пулинга (Рис. 7): эти слои снижают размерность и объем данных путем объединения информации с определенных областей. Это помогает уменьшить количество параметров и вычислений в сети, а также делает признаки инвариантными к небольшим изменениям в данных.

3	3	2	1	0		3.0	3.0	3.0
0	0	1	3	1		3.0	3.0	3.0
3	1	2	2	3		3.0	2.0	3.0
2	0	0	2	2		3.0	2.0	3.0
2	0	0	0	1		3.0	2.0	3.0

Рисунок 7 — Пример пулинга с ядром 3×3

Также распространен глобальный пулинг (Рис. 8): он позволяет обрабатывать изображения произвольных размеров путем агрегации данных вдоль каналов.

Полносвязные слои: последние слои в сети, которые объединяют признаки, выученные предыдущими слоями сверточной нейросети, для принятия решений.

CNN применяются в широком диапазоне задач компьютерного зрения, включая распознавание объектов, классификацию изображений, детекцию объектов и семантическую сегментацию. Их эффек-

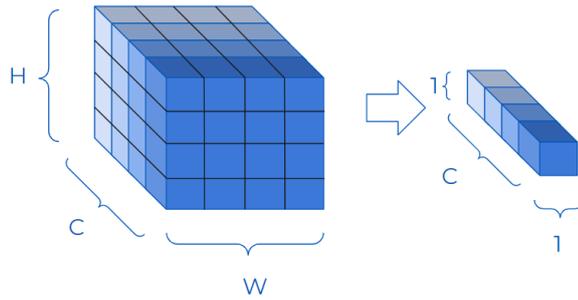


Рисунок 8 — Пример глобального пулинга

тивность объясняется способностью изучать иерархии признаков изображений на различных уровнях абстракции.

В случае необходимости предобработки данных, аугментации и выбора функций потерь, CNN требуют настройки определенных параметров для достижения оптимальных результатов в конкретных задачах.

4.2 ПОСТАНОВКА ЗАДАЧИ

На вход подается матрица сигналов с детектора $M \in R^{N \times N}$, на выходе надо получить значение скорости $v \in R^c$ для всех частиц, вызвавших срабатывание детектора. Координаты центров $C \in R^{2 \times c}$ можно получить либо из другого детектора, либо из матрицы M .

Данную задачу можно разбить на две подзадачи:

- 1) Представление входных данных в виде вектора $V \in R^n$ меньшей размерности (эмбеддинг). В дальнейшем этот вектор может быть использован для нахождения центров.
- 2) Использование полученных центров и эмбеддинга для нахождения скорости, параметров эллипса или конуса.

Можно выделить три возможных варианта решения:

- 1) Обработать входные данные как облако точек.
- 2) Обработать входные данные как обычное изображение.
- 3) Обработать входные данные, разбивая входное изображение на более маленькие и обрабатывать каждое в отдельности.

Далее будет рассматриваться только регрессия. Центры эллипсов берутся из моделирования.

4.3 ОБЛАКО ТОЧЕК

Облако точек представляет собой набор n -мерных векторов, каждый из которых описывает одну точку в пространстве. В данной задаче облако точек задается матрицей размера $N_p \times 3$ (Рис. 9), где N_p — число точек. Преимущество такого подхода в меньшем количестве вычислений и меньших затратах памяти по сравнению с классическими сверточными сетями.

4.3.1 Архитектура нейросети

В общем случае архитектура сети выглядит следующим образом (Рис. 10):

- n слоев из комбинации трех сверток и пулинга;
- глобальный усредняющий пулинг вдоль каналов;
- последовательность полносвязных слоев

В качестве функций активаций используется GELU (Gaussian Error Linear Units function)[9].

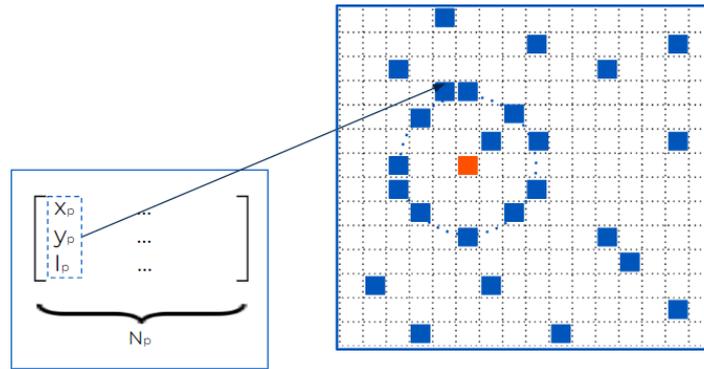


Рисунок 9 — Вид входных данных

Полученный эмбединг подается на вход другой нейросети со следующей архитектурой (Рис. 11):

- несколько билинейных слоев, принимающих на вход координаты и выход предыдущего слоя;
- последовательность полносвязных слоев.

В качестве функций активаций так же используется GELU.

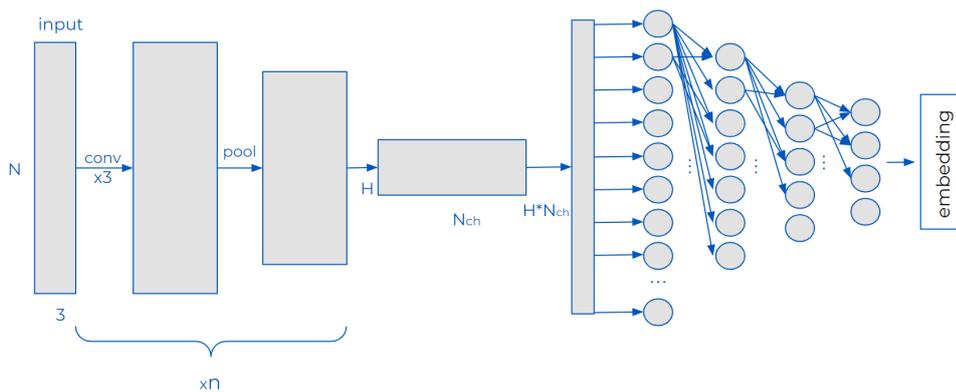


Рисунок 10 — Первый вариант архитектуры

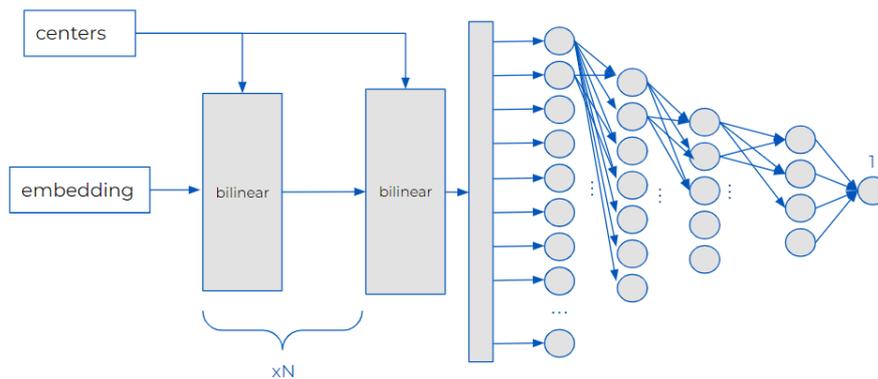


Рисунок 11 — Нейросеть для регрессии

4.3.2 Обучение и результаты

Обучение проводилось на сгенерированных данных размерности 12000 образцов, валидационная выборка имела размерность 3000 образцов. Во время обучения использовался оптимизатор NAdam с различными скоростями в диапазоне от $1 \cdot 10^{-6}$ до $5 \cdot 10^{-2}$. В качестве функции ошибки брались две функции:

- среднеквадратичная ошибка (MSE) $L = (y_{predict} - y_{target})^2$;
- средняя абсолютная ошибка (MAE) $L = |y_{predict} - y_{target}|$.

Критерием останова было либо рост ошибки на валидационной выборке, либо отсутствие изменения ошибки в течение трех эпох.

В качестве целевых переменных выбирались:

- скорость v ;
- логарифм разности $1 - v$;
- угол раствора конуса θ_c ;
- эксцентриситет e и угол падения исследуемой частицы на поверхность аэрогеля θ .

Были получены следующие результаты:

- при обучении на предсказание угла раствора конуса нейросеть выдавала относительно корректные результаты, но предсказать по ним массу частицы (даже при абсолютно точном значении импульса) оказалось практически невозможно (Рис. 12)
- в остальных ошибка падает до некоторой постоянной величины и дальше не меняется;
- выдаваемые значения — константа, примерно равная либо среднему по выборке (для MSE), либо медиане по выборке (MAE).

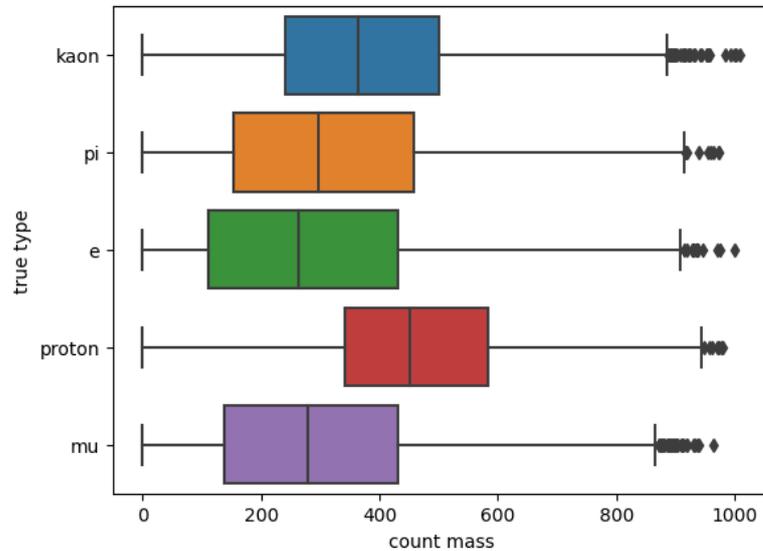


Рисунок 12 — Диаграмма размаха посчитанных масс по предсказаниям угла раствора конуса (входные данные — облако точек)

Ниже приведен график изменения ошибки в ходе обучения нейросети для предсказания угла раствора конуса (Рис. 13)

Для улучшения качества данной модели возможно предпринять следующие шаги:

- 1) построить граф, каким-либо образом задав взаимосвязи между точками, и далее работать уже с ним;
- 2) учесть локальные окрестности каждой точки.

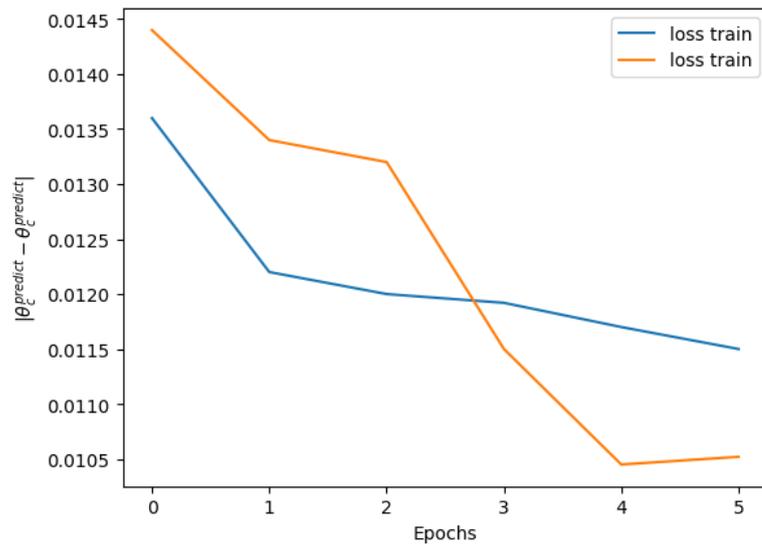


Рисунок 13 — график MAE в ходе обучения нейросети для предсказания угла раствора конуса

4.4 ИЗОБРАЖЕНИЕ

Классический подход, в котором на вход подается матрица размера $H \times W$ (Рис. 14), где каждое число обозначает интенсивность пикселя. В данной задаче $H = 264$, $W = 264$.

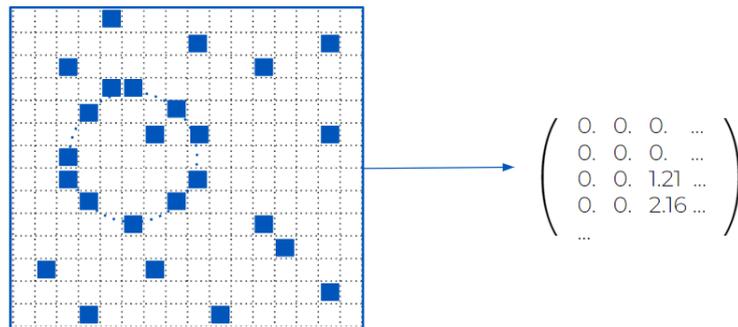


Рисунок 14 — Вид входных данных

4.4.1 Архитектура нейросети

В общем случае архитектура сети выглядит следующим образом (Рис. 15):

- N слоев из комбинации модуля «Inseption» [10] и пулинга;
- сверточный слой, на выходе которого получается эмбединг;
- в дальнейшем последовательными операциями

В качестве функций активаций используется GELU. Полученный эмбединг подается на вход другой нейросети с такой же архитектурой, как и в предыдущем варианте (Рис. 11).

4.4.2 Обучение и результаты

Обучение проводилось на сгенерированных данных размерности 12000 образцов, валидационная выборка имела размерность 3000 образцов. Во время обучения использовался оптимизатор NAdam

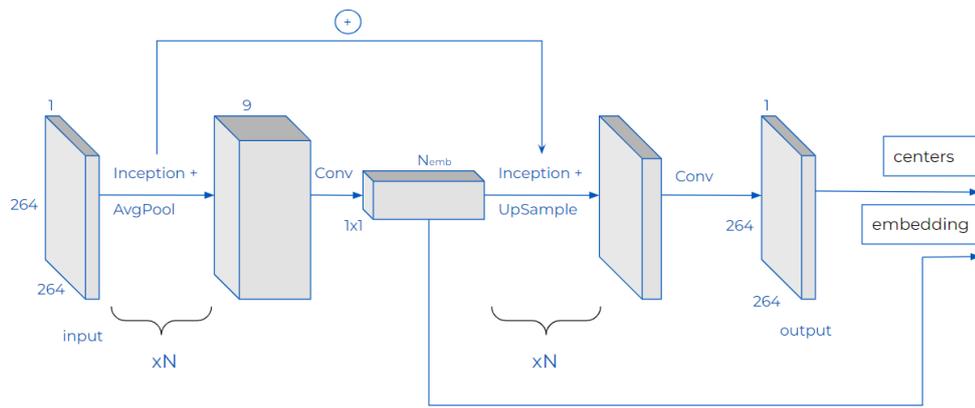


Рисунок 15 — Второй вариант архитектуры

с различными скоростями в диапазоне от $1 \cdot 10^{-6}$ до $5 \cdot 10^{-2}$. В качестве функции ошибки брались среднеквадратичная ошибка и средняя абсолютная ошибка:

- (MSE – Mean Squared Error): $L = (y_{predict} - y_{target})^2$;
- (MAE – Mean Absolute Error): $L = |y_{predict} - y_{target}|$.

Критерием останова было либо рост ошибки на валидационной выборке, либо отсутствие изменения ошибки в течение трех эпох.

В качестве целевой переменной для регрессии выбирались:

- скорость v ;
- логарифм разности $1 - v$;
- угол раствора конуса θ_c ;
- эксцентриситет e и угол падения исследуемой частицы на поверхность аэрогеля θ .

На регрессии были получены следующие результаты:

- ошибка падает до некоторой постоянной величины и дальше не меняется;
- выдаваемые значения — константа, примерно равная либо среднему по выборке (для MSE), либо медиане по выборке (MAE).

Ниже приведен график изменения ошибки в ходе обучения нейросети для предсказания угла раствора конуса (Рис. 16)

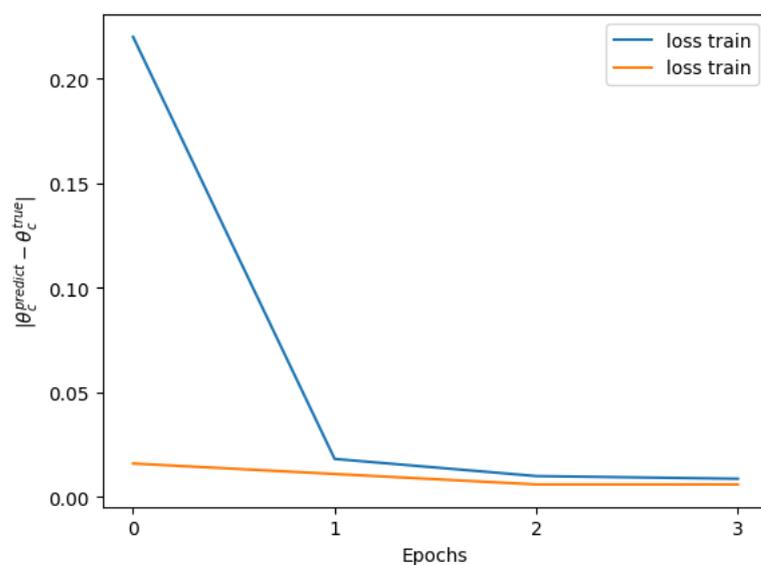


Рисунок 16 — график MAE в ходе обучения нейросети для предсказания угла раствора конуса

4.5 РАЗБИЕНИЕ ВХОДНОГО ИЗОБРАЖЕНИЯ

В основе этого подхода лежит следующая идея: можно натренировать несколько небольших нейросетей работать с изображением одиночного эллипса, а затем параллельно применять их к изображению. Агрегируя некоторым образом выходы отдельных моделей, можно находить необходимые зависимости. Плюсы данного подхода в том, что сильно отдаленные точки, не относящиеся к данному эллипсу, почти или совсем не будут влиять на предсказание параметров данного эллипса.

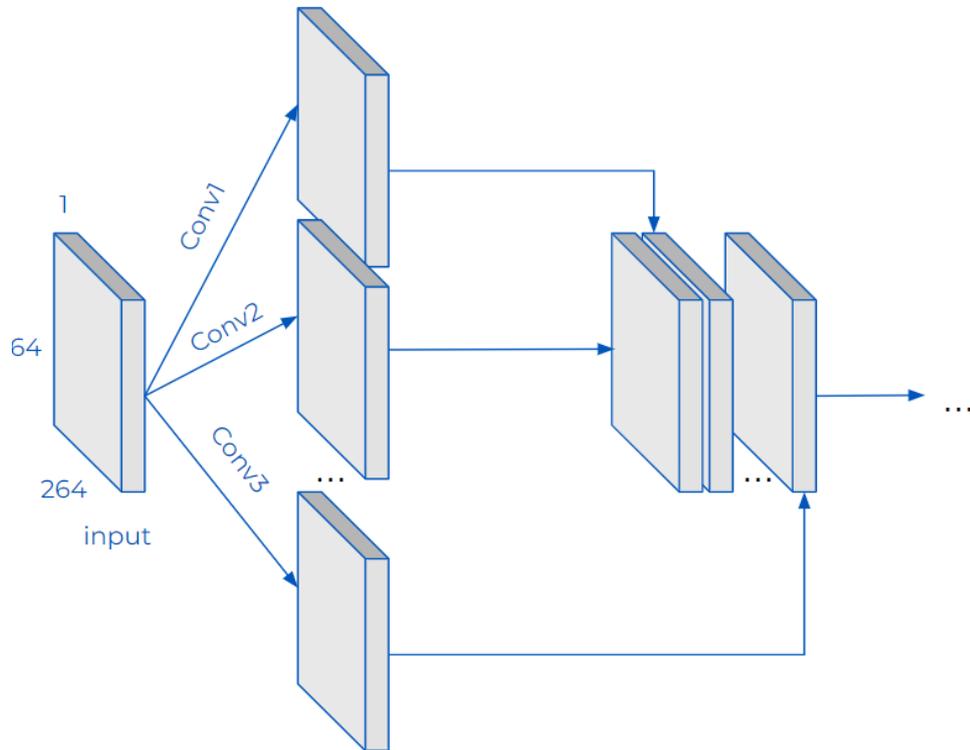


Рисунок 17 — Третий вариант архитектуры

ЗАКЛЮЧЕНИЕ

В ходе работы намечены 3 варианта решения задачи распознавания колец в детекторе FARICH. Сделаны попытки реализовать первые два варианта. Из полученных результатов можно сделать вывод о невозможности использовать эти подходы на данном этапе реализации. Необходима существенная перестройка архитектуры.

Описан третий путь решения задачи. В дальнейшей работе планируется развить третий способ, а также перейти к задаче нахождения центров эллипсов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. RICH with multiple aerogel layers of different refractive index / S. Korpar [и др.] // Nucl. Instrum. Meth. A / под ред. J. Engelfried, G. Paic. — 2005. — Т. 553.
2. *Лебедев С. А.* Математическое обеспечение для реконструкции колец черенковского излучения и идентификации электронов в RICH детекторе эксперимента CBM : дис. ... канд. / Лебедев С. А. — Объединённый институт ядерных исследований, 2011.
3. *Hough V P. C.* Method and means for recognizing complex patterns. — 1962.
4. *Chernov N. I., Ososkov G. A.* Effective Algorithms of Circle Fitting // Comput. Phys. Commun. — 1984. — Т. 33. — С. 329–333.
5. *Lebedev S. A., Ososkov G. A.* Fast algorithms for ring recognition and electron identification in the CBM RICH detector // Phys. Part. Nucl. Lett. — 2009. — Т. 6. — С. 161–176.
6. Recent developments in Geant4 / J. Allison [и др.] // Nucl. Instrum. Meth. A. — 2016. — Т. 835. — С. 186–225.
7. Geant4 developments and applications / J. Allison [и др.] // IEEE Transactions on Nuclear Science. — 2006. — Т. 53, № 1. — С. 270–278.
8. GEANT4—a simulation toolkit / S. Agostinelli [и др.] // Nucl. Instrum. Meth. A. — 2003. — Т. 506. — С. 250–303.
9. *Hendrycks D., Gimpel K.* Gaussian Error Linear Units (GELUs) // arXiv e-prints. — 2016. — arXiv:1606.08415. — arXiv: 1606.08415 [cs.LG].
10. Going Deeper with Convolutions / C. Szegedy [и др.] // arXiv e-prints. — 2014. — arXiv:1409.4842. — arXiv: 1409.4842 [cs.CV].