

ЯДЕРНАЯ ФИЗИКА,
ЭЛЕМЕНТАРНЫЕ ЧАСТИЦЫ И ФИЗИКА ВЫСОКИХ ЭНЕРГИЙ

УДК 53.072

АНАЛИЗ ИЗОБРАЖЕНИЙ RICH-ДЕТЕКТОРА
С ПОМОЩЬЮ НЕЙРОННОЙ СЕТИ

Н. И. Старков

В работе рассмотрена возможность обработки сигналов от RICH-детектора с помощью нейронной сети при прохождении через него частицы высокой энергии под углом к чувствительной плоскости. Показано, что в этом случае для определения заряда, скорости и угла частицы требуются только координаты инициированных ФЭУ.

Ключевые слова: RICH-детектор, обработка изображений RICH-детектора, применение нейронной сети, моделирование работы RICH-детектора.

Введение. В последние десятилетия широкое применение в физике высоких энергий и элементарных частиц нашли нейронные сети. Их применение началось ещё в 90-е годы. С их помощью разделялись кварк-глюонные струи [1], идентифицировались частицы в конечном состоянии [2], осуществлялась режекция нейтронов и гамма-квантов [3]. В современных экспериментах с их помощью реконструируются потоки частиц в детекторе CMS [4] и струи в детекторе ATLAS [5], изучается образование парных частиц на БАК [6] и т. д. В настоящей работе будет показано, что использование нейронной сети позволяет существенно сократить количество необходимых данных при анализе изображений RICH-детектора. В частности, для определения заряда, скорости и угла падения частицы достаточно знать положение сработавших ФЭУ и полную амплитуду сигнала от них.

Принцип работы RICH-детектора. В данной работе представлены результаты анализа изображений RICH-детектора (Ring Imaging Cherenkov detector), предназначенного для получения характеристик заряженной частицы – заряда, скорости и угла относительно плоскости детектора. Как известно, действие RICH-детектора основано на регистрации черенковского излучения, которое возникает при движении заряженной

частицы через среду со скоростью, большей скорости света в данной среде. Условием возникновения черенковского излучения является соотношение $v = \beta \cdot c > c/n$, где v – скорость частицы, β – её относительная скорость, c – скорость света, n – коэффициент преломления среды. Излучаемый свет распространяется по поверхности конуса с углом к оси конуса $\gamma = \arccos(1/(\beta n))$.

Упрощённая схема RICH-детектора изображена на рис. 1. Конструкция детектора включает радиатор – тонкую пластину из прозрачного материала, в которой генерируется черенковское излучение. Дифференциальное количество фотонов с длиной волны λ , испускаемых частицей с зарядом Z на единице длины пути в материале, выражается формулой

$$\frac{d^2N}{d\lambda dx} = \frac{2\pi}{137} \frac{Ze^2}{\lambda^2} \left[1 - \frac{1}{\beta^2 n^2(\lambda)} \right]. \quad (1)$$

При постоянстве показателя преломления $n = 1.334$, что практически выполняется для NaF в диапазоне от $\lambda_1 = 300$ до $\lambda_2 = 650$ нм, для частицы с $Z = 1$ и $\beta = 1$ можно получить следующую оценку для числа фотонов:

$$N = \frac{2\pi}{137} \left(1 - \frac{1}{n} \right) \cdot \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} \right) \cong 360 \text{ см}^{-1}. \quad (2)$$

На пути излучения располагается регистрирующее устройство – фотоприёмник, как правило, это электронные фотоумножители, которые заполняют плоскую матрицу. На границе радиатора фотоны испытывают преломление, частичное отражение и падают на фотоприёмник. В результате этого на матрице ФЭУ ими формируется фигура определенной формы, образованная иницированными ФЭУ. При углах падения частицы, близких к перпендикулярю к плоскости матрицы ФЭУ (под углом $\theta > 78\text{--}82^\circ$ к плоскости, в зависимости от скорости), эта фигура замкнутая (квазиэллипс), при меньших θ – разомкнутая, вследствие полного отражения на границе радиатора. Часть фотонов, попавших в область чувствительности ФЭУ, вызывает выбивание фотоэлектронов и его срабатывание.

В данной работе при моделировании в качестве радиатора взята пластина NaF площадью 1 м^2 толщиной 1 см. Выбор NaF-излучателя с относительно низким показателем преломления $n \approx 1.33$ [5] позволяет минимизировать потери черенковского света из-за полного внутреннего отражения на поверхности излучателя в случае наклонных траекторий частиц. Матрица фотоприемника расположена на расстоянии 10 см от радиатора и выполнена в виде плотной гексагональной упаковки из более чем 1000 торцевых фотоэлектронных умножителей марки ФЭУ-140. Диаметр входного окна ФЭУ 25 мм,

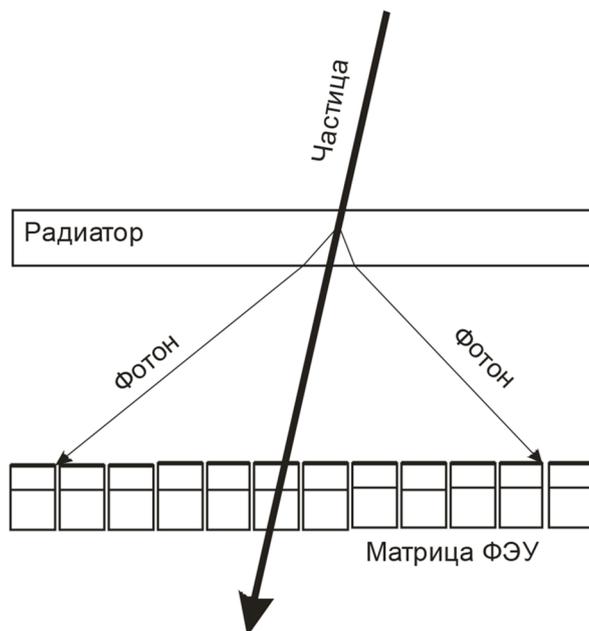


Рис. 1: Схема действия RICH-детектора.

наружный диаметр 32 мм. Использование ФЭУ со стеклянным окном приводит к ограничению спектрального диапазона обнаружения фотонов длинами волн 300–650 нм (при которых $n \approx \text{const}$). Это позволяет значительно уменьшить влияние рассеивания света в NaF в ультрафиолетовой области.

Моделирование отклика детектора. События, необходимые для анализа, были получены в результате Монте-Карло моделирования прохождения частицы через RICH-детектор. С этой целью была разработана программа, в которой были учтены практически все явления, влияющие на конечный результат – срабатывание ФЭУ: поглощение, отражение и преломление света в радиаторе и окнах ФЭУ, спектр черенковских фотонов, спектральная чувствительность ФЭУ, вероятность срабатывания ФЭУ при попадании фотонов на его входное окно.

Целью использования RICH-детектора является получение информации о характеристиках частицы – заряд, скорость, угол падения. Обычно для их получения требуется дополнительная информация от других частей установки, в которую RICH-детектор входит как отдельный элемент. В настоящей работе показано, что при использовании нейронной сети (НС) для получения перечисленной выше информации требуется только знание координат и амплитуды сигнала инициированных ФЭУ или их числа.

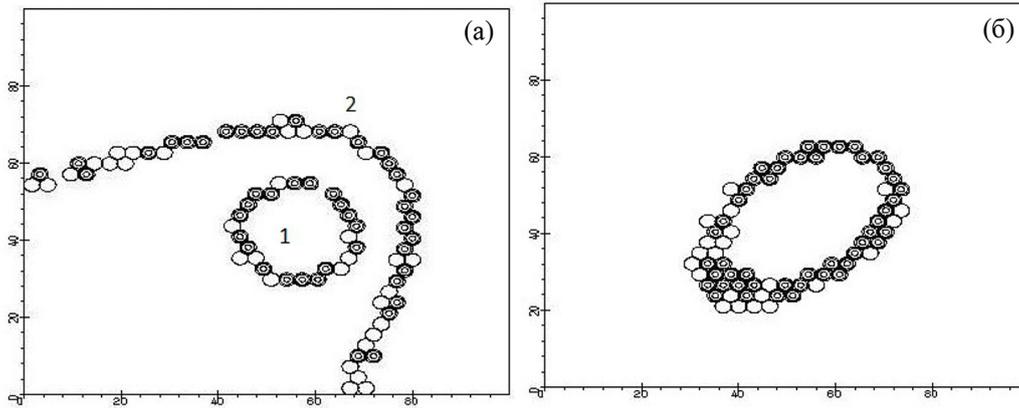


Рис. 2: Изображение на чувствительной плоскости RICH-детектора, созданное засвеченными (пустые кружки) и иницированными (затемненные кружки) ФЭУ при прохождении ядра ${}^4\text{He}$ через установку: (а) $\beta = 0.85$, $\theta = 90^\circ$ (1), $\beta = 0.97$, $\theta = 78^\circ$ (2); (б) $\beta = 0.97$, $\theta = 82^\circ$. Угол $\varphi = 45^\circ$.

Принцип работы нейросети. В данной работе была использована многослойная нейронная сеть – многоуровневый перцептрон, с обратным распространением ошибки. Она используется для установления нелинейного многопараметрического отображения одного многомерного пространства векторов (вектор входа V_{in} , пространство признаков) на другое (вектор выхода V_{out} , пространство значений) с помощью обучения сети. Это означает, что, будучи подготовленная к работе, при подаче на вход сети одного из векторов V_{in} , она должна дать на выходе соответствующий вектор V_{out} .

Устройство и принцип работы нейронных сетей описаны во многих работах (см., напр., [7–9]). Поэтому здесь мы приведем только упрощенную схему её действия.

Нейронная сеть такого типа состоит из трех или большего числа последовательных слоев, каждый из которых включает некоторое количество нейронов. Первый и последний слой являются входом и выходом, соответственно, поэтому количество нейронов в них должно быть равно размерностям соответствующих пространств. Промежуточные слои носят название скрытых. Каждый нейрон одного слоя имеет связь с каждым нейроном предыдущего и последующего слоев. Отдельный нейрон характеризуется числовым значением u (которое, как правило, входит в интервал $[0,1]$), а каждая связь – своим весом w . Нейроны первого слоя имеют значения, равные компонентам входного вектора. Значения нейронов последующих слоев u_j являются нелинейной функцией значений нейронов предыдущего слоя u_i и соответствующих весов w_{ij}

$$u_j = F(\sum w_{ij}u_i), \quad (3)$$

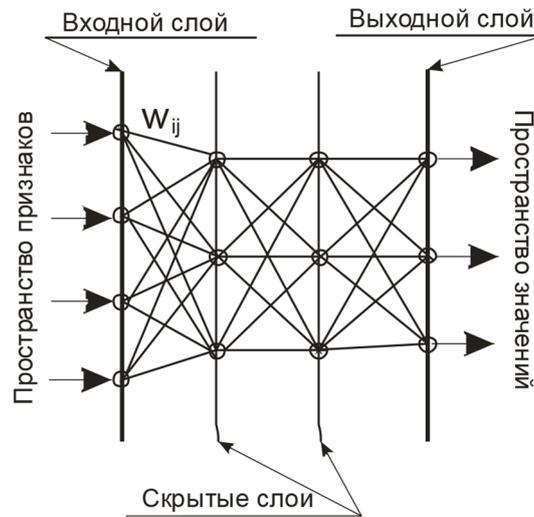


Рис. 3: Схема устройства нейронной сети.

где $F(x) = \frac{1}{1 + \exp(-x)}$ — так называемая функция активации, которая ограничивает значения u интервалом $(0, 1)$.

Работа сети протекает в два этапа. На первом из них происходит ее обучение. Для этого с помощью экспериментов или их программных имитаторов формируется обучающее множество $\{V_{in}; V_{out}\}$, которое представляет собой набор пар входного и соответствующего выходного векторов. Размер обучающего множества (количество пар) зависит от сложности анализируемого отображения и общего числа слоев и нейронов на них. Входные вектора включают непосредственно наблюдаемые величины (показания приборов или их комбинации), а выходные — физические характеристики процесса, которые требуется определить в результате анализа (типы частиц, их заряды, скорости и т. д.). В начале обучения значения w НС иницируются небольшими случайными значениями, и на ее вход по очереди подаются значения входных векторов из обучающего множества. Порядок следования векторов определяется случайным образом, чтобы предотвратить “привыкание” сети к их фиксированной последовательности. С помощью нелинейного преобразования из значений u_j каждого из нейронов входного слоя и величин w_{ji} формируются значения u_i нейронов второго слоя. Затем эти значения используются для вычисления третьего слоя. Такая процедура повторяется для всех слоев вплоть до последнего — выходного. Значения нейронов последнего слоя u_j сравниваются с соответствующими значениями выходного вектора v_i из обучающего множества и вычисляется ошибка $\Delta_j = u_j - v_j$. По величине ошибки с помощью алгоритма обрат-

ного распространения ошибки вычисляются поправки к w_{ij} – значениям весов связей последнего и предпоследнего слоев. Для этого используется метод градиентного спуска, задаваемый производной функции активации $F'(x) = F(x)(1 - F(x))$

$$\Delta w_{ij} = \alpha F'(\Delta_j), \quad (4)$$

где коэффициент α определяет масштаб величины поправки.

По этим поправкам последовательно вычисляются поправки к весам более ранних слоев вплоть до первого скрытого слоя. Затем происходит подача на вход следующего входного вектора, и процедура уточнения весов повторяется до получения приемлемого результата, когда разность между значением нейрона выходного слоя и соответствующей величиной из пространства значений станет достаточно малой. Процедуру обучения, охватывающую весь обучающий набор, принято называть эпохой. По мере обучения ошибка, как правило, уменьшается, и процесс сходится, при этом весь цикл обучения может занять не один десяток, а то и сотни эпох. Успех обучения зависит от целого ряда причин, в том числе от удачного выбора пространства признаков и достаточного объема обучающего множества. Упрощая, можно сказать, что процесс обучения является подгоночной процедурой для описания многопараметрической функцией (параметрами являются величины w_{ij}) упомянутого выше отображения пространства входных векторов на пространство выходных.

На эффективность работы и сходимости процесса обучения НС большое влияние оказывает правильный выбор входных сигналов. Они должны быть такими, чтобы в многомерном пространстве – пространстве признаков, образуемом ими, области с различными значениями выходных векторов разделялись как можно лучше. После обучения проводится процедура верификации с набором входных и выходных векторов, отличным от использованного при обучении. Второй этап представляет собой непосредственную работу обученной НС по анализу физической информации.

Применение нейросети при анализе данных RICH-детектора. Наблюдаемая частица характеризуется набором физических величин: тип частицы, заряд, скорость, направление движения (углы θ и φ). Исходная информация RICH-детектора, из которой можно построить пространство входных векторов, включает количество инициированных ФЭУ и их координаты. Теоретически возможен вариант установки, когда доступна информация об амплитуде сигнала от каждого ФЭУ, но он сложен в технической реализации и здесь рассматриваться не будет.

При обсуждении возможностей нейронной сети ограничимся рассмотрением углов θ , больших 82° , при которых фигура на матрице ФЭУ остается замкнутой. На самом деле

НС можно применить и в случае незамкнутой фигуры при небольшой модификации, связанной с описанием ее геометрии.

Теоретически фигура, создаваемая черенковским светом на светочувствительной матрице симметрична относительно проекции траектории на плоскость матрицы. Это свойство можно было бы использовать для идентификации азимутального угла φ . Однако не все ФЭУ, на которые падают фотоны, иницируются. Кроме того, рассматриваемая конструкция обладает большой дискретностью чувствительной области, которая занимает только 55% площади детектора, что приводит к ломаному контуру фигуры и непостоянству ее формы. Особенно сильно этот эффект проявляется при малых скоростях v , когда размер фигуры мал. Вследствие этого симметрия фигуры в значительной степени нарушается и угол φ определяется с большой погрешностью, которая сильно снижает надежность определения остальных характеристик частицы. По этой причине этот угол не участвовал в выходном наборе.

Формула (1) показывает, что амплитуда сигнала очень чувствительна к заряду частицы. Однако анализ свойств RICH-детектора показал, что для определения величины заряда (вплоть до значений $q = 38-40$) недостаточно знания полной амплитуды снимаемого сигнала N_a (количество фотонов, инициировавших ФЭУ) или полного количества инициированных ФЭУ $N_{\text{ФЭУ}}$, так как эти величины могут быть одинаковыми у частиц с разными зарядами при разных скоростях. Кроме того, при большом количестве черенковских фотонов, то есть при больших зарядах и скоростях, вероятность срабатывания ФЭУ близка к единице, что приводит к насыщению величины N_a и снижению чувствительности RICH-детектора к заряду.

Анализ показал, что геометрические параметры фигуры (длина и ширина) содержат достаточно полную информацию. В итоге в качестве одного из основных параметров, определяющих скорость, был выбран минимальный размер фигуры Y_{max} (ширина, “малая ось” квазиэллипса), который также очень чувствителен к углу θ , однако он испытывает сильный разброс вследствие дискретности чувствительной области. Другим возможным параметром может служить величина максимального размера фигуры X_{max} (длина, “большая ось” квазиэллипса). Этот параметр чувствителен к величине угла θ , однако он испытывает сильный разброс вследствие дискретности чувствительной области и поэтому фиксирует угол с большой ошибкой.

В итоге, после проведенного анализа в качестве входного вектора был выбран набор $(N_a, X_{\text{max}}, Y_{\text{max}})$, а в качестве выходного – набор (q, β, θ) . На рис. 4(а) показаны распределения величины N_a , а на рис. 4(б) – зависимость Y_{max} от скорости, полученные

в результате розыгрыша 2400 событий для трех ядер (${}^4\text{He}$, ${}^9\text{Be}$, ${}^{12}\text{C}$) при 10 значениях скорости и 8 значениях угла θ . Видно, что в пространстве (N_a, Y_{\max}) заряды и скорости довольно хорошо пространственно разделены, что обеспечивает возможность решения поставленной задачи с помощью НС.

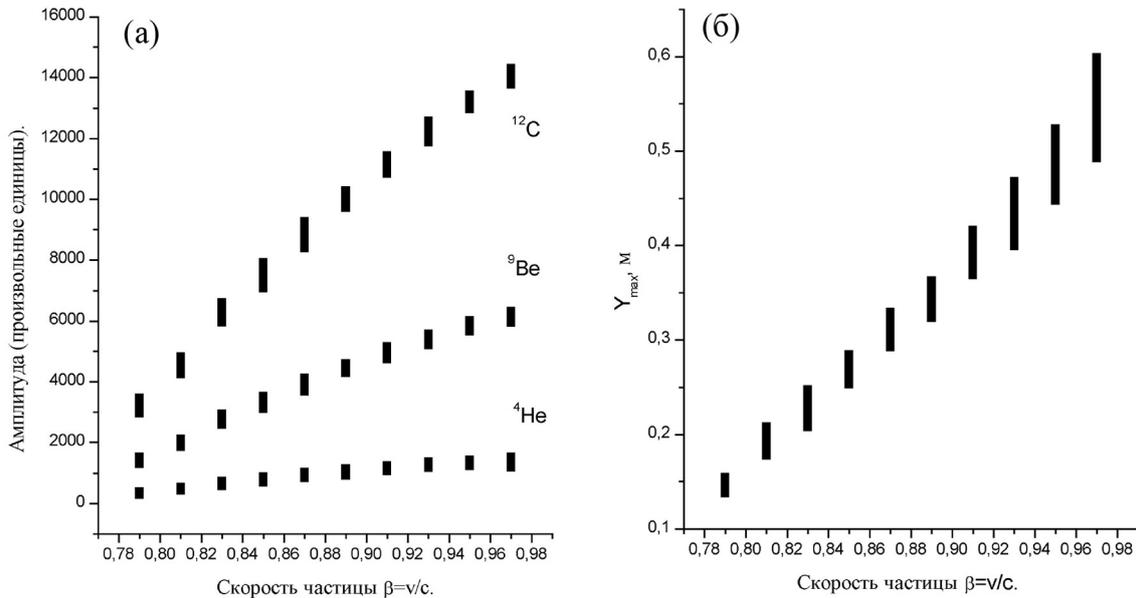


Рис. 4: (а) зависимость от скорости частицы величины сигнала от иницированных ФЭУ для трех ядер. Приведены точки обучающего множества, всего 2400 событий: 3 ядра, 10 скоростей, 8 углов. По 10 событий при каждом наборе значений; (б) зависимость от скорости частицы величины Y_{\max} фигуры, создаваемой иницированными ФЭУ. Приведены точки обучающего множества.

Для имитации работы НС была создана программа, реализующая алгоритм обратного распространения ошибки. В качестве обучающего множества был взят упомянутый выше набор из 2400 событий. В процессе выбора оптимальной конфигурации НС рассматривались варианты с различным числом скрытых слоев и нейронов на них. Увеличение их числа приводит к росту времени обучения и требует большего объема обучающего множества, а уменьшение ухудшает качество описания. Окончательный вариант НС включал три скрытых слоя по 10 нейронов на каждом. Обучение по 2400 образцам потребовало около 2000 эпох.

На рис. 4 показаны диапазоны величин амплитуды сигнала ФЭУ (а) и параметра Y_{\max} (б), полученные при работе имитационной программы розыгрыша событий.

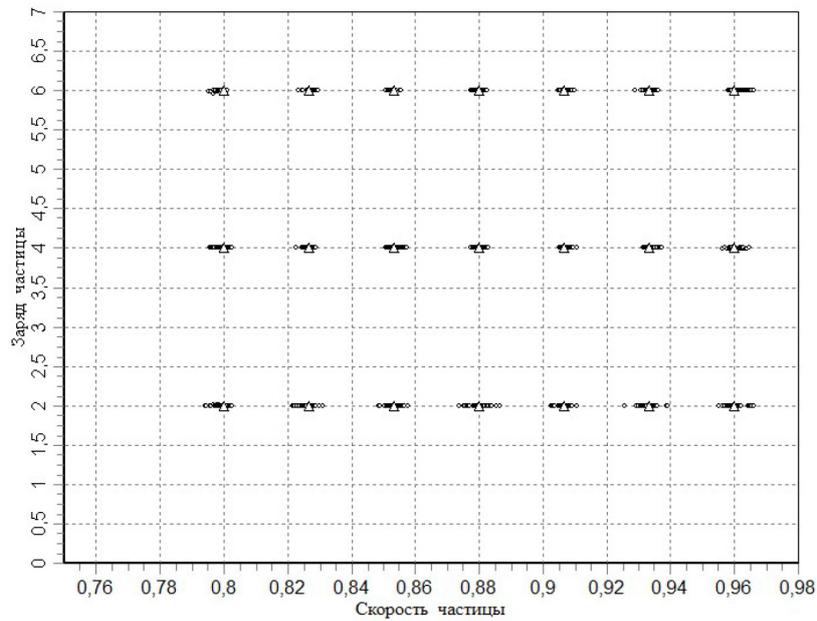


Рис. 5: Разброс скоростей частиц β , полученный в результате работы сети с контрольным набором из 840 событий.

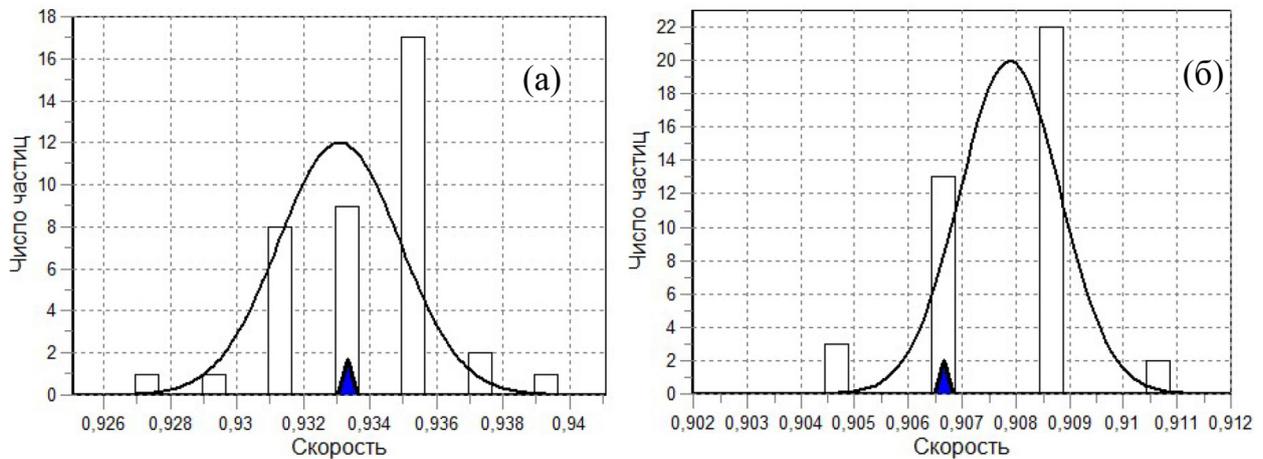


Рис. 6: Распределение скоростей в точках контрольного набора: (а) с наибольшим ($\beta = 0.934 \pm 0.003$); (б) с наименьшим ($\beta = 0.9069 \pm 0.0012$) разбросом. Тёмный треугольник отмечает истинное значение скорости. Кривые – аппроксимация функцией Гаусса.

Для проверки качества работы НС был сгенерирован тестовый набор входных и выходных векторов, включающий 840 событий для тех же ядер, но с промежуточными значениями скорости и угла θ (3 заряда, 7 скоростей и 5 углов по 8 точек для каждого

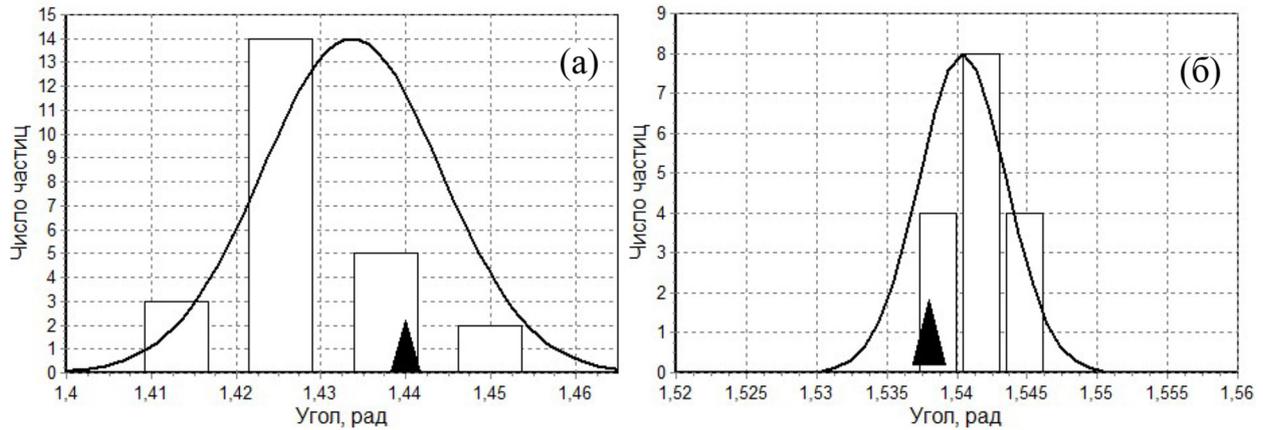


Рис. 7: Распределение углов в точках контрольного набора: (а) с наибольшим ($\beta = 0.8, \theta = 1.44$); (б) с наименьшим ($\beta = 0.907, \theta = 1.538$) разбросом. Тёмный треугольник отмечает истинное значение скорости. Кривые – аппроксимация функцией Гаусса.

набора). Разброс полученных величин скоростей приведён на рис. 5 вместе с контрольными значениями, обозначенными треугольником. Видно, что заряды частиц, будучи дискретными, определяются практически с абсолютной точностью: $Z_{He} = 2 \pm 0.005$, $Z_{Be} = 4 \pm 0.004$, $Z_C = 6 \pm 0.002$.

Точность определения скорости и угла θ частиц колеблется в интервале от 0.1% до 0.3% при 65% уровне достоверности. На рис. 6(а) приведено распределение в одной точке ($q = 2, \beta = 0.933$), где разброс найденных значений β на рис. 5 наибольший: $\beta = 0.934 \pm 0.003$. Темным треугольником обозначено положение истинного значения β . На рис. 6(б) показано распределение скоростей в точке с ошибкой 0.0015 ($q = 6, \beta = 0.9069 \pm 0.0013$, истинное значение $\beta = 0.9067$).

Для увеличения статистики результаты по углу собирались по всем трём зарядам. Разброс распределений для углов наибольший при малых скоростях, что объясняется влиянием дискретности матрицы фотоприёмника и увеличением размера фигуры с ростом скорости. На рис. 7(а) приведено распределение угла θ с наибольшим разбросом точек при угле 1.44 радиана ($\theta = 1.43 \pm 0.02$), а на рис. 7(б) – с наименьшим разбросом точек при угле $\theta = 1.538$ ($\beta = 0.907, \theta = 1.538$).

Заключение. Таким образом, в данной работе показано, что применение нейронной сети намного повышает эффективность обработки изображений RICH-детектора, поскольку не требует дополнительной информации от других элементов регистрирующей установки. При этом полученные результаты имеют достаточно высокую точность.

Работа поддержана Российским научным фондом по программе “Проведение фундаментальных научных исследований и поисковых научных исследований отдельными научными группами” (проект № 23-12-00054).

Л И Т Е Р А Т У Р А

- [1] M. F. A. Graham, L. M. Jones, S. Herbin, Phys. Rev. **51**, 4789 (1995). <https://doi.org/10.1103/PhysRevD.51.4789>.
- [2] R. Bellotti, M. Castellano, C. De Marzo, et al., Nucl. Instr. Meth. **A 350**, 556 (1994). [https://doi.org/10.1016/0168-9002\(94\)91257-2](https://doi.org/10.1016/0168-9002(94)91257-2).
- [3] P. Pavlopoulos, G. Polivka, S. Vlachos, H. Wendler, Nucl. Instr. Meth. **A 359**, 566 (1995). [https://doi.org/10.1016/0168-9002\(95\)00073-9](https://doi.org/10.1016/0168-9002(95)00073-9).
- [4] A. M. Sirunyan, A. Tumasyan (on behalf of CMS Collaboration), J. Instrum. **12**, P10003 (2017). <https://doi.org/10.1088/1748-0221/12/10/P10003>.
- [5] M. Aaboud, G. Aad (on behalf of ATLAS Collaboration), Eur. Phys. J. C **77**, 466 (2017). <https://doi.org/10.1140/epjc/s10052-017-5031-2>.
- [6] Murat Abdughani, Jie Ren, Lei Wu, Jin Min Yang, J. High Energy Phys. **8**, 55 (2019). [https://doi.org/10.1007/JHEP08\(2019\)055](https://doi.org/10.1007/JHEP08(2019)055).
- [7] C. Peterso, T. Rönvaldsson, An Introduction to Artificial Neural Networks, Proceedings of the 1991 Cern School of Computing, (Ystad, Sweden, 1992) CERN Yellow Report 92-02, LU-TP-91-23; LUTP-91-23, 74 p. <https://doi.org/10.5170/CERN-1992-002.113>.
- [8] R. Callan, *The Essence of Neural Network* (Prentice Hall Europe, 1999). ISBN: 013908732X, 9780139087325. – 232 p.
- [9] L.V. Fausett, *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications* (Pearson Education, 2006). ISBN: 8131700534, 9788131700532. – 467 p.

Поступила в редакцию 26 марта 2025 г.

После доработки 9 июня 2025 г.

Принята к публикации 10 июня 2025 г.