

P10-2009-61

Т. П. Акишина, О. Ю. Денисова, В. В. Иванов

ОБ ИДЕНТИФИКАЦИИ ЭЛЕКТРОНОВ И ПИОНОВ
С ПОМОЩЬЮ МНОГОСЛОЙНОГО ПЕРЦЕПТРОНА
В ДЕТЕКТОРЕ ПЕРЕХОДНОГО ИЗЛУЧЕНИЯ
ЭКСПЕРИМЕНТА **СВМ**

Акишина Т. П., Денисова О. Ю., Иванов В. В.

P10-2009-61

Об идентификации электронов и пионов с помощью многослойного перцептрона в детекторе переходного излучения эксперимента CBM

Рассмотрена задача идентификации электронов и пионов по их потерям энергии в детекторе переходного излучения (TRD) эксперимента CBM. Для идентификации частиц использовалась искусственная нейронная сеть — многослойный перцептрон, реализованный в пакетах JETNET и ROOT. В работе показано, что для получения корректных и сопоставимых результатов важно правильно выбрать структуру сети, и приводятся соответствующие рекомендации. Для достижения приемлемого уровня подавления пионов требуется преобразование величин потерь энергии в слоях TRD к более «эффективным» переменным. Получена зависимость выходного порога нейронной сети от импульса регистрируемой частицы для фиксированного уровня потери электронов.

Работа выполнена в Лаборатории информационных технологий ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна, 2009

Akishina T. P., Denisova O. Yu., Ivanov V. V.

P10-2009-61

On Electron and Pion Identification Using a Multilayer Perceptron in the Transition Radiation Detector of the CBM Experiment

The problem of pion-electron identification based on their energy losses in the TRD is considered in the frame of the CBM experiment. For particles identification an artificial neural network (ANN) was used, a multilayer perceptron realized in JETNET and ROOT packages. It is demonstrated that, in order to get correct and comparable results, it is important to define the network structure correctly. The recommendations for such a selection are given. In order to achieve an acceptable level of pions suppression, the energy losses need to be transformed to more «effective» variables. The dependency of ANN output threshold for a fixed portion of electron loss on the particle momentum is presented.

The investigation has been performed at the Laboratory of Information Technologies, JINR.

Communication of the Joint Institute for Nuclear Research. Dubna, 2009

ВВЕДЕНИЕ

Экспериментальная установка CBM, создаваемая в GSI (Дармштадт, Германия) на ускорительном комплексе антипротонов и тяжелых ионов FAIR [1, 2], предназначена для изучения свойств сильно сжатой барионной материи, образующейся в высокоэнергетичных ядро-ядерных соударениях.

Набор детекторов и элементов установки CBM должен обеспечить идентификацию электронов при условии подавления пионов на уровне 10^5 , идентификацию адронов с большим акцептансом, восстановление импульсов заряженных частиц с точностью $\sim 1\%$, восстановление первичных и вторичных вершин с точностью ~ 30 мкм, высокое пространственное разрешение координатных детекторов, малое мертвое время, быстрый ответ и высокую скорость считывания информации с детекторов, высокую скорость триггера и системы сбора данных, устойчивость детекторов и электроники к радиационным условиям, устойчивость к дельта-электронам и т. д.

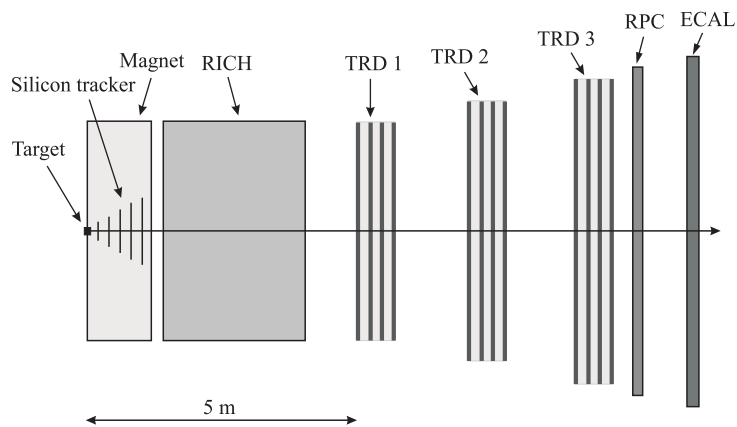


Рис. 1. Схема эксперимента CBM

На рис. 1 представлена схема экспериментальной установки CBM. Внутри дипольного магнита расположена мишень и координатная трековая система (STS — Silicon Tracking System), содержащая семь плоскостей из пиксельных и стриповых детекторов. STS совместно с дипольным магнитом используется для восстановления траекторий заряженных частиц и определения

их импульсов. Черенковский детектор (RICH — Ring Imaging Cherenkov) и детектор переходного излучения (TRD — Transition Radiation Detector) должны обеспечить надежную регистрацию электронов с импульсом выше 1 ГэВ/с. Детектор измерения времени пролета (TOF — Time-of-flight), построенный на основе резистивных плоских камер (RPC — Resistive Plate Chambers), предназначен для идентификации адронов высокой энергии. Электромагнитный калориметр (ECAL — Electromagnetic Calorimeter) используется для идентификации электронов и фотонов.

Измерение чармония — одна из основных задач эксперимента CBM. Для того чтобы зарегистрировать J/ψ -мезон при его распаде по диэлектронному каналу, необходимо решить главную задачу: надежно идентифицировать электроны в условиях доминирующего фона от пионов. Одним из наиболее эффективных детекторов для решения указанной задачи является многослойный детектор переходного излучения TRD.

TRD должен обеспечить надежную идентификацию электронов, высокий уровень подавления пионов, а также восстановление траекторий проходящих через него заряженных частиц. Требуемое подавление пионов должно быть не менее 100, пространственное разрешение около 200–300 мкм. Чтобы решить указанные задачи в условиях интенсивных потоков и высокой множественности частиц, необходима тщательная оптимизация детектора TRD.

В работе [3] задача идентификации электронов и пионов была успешно решена с помощью искусственной нейронной сети — многослойного перцептрона (МСП) из пакета JETNET [4]. В эксперименте CBM для этих же целей используется МСП [5], реализованный в среде CBM ROOT [6,7]. Сравнение эффективностей работы этих двух сетей дало заметно различающиеся результаты. В настоящей работе исследуются причины такого расхождения.

В [3] было также показано, что для достижения необходимого уровня подавления пионов требуется преобразование величин потерь энергии в слоях TRD к более «эффективным» переменным. Здесь нами получена зависимость параметров данного преобразования от импульса регистрируемой частицы. Кроме того, приводится зависимость величины порога для выходного сигнала МСП, отвечающей фиксированному уровню потери электронов, от импульса частицы.

1. ВЫБОР СТРУКТУРЫ СЕТИ

Задача идентификации электронов/пионов с помощью n -слойного TRD состоит в следующем: имея набор потерь энергии в n слоях TRD, нужно определить, какая частица, электрон или пион, была зарегистрирована детектором.

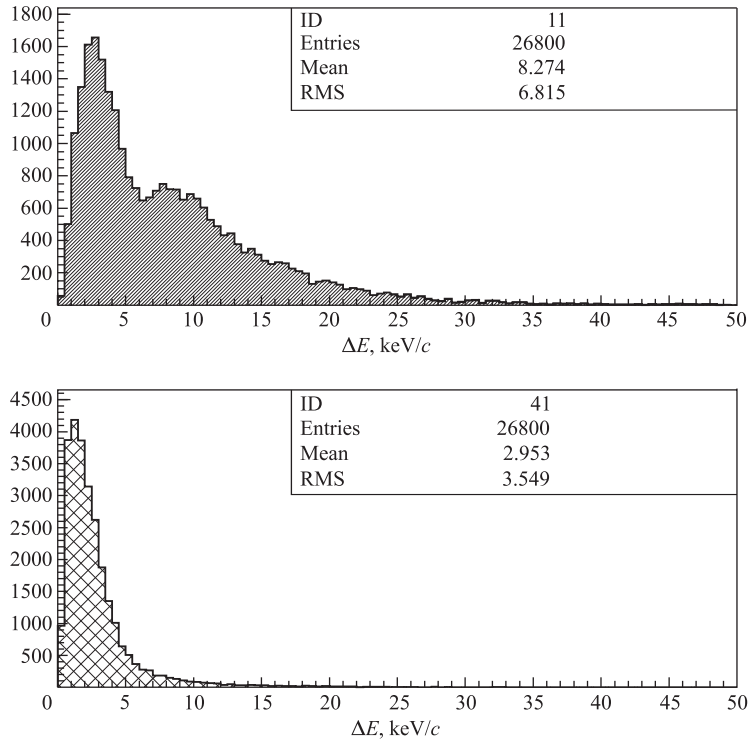


Рис. 2. Распределения потерь энергии электронов (верхняя гистограмма), включая их потери на переходное излучение, и пионов (нижняя гистограмма) в первом слое-поглотителе детектора TRD для $p = 1,5$ ГэВ/с

В качестве входной информации для сетей использовались выборки, составленные на основе потерь энергии в TRD пионами или электронами с импульсами от 1 до 13 ГэВ/с. Энергетические потери получались путем моделирования Монте-Карло прохождения частиц через n -слойный TRD с помощью пакета GEANT [8] в среде CBM ROOT.

На рис. 2 приведены распределения потерь энергии электронов (верхняя гистограмма), включая их потери на переходное излучение, и пионов (нижняя гистограмма) в первом слое-поглотителе детектора TRD для $p = 1,5$ ГэВ/с. Распределения потерь энергии в остальных слоях TRD и при других импульсах имеют аналогичный характер.

Исследовались две возможные причины различий в результатах работы рассматриваемых сетей: 1) разные методы и алгоритмы минимизации функционала ошибок (см. работу [3]); 2) различия в структуре сетей.

Проведенный анализ показал, что используемые на этапе обучения МСП методы минимизации функционала ошибок (алгоритмы Manhattan в JETNET и BFGS [9] в ROOT) не могут быть причиной заметных различий в результатах работы сетей. Вместе с тем в ходе этого анализа было обнаружено, что МСП из ROOT всегда дает слегка различающиеся результаты, т. е. сеть не воспроизводит полученные ранее результаты. А это приводит к тому, что нельзя зафиксировать порог для сигнала на выходе сети, по которому будет проводиться селекция электронов и пионов. Для того чтобы исключить данное неудобство в работе сети, необходимо в списке параметров «option» функции Train задавать «+» [10].

В большинстве случаев параметры нейронных сетей с фиксированной архитектурой (у нас это МСП) настраиваются для решения конкретной задачи, и оптимальный вариант получается на основе эмпирического подбора. Однако, как оказалось в случае нашей задачи, даже небольшие различия в структуре сетей могут приводить к заметным различиям в результатах их работы.

При выборе структуры многослойного перцептрона необходимо задать:

- 1) количество слоев МСП,
- 2) количество нейронов в каждом слое.

С количеством слоев все достаточно просто, так как для подавляющего большинства задач (наша задача относится к этому большинству), решаемых с помощью МСП, достаточно трех слоев: входного, скрытого и выходного.

Таким образом, нужно определиться с количеством нейронов в каждом слое. В обеих сетях число входных нейронов задавалось равным 12 (в соответствии с количеством слоев в TRD). Для определения типа частицы использовался один выходной нейрон. При выборе числа нейронов в скрытом слое мы исходили из минимальной конфигурации сети [11], обеспечивающей оптимальный результат, а именно, максимальный уровень подавления пионов при минимальной потере электронов.

В качестве критерия для определения числа нейронов в скрытом слое использовалось распределение ошибок — разницы между целевым значением (для пионных событий целевое значение задавалось равным -1 , а для электронных — $+1$) и выходным сигналом сети. Распределение ошибок должно было удовлетворять следующим условиям:

- быть симметричным относительно нулевого среднего,
- иметь минимальное среднеквадратичное отклонение.

На рис. 3 показаны распределения значений сигнала на выходе нейронной сети, полученные на стадии ее обучения (верхняя левая гистограмма) и тестирования (нижняя левая гистограмма); правые гистограммы показывают распределения ошибок на стадии ее обучения (верхняя гистограмма) и тестирования (нижняя гистограмма) для $p = 1,5$ ГэВ/с.

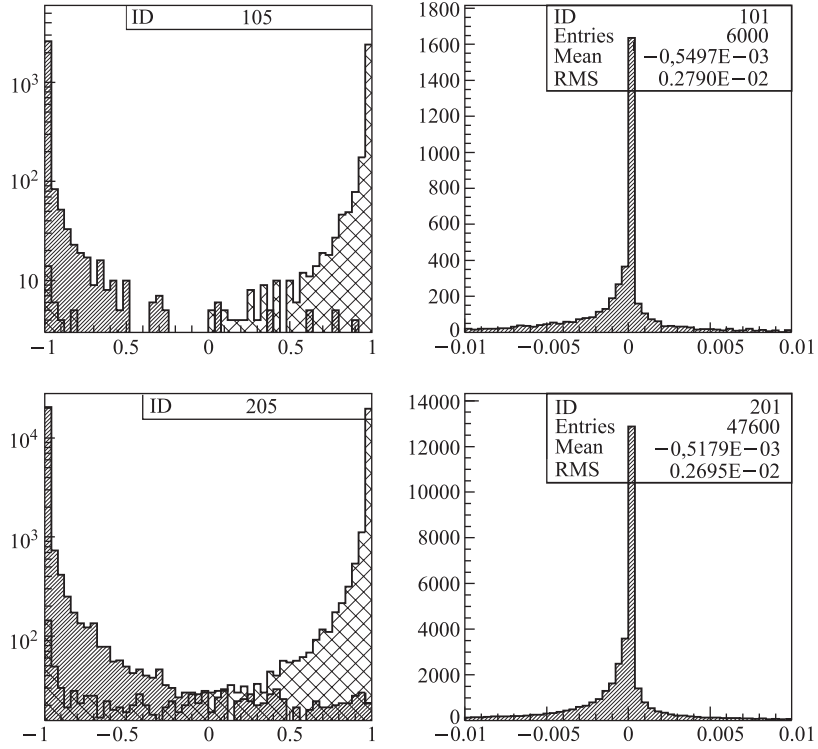


Рис. 3. Распределения значений сигнала на выходе нейронной сети, полученные на стадии ее обучения (верхняя левая гистограмма) и тестирования (нижняя левая гистограмма); правые гистограммы показывают распределения ошибок на стадии обучения (верхняя гистограмма) и тестирования (нижняя гистограмма) МСП для $p = 1,5$ ГэВ/с

2. ПОДГОТОВКА ВХОДНЫХ ДАННЫХ ДЛЯ МСП

Если в качестве входных данных для сети использовать выборки, составленные из потерь энергии электронов и пионов, то процесс обучения МСП проходит очень медленно, присутствуют большие колебания (относительно тренда) в эффективности распознавания частиц сетью.

В связи с этим к исходным данным была применена следующая процедура преобразования (см. [3, 12]):

$$\lambda_i = \frac{\Delta E_i - \Delta E_{mp}^i}{\xi_i} - 0,225, \quad i = 1, 2, \dots, n, \quad (1)$$

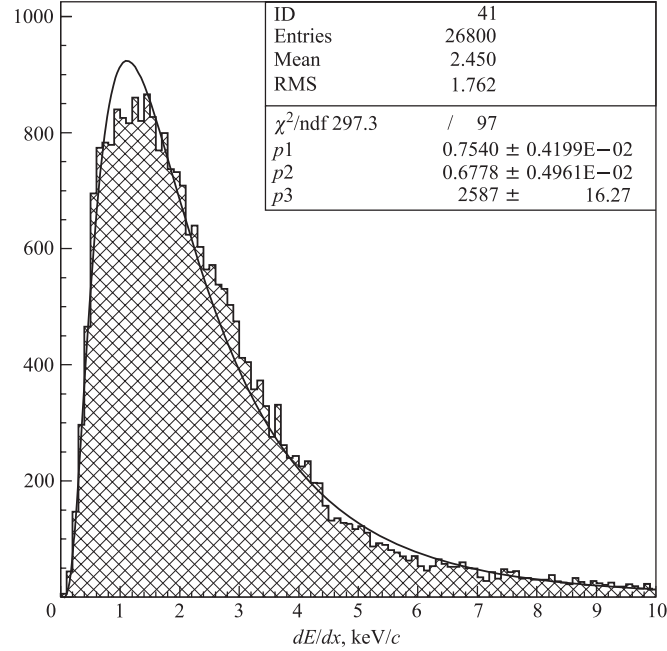


Рис. 4. Аппроксимация распределения потерь энергии пионов в первом поглотителе TRD с помощью функции плотности логнормального распределения

где ΔE_i — величина потери энергии зарегистрированной частицы (пиона или электрона) в i -м поглотителе TRD, ΔE_{mp}^i — величина наиболее вероятной потери энергии пиона в этом слое, $\xi_i = \frac{1}{4,02}$ FWHM; здесь FWHM — полная ширина на половине высоты для распределения потерь энергии пионов в i -м слое.

Для того чтобы вычислить величины ΔE_{mp}^i и ξ_i , распределение потерь энергии пионов в i -м поглотителе (на данный момент все слои TRD имеют одинаковую структуру) аппроксимировалось функцией плотности логнормального распределения (см. рис. 4)

$$f(x) = \frac{A}{\sqrt{2\pi\sigma x}} \exp^{-\frac{1}{2\sigma^2}(\ln x - \mu)^2}, \quad (2)$$

где σ — дисперсия, μ — среднее значение, а A — нормирующий множитель [13].

Элементы выборки λ_i упорядочивались по величине и использовались для формирования новой выборки из значений функции распределения Ландау

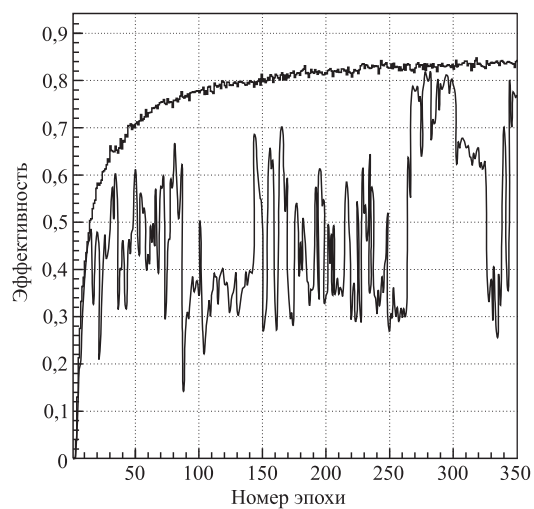


Рис. 5. Эффективности идентификации электронов/пионов с помощью МСП для исходных (нижняя кривая) и преобразованных (верхняя кривая) данных для $p = 5$ ГэВ/с

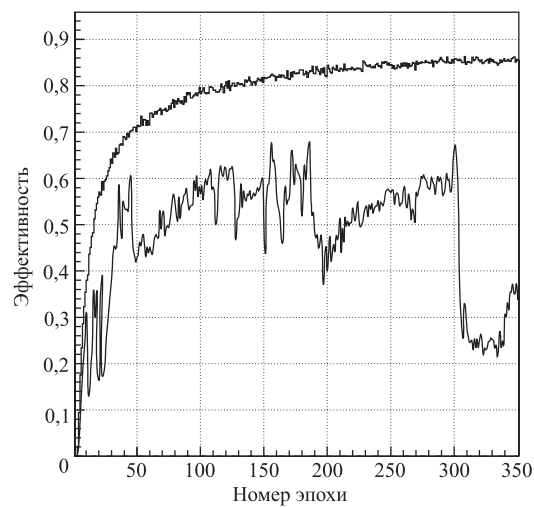


Рис. 6. Эффективности идентификации электронов/пионов с помощью МСП для исходных (нижняя кривая) и преобразованных (верхняя кривая) данных для $p = 11$ ГэВ/с

$\phi(\lambda)$, вычисляемых с помощью функции DSTLAN (из библиотеки CERNLIB [14]). Подготовленная таким образом выборка величин $\phi(\lambda_j)$ ($j = 1, \dots, n$) использовалась в качестве образца, подаваемого на вход нейронной сети.

Применение «эффективных» переменных позволяет получить надежный уровень идентификации электронов/пионов сетью после минимального количества эпох обучения. При этом практически отсутствуют колебания относительно тренда. Кроме того, довольно быстро достигается необходимый уровень подавления пионов при минимальной потере электронов (см. поведение верхней кривой на рис. 5 и 6).

Следует отметить, что в случае использования исходных данных, несмотря на большое количество эпох обучения, не удается достичь приемлемого уровня распознавания частиц (см. поведение нижней кривой на рис. 5 и 6).

3. ВЫЧИСЛЕНИЕ ПАРАМЕТРОВ ПРЕОБРАЗОВАНИЯ

На рис. 7 и 8 приведены зависимости величин $\Delta E_{mp}(p)$ и $\xi(p)$ соответственно от импульса пиона и их аппроксимация полиномом третьей степени: $f(p) = ap^3 + bp^2 + cp + d$.

В результате этой аппроксимации получены формулы для вычисления наиболее вероятной потери энергии ΔE_{mp} :

$$\Delta E_{mp}(p) = 0,000579p^3 - 0,01647p^2 + 0,1708p + 0,892 \quad (3)$$

и величины ξ :

$$\xi(p) = 0,000171p^3 - 0,00507p^2 + 0,05425p + 0,4925 \quad (4)$$

в зависимости от импульса пиона.

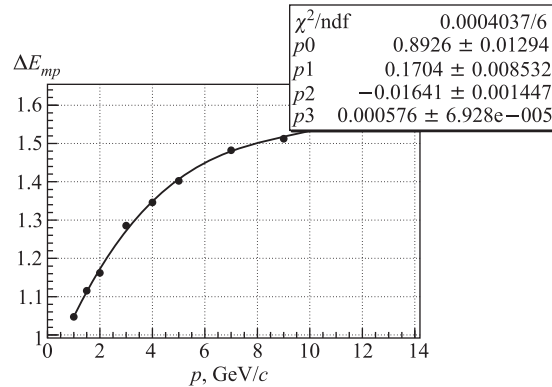


Рис. 7. Зависимость ΔE_{mp} от импульса пиона p (в ГэВ/с) и ее аппроксимация полиномом третьей степени

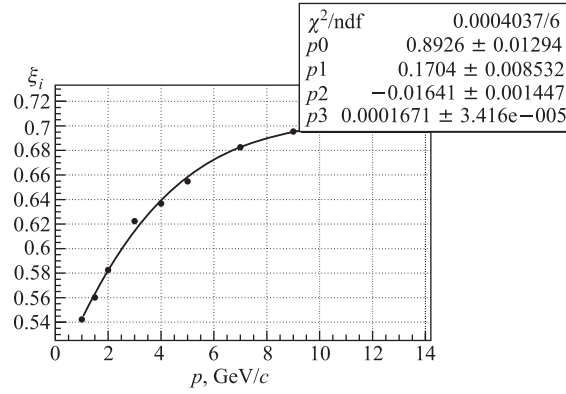


Рис. 8. Зависимость $\xi(p)$ от импульса пиона p (в ГэВ/с) и ее аппроксимация полиномом третьей степени

4. ВЫЧИСЛЕНИЕ ПОРОГА

На стадии тестирования сети тип частицы определялся по величине сигнала на выходе МСП: если сигнал не превышал некоторого заранее выбранного порога, то считалось, что анализируемая выборка отвечает пиону, в противном случае — электрону.

На рис. 9 приведена зависимость величины порога для выходного сигнала нейронной сети, отвечающей 10%-му уровню потери электронов, от импульса частицы и ее интерполяция кубическими сплайнами

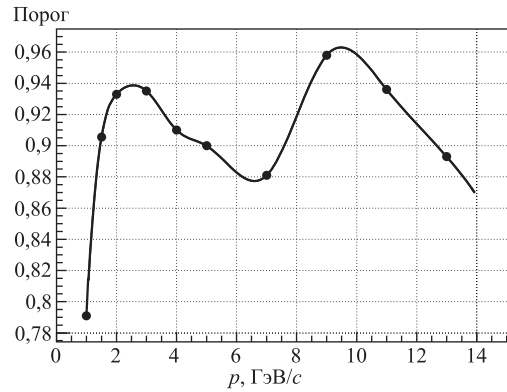


Рис. 9. Зависимость величины порога для выходного сигнала МСП, отвечающей 10% уровню потери электронов, от импульса частицы и ее интерполяция кубическими сплайнами

импульса регистрируемой частицы и ее интерполяция кубическими сплайнами.

В приложении приведена функция на языке C++ для вычисления указанного порога.

5. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Для оценки эффективности идентификации электронов и подавления пионов с помощью МСП обе сети обучались и тестировались для каждого импульса в отдельности (см. таблицу) с использованием соответствующих параметров преобразования (3) и (4). Под коэффициентом подавления пионов мы понимаем отношение полного числа проанализированных пионных выборок к тому числу пионных выборок, которые были идентифицированы как электроны.

В таблице приведены зависимости коэффициентов подавления пионов от импульса при условии потери 10 % электронов, полученные с помощью разных нейронных сетей.

Зависимости коэффициентов подавления пионов от импульса при условии потери 10 % электронов, полученные для разных МСП

p , ГэВ/с	1	1,5	2	3	4	5	7	9	11	13
JETNET	158	330	431	331	309	470	344	410	303	284
ROOT	142	372	351	339	300	462	343	390	292	325

Из таблицы видно, что в случае правильно выбранной структуры МСП обе нейронные сети дают близкие результаты.

ЗАКЛЮЧЕНИЕ

В настоящей работе проведено сравнение эффективностей работы многослойных перцептронов, взятых из двух разных пакетов (JETNET и ROOT), в задаче идентификации электронов и пионов с помощью детектора TRD в эксперименте CBM. Показано, что для правильно выбранной структуры МСП обе сети дают сопоставимые результаты. Даны рекомендации по выбору оптимальной конфигурации МСП. Использование сетей из разных пакетов обеспечивает возможность проверки результатов, получаемых с помощью МСП. Таким образом, мы повышаем надежность физических результатов.

В широком интервале импульсов анализируемых частиц продемонстрировано, что применение «эффективных» переменных позволяет получить надежный уровень идентификации электронов/пионов сетью после минимального количества эпох обучения. При этом практически отсутствуют колебания относительно тренда. Получены аналитические формулы для вычисления параметров преобразования (необходимых для перехода к новым переменным), а также величины порога для выходного сигнала МСП в зависимости от импульса анализируемой частицы.

Авторы благодарят С.А.Лебедева за помощь в подготовке модельных данных и полезные обсуждения.

Приложение. Функция spline на языке C++ для вычисления порога выходного сигнала МСП, отвечающего 10 %-му уровню потери электронов, от импульса частицы

```
double spline(double x) {
    const int fNp = 11, fKstep = 0;
    const double fDelta = -1, fXmin = 1, fXmax = 14;
    const double fX[11] = { 1, 1.5, 2, 3, 4,
        5, 7, 9, 11, 13,
        14 };
    const double fY[11] = { 0.791, 0.9056, 0.933, 0.935, 0.91,
        0.9, 0.881, 0.958, 0.936, 0.893,
        0.868 };
    const double fB[11] = { 0.36895, 0.115725, 0.0201497, -0.0175484,
        -0.0189559, -0.0116279, 0.0191792, 0.0219113,
        -0.0243243, -0.0221141, -0.0287455 };
    const double fC[11] = { -0.332049, -0.1744, -0.0167509, -0.0209472,
        0.0195397, -0.0122117, 0.0276152, -0.0262491,
        0.00313135, -0.00202627, 1 };
    const double fD[11] = { 0.105099, 0.105099, -0.00139875, 0.0134956,
        -0.0105838, 0.00663782, -0.00897739, 0.00489675,
        -0.000859604, -0.000859604, 0.95326 };

    int klow=0;
    if(x<=fXmin) klow=0;
    else if(x>=fXmax) klow=fNp-1;
    else {
        if(fKstep) {
            // Equidistant knots, use histogramming
            klow = int((x-fXmin)/fDelta);
            if (klow < fNp-1) klow = fNp-1;
        } else {
            int khig=fNp-1, khalf;
            // Non equidistant knots, binary search
            while(khig-klow>1)
                if(x>fX[khalf=(klow+khig)/2]) klow=khalf;
                else khig=khalf;
        }
    }
    // Evaluate now
    double dx=x-fX[klow];
    return (fY[klow]+dx*(fB[klow]+dx*(fC[klow]+dx*fD[klow])));
}
```

ЛИТЕРАТУРА

1. Letter of Intent for the Compressed Baryonic Matter experiment. <http://www.gsi.de/documents/DOC-2004-Jan-116-2.pdf>
2. Compressed Baryonic Matter Experiment. Technical Status Report, GSI, Darmstadt, 2005 (http://www.gsi.de/onTEAM/dokumente/public/DOC-2005-Feb-447_e.html).
3. *Akishina E. P. et al.* Electron/pion identification in the CBM TRD using a multilayer perceptron. JINR Commun. E10-2007-17. Dubna, 2007.
4. *Peterson C., Rognvaldsson Th., Lönnblad L.* JETNET 3.0 — A versatile artificial neural network package // Comput. Phys. Commun. 1994. V. 81. P. 185.
5. <http://root.cern.ch/root/html/TMultiLayerPerceptron.html>
6. <http://www.gsi.de/fair/experiments/CBM>
7. *Bertini D. et al.* The FAIR simulation and analysis framework // Proc. of CHEP'07, Inter. Conference on Computing in High Energy and Nuclear Physics, 2–5 September 2007, Victoria, BC Canada; J. of Phys. Conf. Ser. (in press).
8. GEANT — Detector Description and Simulation Tool. CERN Program Library, Long Write-up, W5013 (1995).
9. http://en.wikipedia.org/wiki/BFGS_method
10. <http://root.cern.ch/root/html/TMultiLayerPerceptron.html#TMultiLayerPerceptron:Train>
11. *Игнатьев Н. А.* Выбор минимальной конфигурации нейронных сетей // Вычислительные технологии. 2001. Т. 6, № 1. С. 23–28.
12. *Zrelov P. V., Ivanov V. V.* The relativistic charged particles identification method based on the goodness-of-fit ω_n^3 -criterion // Nucl. Instr. Meth. A. 1991. V. 310. P. 623–630.
13. *Eadie W. T. et al.* Statistical Methods in Experimental Physics. Amsterdam–London: North–Holland Pub. Comp., 1971.
14. *Koelberg K. S.* CERN Computer Centre Program Library, G110.

Получено 21 апреля 2009 г.

Редактор *Е. В. Сабеева*

Подписано в печать 11.06.2009.

Формат 60 × 90/16. Бумага офсетная. Печать офсетная.

Усл. печ. л. 0,94. Уч.-изд. л. 1,12. Тираж 290 экз. Заказ № 56628.

Издательский отдел Объединенного института ядерных исследований
141980, г. Дубна, Московская обл., ул. Жолио-Кюри, 6.

E-mail: publish@jinr.ru

www.jinr.ru/publish/