

Министерство науки и высшего образования Российской Федерации  
Федеральное государственное автономное образовательное учреждение  
высшего профессионального образования «Национальный  
исследовательский ядерный университет «МИФИ»

УДК 53-08, 53-05

ОТЧЕТ  
О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

Пространственная реконструкция событий в детекторе  
DarkSide-20k

Научный руководитель  
к.ф.-м.н., доцент каф.40

\_\_\_\_\_ А. В. Гробов

Выполнила

\_\_\_\_\_ Н. А. Полякова

Москва 2021

# Содержание

1	Цель работы . . . . .	2
2	Введение . . . . .	2
3	Эксперимент DarkSide . . . . .	3
4	Сверточные нейронные сети . . . . .	4
5	Инструменты, использовавшиеся в работе . . . . .	5
6	Обработка данных . . . . .	6
7	Реорганизация данных . . . . .	12
8	Определение максимального числа SIPM, вышедших из строя . . . . .	13
9	Результаты . . . . .	15
	Список литературы . . . . .	16

# 1 Цель работы

- Изучить принципы работы и основную информацию об эксперименте и установке, данные с которых подлежат изучению.
- Провести анализ и обработку предоставленных данных.
- Оптимизировать параметры нейронной сети и алгоритм реконструкции событий.
- Реконструировать положения событий в пространстве внутри TPC - time projection chamber - время-проекционной камеры в DS-20k.
- Определить максимальное допустимое число выключенных ФЭУ, при котором реконструкция остается точной ( $\approx 1cm$ )

# 2 Введение

Ученые занимающиеся астрофизикой неоднократно сталкивались с противоречивыми данными. Например, при расчете скорости движения звезд на границах галактики относительно ее центра, согласно законам Кеплера, она должна уменьшаться. Однако в действительности этого не происходит, скорость на достаточно больших расстояниях становится практически неизменной, что противоречит предположениям ньютоновской механики, поскольку плотность видимого вещества уменьшается при отдалении от центра галактики. Данный эффект может быть объяснен наличием «невидимой», или как ее позже называли, темной материей, которая обеспечивает постоянство скорости [4].

Существует множество частиц-кандидатов на роль темной материи, но наиболее распространенными являются так называемые вимпы — — — Weakly Interacting Massive Particles (WIMPs), что в переводе означает слабо взаимодействующие массивные частицы. Соответственно, эти частицы слабо взаимодействуют с обычным барионным веществом, поэтому для их обнаружения используются детекторы, которые имеют низкий фон и обладают отличными

свойствами для их детектирования, например, детекторы на жидких благородных газах, в частности, аргоне.

### 3 Эксперимент DarkSide

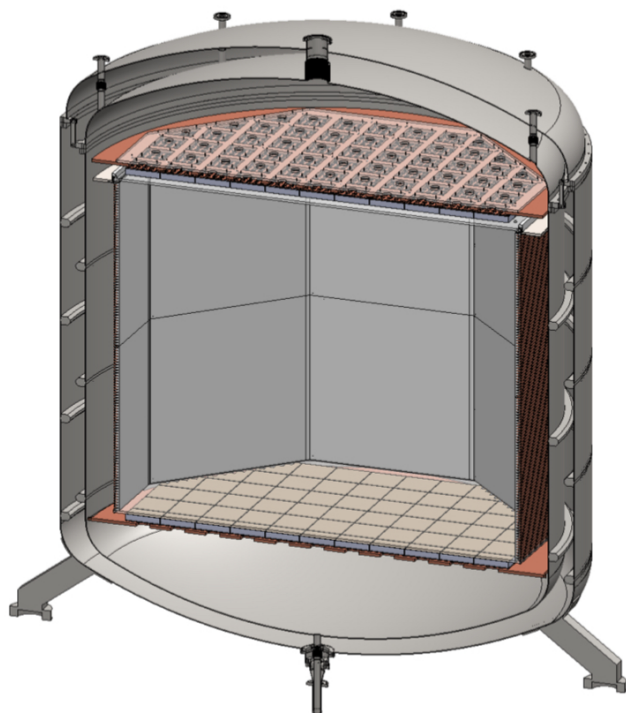


Рисунок 1. Двухфазная время-проекционная камера (TPC) эксперимента DarkSide-20k [3].

Полученные, в первом случае, от возбужденных димеров аргона, а во втором от электронов.

В эксперименте DarkSide электрическое поле направлено вдоль оси  $Z$  детектора, и поэтому глубина события может быть с большой точностью определена путем измерения времени дрейфа — временного промежутка между сцинтилляционными сигналами  $S1$  и  $S2$  в жидкой и газообразной фазе. Что же касается координат  $X$  и  $Y$ , то это является уже более трудоемкой задачей.

Решением этой задачи может послужить алгоритм с использованием свер-

Проект DarkSide - это обширная программа прямого поиска темной материи, которая ориентирована на серию экспериментов на основе жидкого аргона, но также включает в себя дополнительные проекты для разработки и оптимизации новых технологий, необходимых для повышения эффективности работы используемых детекторов [1].

Работая в двухфазном режиме, такие детекторы позволяют измерять как сцинтилляционный свет, так и ионизацию. В связи с тем, что взаимодействия с WIMP крайне редки, необходимо минимизировать всевозможные фоны.

В аргонных детекторах важную роль играют следующие два детектируемого сигнала:  $S1$  и  $S2$ , сцинтилляции,

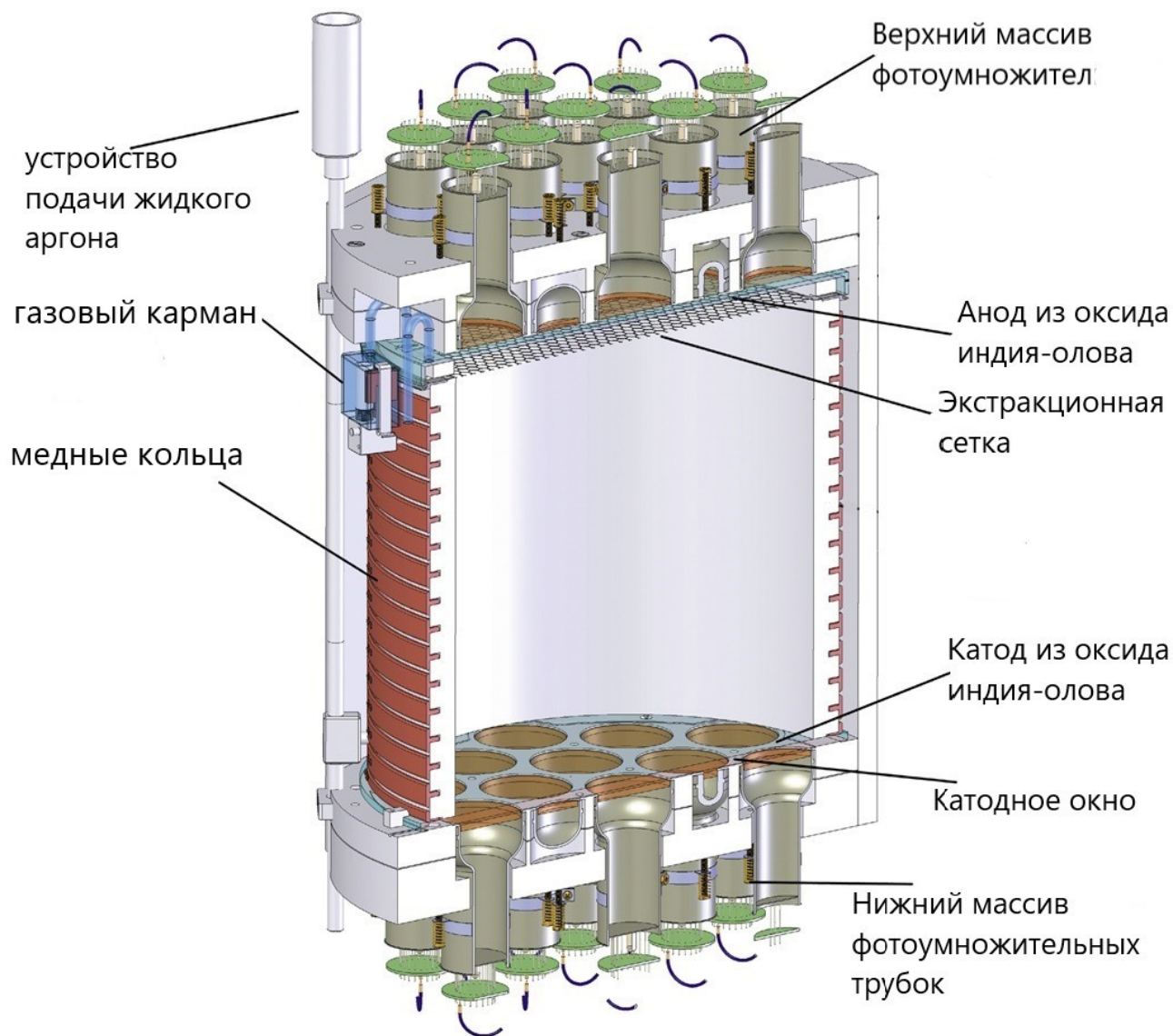


Рисунок 2. Устройство двухфазной время-проекционной камера (TPC) на примере эксперимента DarkSide-50.

точных нейронных сетей. Для этого необходимо обработать данные, получаемые в ходе эксперимента, и понять как их можно использовать для реконструкции положения события в плоскости X и Y.

## 4 Сверточные нейронные сети

Сверточные нейронные сети — Convolutional Neural Networks — это мощный класс искусственных нейронных сетей, которые очень хорошо справляются с проблемами восстановления изображений.

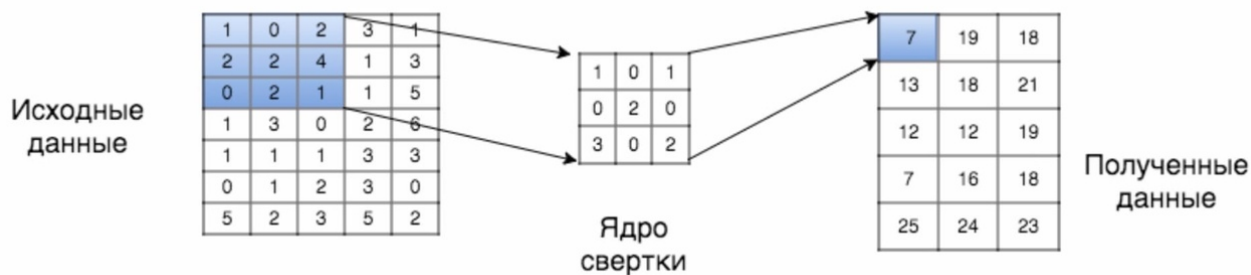


Рисунок 6. Пример работы сверточного слоя.

Принцип работы этих сетей основан на обработке входных данных несколькими слоями, в частности обработки изображений, каждый из которых состоит из линейного оператора и действует как свертка, за которым следует нелинейное преобразование с некоторой функцией активации.

Каждый фрагмент входного изображения умножается на матрицу (ядро) свертки поэлементно, а результат суммируется и записывается в аналогичную позицию выходного изображения [2].

Форма вывода у СНС выражается как:

$$y = F(\Theta, x) = f_n(W_n f_{n-1}(\dots(f_2(W_2 f_1(W_1 x + b_1) + b_2)\dots) + b_n), \quad (1)$$

где  $x$  — это входные данные, а  $y$  выходные,  $W_i$  — матрица свертки  $i$ -го слоя,  $b_i$  — смещение  $i$ -го слоя свертки,  $f_i$  — нелинейная функция,  $\Theta$  — набор всех настраиваемых параметров.

Основная цель работы СНС состоит в том, чтобы найти оптимальный набор параметров  $\Theta$  и количество эпох  $K$  и в последующем минимизировать потери, выражающиеся следующей формулой:

$$\sum_{k=1}^K L(y_k, F(\Theta, x_k)), \quad (2)$$

где  $L(y_k, F(\Theta, x_k))$  — функция потерь для  $k$ -ой эпохи.

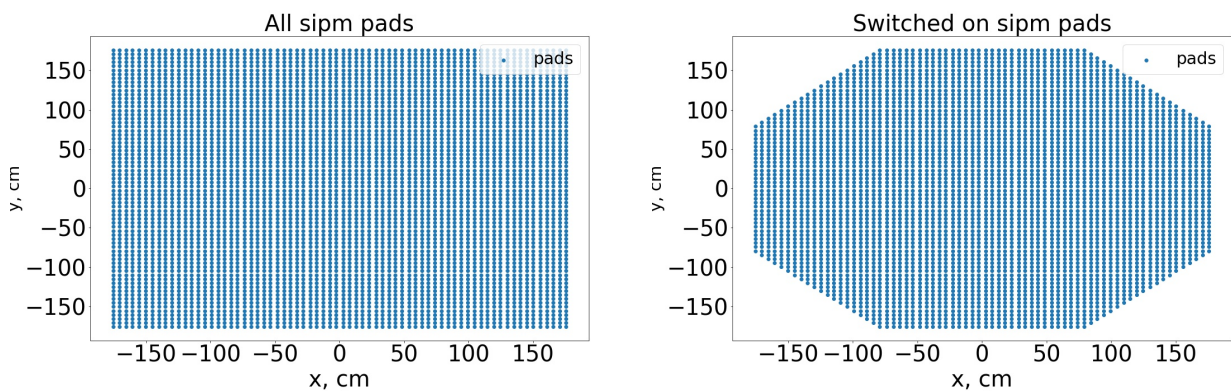
## 5 Инструменты, использовавшиеся в работе

- программный пакет ROOT

- Geant 4 for DarkSide
- Jupyter
- Google Colab

## 6 Обработка данных

Для сравнительного анализа были предоставлены данные двух типов, смоделированные методом Монте-Карло. Одни были заранее подготовлены для обучения нейронной сети, а другие предстояло обработать и привести в вид, подходящий для обучения. Оба набора данных содержали  $X$  и  $Y$  координаты событий. Подготовленные данные представляли собой 3 сорта информации для каждого события соответственно: пара координат  $x$ ,  $y$  и двумерный массив, содержащий номера сработавших фотоумножителей и число фотоэлектронов дошедших до него, отнормированное на общее число фотоэлектронов данного события. Второй набор содержал информацию в таком виде, в каком она формируется в ходе выполнения эксперимента.



(а) Расположение всех ФЭУ.

(б) Расположение включенных ФЭУ.

Рисунок 7. Расположение фотоумножителей в детекторе DarkSide-20k

Также было известно заранее истинное расположение реальных участвующих в работе фотоумножителей, которое изображено на рисунке 7(б), и общее расположение всех фотоумножителей, которое изображено на рисунке 7(а).

Необходимо было оценить корректность и пригодность переорганизованных "сырых" данных. Для этого на готовых "правильных" данных заранее обучалась нейросеть, чтобы после этого на новых данных можно было сравнить результаты с предыдущими и сделать выводы о качестве подготовленного материала. Графики и гистограммы результатов обучения представлены ниже. Основными из них являются: графики потерь, распределение реальных и предсказанных нейросетью значений координат, гистограммы ошибок для каждой из координат (рисунки 9 и 10).

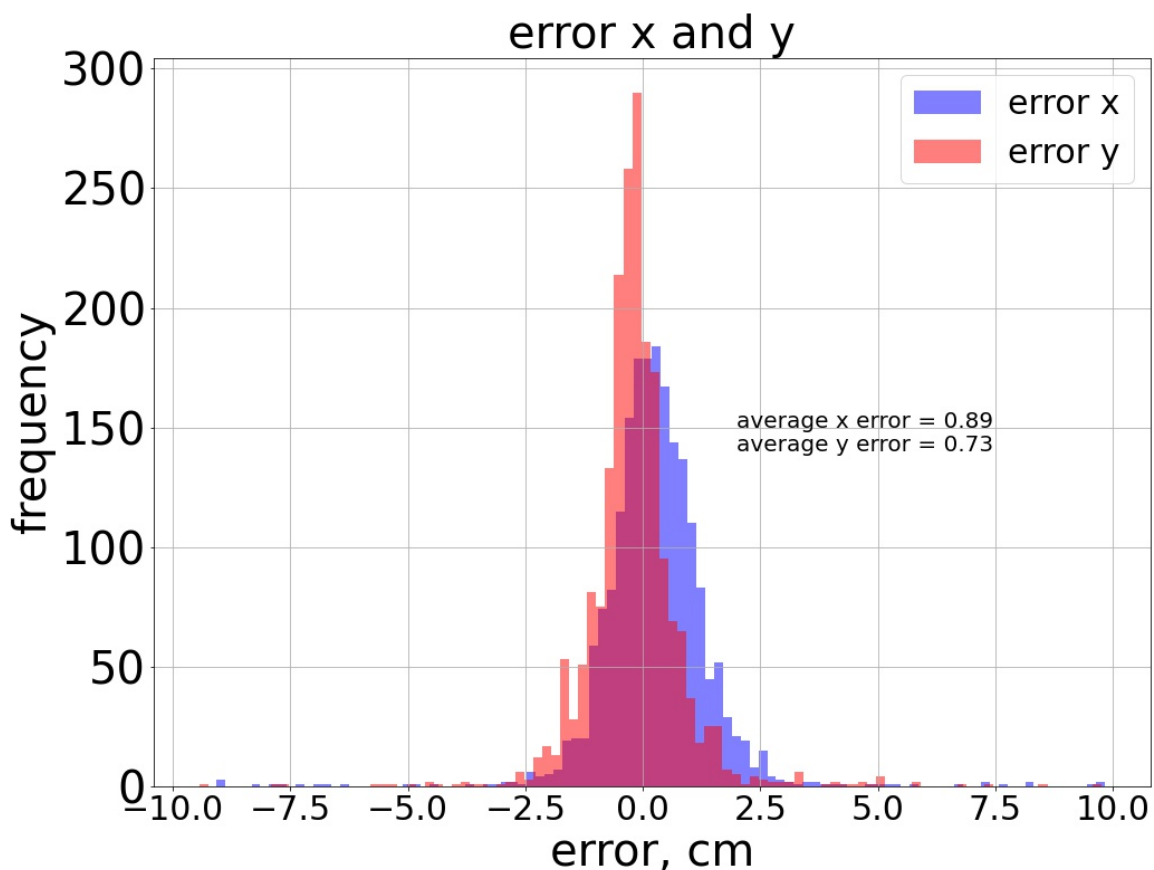
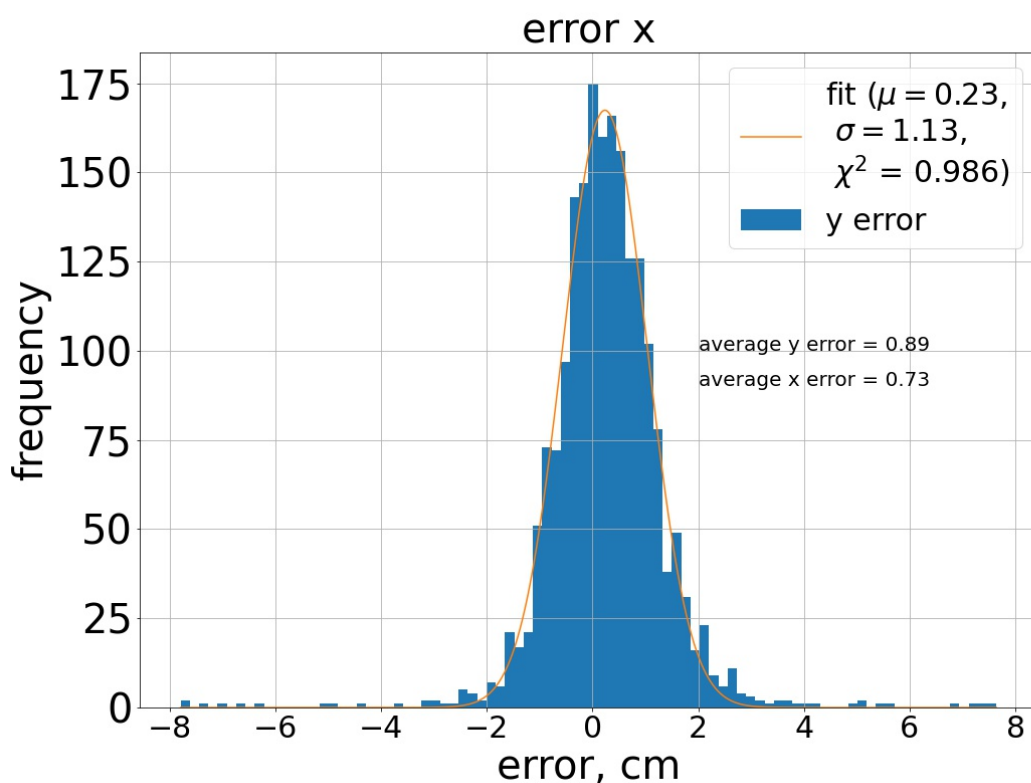


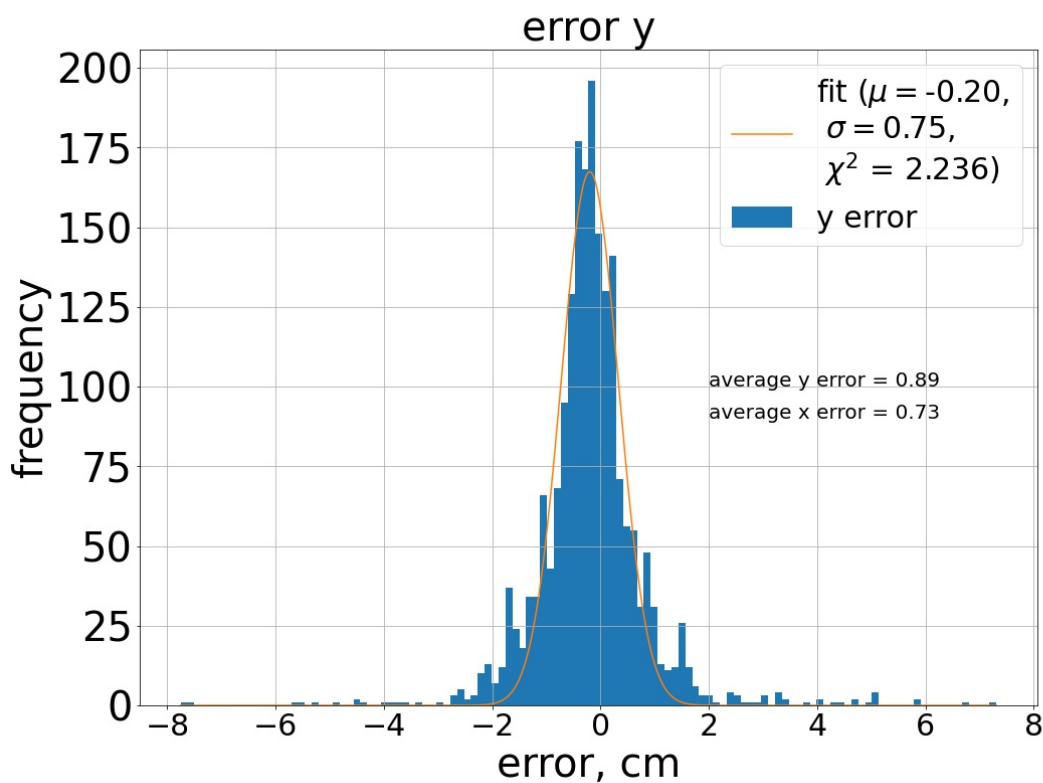
Рисунок 9. Распределение ошибки — разности предсказанного и реального значений для обеих координат.

Важно было не допустить "переобучения" нейронной сети, когда результаты становятся слишком идеальными, потому что тогда на реальных данных с эксперимента обученная нейросеть может неверно предсказывать значения.





(a) Распределение ошибки (разности предсказанного и реального значений) координаты x.



(б) Распределение ошибки (разности предсказанного и реального значений) координаты y.

Рисунок 10. Распределение ошибки — разности предсказанного и реального значений координат для 30 эпох. Подготовленные заранее данные. 8

Таблица 1: Полученные средние ошибки для координат

	$x, cm$	$y, cm$
Ошибка	0.89	0.73

Ранее было установлено, что оптимальное число эпох для существующей конфигурации нейронной сети составляет 30 эпох. Это также подтверждает график потерь для координаты  $y$  для 20000 событий на рисунке 11. Нейросеть еще не успевает переобучиться, но при этом точность реконструкции так же остается приемлемой ( $\approx 1cm$ ).

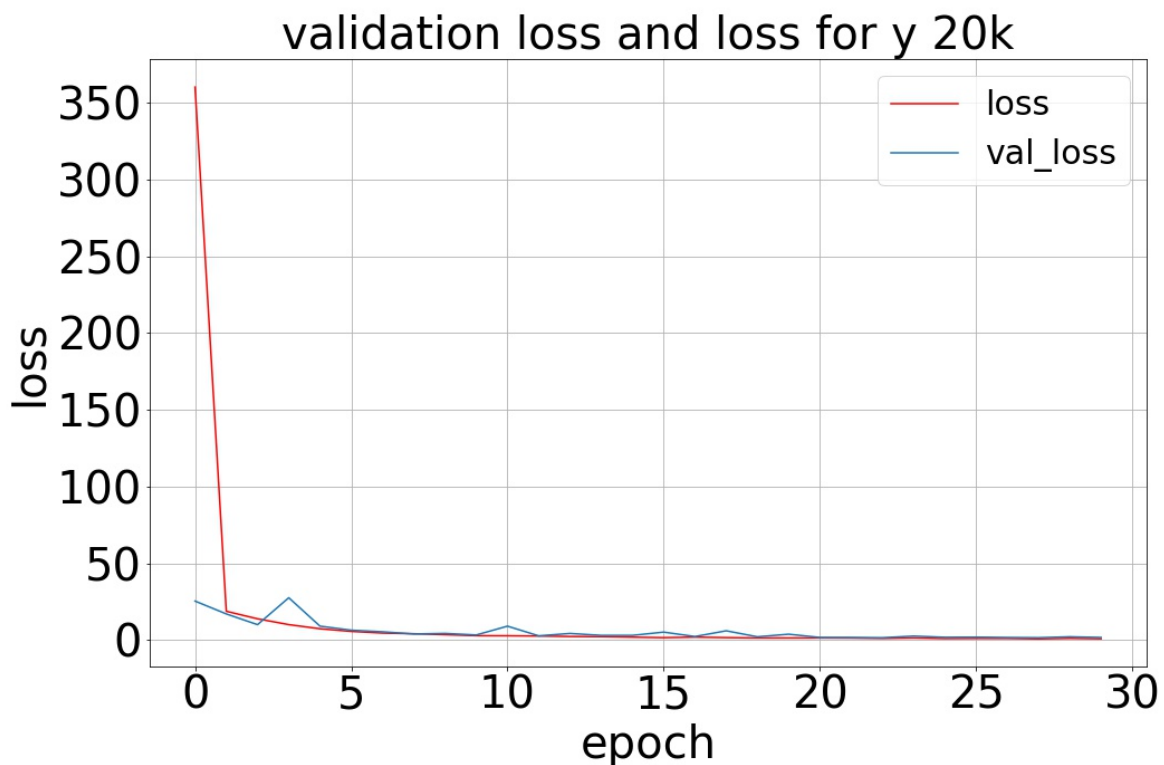
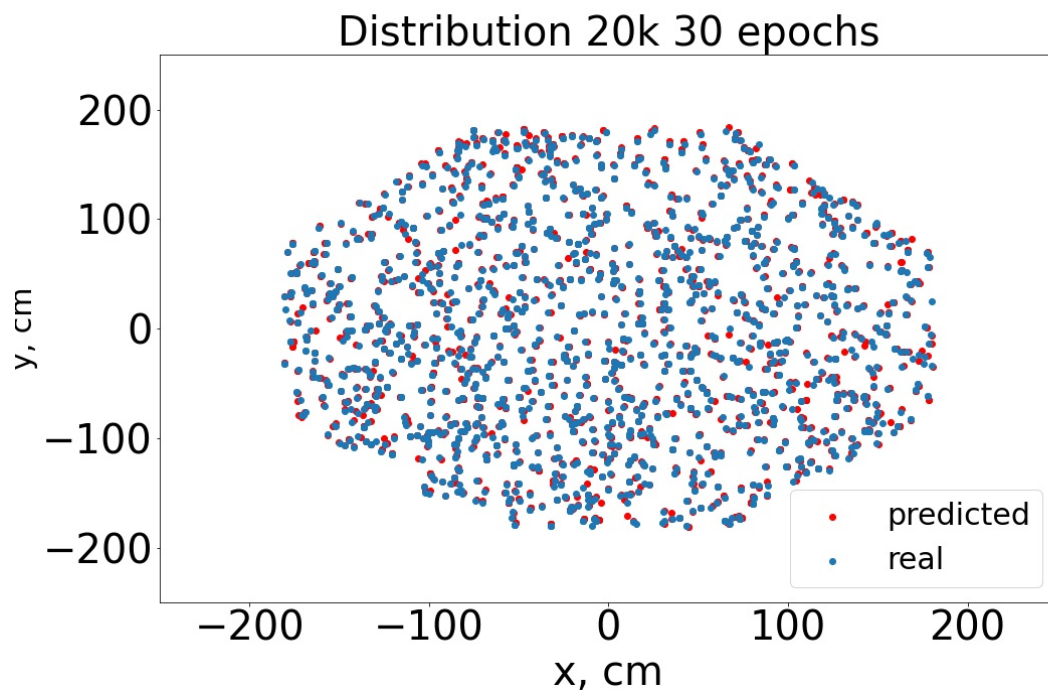


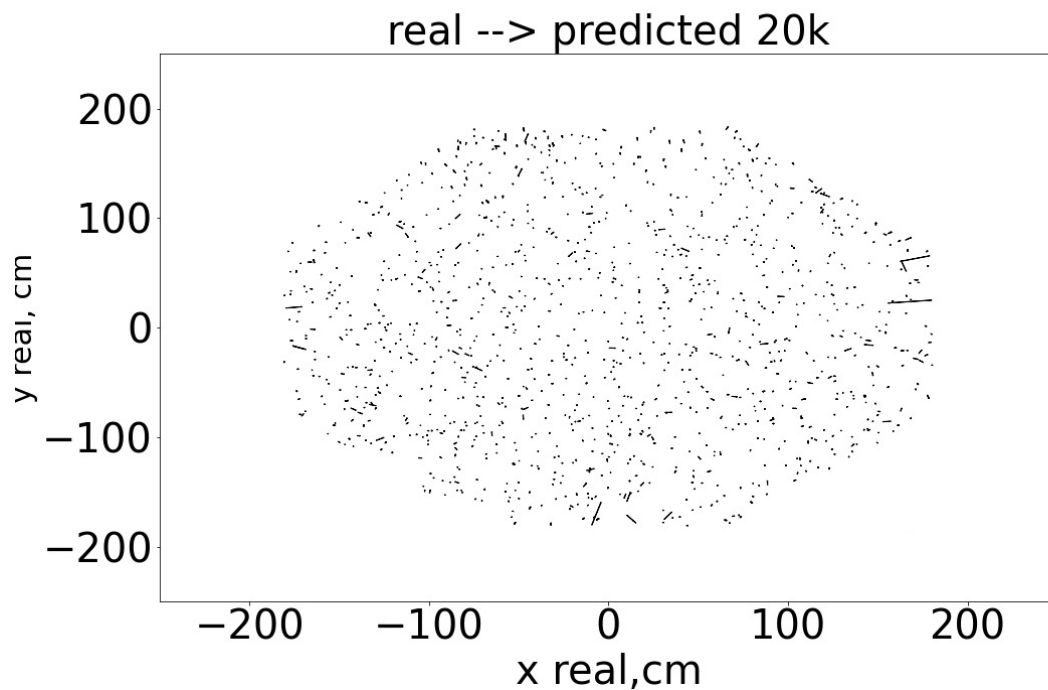
Рисунок 11. График потерь – validation loss и loss для 30 эпох обучения

На рисунке 12 представлены распределения предсказанных и реальных координат (а) и разности этих двух значений (б) полученные после обучения нейросети. На втором распределении видно насколько точным является предсказание координат, т.к. практически отсутствуют "длинные" фрагменты, показывающие большое различие между предсказанными и реальными координатами.

Эти результаты в ходе дальнейшей работы будут считаться эталонными, именно сравнивая с ними будет определяться корректность обработанных данных.



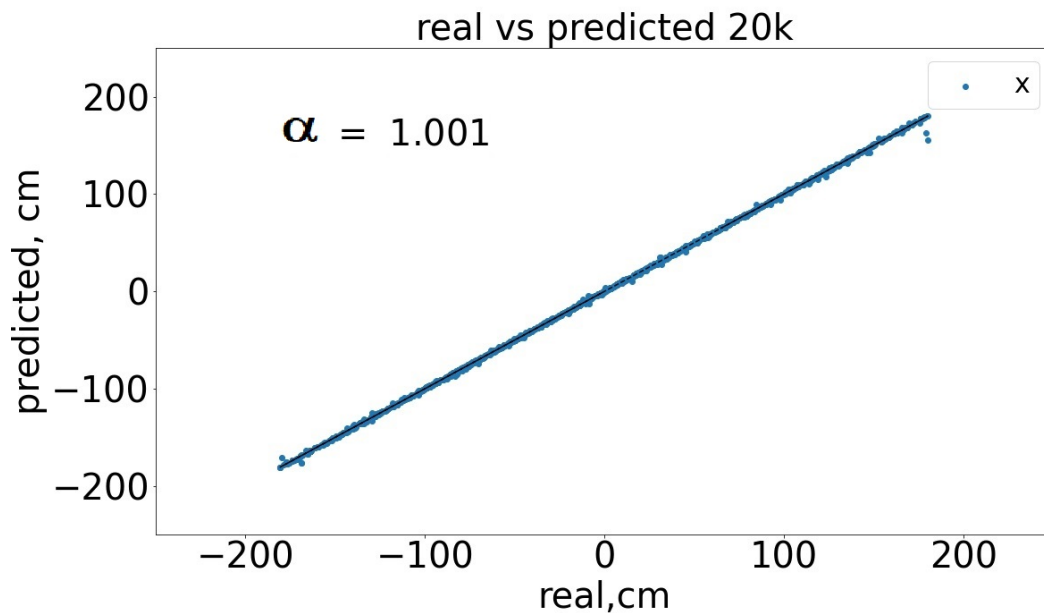
(а) Распределение предсказанных и реальных координат



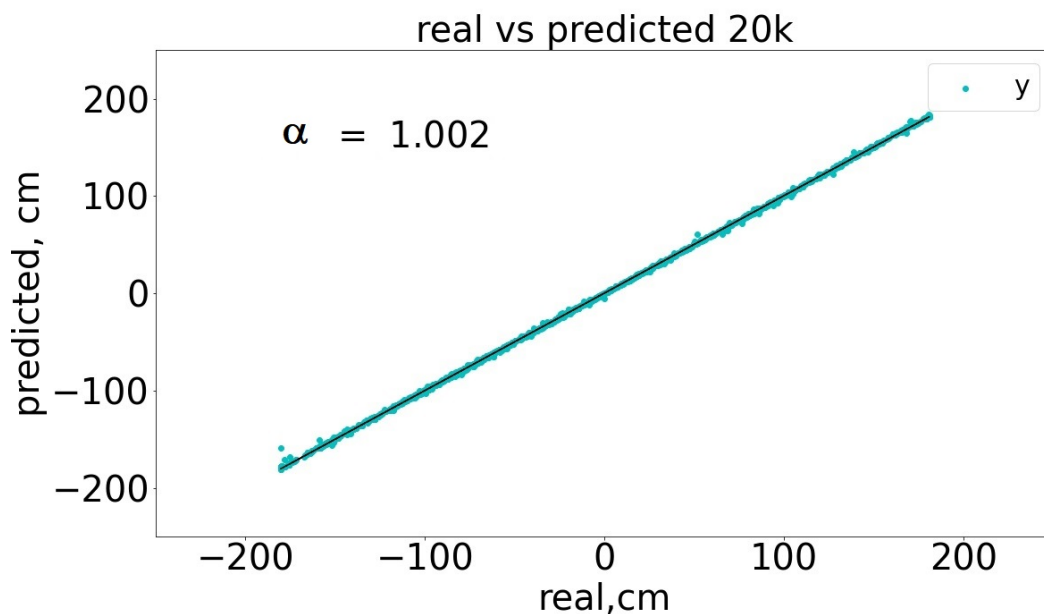
(б) Распределение разности между предсказанными и реальными координатами

Рисунок 12. Распределение координат

Для дополнительного анализа была также построена зависимость реального значения координат от предсказанного. Линейный коэффициент, определенный по МНК, для каждой из координат практически равен единице (см. рисунок 13 а и б). А также можно заметить, что вблизи концов прямых практически отсутствуют различного рода нелинейности, что в свою очередь говорит об корректной работе нейронной сети.



(а) Зависимость реального значения от предсказанного для координаты  $x$ .



(б) Зависимость реального значения от предсказанного для координаты  $y$ .

Рисунок 13. Зависимость реального значения от предсказанного.

Данное исследование было проведено на различных выборках по 20 тысяч событий из 200 тысяч событий. На рисунке 14 представлены ошибки для 7 выборок для каждой из координат, каждая выборка обозначается соответствующим индексом на графике.

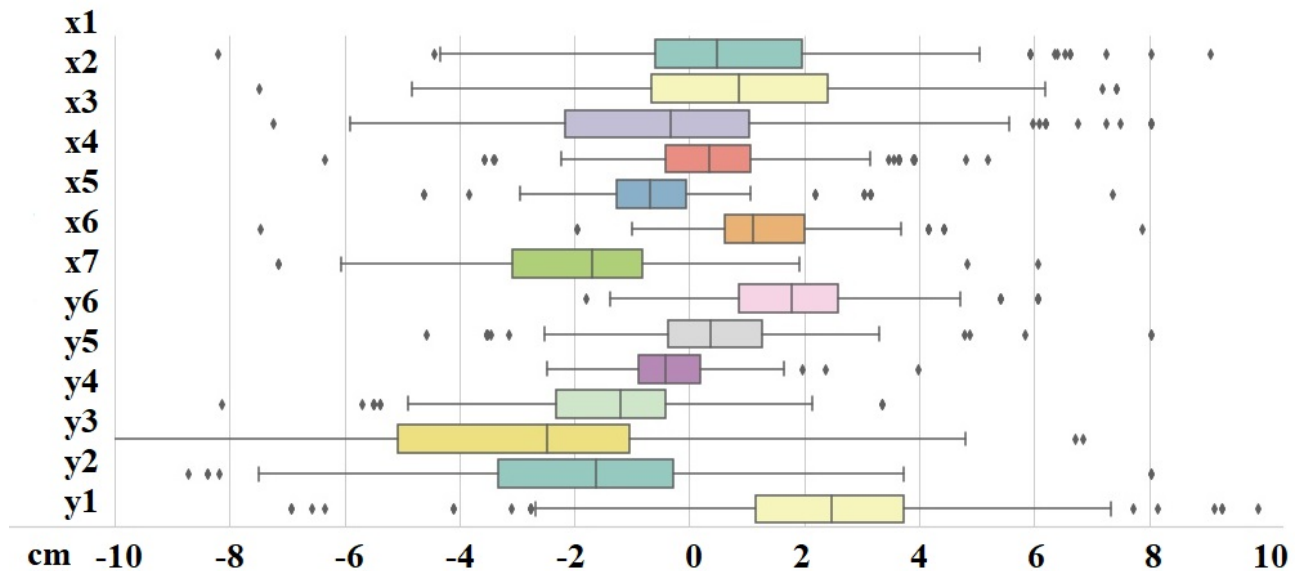


Рисунок 14. Распределение ошибки — разности предсказанного и реального значений для координат для различных выборок.

## 7 Реорганизация данных

Из "сырых" данных была сформированна таблица пригодная для обучения нейронной сети по подобию "готовых" данных. На этих данных обучалась нейросеть с теми же параметрами. Наименьшая ошибка, которую удалось получить по обоим координатам,  $\approx 2\text{cm}$ .

Основные распределения, графики и гистограммы представлены на рисунках 15 и 16.

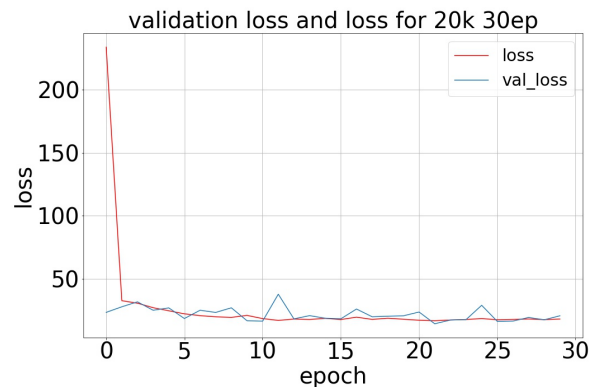
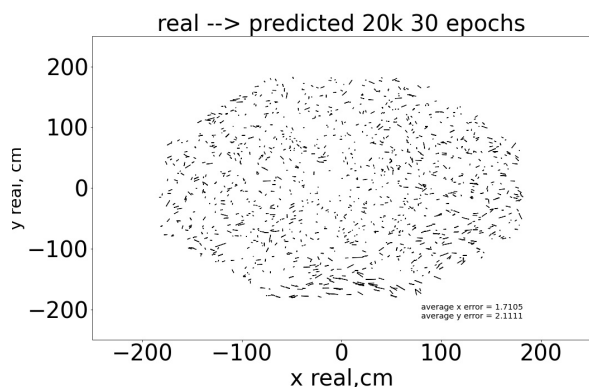
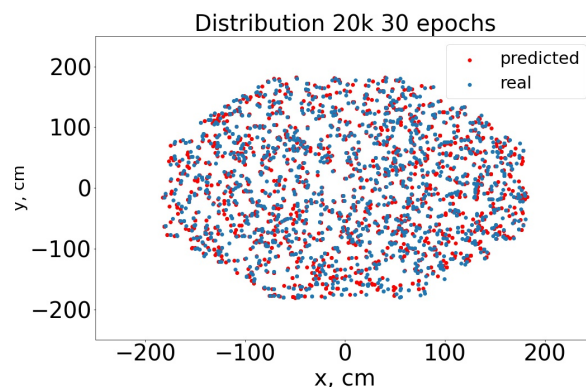


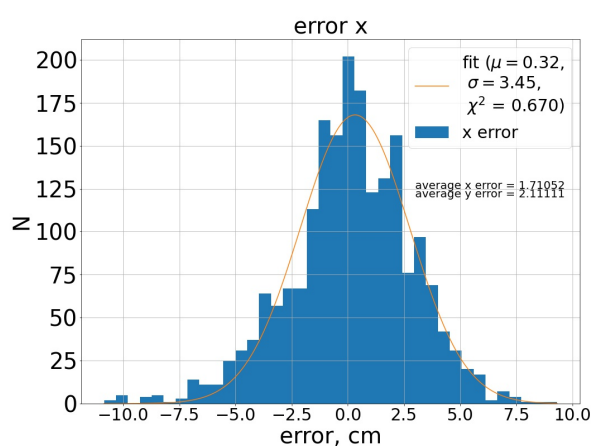
Рисунок 15. График потерь — validation loss и loss



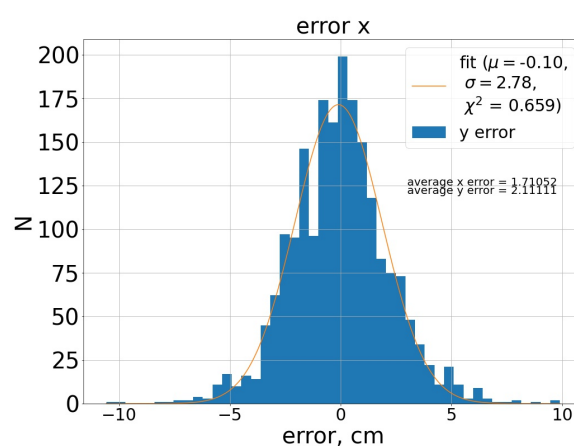
(а) Разность реального и предсказанного



(б) Распределение реал. и пред.



(в) Гистограмма ошибки  $x$



(г) Гистограмма ошибки  $y$

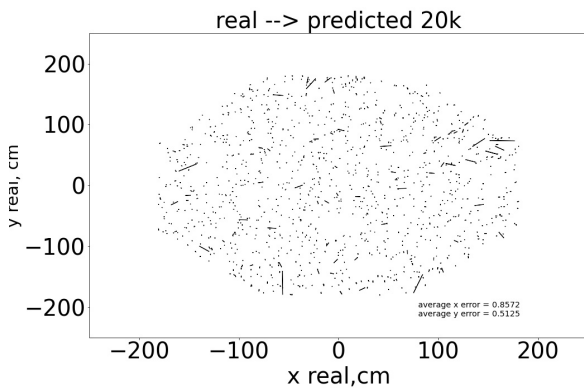
Рисунок 16. Основные гистограммы для переорганизованных данных

## 8 Определение максимального числа SIPM, вышедших из строя

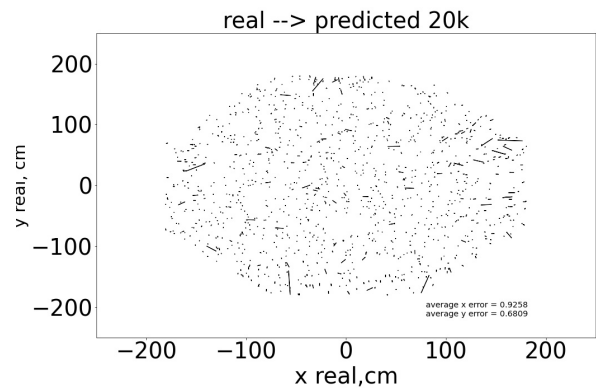
Как показывает практика, в ходе реального эксперимента довольно часто кремнивые ФЭУ выходят из строя. Их замена является трудоёмкой для реализации задачей, в связи с особенностями технической конструкции детекторов DarkSide.

Поэтому одной из задач данной исследовательской работы было определить максимальное количество SiPM, которые могут выйти из строя, при котором реконструкция остается точной. В качестве критерия оценки точности было принято, что точной является реконструкция, ошибка которой остается  $\approx 2\text{cm}$  для каждой из координат.

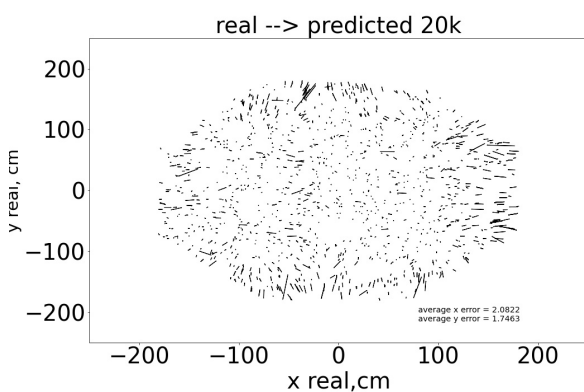
Для подготовленных данных достаточное число эпох для обучения равняется 30. Нагляднее в этом можно убедиться на рисунке 17(а), где для 20 тысяч событий изображена разность предсказанных и реальных координат при всех работающих sipm. На последнем изображении (г) разброс значительно больше, чем при всех включенных ФЭУ на рисунке (а).



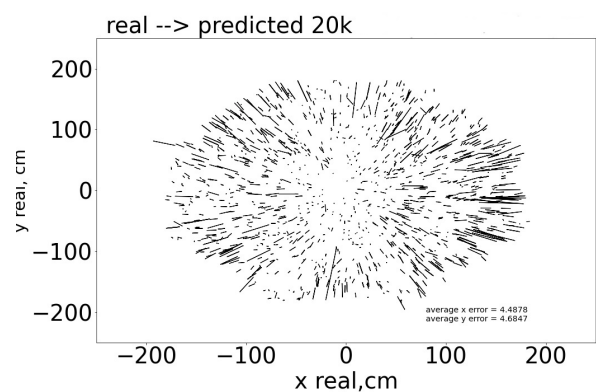
(а) все sipm работают



(б) 250 sipm выключены



(в) 500 sipm выключены



(г) 1000 sipm выключены

Рисунок 17. Разность между предсказанным и реальным значением для разного количества выключенных ФЭУ

В таблице 2 представлены основные полученные значения средних ошибок для  $x$  и  $y$  при разном числе SiPM, вышедших из строя. При всех работаю-

щих фотоумножителях средняя ошибка  $< 1\text{см}$  (рисунок 17(а)). Оптимальным числом выключенных  $\text{simr}$ , при котором точность реконструкции удовлетворяет критериям, оказалось 500 из 4120 общего числа всех ФЭУ (рисунок 17(в)). Для 1000 неработающих ФЭУ разброс становится очень большим  $> 4\text{см}$  (рисунок 17(г)).

Таблица 2: Средние ошибки для разного количества выключенных фотоумножителей для координат

выключенные ФЭУ	0	250	500	1000
$x, \text{см}$	0.86	0.93	2.08	4.49
$y, \text{см}$	0.51	0.68	1.75	4.69

## 9 Результаты

- При выполнении данной научно-исследовательской работы были изучены основные принципы работы эксперимента DarkSide-20k и сверточных нейронных сетей.
- Разработана собственная нейронная сеть, которая с необходимой точностью реконструирует положение события в пространстве в плоскости  $x$  и  $y$ .
- Обработаны данные, смоделированные методом Монте-Карло, отображена необходимая для реконструкции информация о событиях, с заведомо известными значениями координат  $x$  и  $y$ , на основе которой была обучена нейросеть.
- Подобраны оптимальные параметры обучения, такие чтобы реконструкция была точной.
- Определено число выключенных ФЭУ при работе детектора, при котором реконструкция событий остается возможной.



## Список литературы

1. Bottino B. Dark matter search with liquid argon in DarkSide: results with scientific and technological prototypes. — 2019.
2. Cullo A. Using Convolutional Neural Networks for DarkSide-20k Analysis. — 2018.
3. DarkSide-20k: A 20 Tonne Two-Phase LAr TPC for Direct Dark Matter Detection at LNGS / C. E. Aalseth [et al.]. — 2017.
4. Marini L. Dark Matter Direct Detection with DarkSide-50: Analysis of Early Data. — 2018.