

Предварительные результаты по применению машинного обучения для анализа изображений TAIGA-IACT

Докладчик: Гресь Елизавета, магистрант 1 курса, ИГУ

Научный руководитель: Крюков А.П., к.ф.-м.н., зав. ЛКФВЭ НИИЯФ МГУ

1 апреля 2021

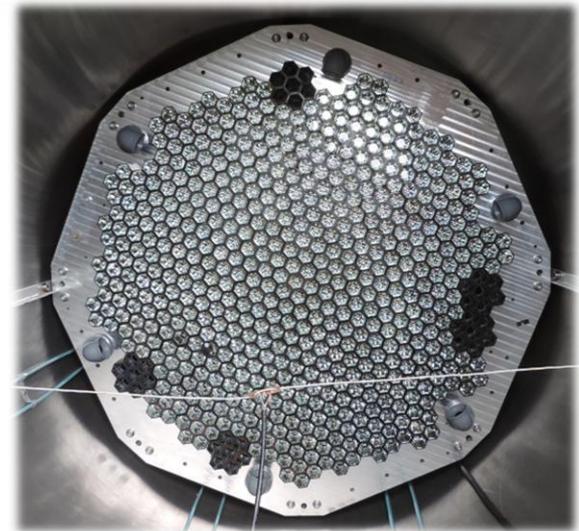
Телескопы TAIGA-IACT

- Местонахождение: Тункинская долина, р. Бурятия
- Диаметр зеркала — 4,3 м;
- Фокусное расстояние — 4,75 м;
- ФЭУ в матрице — 560;
- Угол обзора камеры — $9,6^\circ$.
- Основная цель TAIGA-IACT — селекция событий от гамма-квантов и от адронов ШАЛ.



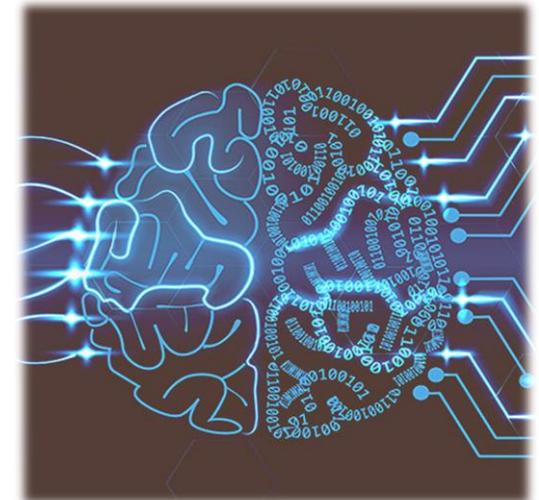
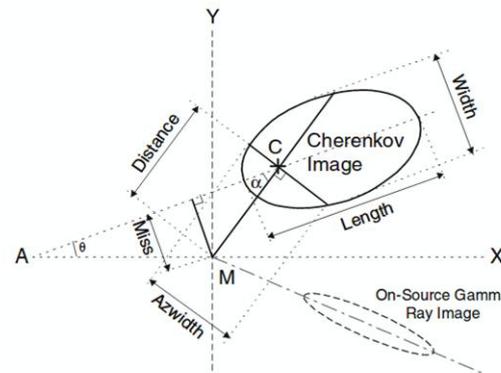
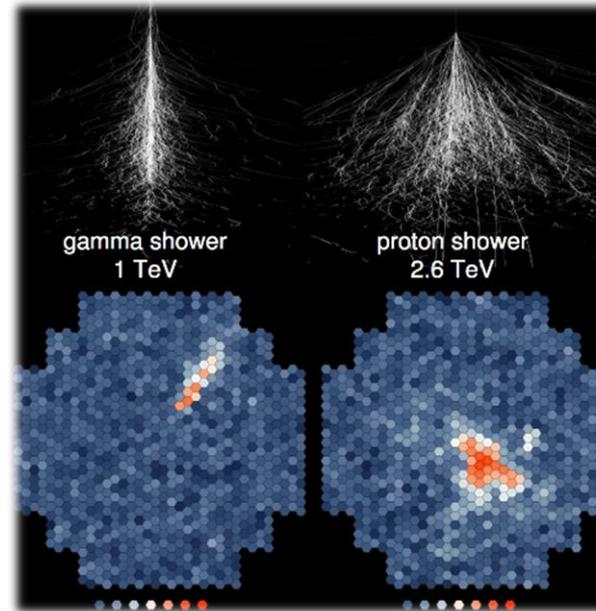
IACT02

IACT01



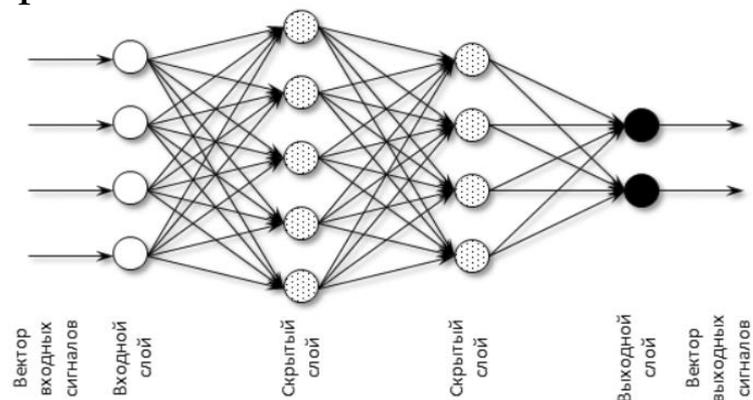
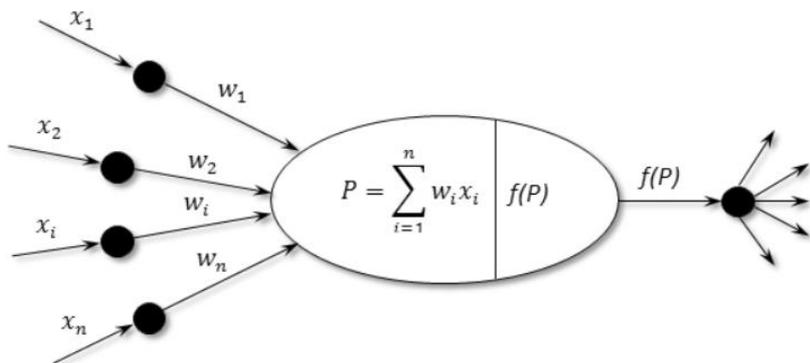
Способы анализа изображений IACT

- *Метод параметров Хилласа*: изображения описываются эллипсом.
- *Машинное обучение*: свёрточные нейронные сети (CNN).

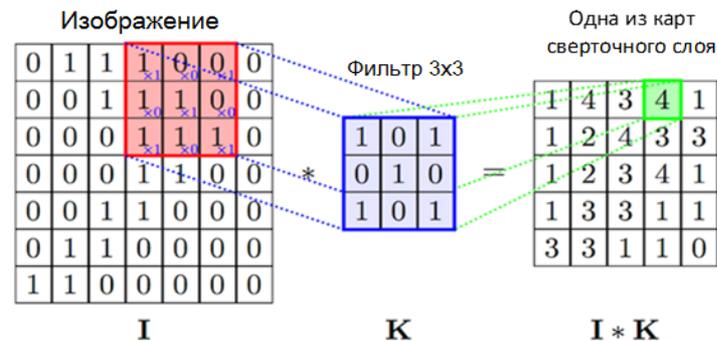
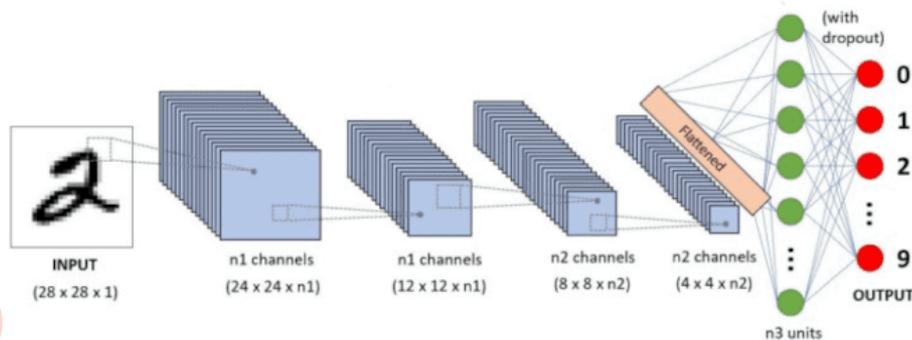


Немного о нейронных сетях

Математическая модель нейрона (*парсептрон*) и многослойная структура искусственной нейронной сети:



Свёрточная нейронная сеть (CNN) - наличие свёрточных слоев, где применяются фильтры, которые выделяют в изображении более общие структуры и признаки:



Постановка задачи:

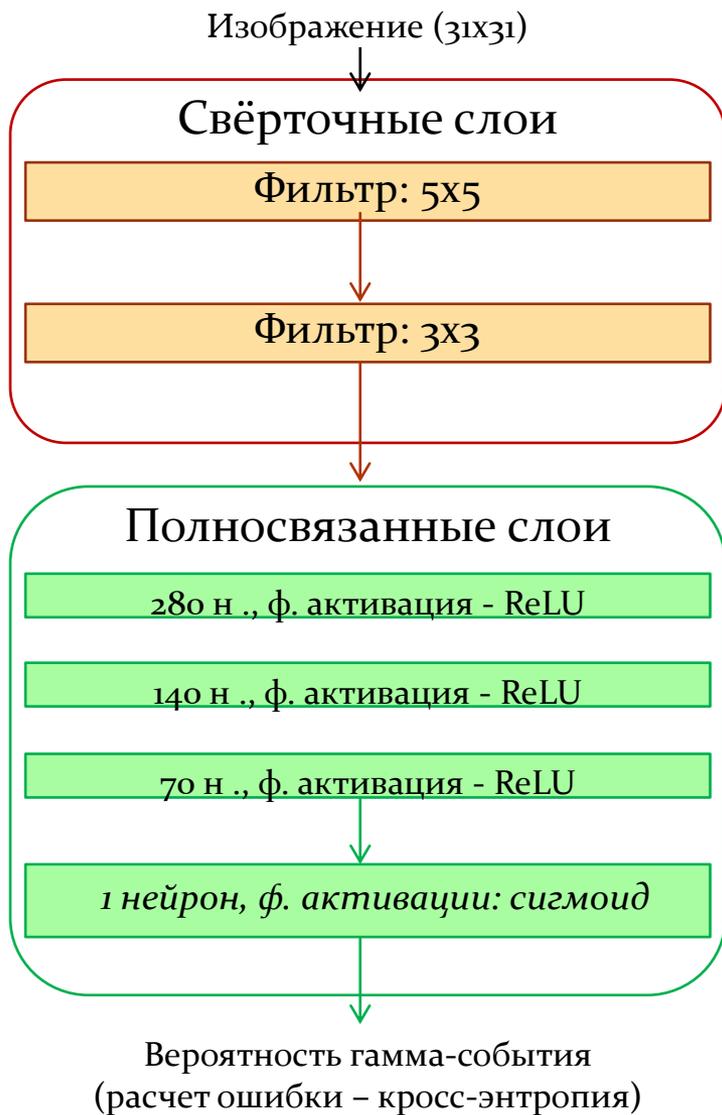
Решить задачу классификации и регрессии для одного и нескольких телескопов с помощью свёрточных нейронных сетей.

▶ *Классификация* –
определение
вероятности того, что
событие является от
гамма-фотона.

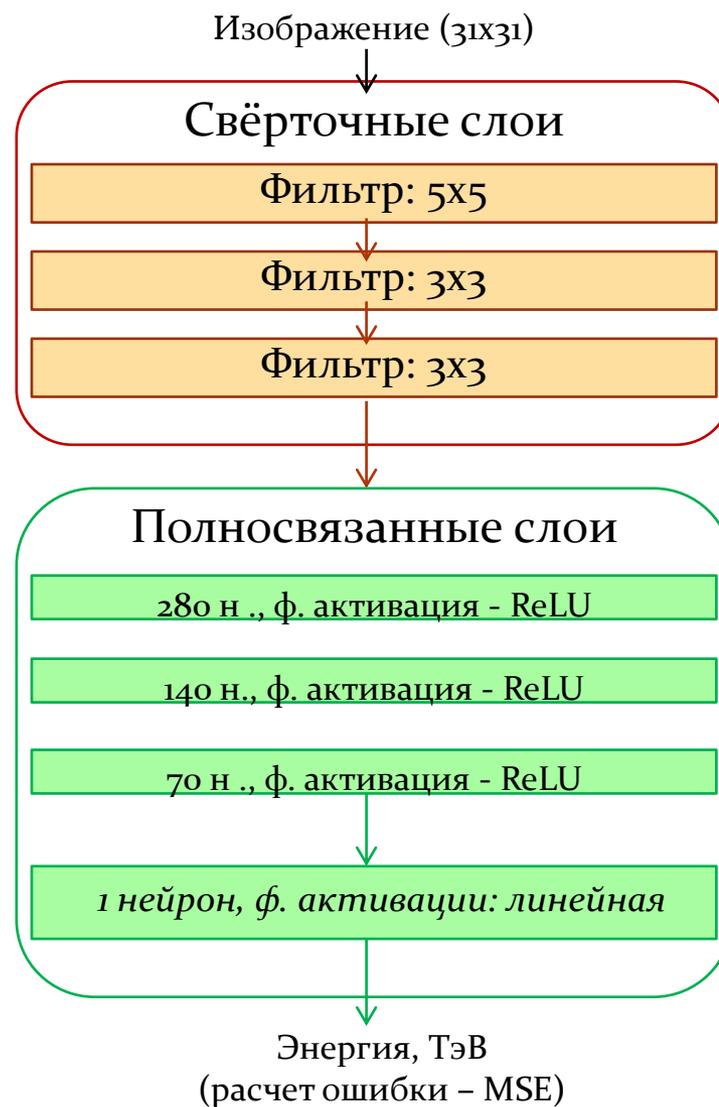
▶ *Регрессия* –
восстановление энергии
первоначальной
частицы.

Структура свёрточных нейронных сетей:

- **Классификация:**

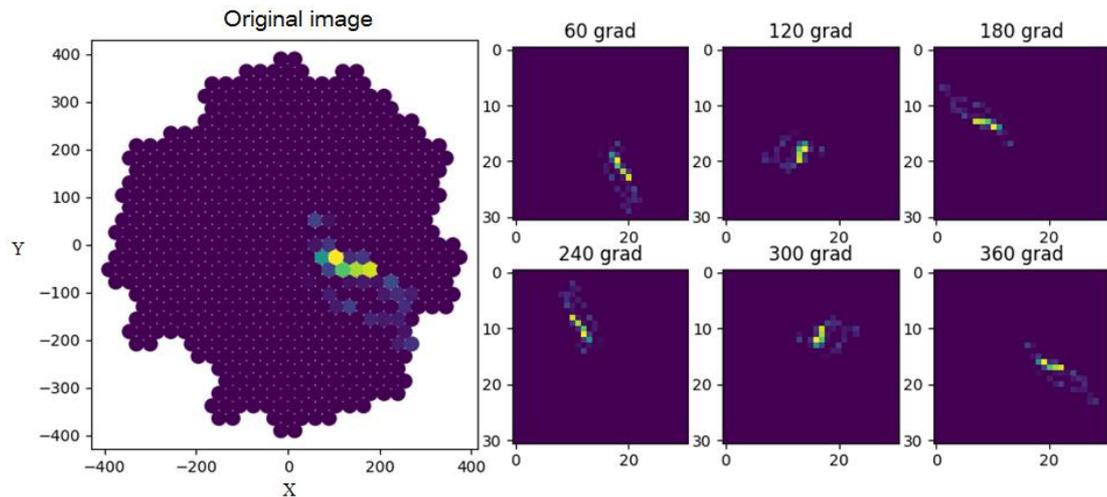
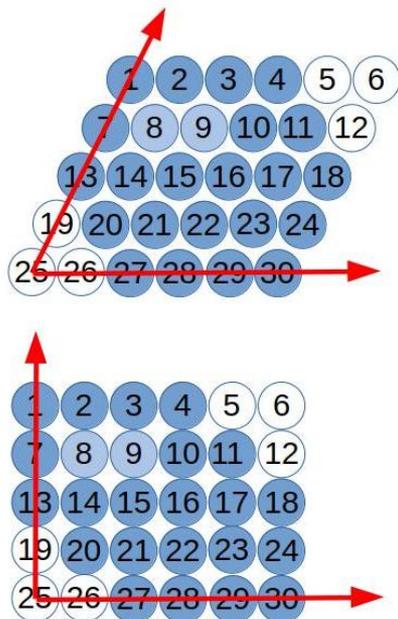


- ▶ **Регрессия:**



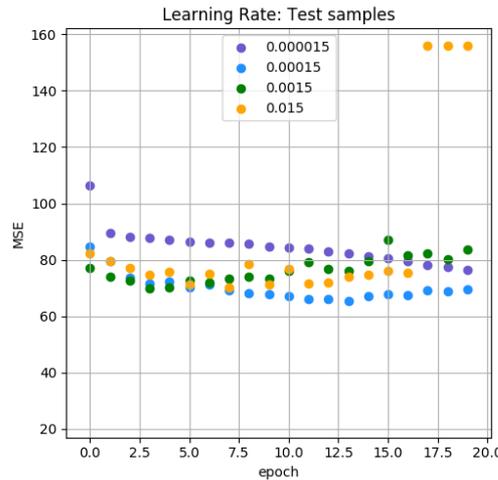
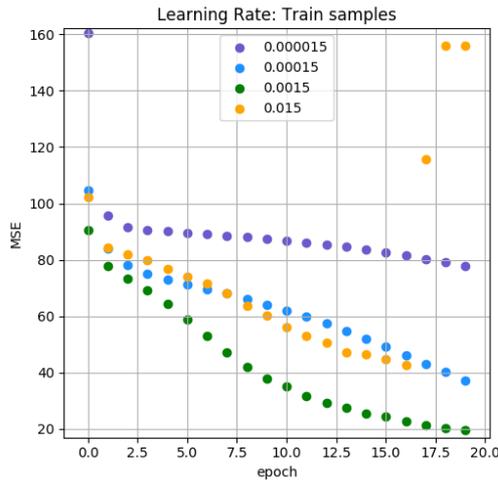
Используемые модельные данные

- **Модельные данные:**
 - CORSIKA (предоставлены НИИЯФ МГУ)
 - Всего 30 000 событий:
 - 13 000 – адронные события
 - 17 000 – гамма события
- **Обучение/тестирование:**
20000/10000
- ▶ Перед обучением каждое изображение преобразовалось к квадратному виду.
- ▶ Расширение выборки: поворот на 60° относительно центра камеры.



Пример пикселизации изображения и поворотов по 60°

Влияние гиперпараметров: скорость обучения и число фильтров



Удовлетворительные скорости обучения, при которых CNN не успевает переобучиться и при этом ошибка обучения достаточно мала – 0,001-0,0001

График зависимости ошибки обучения и тестирования в зависимости от эпохи для разных скоростей обучения

Оптимальное количество фильтров – 5. Дальнейшее увеличение мало влияет на ошибку, однако возрастает время обучения нейронной сети

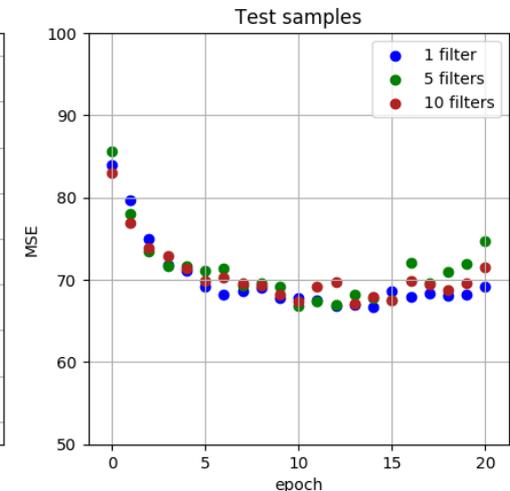
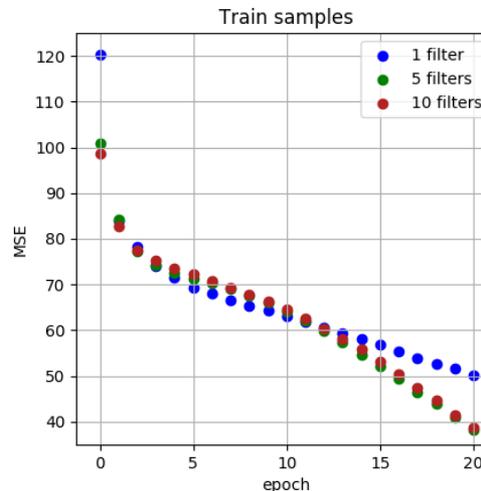
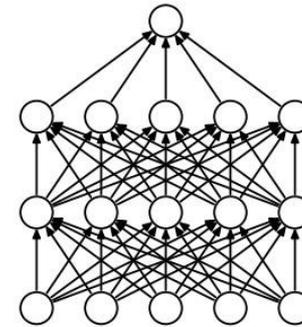


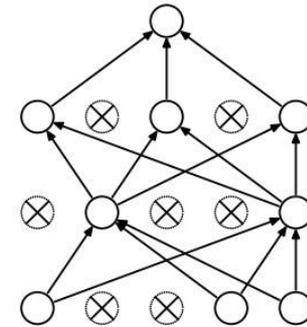
График зависимости ошибки обучения и тестирования в зависимости от эпохи для скоростей обучения

Влияние регуляризации: drop out

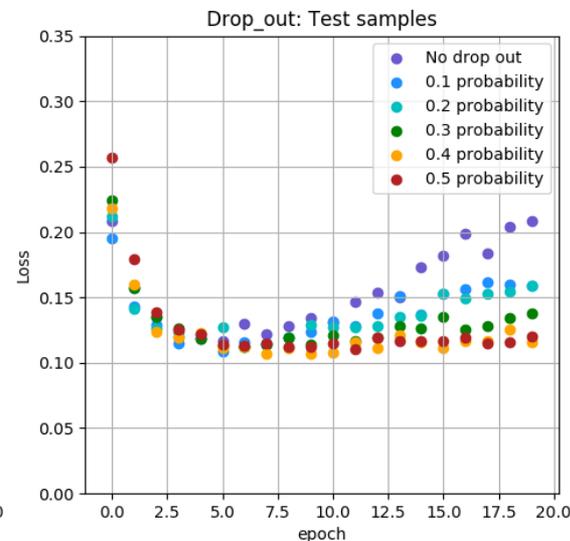
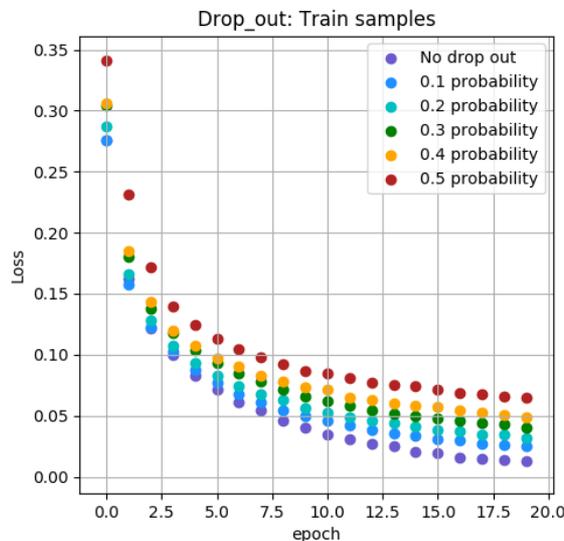
- Drop out - метод регуляризации, предназначенный для уменьшения переобучения нейронной сети путем исключения нейронов из сети с вероятностью p .
- Вероятность p в нашем случае равна 0,3-0,4



(a) Standard Neural Net



(b) After applying dropout.



Поведение ошибки обучения и тестирования в зависимости от эпохи для разных вероятностей p

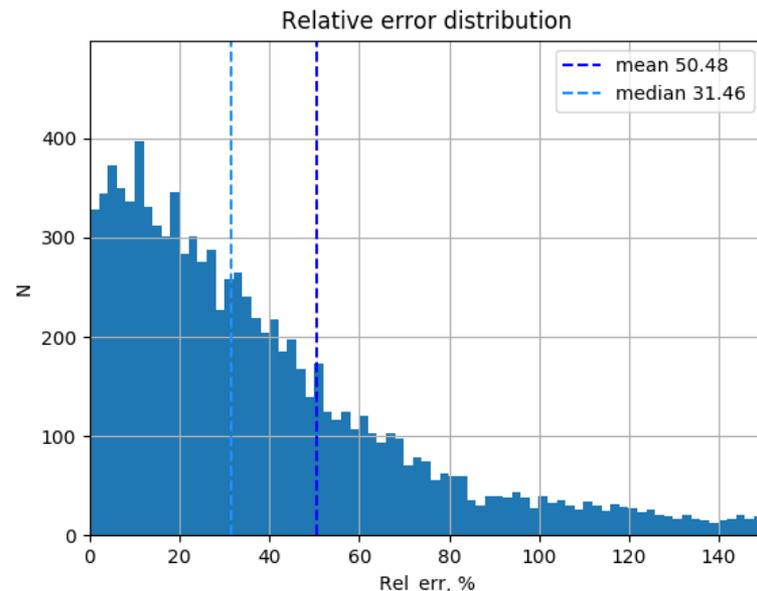
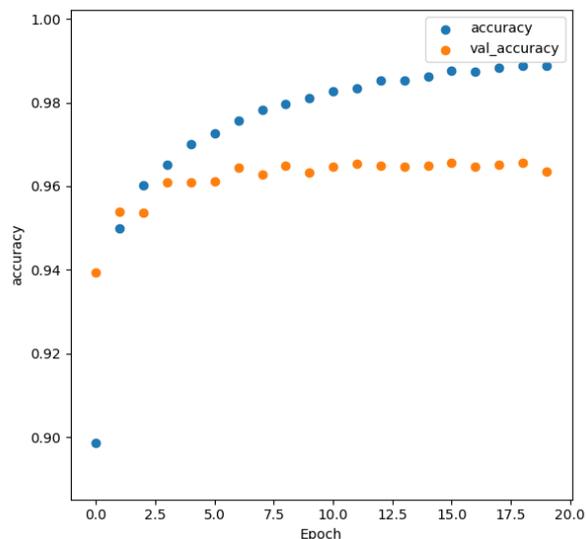
Результат применения CNN

- Классификация:
 - Точность определения событий – 96%;
 - Параметр Q больше в 2 раза.
- Регрессия:
 - Среднеквадратичная ошибка на тестовом наборе – 70 ТэВ;
 - Средняя относительная погрешность – ~50%.

Отбор	Q
По параметрам Хилласа	1,8 – 2,0
CNN	3,99

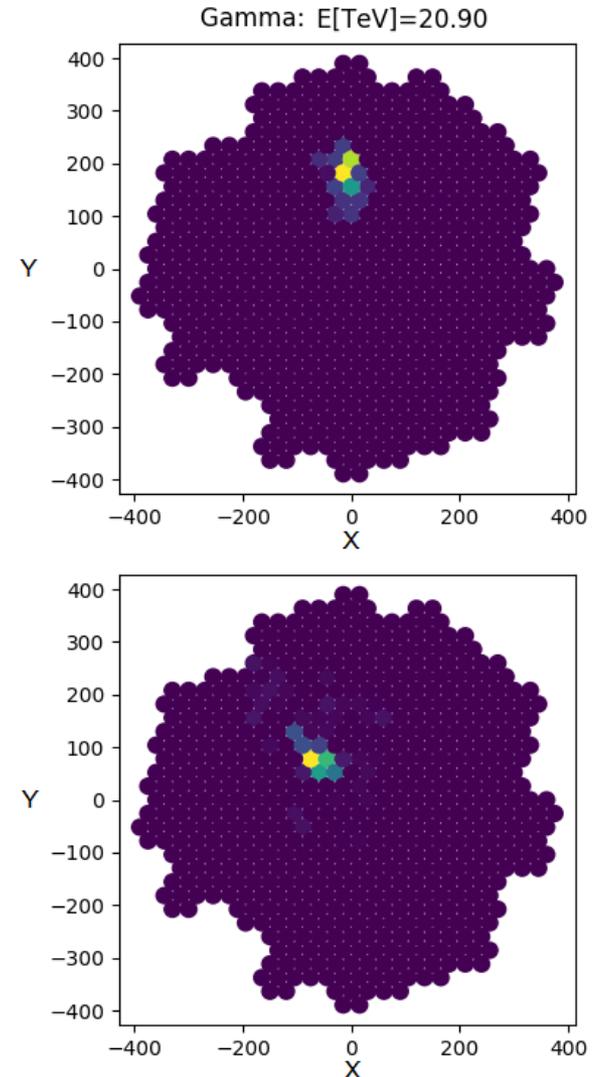
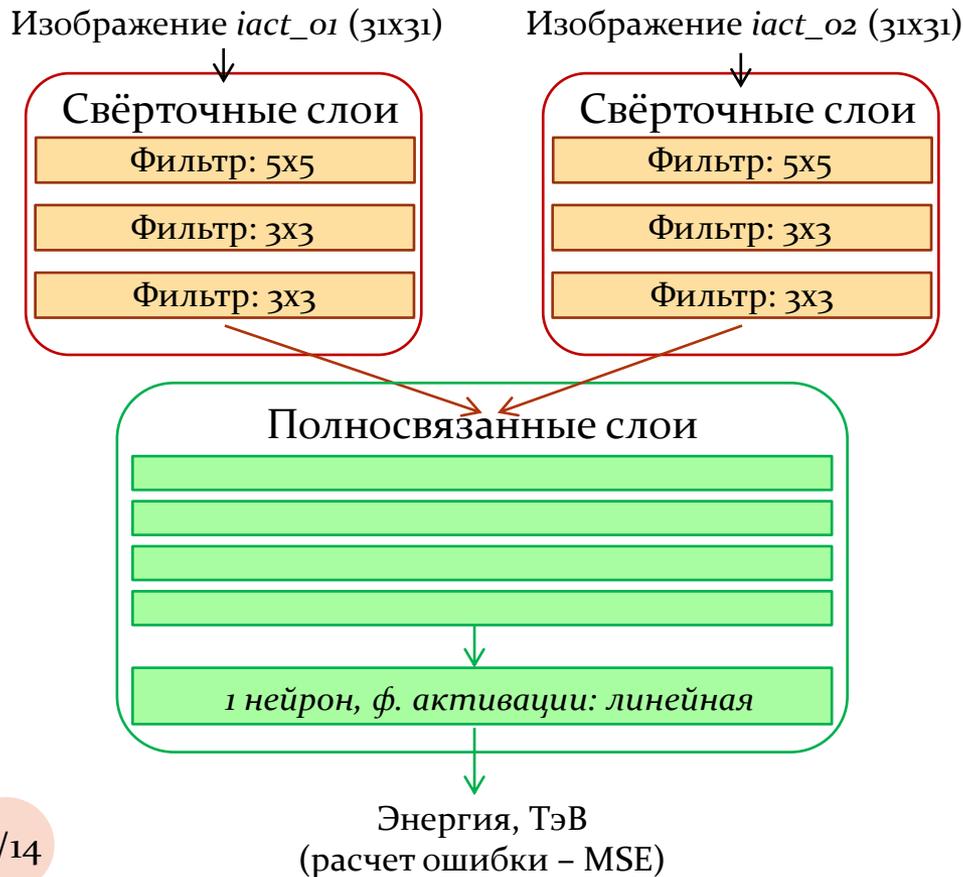
$$Q = \varepsilon_\gamma / \sqrt{\varepsilon_{CR}}, \text{ где } \varepsilon_i = \tilde{N}_i / N_i$$

N_i – число событий до отбора, \tilde{N}_i – число событий, прошедшие отбор



Стерео-режим: первые результаты

- Модельные данные:
 - CORSIKA (предоставлены НИИЯФ МГУ)
 - Всего примерно 18 000 гамма-событий:
- Обучение/тестирование: 12000/6000



Стерео-режим: первые результаты

- Среднеквадратичная ошибка на тестовом наборе уменьшилась в 2 раза.

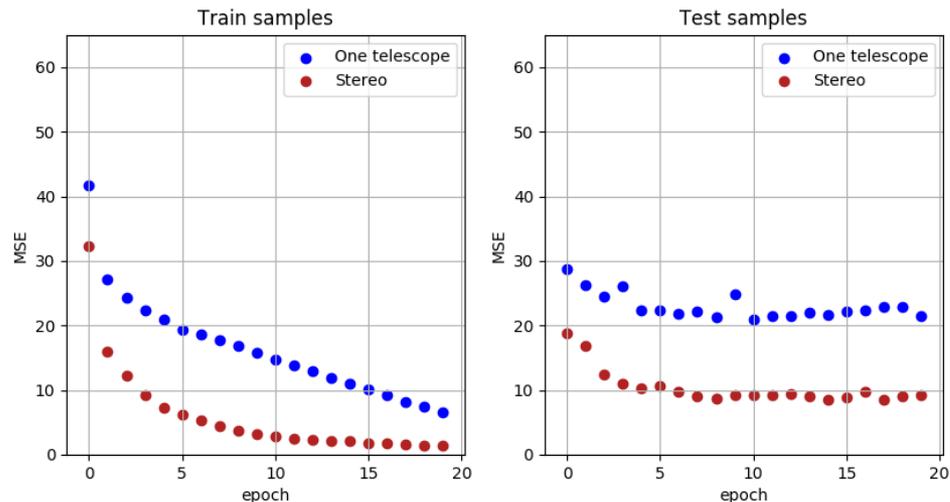
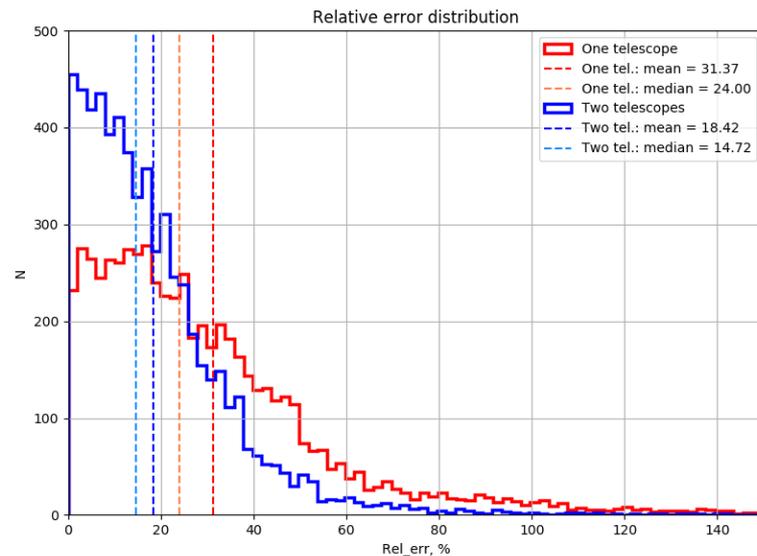


График зависимости ошибки обучения и тестирования в зависимости от эпохи для моно- и стерео-режимов

- Медиана и средняя относительная погрешность уменьшились примерно в 1,5 раза.

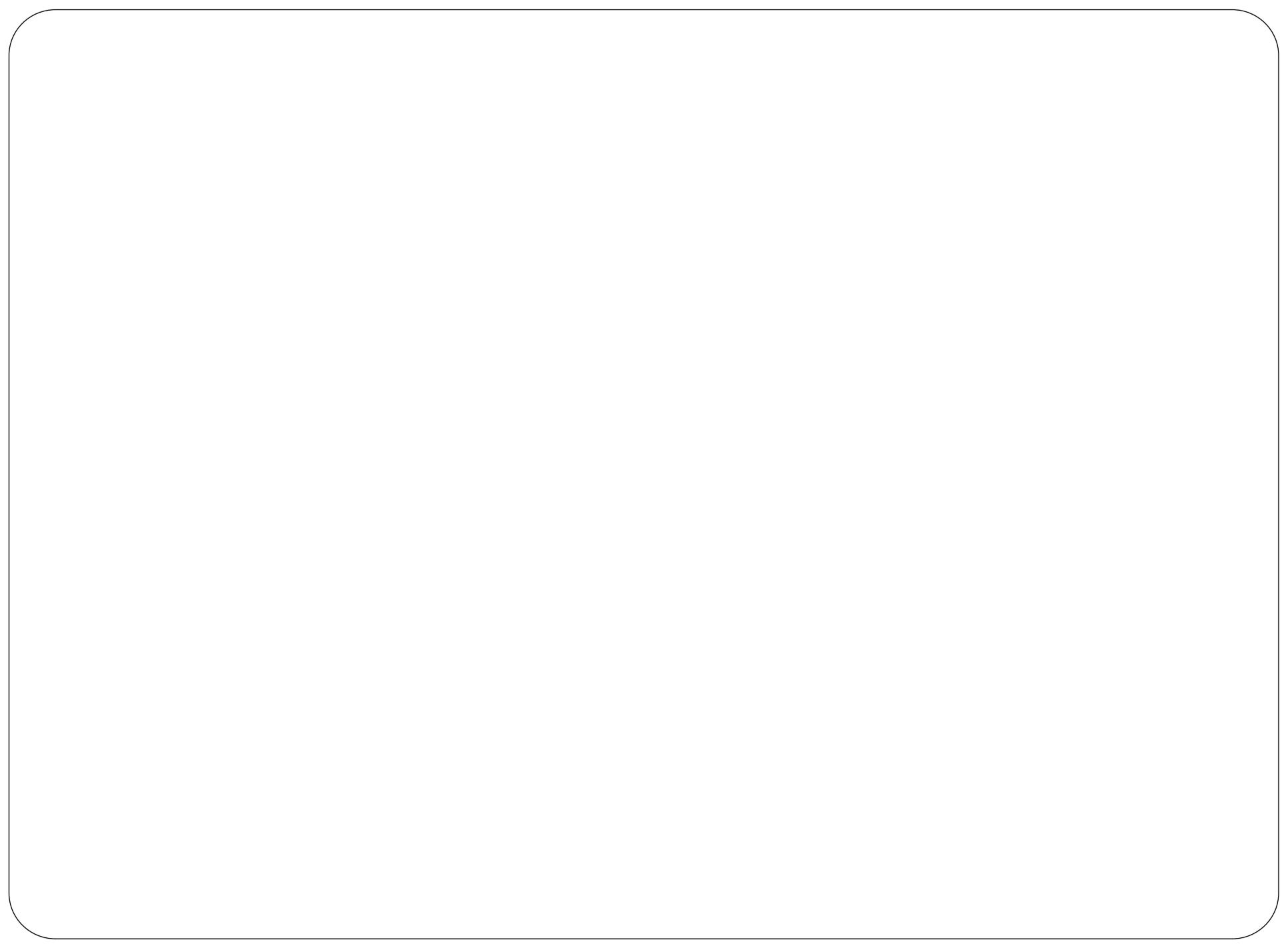


Распределение событий по относительным погрешностям в случае моно- и стерео-режимов

Заключение

- Применение CNN на модельных данных дают хороший результат селекции событий (96%).
- В случае определения энергии первичного гамма-кванта получается удовлетворительная относительная погрешность (30%). Данная погрешность почти линейно уменьшается с увеличением числа телескопов.
- В дальнейших планах:
 - изучение применения стерео-режима с двумя и более телескопами;
 - более тщательное сравнение с традиционными методами классификации и регрессии энергий первичных частиц.

Спасибо за внимание!



Дополнительные слайды

Современные установки ИАСТ

MAGIC

(о. Ла-Пальма, Испания)

Энергии: 0,025–30 ТэВ

$D_{\text{зеркала}}$: 17 м

Всего телескопов: 2



H.E.S.S.

(г. Гамсберг, Намибия)

Энергии: 0,01–10 ТэВ

$D_{\text{зеркала}}$: 12–28 м

Всего телескопов: 5



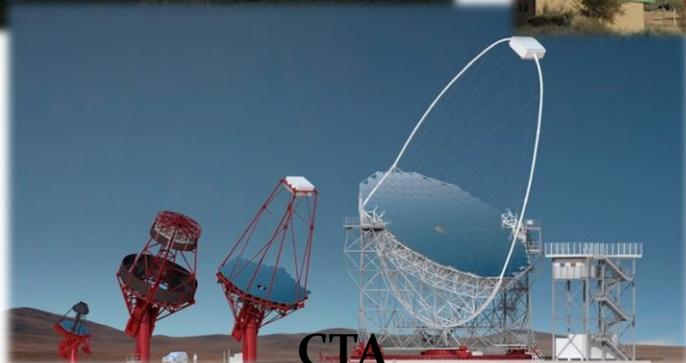
VERITAS

(гора Хопкинс, США)

Энергии: 0,085–30 ТэВ

$D_{\text{зеркала}}$: 12 м

Всего телескопов: 4



CTA

(о. Ла-Пальма, Испания и п. Атакама, Чили)

Энергии: 0,01–10 ТэВ

$D_{\text{зеркала}}$: 4–23 м

Всего телескопов (план): 118



TAIGA-IACT

(Тункинская долина, респ. Бурятия, РФ)

Энергии: 1–100 ТэВ

$D_{\text{зеркала}}$: 4,3 м

Всего телескопов (план): 16

Таблица по расчету параметра качества Q

Ng - полное число настоящих гамма в наборе.

Nh - полное число настоящих адронов

Ng_g - число настоящих гамма, которые были идентифицированы как гамма (True Positive).

Nh_g - число адронов, которые были идентифицированы как гамма (фейковые гамма) (False Positive)

Ng_h - число гамма, идентифицированные как адроны (фейковые адроны) (False Negative)

Nh_h - число адронов, которые были идентифицированы как адроны (True Negative)

Ntot - общее число всех событий в наборе

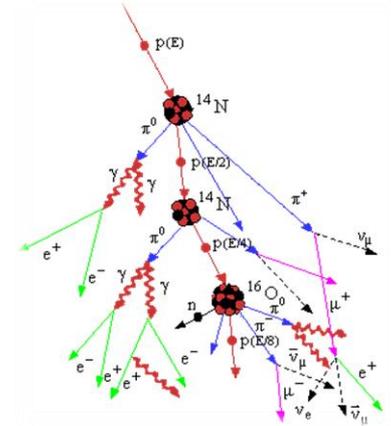
Run - отбор по определенным критериям

Run	Criteria		
1	log(Size)	>	2,05
	Concentration	>	0,55
	Ellipticity	>	0,5
	Density	>	45
	Distance	∈	(0,7; 1,8)
2	log(Size)	>	1,9
	Concentration	>	0,55
	Ellipticity	>	0,5
	Distance	∈	(0,3; 2,0)
3	CNN		

Run	Ntot	Ng	Nh	Ng_g	Nh_g	Ng_h	Nh_h	Q
1	30000	17513	12487	281	1	17233	12486	1,79
2	30000	17513	12487	2631	69	14883	12418	2,02
3	10000	5876	4124	5754	248	122	3876	3,99

Некоторые понятия

- Широкий атмосферный ливень – каскад вторичных частиц, образующийся в результате ядерного взаимодействия первичной частицы (космические лучи) с ядрами атмосферы



- Гиперпараметр – это параметр обучающего алгоритма, который задается вручную

- Кросс-энтропийная функция: $J(\omega) = -\frac{1}{m} \sum_i (y_i \log(p_i) + (1 - y_i) \log(1 - p_i))$

- MSE: $MSE = \frac{1}{m} \sum_i (\hat{y}_i - y_i)^2$