

Статус задачи классификации событий и
определения энергии порождающих
гамма-квантов на основании изображений
черенковских телескопов в эксперименте
TAIGA

Станислав Поляков
НИИЯФ МГУ

Данные наземных телескопов для регистрации черенковского излучения (IACT) позволяют классифицировать события, вызвавшие широкие атмосферные ливни (ШАЛ), а также определять их параметры.

Основная задача, которую мы рассматривали — оценка энергии гамма-событий по изображениям в стереорежиме посредством нейронных сетей.

События

Рассматривались три выборки гамма-событий, сгенерированные посредством алгоритмов Монте-Карло для черенковских телескопов TAIGA:

- m1: 40583 событий, изображения с одного телескопа, диапазон энергий от 1.5 до 60 ТэВ;
- s1: 3400 событий, изображения с двух телескопов на переменном расстоянии от 300 до 350 метров, диапазон энергий от 1 до 45 ТэВ;
- s2: 18359 событий, изображения с двух телескопов на фиксированном расстоянии 324 метра, диапазон энергий от 1 до 50 ТэВ.

Обучающие и тестовые выборки

Энергия гамма-событий оценивалась посредством сверточных нейронных сетей. Обучающие выборки составляли 80% от соответствующей выборки ($m1$, $s1$ или $s2$), тестовые выборки — оставшиеся 20% событий. Результаты (средняя относительная ошибка) усреднялись по 10 итерациям с различными разбиениями исходной выборки.

Результаты: сравнение по выборкам

Если при обучении использовалось только одно изображение (единственное в выборке $m1$ и первое в выборках $s1$ и $s2$), то при наилучшей найденной точности определения энергии относительная ошибка составляла: 23.3% для выборки $m1$, 20.8% для выборки $s1$ (первые изображения), 24.0% для выборки $s2$ (первые изображения).

Результаты: сравнение точности оценок при добавлении данных второго телескопа

Для выборки s1 наилучшие результаты — ошибка 20.8% при оценке с использованием только изображений с первого телескопа, 19.5% при оценке с использованием одного изображения, выбранного по суммарной яркости пикселей, 15.5% при оценке с использованием обоих изображений.

Для выборки s2 наилучшие результаты — ошибка 24.0% при оценке с использованием только изображений с первого телескопа, 20.0% при оценке с использованием одного изображения, выбранного по суммарной яркости пикселей, 12.5% при оценке с использованием обоих изображений.

Предварительная обработка данных

Рассматривались несколько вариантов предварительной обработки данных, например, обнуление отрицательных значений, взятие логарифма (точнее, $\log(\max(x,0)+1)$), сортировка изображений телескопов по суммарной яркости.

Систематическое сравнение по большому числу вариантов показывает, что эти решения влияют на результат, но весьма слабо, меняя среднюю ошибку на $\sim 0.1-0.2\%$

(Наилучшие варианты: использовать для каждого изображения два входных канала, один с обнулением отрицательных сигналов и другой с логарифмом, сортировку по яркости не использовать.)

Архитектура нейронных сетей: объединение изображений

Систематически сравнивались три варианта обработки нейросетями двух изображений: с использованием общего набора сверток, применяющихся параллельно, с использованием двух независимых наборов сверток, и с использованием общего набора сверток, которому доступны одновременно данные с обоих изображений в разных входных каналах. Это также мало влияет на результат: изменение средней ошибки $\sim 0.1-0.2\%$. (Лучшим оказался последний вариант.)

Архитектура нейронных сетей

Число слоев и нейронов в каждом слое, размер фильтров в конволютивных слоях варьировались существенно, но недостаточно для обобщенных выводов о преимуществах тех или иных вариантов. (И эти выводы, возможно, различались бы в зависимости от выборки: среди рассмотренных сетей для выборки m_1 наилучшие результаты дают более простые сети с меньшим числом нейронов, для s_1 — наоборот, более сложные, для s_2 — средние.)

Среди рассмотренных вариантов архитектуры сетей средняя ошибка для выборки s_2 менялась от 12,5% до 13,9%, для s_1 — от 15.5% до 18.4%.

Пример архитектуры нейронной сети

Сеть, дающая наилучшие результаты на выборке s2:

Input: 30x31x4

Conv2D 5x5x12

AvgPool 2x2

Conv2D 5x5x25

AvgPool 2x2

Conv2D 3x3x25

AvgPool 2x2

Flatten 3x3x25 → 225

Fully connected 50

Fully connected 50

Fully connected 50

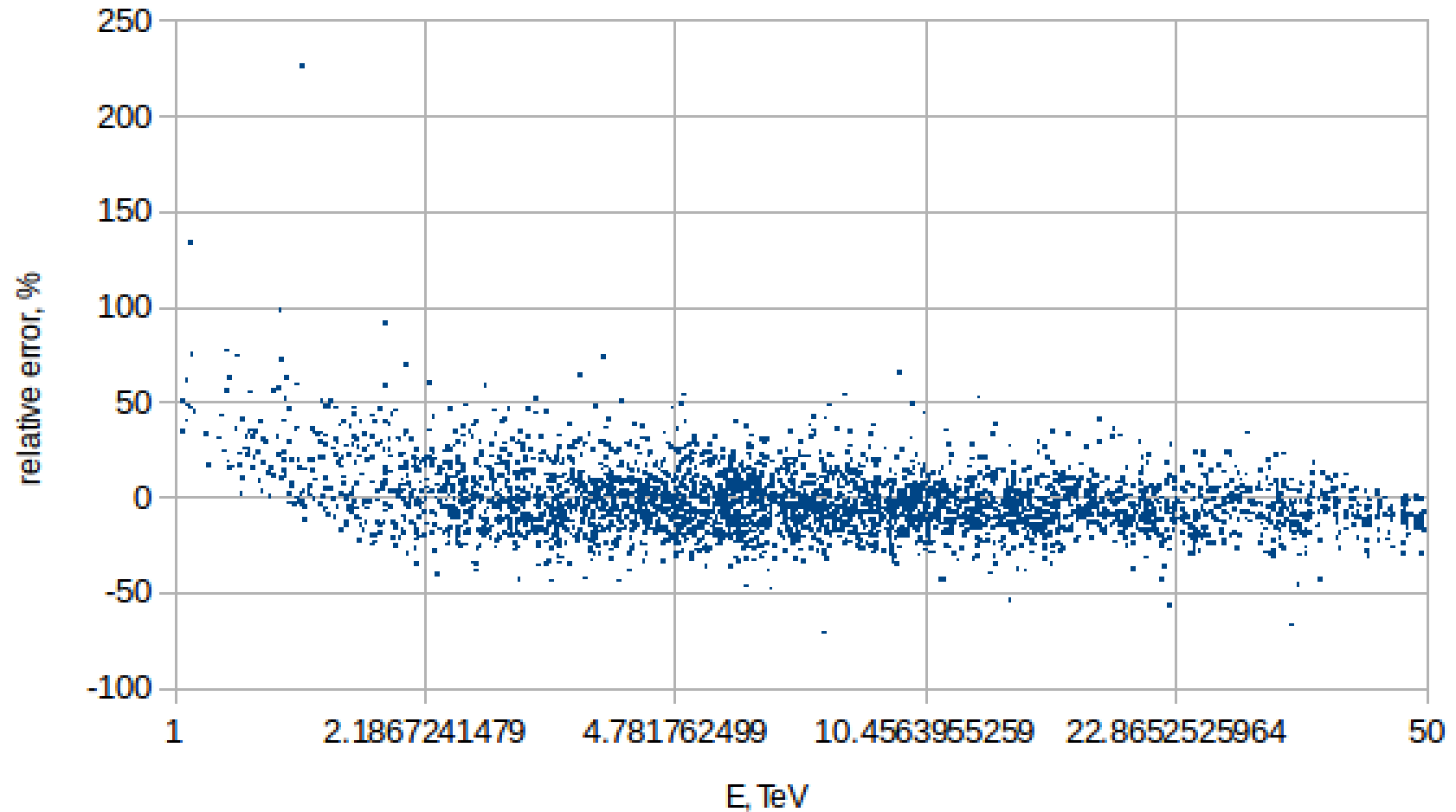
Output 1

Dropout

Дропаут (случайное обнуление выходных значений нейронов) не упомянут как часть архитектуры сети, поскольку при применении обученной нейросети он не используется. Однако при обучении он использовался с одним и тем же параметром вероятности обнуления p после каждого скрытого слоя. Выбор параметра p значительно влияет на результаты: так, для одной и той же сети можно получить среднюю ошибку 13.0% при $p=0.025$ и 31.1% при $p=0.2$.

Для каждой сети обучение выполнялось с несколькими положительными значениями параметра p (до 0.2 с шагом 0.025) и выбирался лучший результат.

Относительная ошибка оценки в зависимости от энергии события



Основные результаты

Были обучены нейросети для оценки энергии по изображениям с черенковских телескопов проекта TAIGA. Показано, что точность оценки можно значительно повысить, добавив изображение со второго телескопа, что влияние архитектуры сети на точность оценки также существенно. В то же время рассмотренные варианты предварительной обработки данных и способов объединения изображений с телескопов влияют на точность оценок незначительно. Лучший из найденных вариантов дает среднюю относительную ошибку 12.5% на большей из двух выборок, для которых доступны изображения двух телескопов.

Упрощенные оценки

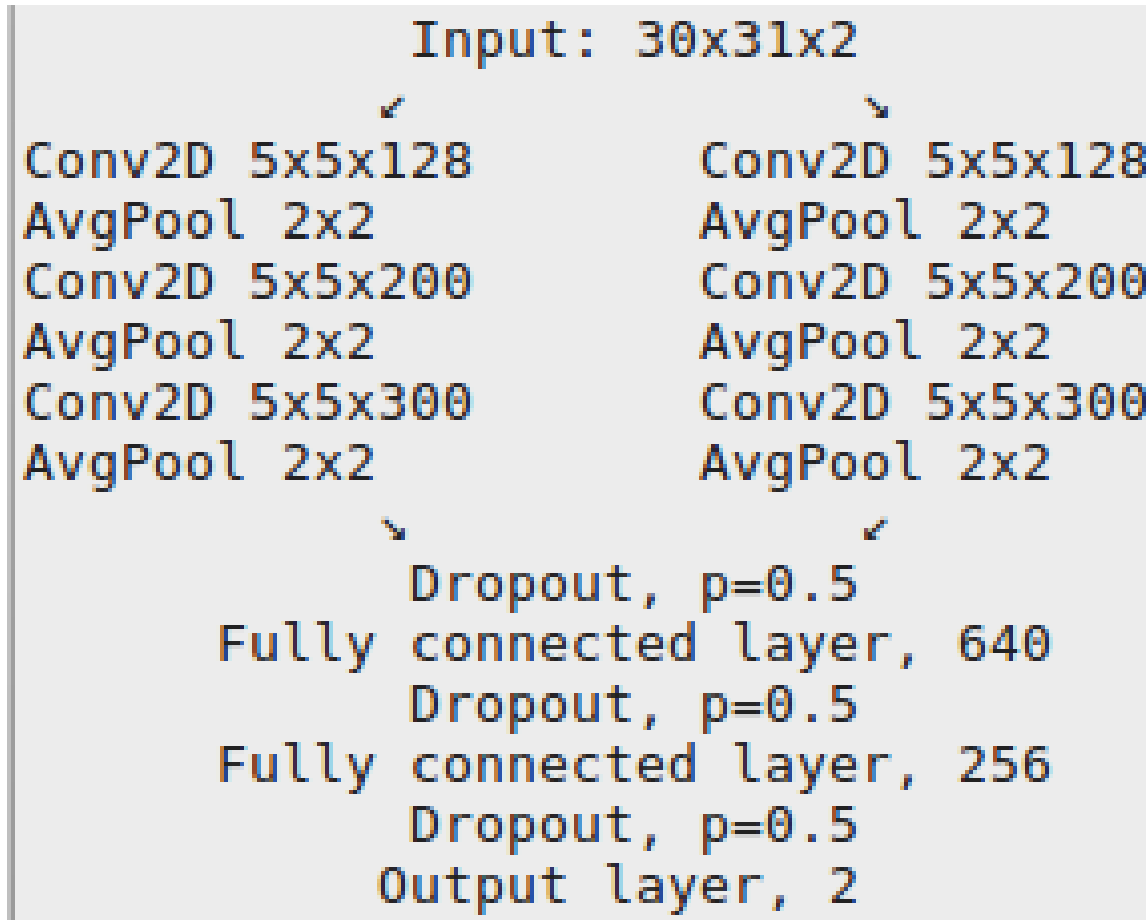
Если всегда давать одну и ту же оценку (учитывая распределение энергии в выборке), можно получить среднюю ошибку 49% для выборки $m1$ (оценка 3.86 ТэВ), 46.5% для $s1$ (оценка 2.81 ТэВ), 57.2% для $s2$ (оценка 3.34 ТэВ).

Можно использовать всего два значения: усредненную яркость с первого и второго телескопов. Этого достаточно, чтобы получить среднюю ошибку 23% для $s1$ и 36.9% для $s2$.

Классификация событий

Задача классификации событий на основании изображений двух телескопов рассматривалась только для выборки s_1 (для s_2 не были доступны изображения для адронных событий).

Архитектура сети для классификации событий



Результаты классификации событий (предварительные)

Сеть обучалась 5 раз. Q-фактор был равен 17.3, 24.2 и в трех случаях бесконечный (т.е. адронных событий, ошибочно распознанных как гамма-события, не было). Для сравнения та же сеть обучалась с использованием только первого (продублированного) изображения, в этом случае Q-фактор получился 6.2, 7.5, 7.7, 8.2, 10.3.

Q-фактор определяется как $Q = \frac{\Gamma/N_\Gamma}{\sqrt{\Gamma'/N_P}}$,

где Γ и Γ' — число гамма- и адронных событий, распознанных как гамма-события, N_Γ и N_P — полное число гамма- и адронных событий.

Необходимое условие: $\Gamma \geq N_\Gamma/2$.