



Лаборатория космических лучей  
предельно высоких энергий НИИЯФ МГУ



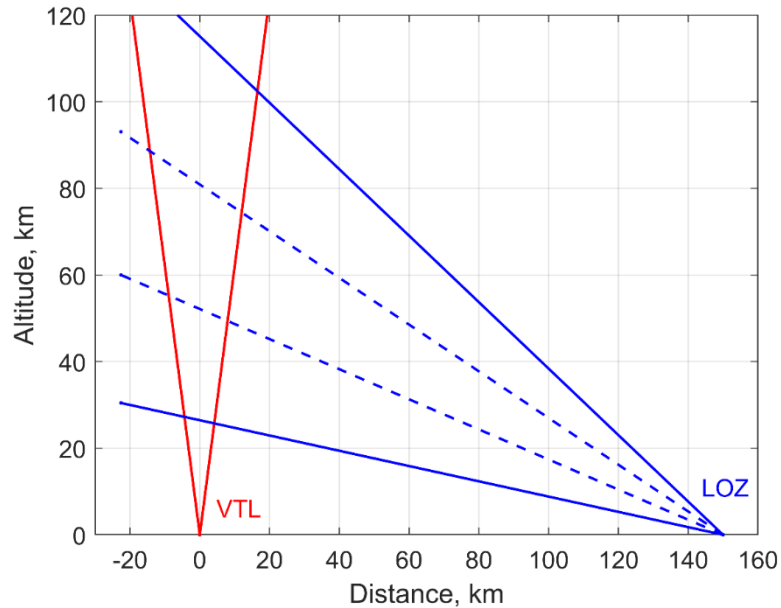
# НЕЙРОСЕТЕВЫЕ АЛГОРИТМЫ ПОИСКА ОПТИЧЕСКИХ МИКРОВСПЛЕСКОВ

Щелканов К.Д.<sup>1,2</sup>, Николаева В.Д.<sup>1</sup>, Климов П.А.<sup>1,2</sup>

*<sup>1</sup>Московский государственный университет имени М.В.Ломоносова, Научно-исследовательский институт ядерной физики имени Д.В.Скобельцына, Москва, Россия*

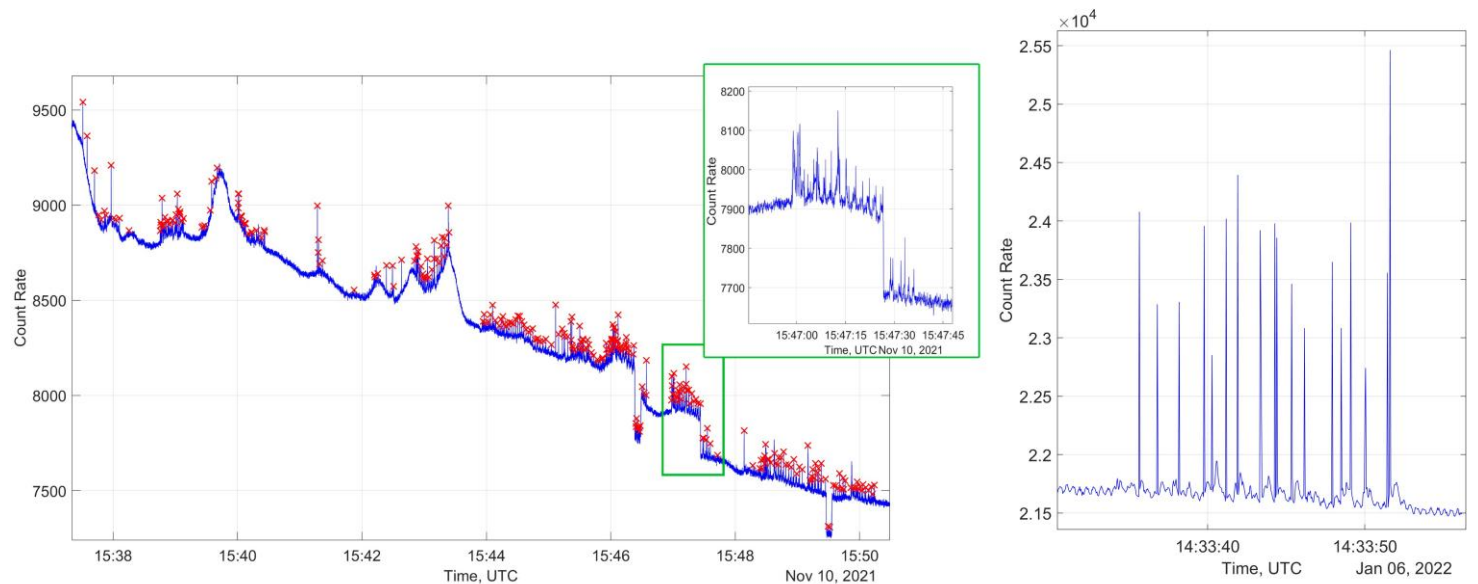
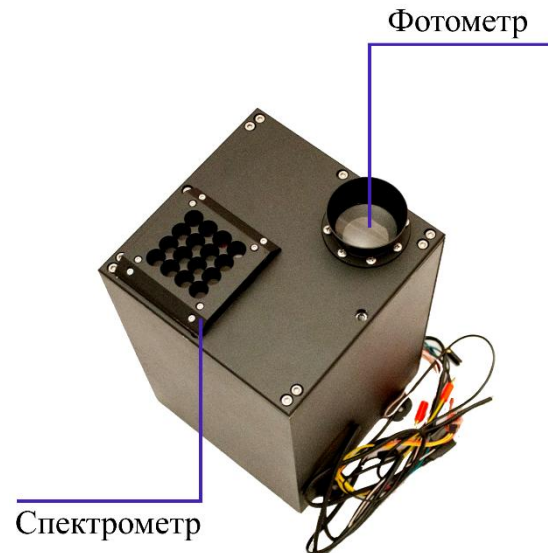
*<sup>2</sup>Московский государственный университет имени М.В. Ломоносова, физический факультет, Москва, Россия*

# Микровсплески оптического излучения в авроральной зоне

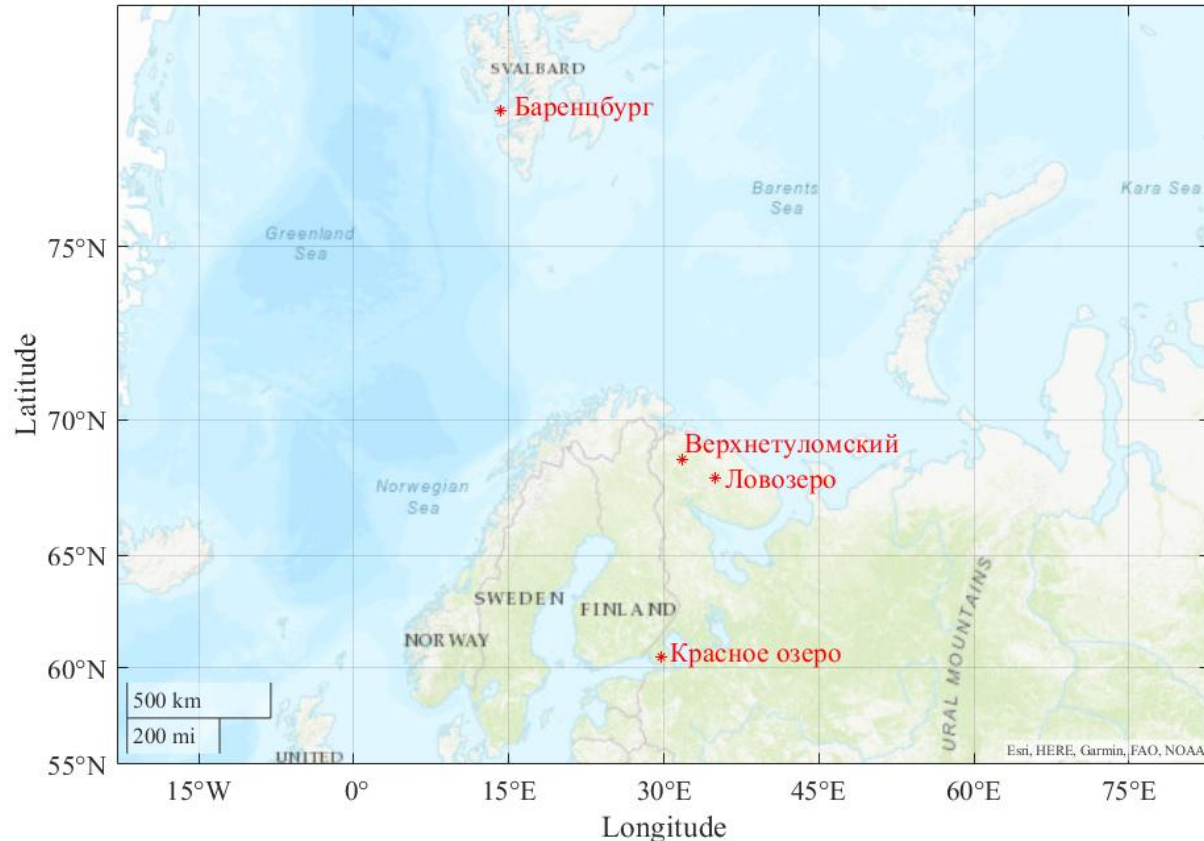


- Детектором RAIP-V (обсерватория VTL) обнаружены микровсплески оптического излучения в авроральной зоне.
- Все обнаруженные события представляют собой группы микровсплесков, состоящие из серий вспышек – несколько подряд идущих импульсов оптического излучения.
- Характерная длительность отдельной вспышки порядка 100 мс.

## Примеры группы и серии УФ-микровсплесков



# Система фотометров PAIPS



В будущем планируется существенное расширение сети фотометров, что сильно усложнит и без того нетривиальную процедуру поиска событий различных типов, в частности, ОМВ.

Уже предпринимались попытки автоматизировать поиск ОМВ в данных с помощью порогового триггера, однако этот подход не показал достаточной эффективности. Причина – высокое разнообразие серий ОМВ по параметрам, наличие событий, внешне похожих на ОМВ, но не являющихся ими (например самолеты), а также зависимость результатов от условий регистрации; в итоге пороговая схема дает много ложных срабатываний и/или пропусков.

С учетом этих ограничений было принято решение разработать нейросетевую модель, способную быстро классифицировать фрагменты сигналов детекторов, по возможности отсеивать ОМВ-подобные фоновые события и обрабатывать большие массивы данных.

# Архитектура нейросети

Вход: 1500 точек, разрешение 40 мс – 1-минутный сигнал

**Conv1D: 32 фильтра**

**Conv1D: 64 фильтра**

**Conv1D: 128 фильтров**

**LSTM: 128 нейронов (return\_sequences=True)**

**LSTM: 64 нейрона (return\_sequences=False)**

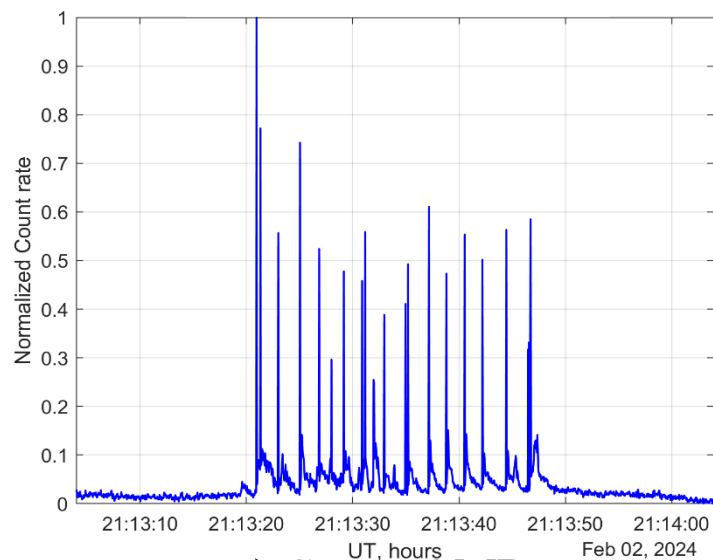
**Dense: 64 нейрона**

**Dense: 32 нейрона**

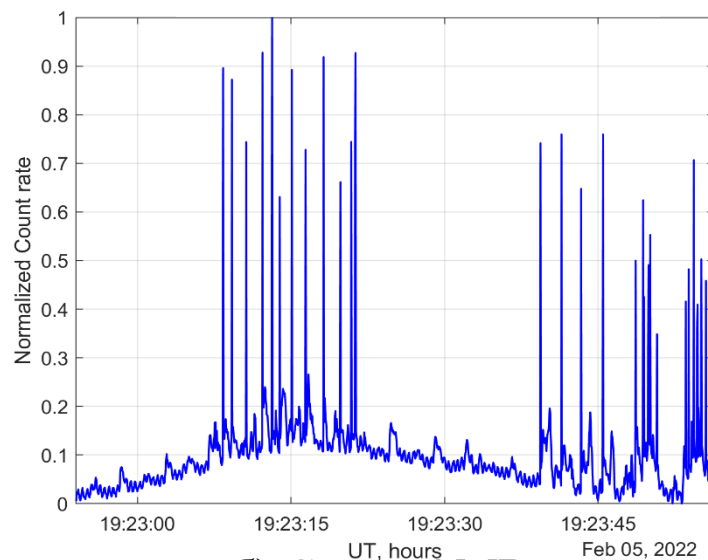
**Dense: 1 нейрон, Sigmoid**

Выход: вероятность микровсплеска [0, 1]

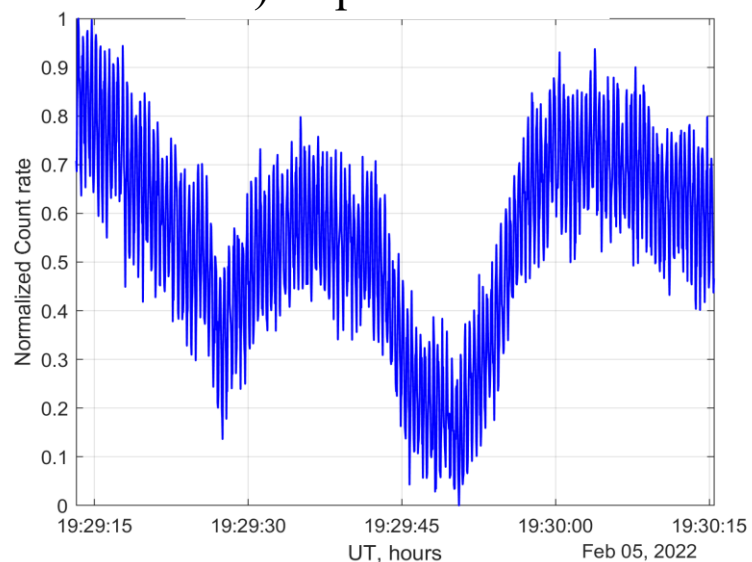
# Обучающая выборка



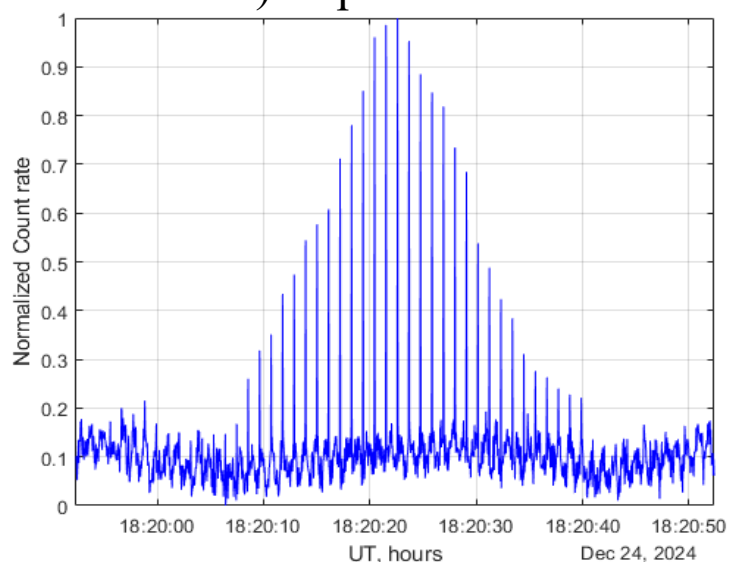
а) Серия ОМВ



б) Серия ОМВ



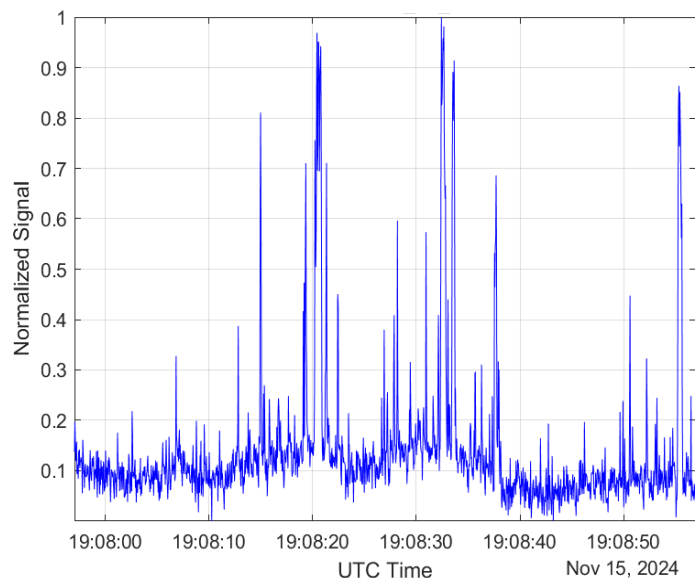
в) Фоновый сигнал



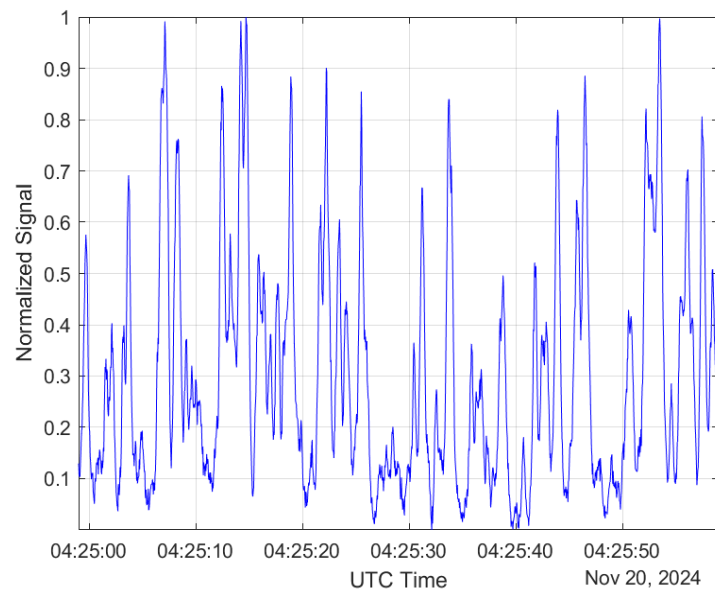
г) Самолет

- Исходные данные предварительно обработаны следующим образом: временные ряды разбиты на минутные интервалы, из которых вычтен фоновый уровень, а сигнал нормирован к единице.
- Аналогичная процедура применяется к данным и для анализа.
- Нормализация необходима для предотвращения смещения модели на конкретные значения фона и амплитуды событий. Обучающая выборка содержит ~2000 событий, сбалансированно распределённых по двум классам.
- На рисунке а) и б) – примеры ОМВ  
в) – пример фонового сигнала  
г) – самолет

# Результаты. Набор 1



a) True positive



б) False positive

True Positive	True Negative	False Positive	False Negative
5	13523	50	0

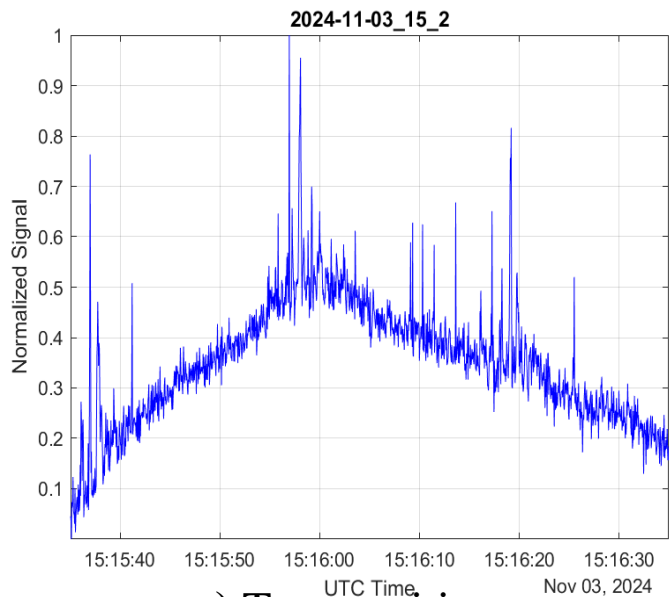
Первый тест – данные измерений детектора RAIPS-V, установленного в Верхнетуломской обсерватории ПГИ с вечера 14 ноября 2024 года по утро 23 ноября 2024 года. Всего в классификации участвовали 13578 минутных отрезков данных. Показатели результата в таблице.

**Accuracy:** 99,63% – модель правильно классифицировала 99.63% всех объектов, но данная метрика непоказательная из-за редкости ОМВ

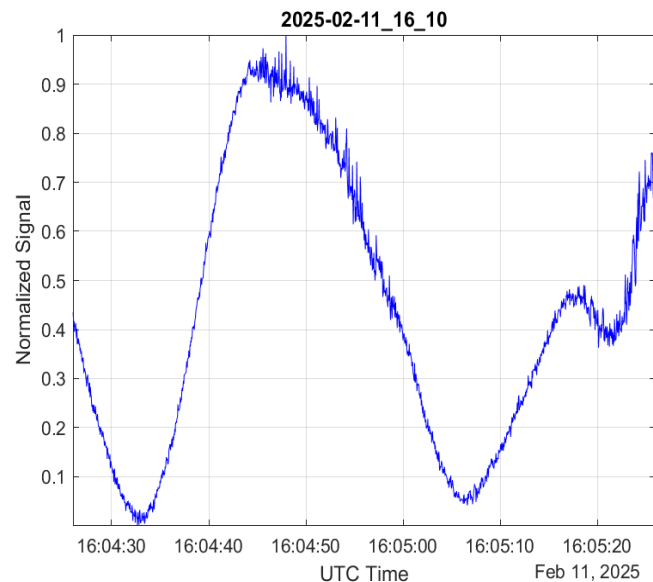
**Precision:** 9,09% – из 55 событий определенных как ОМВ, только 5 действительно ими являются

**Recall:** 100% – модель обнаружила все ОМВ, которые были в выборке

# Результаты. Набор 2



а) True positive



б) False negative

True Positive	True Negative	False Positive	False Negative
14	3887	13	1

Второй тест – набор из нескольких отдельных ночей измерений: 4 ноября 2024, 16 ноября 2024, 12 февраля 2025 и 15 февраля 2025. Примечательны эти данные тем, что в них визуально были обнаружены слабые по соотношению сигнал-шум события. Результаты следующие:

**Accuracy:** 99,64%; **Precision:** 51,85%;  
**Recall:** 93,33%

В данном наборе данных нейросеть не смогла обнаружить одно событие с низким соотношением сигнал-шум, но все остальные были классифицированы верно.

# Заключение

- Гибридная архитектура CNN-LSTM достаточно успешно решает задачу классификации оптических микровсплесков в данных фотометров.
- Высокий Recall (100% на первом наборе, 93.33% на втором) практически гарантирует, что модель не пропускает искомые события, что критически важно для редких явлений.
- Автоматизация поиска кратно ускоряет обработку больших объемов данных RAIPS и позволяет справиться с планируемым расширением системы наблюдений.
- Большая доля ложноположительных срабатываний требует дальнейшей модернизации алгоритмов пред-/постобработки, параметров обучения или обучающей выборки.

Спасибо за внимание