

УДК 004.9

DOI:10.25729/ESI.2025.37.1.003

Прогноз пространственно-временной динамики аврорального овала с применением машинного обучения

Лебедева Анастасия Александровна¹, Гарашенко Александр Алексеевич², Сидоров Денис Николаевич³

¹ПАО Сбербанк,
Россия, Иркутск

²Иркутский национальный исследовательский технический университет,
Россия, Иркутск, *garashchenkoaa@mail.ru*

³Институт систем энергетики им. Л.А. Мелентьева СО РАН,
Россия, Иркутск

Аннотация. Ионосфера – часть атмосферы Земли с высокой концентрацией свободных электронов и ионов. К характерным чертам ионосферы относятся изменчивость и неоднородность. Одной из неоднородностей является так называемый авроральный овал, который определяет диапазон полярного сияния. Распознавание аврорального овала – важная задача для прогнозирования авроральных бурь, так как они влияют на работу систем связи на дальние расстояния, навигацию, связь между спутниками и землей, затрудняя или делая ее невозможной. Таким образом, возникает потребность в обнаружении и прогнозировании перемещения аврорального овала, чтобы быть осведомленным в области их возможного влияния в определенные периоды времени. На основе имеющегося набора изображений, полученных в системе SIMuRG, которые основаны на наборах данных GNSS, предлагается использовать модель LSTM и архитектуру CNN. В статье проводится обзор уже существующих реализаций и впервые предлагается способ применения Convolutional LSTM-архитектуры, которая объединяет обработку временных рядов и компьютерное зрение, для прогнозирования перемещений авроральных овалов на изображениях. Результатом работы является модель машинного обучения, которая может делать прогнозы, основываясь даже на небольших объемах данных.

Ключевые слова: архитектура предсказания кадра, компьютерное зрение, машинное обучение, исследование операций

Цитирование: Лебедева А.А. Прогноз пространственно-временной динамики аврорального овала с применением машинного обучения / А.А. Лебедева, А.А. Гарашенко, Д.Н. Сидоров // Информационные и математические технологии в науке и управлении, 2025. – № 1(37). – С. 25-33 – DOI:10.25729/ESI.2025.37.1.003.

Введение. Магнитосфера Земли эффективно защищает ее поверхность от высокоэнергетических частиц, за исключением полярных областей, где геомагнитное поле вновь соединяется с полем солнечного ветра. Это повторное соединение вызывает проникновение заряженных частиц в атмосферу, нагревая и заряжая нейтральные атомы, производя ионы и электроны в ионосфере. Во время различных событий на Солнце (вспышки, выбросы корональной массы и т.д.) Земли достигает больше заряженных частиц. Кроме того, давление солнечного ветра искажает форму магнитосферы, позволяя большему количеству заряженных частиц проникать в атмосферу и вызывая магнитные бури и подбури. И те, и другие проявляются в полярных областях, как авроры (свечение заряженных атомов атмосферы), нагревая ионосферу и порождая электрические токи и неоднородности различных масштабов в ней [1].

Ионосферные неоднородности влияют на работу систем связи на дальние расстояния, навигацию, локацию, связь между спутниками и землей, затрудняя или делая ее невозможной. Также они оказывают значительное влияние на изменение электромагнитной индукции в Земле, в частности на поведение геомагнитно-индуцированных токов (ГИТ) [2, 3]. Влияние ионосферных неоднородностей различно на разных широтах, и области авроральных овалов под воздействием солнечного ветра индуцируют потоки поверхностных электрических полей,

движущихся в направлении с востока на запад [4]. Таким образом, возникает потребность в обнаружении и прогнозировании перемещения аврорального овала, чтобы быть осведомленными в области их возможного влияния в определенные периоды времени.

Неоднородности малого масштаба, среди прочего, вызывают частые изменения полного электронного содержания (total electron content, TEC), измеряемого на пути от спутника к приемнику. Характер изменений TEC рассчитывается по коэффициенту, именуемому ROTI (Rate of TEC Index). ROTI определяется по следующей формуле:

$$ROTI = \sqrt{\frac{1}{N} \sum_{m=n-\frac{N}{2}}^{n+\frac{N}{2}-1} (ROT_m - \overline{ROT})^2},$$

$$ROT_m = \Delta I / \Delta t.$$

Здесь ΔI обозначает изменение TEC во времени Δt , которое представляет собой временное разрешение измерений (обычно 30 с). Расчет ROTI выполняется на 5-минутном временном интервале. ROTI для определенного времени включает ROTI из интервала $\pm 2,5$ мин, примыкающего к этому времени.

На основе TEC и карт полного ночного неба был создан метод, результатом которого является изображение с обнаруженным на нем авроральным овалом [5]. Однако неоднородности не являются статичными и через определенный промежуток времени могут создать помехи в других областях или выйти из зоны покрытия спутников. Таким образом, целью работы является прогнозирование перемещения авроральных овалов, используя набор изображений на определенном промежутке времени.

Для решения подобных задач используются методы компьютерного зрения, в частности, данная задача может рассматриваться как задача прогнозирования следующего кадра (Next-Frame Prediction) [6].

Прогнозирующее обучение предсказывает возможные события путем анализа имеющейся информации. Существуют несколько практических приложений, где используется предиктивное обучение [7]. Для подобного типа обучения используют рекуррентные сети. Рекуррентные сети подходят для поиска закономерностей в последовательных данных, например, для автомобильного трафика [8], прогнозирования погоды [9] или движений человека [10].

Обучение сети происходит последовательно, согласно временной последовательности изображений. Сеть находит последовательность изменений пикселей между изображениями. Кроме того, для конкретного изображения связь между пикселями является важным фактором, который необходимо учитывать при прогнозировании следующего кадра.

Ключевой признак может быть извлечен из пространственной структуры изображения по положению, внешнему виду и форме объекта. В работе [7] предлагается разделение сетей для предсказания следующего кадра на две архитектуры: «последовательность-один» и «последовательность-последовательность».

В первой архитектуре входными данными для модели глубокого обучения является набор кадров в порядке временных шагов от t до $t + k$. Предсказанием является следующий кадр [11-14].

Для второй архитектуры входными данными являются временные кадры, которые отдельно подаются в нейронную сеть [15-17]. В частности, кадр с временным шагом t подается на вход модели глубокого обучения, а предсказанием является следующий кадр с временным

шагом $t + k$ Эта операция выполняется непрерывно, пока модель глубокого обучения не получит кадр на $(t + k)$ -м временном шаге.

Иными словами, архитектура «последовательность-один» фокусируются на пространственной структуре из набора входных кадров, в то время как архитектуры «последовательность-последовательность» фокусируются на факторе временной последовательности.

Были рассмотрены несколько реализаций на основе сети CNN (Convolutional neural networks) и архитектуры LSTM (Long short-term memory).

Работа «VideoPredictionICLR2016» является представителем архитектуры «последовательность-один». Авторы этой работы [13] обучают сверточную сеть генерировать будущие кадры с учетом входной последовательности. Чтобы справиться с изначально размытыми прогнозами, полученными с помощью стандартной функции потерь среднеквадратичной ошибки (MSE), авторы предлагают три различных и дополнительных стратегии обучения функциям: многомасштабную архитектуру, состязательный метод обучения и функцию потери разницы градиентов изображения. Также авторы сравнивают прогнозы с различными опубликованными результатами, основанными на повторяющихся нейронных сетях в наборе данных UCF101.

Другая работа «Video Frame Synthesis using Deep Voxel Flow» [12] также имеет архитектуру «последовательность-один» и решает проблему синтеза новых видеок кадров в существующем видео либо между существующими кадрами (интерполяция), либо после них (экстраполяция). Авторы обучают сеть, которая учится синтезировать видеок кадры, передавая значения пикселей из существующих, что они называют глубоким потоком вокселей.

Принцип работы состоит в том, что DVF учится синтезировать целевой кадр из входного видео. Целевой кадр может быть промежуточным (интерполяция) или последующим (экстраполяция) входным видео. DVF использует архитектуру полностью сверточного кодера-декодера, содержащую три сверточных слоя, три уровня деконволюции и один слой узкого места. Единственное, что нужно для контроля DVF, – это синтезируемый целевой кадр.

В следующей работе PredRNN++ [18], которая является представителем архитектуры «последовательность-последовательность», предложена рекуррентная сеть для пространственно-временного прогнозирующего обучения. Авторы в своей работе использовали рекуррентную структуру под названием Causal LSTM с каскадной двойной памятью, а также блок Gradient Highway Unit, который обеспечивает альтернативные быстрые маршруты для градиентных потоков от выходов обратно к длинным предыдущим входам.

В работе Folded Recurrent Neural Networks [19], также придерживающейся архитектуры «последовательность-последовательность», использованы биективные рекуррентные блоки с управлением (bGRU). Стандартные GRU обновляют состояние, выставляемое, как выход с учетом входа. Авторы работы расширяют их, рассматривая вход, как другое рекуррентное состояние, и обновляют его с учетом выхода, используя дополнительный набор логических ворот. Укладка нескольких таких слоев приводит к рекуррентному автокодировщику: операторы, обновляющие выходы, составляют кодер, а те, что обновляют входы, образуют декодер. Состояния кодера и декодера являются общими, представление расслаивается в процессе обучения: некоторая информация не передается следующим слоям. Авторы используют только кодер или декодер для кодирования или предсказания. Это снижает вычислительные затраты и позволяет избежать повторного кодирования предсказаний при генерации нескольких кадров. Кроме того, можно удалять слои из обученной модели, что позволяет понять роль каждого слоя.

Несмотря на имеющиеся преимущества у этих работ, они также имеют ряд недостатков. Во-первых, в них используются устаревшая версия библиотеки TensorFlow, что делает невозможным доработку этих работ без полного или частичного переписывания изначального кода. Во-вторых, наборы данных, на которых тестировались работы, рассчитаны на определение человеческих движений или физики объекта, что не совсем подходит для решения нашей задачи.

Однако, методы, на которых основаны эти работы, могут быть использованы и в этом случае, но с определенной спецификой. Поскольку данные, которыми мы располагаем, являются своего рода временным рядом, то для них существует подходящий метод – ConvLSTM [9].

Convolutional LSTM-архитектуры объединяют обработку временных рядов и компьютерное зрение путем внедрения конволюционной рекуррентной ячейки в слой LSTM.

Методы. Для данного исследования использовались данные в виде набора изображений за определенные дни, которые являются совмещением данных ТЕС и данных карт полного ночного неба [5]. Эти данные, собранные за последовательные периоды времени, можно охарактеризовать, как временной ряд. В таких случаях подходящим подходом является использование модели на основе LSTM (Long Short Term Memory) [20], архитектуры рекуррентной нейронной сети. Математически ее можно представить следующим образом:

$$\begin{aligned} f_t &= \sigma_g(W_f x_t + U_f h_{t-1} + b_f), \\ i_t &= \sigma_g(W_i x_t + U_i h_{t-1} + b_i), \\ o_t &= \sigma_g(W_o x_t + U_o h_{t-1} + b_o), \\ c_t &= f_t \circ c_{t-1} + i_t \circ \sigma_c(W_c x_t + U_c h_{t-1} + b_c), \\ h_t &= o_t \circ \sigma_{c_t}, \end{aligned}$$

где x_t – входной вектор, h_t – выходной вектор, c_t – вектор состояний, W, U, b – матрицы и вектор параметров, f_t – вес забывания старой информации, i_t – вес получения новой информации, o_t – кандидат на выход, σ_g – функция активации на основе сигмоиды, σ_c – на основе гиперболического тангенса. В такой архитектуре модель передает предыдущее скрытое состояние на следующий шаг последовательности. Таким образом, сеть хранит информацию о предыдущих данных и использует их для принятия решений.

При работе с изображениями лучшим подходом является архитектура CNN (Convolutional Neural Network) [21]. Изображение проходит через конволюционные слои, в которых несколько фильтров извлекают важные особенности. После последовательного прохождения нескольких конволюционных слоев выходной сигнал подключается к полносвязной плотной сети.

С такими последовательными изображениями одним из использованных подходов являлось использование слоев ConvLSTM. Математическое описание модели LSTM со сверточным слоем можно представить следующим образом:

$$\begin{aligned} f_t &= \sigma_g(W_f * x_t + U_f * h_{t-1} + V_f \circ c_{t-1} + b_f), \\ i_t &= \sigma_g(W_i * x_t + U_i * h_{t-1} + V_i \circ c_{t-1} + b_i), \\ o_t &= \sigma_g(W_o * x_t + U_o * h_{t-1} + V_o \circ c_{t-1} + b_o), \\ c_t &= f_t \circ c_{t-1} + i_t \circ \sigma_c(W_c * x_t + U_c * h_{t-1} + b_c), \\ h_t &= o_t \circ \sigma_{c_t}, \end{aligned}$$

где $*$ обозначает свертку. Это рекуррентный слой, как и LSTM, но внутренние матричные умножения заменены операциями свертки. В результате данные, проходящие через ячейки ConvLSTM, сохраняют размерность входа, а не являются просто одномерным вектором с признаками.

Для обработки данных была решено использовать реализацию этих методов в библиотеке Keras. Для создания датасета используются несколько наборов из шестнадцати кадров с разрешением 320x320 пикселей. Наборы данных загружались в фреймворк с последующим созданием и предварительной обработкой обучающих и проверочных датасетов, которые имели соотношение 90 обучающих к 10 тестовым. На рисунке 1 представлены подготовленные данные.

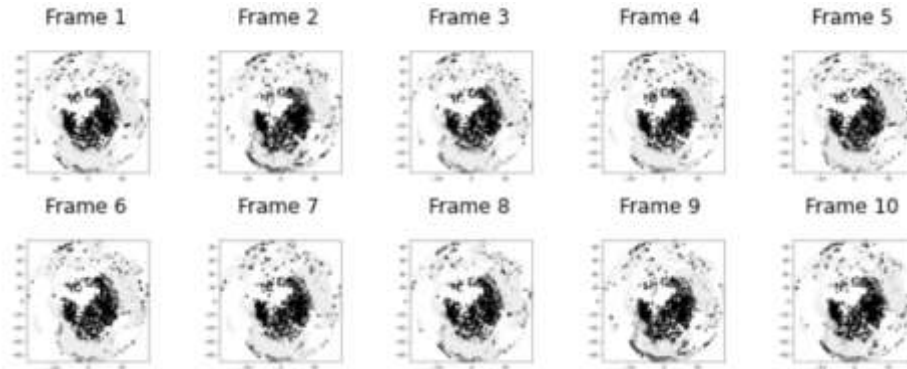


Рис. 1. Подготовленные данные для обучения сети

Для предсказания следующего кадра модель использует предыдущий кадр, который называется f_n , для прогнозирования нового кадра, называемого f_{n+1} . Чтобы модель могла создавать эти прогнозы, данные обработаны так, чтобы были «смещенные» входы и выходы, где входные данные – это кадр x_n , используемый для прогнозирования кадра y_{n+1} .

Для построения конволюционной LSTM-модели используется слой ConvLSTM2D, который принимает входные данные и возвращает прогноз кадра той же формы, то есть сначала строятся 3 слоя «ConvLSTM2D» с нормализацией пакетов, затем слой «Conv3D» для пространственно-временных выходов. После построения модели и подготовки данных определяются параметры и начинается обучение модели.

Для оценки точности полученных результатов было решено использовать Индекс структурного сходства (SSIM). Метрика индекса структурного сходства извлекает из изображения 3 ключевые характеристики: яркость, контрастность и структуру [22]. Сравнение двух изображений выполняется на основе этих трех характеристик.

Эта система рассчитывает индекс структурного сходства между двумя заданными изображениями, который представляет собой значение от -1 до $+1$. Значение $+1$ означает, что два заданных изображения очень похожи или одинаковы, а значение -1 означает, что эти два изображения очень разные. Часто эти значения корректируются, чтобы быть в диапазоне $[0, 1]$, где крайние значения имеют одинаковое значение.

Результаты. После построения и обучения модели, генерируется несколько примеров прогнозирования кадров на основе набора кадров. Выбирается случайный пример из проверочного набора и из него первые два кадра. После этого модель спрогнозирует два новых кадра, которые мы можем сравнить с реальными. Результаты представлены в таблице 1 и на рисунке 2.

Таблица 1. Результаты прогнозирования

Количество наборов	Количество кадров в наборе	Средняя квадратичная ошибка	Индекс структурного сходства
2	16	0.08804311	0.64
10	16	0.071662135	0.76
20	16	0.03555805	0.81

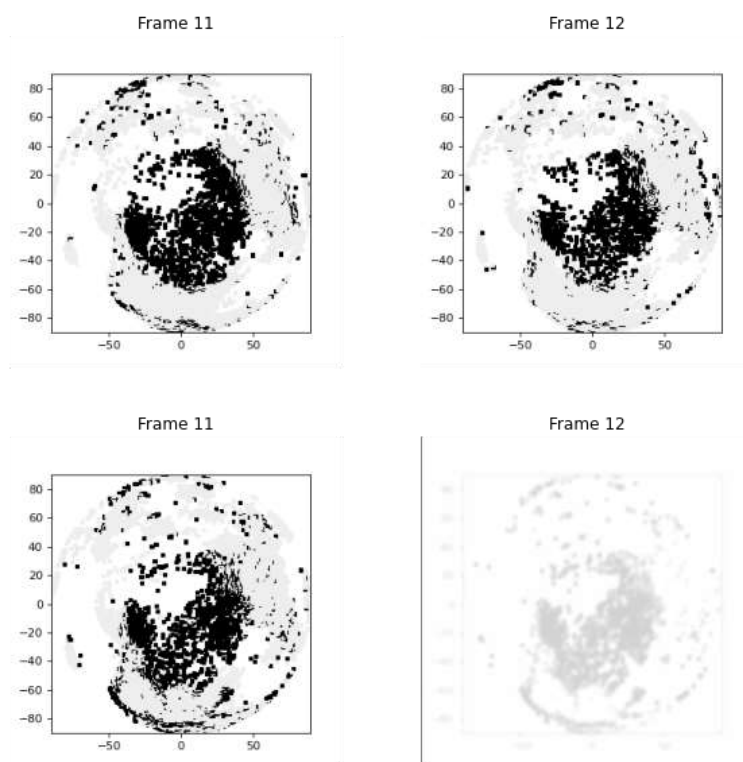


Рис. 2. Реальные кадры (сверху) и прогнозные кадры (снизу)

Заключение. В результате работы была выявлена высокая практическая эффективность в применении Convolutional LSTM-архитектуры для прогнозирования перемещения авроральных овалов на основе изображений, полученных в системе SIMuRG. В работе был использован сравнительно небольшой набор данных, в дальнейшем планируется его увеличение, что приведет к повышению точности модели.

Чтобы результат прогнозирования имел более очевидное практическое применение, требуется разработка методики обнаружения овала на прогнозируемых изображениях, что является следующим этапом этой работы.

Выполненная работа в сочетании с методом определения овалов может способствовать развитию стабильного подхода к определению и прогнозированию границ районов, характеризующихся возникновением сбоев в работе навигационных систем, большими ошибками позиционирования, которые могут возникать даже в средних широтах во время сильных магнитных бурь, когда авроральный овал смещен к экватору.

Благодарности. Работа выполнена в рамках государственного задания FZZS-2024-0003.

Список источников

1. Opgehoorth H.J. et al. Earth's geomagnetic environment – progress and gaps in understanding, prediction, and impacts. *Advances in space research*, 2024, DOI:10.1016/j.asr.2024.05.016.
2. Pulkkinen A. et al. Geomagnetically induced currents: science, engineering, and applications readiness. *Space weather*, 2017, vol. 15, no. 7, p. 828-856, DOI:10.1002/2016SW001501.
3. Rajput V.N. et al. Insight into impact of geomagnetically induced currents on power systems: overview, challenges and mitigation. *Electric power systems research*, 2021, vol. 192, p. 106927, DOI:10.1016/j.epsr.2020.106927.
4. Thomson A. W. P. et al. Present day challenges in understanding the geomagnetic hazard to national power grids. *Advances in space research*, 2010, vol. 45, no. 9, p. 1182-1190, DOI:10.1016/j.asr.2009.11.023.
5. Vasiliev A., Yasyukevich Y., Garashchenko A., et al. Computer vision for GNSS-based detection of the auroral oval boundary. *International journal of Artificial Intelligence*, vol. 19, no. 2, 2021.
6. Yue B., Liang J. Frame prediction using recurrent convolutional encoder with residual learning. *AIP Conference Proceedings*, AIP Publishing, 2018, vol. 1967, no.1, DOI:10.1063/1.5039073.
7. Zhou Y., Dong H., El Saddik A. Deep learning in next-frame prediction: a benchmark review. *IEEE Access*, 2020, vol. 8, pp. 69273-69283, DOI:10.1109/ACCESS.2020.2987281.

8. Zhang J. et al. DNN-based prediction model for spatio-temporal data. Proceedings of the 24th ACM SIGSPATIAL international conference on advances in geographic information systems, 2016, pp. 1-4, DOI:10.1145/2996913.2997016.
9. Shi X. et al. Convolutional LSTM network: A machine learning approach for precipitation nowcasting //Advances in neural information processing systems, 2015, vol. 28, DOI:10.48550/arXiv.1506.04214.
10. Walker J., Gupta A., Hebert M. Dense optical flow prediction from a static image. Proceedings of the IEEE international conference on computer vision, 2015, pp. 2443-2451, DOI:10.1109/ICCV.2015.281.
11. Liang X. et al. Dual motion GAN for future-flow embedded video prediction. Proceedings of the IEEE international conference on computer vision, 2017, pp. 1744-1752, DOI:10.48550/arXiv.1708.00284.
12. Liu Z. et al. Video frame synthesis using deep voxel flow. Proceedings of the IEEE international conference on computer vision, 2017, pp. 4463-4471, DOI:10.1109/ICCV.2017.478.
13. Mathieu M., Couprie C., LeCun Y. Deep multi-scale video prediction beyond mean square error. arXiv preprint, 2015, DOI:10.48550/arXiv.1511.05440.
14. Michalski V., Memisevic R., Konda K. Modeling deep temporal dependencies with recurrent grammar cells. Advances in neural information processing systems, 2014, vol. 27.
15. Lotter W., Kreiman G., Cox D. Deep predictive coding networks for video prediction and unsupervised learning. arXiv preprint, 2016, DOI:10.48550/arXiv.1605.08104.
16. Srivastava N., Mansimov E., Salakhudinov R. Unsupervised learning of video representations using LSTMs. International conference on machine learning. PMLR, 2015, pp. 843-852, DOI:10.48550/arXiv.1502.04681.
17. Finn C., Goodfellow I., Levine S. Unsupervised learning for physical interaction through video prediction. Advances in neural information processing systems, 2016, vol. 29, DOI:10.48550/arXiv.1605.07157.
18. Wang Y. et al. Predrnn++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning. International conference on machine learning. PMLR, 2018, pp. 5123-5132.
19. Oliu M., Selva J., Escalera S. Folded recurrent neural networks for future video prediction. Proceedings of the European conference on computer vision (ECCV), 2018, pp. 716-731.
20. Schmidhuber J. et al. Long short-term memory. Neural Computation, 1997, vol. 9, no. 8, pp. 1735-1780, DOI:10.1162/neco.1997.9.8.1735.
21. Yamashita R. et al. Convolutional neural networks: an overview and application in radiology. Insights into imaging, 2018, vol. 9, pp. 611-629, DOI:10.1007/s13244-018-0639-9.
22. Wang Z. et al. Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing, 2004, vol. 13, no. 4, pp. 600-612, DOI:10.1109/TIP.2003.819861.

Лебедева Анастасия Александровна. ПАО Сбербанк, эксперт по цифровым технологиям аудита отдела аудита обеспечивающих процессов.

Гаращенко Александр Алексеевич. Младший научный сотрудник Иркутского национального исследовательского технического университета. SPIN: 7466-9182, AuthorID: 1016551, garashchenkoa@mail.ru.

Сидоров Денис Николаевич. Институт систем энергетики им. Л.А. Мелентьева СО РАН, Отдел прикладной математики, главный научный сотрудник. SPIN: 6049-9314, AuthorID: 14218, contact.dns@gmail.com, 664074, ул. Лермонтова, 83.

UDC 004.9

DOI:10.25729/ESI.2025.37.1.003

Forecasting the spatiotemporal dynamics of the auroral oval using machine learning

Anastasiia A. Lebedeva¹, Aleksandr A. Garashchenko², Denis N. Sidorov³

¹PJSC Sberbank, Russia, Irkutsk

²National research Irkutsk state technical university, Russia, Irkutsk, *garashchenkoa@mail.ru*

³Melentiev energy systems institute SB RAS, Russia, Irkutsk

Abstract. The ionosphere is a part of the Earth's atmosphere with a high concentration of free electrons and ions. The characteristic features of the ionosphere include variability and heterogeneity. One of the heterogeneities is the so-called auroral oval, which determines the range of the polar lights. Recognition of the auroral oval is an important task for predicting auroral storms, since they affect the operation of long-range communication systems, navigation, communication between satellites and the ground, making it difficult or impossible. Thus, there is a need to detect and predict the movement of the auroral oval in order to be aware of the area of their possible influence in certain periods of time. Based on the available set of images obtained in the SIMuRG system, which are based on GNSS datasets, it is proposed to use the LSTM model and the CNN architecture. The paper reviews existing implementations and proposes a method for predicting auroral oval movements in images using a Convolutional LSTM architecture that combines time series processing and computer vision. The result is a machine learning model that can make predictions based on even small amounts of data.

Keywords: frame prediction architecture, computer vision, machine learning, operations research

Acknowledgements: the work was carried out within the framework of the state assignment FZZS-2024-0003.

References

1. Opgenoorth H.J. et al. Earth's geomagnetic environment – progress and gaps in understanding, prediction, and impacts. *Advances in space research*, 2024, DOI:10.1016/j.asr.2024.05.016.
2. Pulkkinen A. et al. Geomagnetically induced currents: science, engineering, and applications readiness. *Space weather*, 2017, vol. 15, no. 7, p. 828-856, DOI:10.1002/2016SW001501.
3. Rajput V.N. et al. Insight into impact of geomagnetically induced currents on power systems: overview, challenges and mitigation. *Electric power systems research*, 2021, vol. 192, p. 106927, DOI:10.1016/j.epsr.2020.106927.
4. Thomson A. W. P. et al. Present day challenges in understanding the geomagnetic hazard to national power grids. *Advances in space research*, 2010, vol. 45, no. 9, p. 1182-1190, DOI:10.1016/j.asr.2009.11.023.
5. Vasiliev A., Yasyukevich Y., Garashchenko A., et al. Computer vision for GNSS-based detection of the auroral oval boundary. *International journal of Artificial Intelligence*, vol. 19, no. 2, 2021.
6. Yue B., Liang J. Frame prediction using recurrent convolutional encoder with residual learning. *AIP Conference Proceedings*, AIP Publishing, 2018, vol. 1967, no.1, DOI:10.1063/1.5039073.
7. Zhou Y., Dong H., El Saddik A. Deep learning in next-frame prediction: a benchmark review. *IEEE Access*, 2020, vol. 8, pp. 69273-69283, DOI:10.1109/ACCESS.2020.2987281.
8. Zhang J. et al. DNN-based prediction model for spatio-temporal data. *Proceedings of the 24th ACM SIGSPATIAL international conference on advances in geographic information systems*, 2016, pp. 1-4, DOI:10.1145/2996913.2997016.
9. Shi X. et al. Convolutional LSTM network: A machine learning approach for precipitation nowcasting // *Advances in neural information processing systems*, 2015, vol. 28, DOI:10.48550/arXiv.1506.04214.
10. Walker J., Gupta A., Hebert M. Dense optical flow prediction from a static image. *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 2443-2451, DOI:10.1109/ICCV.2015.281.
11. Liang X. et al. Dual motion GAN for future-flow embedded video prediction. *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 1744-1752, DOI:10.48550/arXiv.1708.00284.
12. Liu Z. et al. Video frame synthesis using deep voxel flow. *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 4463-4471, DOI:10.1109/ICCV.2017.478.
13. Mathieu M., Couprie C., LeCun Y. Deep multi-scale video prediction beyond mean square error. *arXiv preprint*, 2015, DOI:10.48550/arXiv.1511.05440.
14. Michalski V., Memisevic R., Konda K. Modeling deep temporal dependencies with recurrent grammar cells. *Advances in neural information processing systems*, 2014, vol. 27.
15. Lotter W., Kreiman G., Cox D. Deep predictive coding networks for video prediction and unsupervised learning. *arXiv preprint*, 2016, DOI:10.48550/arXiv.1605.08104.
16. Srivastava N., Mansimov E., Salakhudinov R. Unsupervised learning of video representations using LSTMs. *International conference on machine learning*. PMLR, 2015, pp. 843-852, DOI:10.48550/arXiv.1502.04681.
17. Finn C., Goodfellow I., Levine S. Unsupervised learning for physical interaction through video prediction. *Advances in neural information processing systems*, 2016, vol. 29, DOI:10.48550/arXiv.1605.07157.
18. Wang Y. et al. Predrnn++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning. *International conference on machine learning*. PMLR, 2018, pp. 5123-5132.
19. Oliu M., Selva J., Escalera S. Folded recurrent neural networks for future video prediction. *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 716-731.
20. Schmidhuber J. et al. Long short-term memory. *Neural Computation*, 1997, vol. 9, no. 8, pp. 1735-1780, DOI:10.1162/neco.1997.9.8.1735.

21. Yamashita R. et al. Convolutional neural networks: an overview and application in radiology. Insights into imaging, 2018, vol. 9, pp. 611-629, DOI:10.1007/s13244-018-0639-9.
22. Wang Z. et al. Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing, 2004, vol. 13, no. 4, pp. 600-612, DOI:10.1109/TIP.2003.819861.

Lebedeva Anastasiia Aleksandrovna. Sberbank PJSC, expert in digital audit technologies of the audit department of supporting processes.

Garashchenko Aleksandr Alekseevich. Junior research fellow at Irkutsk national research technical university. SPIN: 7466-9182, AuthorID: 1016551, garashchenkoa@mail.ru.

Sidorov Denis Nikolaevich. Melentiev energy systems institute SB RAS, Department of applied mathematics, chief research scientist. SPIN: 6049-9314, AuthorID: 14218, contact.dns@gmail.com, 664074, Lermontov str., 83.

Статья поступила в редакцию 21.11.2024; одобрена после рецензирования 26.02.2025; принята к публикации 05.03.2025.

The article was submitted 11/21/2024; approved after reviewing 02/26/2025; accepted for publication 03/05/2025.