

Применение нейронных сетей для обнаружения аномалий в развитии растений

А.А. Кондратьев, С.Б. Березин

Аннотация—Машинное обучение и нейросети стали привычным инструментом в задачах, связанных с обработкой и анализом изображений. Одной из таких задач является обнаружение аномалий на изображениях или аномальных изображений. Все чаще и чаще компьютерное зрение помогает цифровизировать производство и выявить закономерности в сельском хозяйстве. Технология позволяет человеку ускорить анализ, но полностью заменить опытного специалиста не представляется возможным. В настоящей работе методы машинного обучения применяются для обнаружения аномальных последовательностей изображений растений. Полученные результаты могут найти применение в задачах автоматизации сельского хозяйства, где одним из основных средств мониторинга состояния растений является фото или видеосъемка. Разработанный программный комплекс позволит автоматически обнаруживать проблемы со здоровьем растений или окружающей их инфраструктурой, что поможет вовремя принимать соответствующие меры.

В рамках данной статьи реализован автоматический сбор и предобработка данных, формирование из них датасетов для обучения, а также два нейросетевых подхода к обнаружению аномалий в процессе развития растений с помощью изображений. Проведены эксперименты с реальными данными, произведен сравнительный анализ работоспособности методов в конкретной задаче.

Ключевые слова—Anomaly detection, Computer Vision, GAN, Machine learning

I. ВВЕДЕНИЕ

Агропромышленный комплекс считается одним из самых консервативных в плане внедрения инноваций. Тем не менее, в настоящее время одной из самых актуальных тенденций в сельском хозяйстве стало применение искусственного интеллекта, позволяющего принимать решения на основе поступающих данных, а также частично – а иногда и полностью – заменять человека в производственном процессе. Основные области применения методов искусственного интеллекта связаны с обнаружением болезней растений, классификацией и идентификацией сорняков, определением, подсчетом и сбором урожая, управлением водными ресурсами и почвой, прогнозированием погоды, определением поведения животных. Для большинства сельскохозяйственных процессов, требующих оптимизации, необходимо проведение предварительных исследований и экспериментов. Исходя из этого, первостепенной является задача сбора данных. Развитие телеметрии и распространение интернета вещей (internet of things, IoT) позволяет соби-

рать большое количество данных и своевременно передавать их для принятия решений.

Зачастую исходными данными для принятия решения являются результаты аудио- и видео-наблюдения. За последние годы нейронные сети стали привычным методом для обработки и интерпретации этого типа данных. С повсеместным ростом количества камер и повышением качества их съемки, растут и размеры видеоданных.

В случае автоматического контроля процесса роста растений, естественно сформулировать задачу, как обнаружение аномалий во временных рядах, где элементами ряда являются изображения растений. Аномалия – момент или период, когда наблюдается необычное поведение системы. Необходимость своевременного обнаружения таких ситуаций наблюдается в самых различных сферах: здравоохранении [1], энергетике [2], вычислительных технологиях [3] и сельском хозяйстве [4].

В большинстве случаев возможные аномалии в видеоряде могут обладать большим разнообразием вместе с отсутствием какой-либо априорной информации.

Предполагая, что модель может обобщаться только на данные из того же распределения, что и обучающая выборка, аномальные действия в тестовом наборе будут проявляться как отклонение от обычных шаблонов. Общий подход к решению этой проблемы состоит в том, чтобы изучить модель, которая может захватывать обычные шаблоны в последовательностях кадров во время обучения и проверить, существует ли какой-либо нерегулярный шаблон, который расходится с обычными закономерностями в тестовых видео-фрагментах.

Ранние исследования использовали признаки, сформированные вручную, для представления видео-ряда. Например, в работе [5] авторы представили смеси динамических текстур и определили выбросы в рамках этой модели как аномалии закономерности.

Гораздо лучшие результаты показывают подходы на основе состязательного обучения. Например, в [6] разработали подход к обнаружению аномалий на основе генеративно состязательных сетей (GAN), следуя условной структуре GAN. Некоторые работы используют генеративные модели на основании U-Net [7].

В настоящей работе используется генеративно-состязательная сеть TadGAN [8] и алгоритм One-class Deep SVDD [9]. Алгоритмы являются многообещающими, так как они успешно показали себя в задачах обнаружения аномалий, следовательно, имеет смысл применить их к видеоряду.

II. ЦЕЛЬ РАБОТЫ

Целью настоящей работы является применение нейронных сетей двух типов для обнаружения аномалий в процессе развития растений, определение архитектур сетей, гиперпараметров каждой сети и сравнение качества обнаружения аномалий.

III. ИСХОДНЫЕ ДАННЫЕ

Источник данных – видеокамера, расположенная над полкой с растениями, которая один раз в час делает снимок и отправляет его в облачное хранилище. На Рис.1 изображен кадр с камеры после компенсации искажений.



Рис. 1: Рабочая область

Далее из кадров, полученных с камеры формируются индивидуальные изображения растений, на которых каждое растение находится в центре. Общее число таких изображений – около 34 тысяч. В используемом датасете присутствуют четыре вида растений (Рис. 2).



Рис. 2: Растения разных видов

Изображения отдельных растений снимки были вручную размечены на нормальные и аномальные. В датасете можно выделить следующие типы аномалий (Рис. 3):

- 1) Связанные со здоровьем растений
 - Неожиданное изменение цвета листьев
 - Увядание листьев
 - Отсутствие растения
- 2) Относящиеся к несовершенству установки
 - Листья соседнего растения в кадре
 - Лужи воды в кадре



Рис. 3 Примеры аномалий

Разумеется, в плане обнаружения аномалий нас больше интересует именно первый тип – связанный с состоянием растений. Однако это не значит, что аномалии, связанные с параметрами установки можно игнорировать. Выключенный свет или лужи воды в рабочей области

могут сыграть важную роль в дальнейшем развитии растения, поэтому являются серьезным основанием для превентивных мер.

Важно заметить, что в процессе обучения задействованы не отдельные кадры, а их последовательности. Поэтому проводится анализ последовательностей из 8 изображений с шагом в 10 часов. Это позволяет обнаруживать контекстные аномалии. Так, например, на последовательности изображений на Рис. 4 ни один из кадров не является аномальным, однако посередине растение как будто подменили.



Рис. 4: Аномальная последовательность

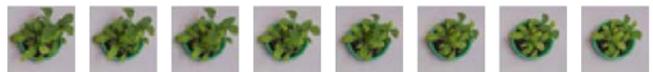


Рис. 5: Деградация растения

Еще одним примером является увядание листьев (Рис. 5). Если во время такого процесса цвет изменяется несильно, то кадры с точки зрения сети будут являться нормальными, в то время как происходит деградация, а не развитие. Анализ последовательностей кадров дает возможность обнаружения таких ситуаций.

IV. ИСПОЛЬЗУЕМЫЕ МЕТОДЫ

A. TadGAN

В основе генеративно-состязательных сетей (GAN [10]) лежит комбинация двух нейронных сетей, одна из которых генерирует образцы, а другая старается отличить сгенерированные образцы от реальных. Такие сети хорошо зарекомендовали себя во множестве задач связанных с изображениями: генерация изображений, их преобразование, предсказание видеоряда. В статье [8] был предложен алгоритм TadGAN, продемонстрировавший в области обнаружения аномалий во временных рядах лучшую производительность среди прочих конкурирующих подходов. В связи с этим было решено реализовать этот метод и применить его к исследуемой задаче.

TadGAN является примером метода обнаружения аномалий, основанного на реконструкции данных (reconstruction-based). Суть таких методов в том, чтобы некоторым образом закодировать исходные данные, а затем восстановить. Модель, обученная на реальных данных без аномалий, не сможет произвести восстановление аномалий так же хорошо, как нормальных образцов, т.к. аномалии теряют часть информации во время преобразований.

В основе алгоритма TadGAN лежит обучение двух функций (нейронных сетей) $\mathcal{E}: X \rightarrow Z$ и $\mathcal{G}: Z \rightarrow X$, где X – пространство входных данных, состоящее из последовательностей исходных изображений

$$x_i \in X = [0,1]^{C \times D \times H \times W},$$

где C – число каналов, D – количество кадров в последовательности, H, W – высота и ширина каждого кадра.

Z – многомерное нормальное распределение, т.е. $z \sim P_z = \mathcal{N}(0,1)$. С помощью данных функций мы мо-

жем закодировать, а затем восстановить последовательность изображений

$$x_i \rightarrow \mathcal{E}(x_i) \rightarrow \mathcal{G}(\mathcal{E}(x_i)) \approx \hat{x}_i$$

Параллельно также обучаются дискриминаторы \mathcal{C}_x и \mathcal{C}_z . Задача \mathcal{C}_x - отличить реальную временную последовательность от сгенерированной \mathcal{G}_z . \mathcal{C}_z , в свою очередь, оценивает качество отображения в пространство скрытых признаков. В статье [8] эти задачи описаны с помощью следующих формул.

$$V_X(\mathcal{C}_x, \mathcal{G}) = E_{x \sim P_X}[\mathcal{C}_x(x)] - E_{z \sim P_Z}[\mathcal{C}_x(\mathcal{G}(z))]$$

$$V_Z(\mathcal{C}_z, \mathcal{E}) = E_{z \sim P_Z}[\mathcal{C}_z(z)] - E_{x \sim P_X}[\mathcal{C}_z(\mathcal{E}(x))]$$

$$V_{L2}(\mathcal{E}, \mathcal{G}) = E_{x \sim P_X}[\|x - \mathcal{G}(\mathcal{E}(x))\|_2]$$

Таким образом, в процессе обучения решается следующая задача оптимизации.

$$\min_{\{\mathcal{E}, \mathcal{G}\}} \max_{\{\mathcal{C}_x \in \mathcal{C}_x, \mathcal{C}_z \in \mathcal{C}_z\}} V_X(\mathcal{C}_x, \mathcal{G}) + V_Z(\mathcal{C}_z, \mathcal{E}) + V_{L2}(\mathcal{E}, \mathcal{G})$$

Перед подачей на вход нейронным сетям исходные изображения были уменьшены до размера 32 x 32 пикселя. Таким образом, $x_i \in [0, 1]^{3 \times 8 \times 32 \times 32}$. В качестве архитектур использовались сверточные нейронные сети с 3D свертками. Схематично изображены на Рис. 6 - 8

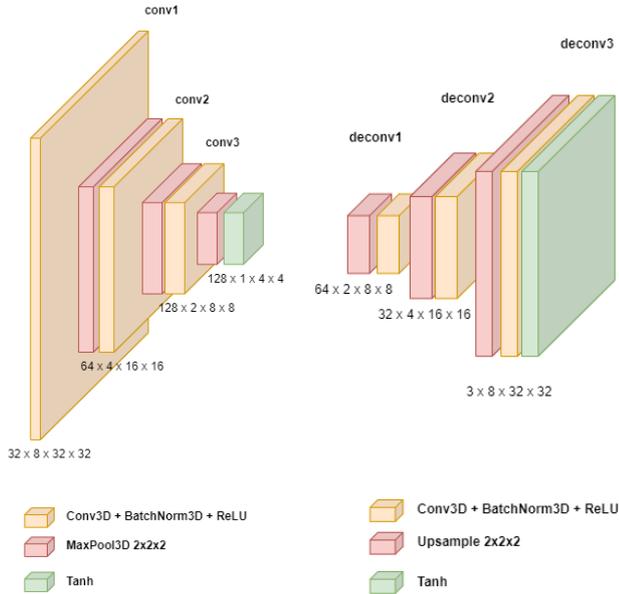


Рис. 6: Архитектура \mathcal{E} и \mathcal{G}

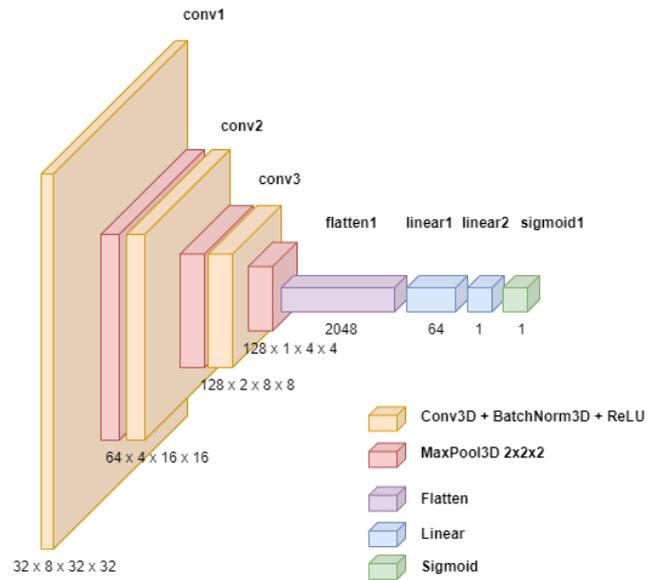


Рис. 7: Архитектура \mathcal{C}_x

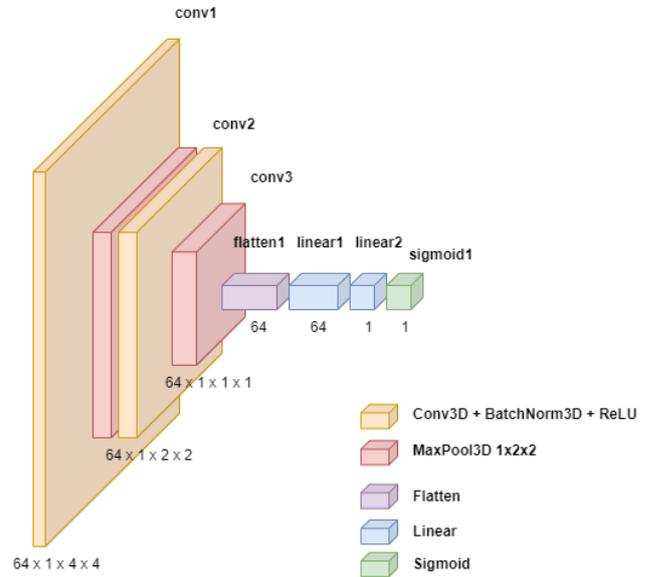


Рис. 8: Архитектура \mathcal{C}_z

B. One-class Deep SVDD

Deep SVDD (Deep Support Vector Data Description) [9] - метод машинного обучения, который также используется для обнаружения аномалий в данных. Он основан на алгоритме Support Vector Data Description (SVDD), который используется для построения границы вокруг нормальных данных. Deep SVDD, в свою очередь, использует нейронную сеть для преобразования данных в пространство скрытых признаков и построения в этом пространстве гиперсферы минимального радиуса, заключающей в себя все нормальные образцы. Те образцы, которые после преобразования не попадают в эту гиперсферу, считаются аномальными.

В алгоритме требуется обучить нейросеть $\phi(\cdot; \mathcal{W}): \mathcal{X} \rightarrow \mathcal{F} \in R^p$ с $L \in N$ скрытыми слоями и набором весов $\mathcal{W} = \{\mathcal{W}^1, \dots, \mathcal{W}^L\}$, где \mathcal{W}^l - веса слоя под номером $l \in \{1, \dots, L\}$. Помимо этого, требуется найти некоторый центроид $c \in \mathcal{F}$, и ставится следующая задача минимизации.

$$\min_{\mathcal{W}} \frac{1}{n} \sum_{i=1}^n \|\phi(x_i; \mathcal{W}) - c\|^2 + \frac{\lambda}{2} \sum_{l=1}^L \|\mathcal{W}^l\|_F^2$$

В оригинальной статье алгоритм был протестирован на датасете CIFAR-10, для чего использовалась сверточная нейронная сеть. В нашей работе в нее были внесены изменения: 2D слои (Conv, MaxPool, BatchNorm и т.п.) были заменены на соответствующие 3D аналоги, а также добавлен Positional Encoding [11]. Полученная архитектура изображена на Рис. 9

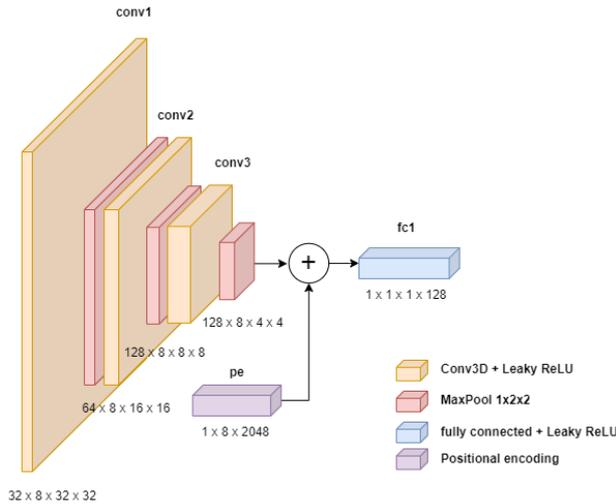


Рис. 9 Архитектура Deep SVDD

V. РЕЗУЛЬТАТЫ

A. TadGAN

На выходе алгоритма мы получаем восстановленную последовательность изображений, а также выход C_x – вероятность того, что изображение реальное. Далее происходит вычисление различных метрик между входными и восстановленными изображениями: среднеквадратическое отклонение MSE , MSE по отдельным каналам (MSE_R, MSE_G, MSE_B), индекс структурного сходства SSIM [12], отношение сигнала к ошибке восстановления (SRE). Полученные значения в дальнейшем используются как признаки для алгоритма Isolation Forest [13], с помощью которого считается вероятность принадлежности к каждому классу. На Рис. 10 проиллюстрировано расположение нормальных и аномальных образцов на плоскости. В качестве координат используются значения MSE и C_x score. Видно, что в большинстве случаев классы разделяются достаточно хорошо.

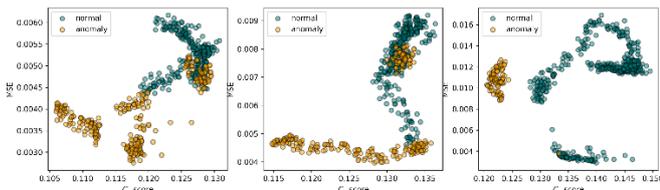


Рис. 10 Расположение образцов на плоскости

На Рис.11 изображены последовательные значения вышеперечисленных метрик для одного из поколений растения. Поколением в данном случае является набор изображений одного растения с начала и до завершения его выращивания в вертикальной ферме.

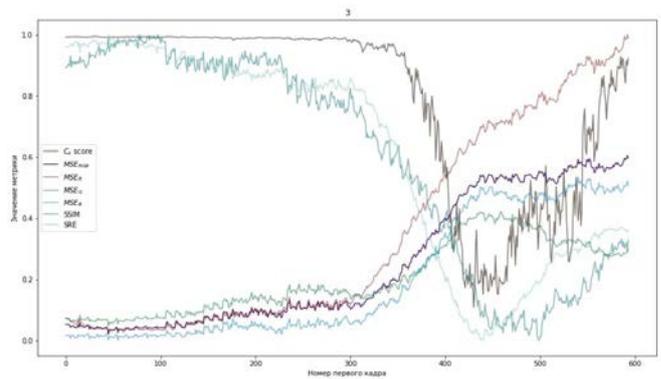


Рис. 11 Обнаружение аномалий на тестовом датасете

Наблюдается, что, начиная с ~400 кадра, с появлением аномальных кадров во входной последовательности, выбранные метрики ведут себя соответствующе. Выход C_x падает, считая x_i менее нормальным, MSE увеличивается, а индекс структурного сходства уменьшается вместе с отношением сигнала к ошибке восстановления. Происходит это потому, что из-за аномалии восстановление получается неточным. Эта неточность выражается в попытке нейросети восстановить засохшие листья, как зеленые или просто отсутствующие. В данном случае в качестве тестовых аномалий были взяты моменты попадания в кадр засохших листьев (Рис. 12)

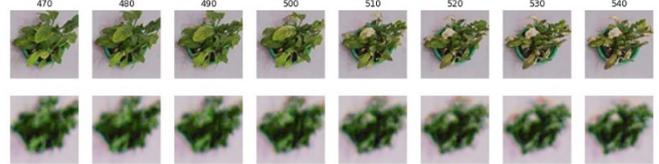


Рис. 12 Исходная и восстановленная последовательности

B. One-class Deep SVDD

На выходе алгоритма имеем расстояние от образца до центра гиперсферы. Принадлежность к тому или иному классу определяется с помощью порога для этого значения. В Таблице 1. приведены значения площадей под ROC-AUC кривой для используемых алгоритмов. Растения в датасетах сгруппированы по видам, то есть каждый датасет соответствует одному из видов растений на Рис. 2

Алгоритм	Номер датасета			
	1	2	3	4
One-Class Deep SVDD	0.74	0.61	0.78	0.94
TadGAN	0.70	0.59	0.70	0.83

Таблица 1. ROC-AUC score

Наблюдается, что в большинстве случаев качество Deep SVDD лучше. Стоит принять во внимание тот факт, что при обучении результат Deep SVDD сильнее зависит от гиперпараметров, чем TadGAN.

VI. ЗАКЛЮЧЕНИЕ

В ходе работы был реализованы методы TadGAN и Deep SVDD на основе сверточных нейронных сетей. Указанный метод был применен для задачи обнаружения аномалий в процессе роста растений. Программная реализация выполнена на языке Python3 с использованием открытой библиотеки машинного обучения PyTorch и

доступна по ссылке <https://github.com/soapun/growvision>

БИБЛИОГРАФИЯ

- [1] S. Festag и C. Spreckelsen, «Semantic Anomaly Detection in Medical Time Series,» т. 278, 2021.
- [2] J. E. Zhang, D. Wu и B. Boulet, «Time Series Anomaly Detection for Smart Grids: A Survey,» в *2021 IEEE Electrical Power and Energy Conference (EPEC)*, 2021.
- [3] G. Colò, «Anomaly detection for cyber security: time series forecasting and deep learning,» *Int. J. Sci. Res. Math. Stat. Sci.*, т. 7, р. 40–52, 2020.
- [4] M. Abdallah, W. J. Lee, N. Raghunathan, C. Mousoulis, J. W. Sutherland и S. Bagchi, «Anomaly detection through transfer learning in agriculture and manufacturing IoT systems,» *arXiv preprint arXiv:2102.05814*, 2021.
- [5] W. Li, V. Mahadevan и N. Vasconcelos, «Anomaly detection and localization in crowded scenes,» *IEEE transactions on pattern analysis and machine intelligence*, т. 36, р. 18–32, 2013.
- [6] M. Ravanbakhsh, M. Nabi, E. Sangineto, L. Marcenaro, C. Regazzoni и N. Sebe, «Abnormal event detection in videos using generative adversarial nets,» в *2017 IEEE international conference on image processing (ICIP)*, 2017.
- [7] O. Ronneberger, P. Fischer и T. Brox, «U-net: Convolutional networks for biomedical image segmentation,» в *International Conference on Medical image computing and computer-assisted intervention*, 2015.
- [8] A. Geiger, D. Liu, S. Alnegheimish, A. Cuesta-Infante и K. Veeramachaneni, «TadGAN: Time series anomaly detection using generative adversarial networks,» в *2020 IEEE International Conference on Big Data (Big Data)*, 2020.
- [9] L. Ruff, R. Vandermeulen, N. Goernitz, L. Deecke, S. A. Siddiqui, A. Binder, E. Müller и M. Kloft, «Deep One-Class Classification,» в *Proceedings of the 35th International Conference on Machine Learning*, 2018.
- [10] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville и Y. Bengio, «Generative adversarial nets,» *Advances in neural information processing systems*, т. 27, 2014.
- [11] S. Takase и N. Okazaki, «Positional Encoding to Control Output Sequence Length,» *CoRR*, т. abs/1904.07418, 2019.
- [12] U. Sara, M. Akter и M. S. Uddin, «Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study,» *Journal of Computer and Communications*, т. 7, р. 8–18, 2019.
- [13] F. T. Liu, K. Ting и Z.-H. Zhou, «Isolation Forest,» 2009.

Кондратьев Алексей Алексеевич
МГУ имени М.В. Ломоносова, факультет ВМК
soapun@ya.ru

Березин Сергей Борисович
МГУ имени М.В. Ломоносова, факультет ВМК
sergey@itis.cs.msu.ru

The Application of Neural Networks for the Anomaly Detection in Plant Development

Aleksei Kondratev, Sergei Berezin

Abstract—Machine learning and neural networks have become a common tool in tasks related to image processing and analysis. One of these tasks is to detect anomalies in images or abnormal images. Computer vision is increasingly helping to digitalize production and reveal patterns in agriculture. Technology allows a human to speed up the analysis, but it is not possible to completely replace an experienced specialist. In this paper, machine learning methods are used to detect abnormal sequences of plant images. The results obtained can be used in agricultural automation tasks, where one of the main means of monitoring the condition of plants is photo or video. The developed software package will be able to automatically detect problems with the health of plants or surrounding infrastructure, which will help to take appropriate measures in time.

In this article, two neural network approaches to detecting anomalies in the process of plant development using images are implemented. Experiments with real data were done as well as a comparative analysis of the efficiency of methods in a specific task.

Keywords—Anomaly detection, Computer Vision, GAN, Machine learning

BIBLIOGRAPHY

- [1] S. Festag и C. Spreckelsen, «Semantic Anomaly Detection in Medical Time Series,» т. 278, 2021.
- [2] J. E. Zhang, D. Wu и B. Boulet, «Time Series Anomaly Detection for Smart Grids: A Survey,» в *2021 IEEE Electrical Power and Energy Conference (EPEC)*, 2021.
- [3] G. Colò, «Anomaly detection for cyber security: time series forecasting and deep learning,» *Int. J. Sci. Res. Math. Stat. Sci.*, т. 7, p. 40–52, 2020.
- [4] M. Abdallah, W. J. Lee, N. Raghunathan, C. Mousoulis, J. W. Sutherland и S. Bagchi, «Anomaly detection through transfer learning in agriculture and manufacturing IoT systems,» *arXiv preprint arXiv:2102.05814*, 2021.
- [5] W. Li, V. Mahadevan и N. Vasconcelos, «Anomaly detection and localization in crowded scenes,» *IEEE transactions on pattern analysis and machine intelligence*, т. 36, p. 18–32, 2013.
- [6] M. Ravanbakhsh, M. Nabi, E. Sangineto, L. Marcenaro, C. Regazzoni и N. Sebe, «Abnormal event detection in videos using generative adversarial nets,» в *2017 IEEE international conference on image processing (ICIP)*, 2017.
- [7] O. Ronneberger, P. Fischer и T. Brox, «U-net: Convolutional networks for biomedical image segmentation,» в *International Conference on Medical image computing and computer-assisted intervention*, 2015.
- [8] A. Geiger, D. Liu, S. Alnegheimish, A. Cuesta-Infante и K. Veeramachaneni, «TadGAN: Time series anomaly detection using generative adversarial networks,» в *2020 IEEE International Conference on Big Data (Big Data)*, 2020.
- [9] L. Ruff, R. Vandermeulen, N. Goernitz, L. Deecke, S. A. Siddiqui, A. Binder, E. Müller и M. Kloft, «Deep One-Class Classification,» в *Proceedings of the 35th International Conference on Machine Learning*, 2018.
- [10] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville и Y. Bengio, «Generative adversarial nets,» *Advances in neural information processing systems*, т. 27, 2014.
- [11] S. Takase и N. Okazaki, «Positional Encoding to Control Output Sequence Length,» *CoRR*, т. abs/1904.07418, 2019.
- [12] U. Sara, M. Akter и M. S. Uddin, «Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study,» *Journal of Computer and Communications*, т. 7, p. 8–18, 2019.
- [13] F. T. Liu, K. Ting и Z.-H. Zhou, «Isolation Forest,» 2009.