

DOI 10.66032/2221-2574-2026-1-2-46-52

УДК 004.932.72

СЕГМЕНТАЦИЯ СНИМКОВ НАНОМАТЕРИАЛОВ С ПОМОЩЬЮ ГИБРИДНОЙ АРХИТЕКТУРЫ СИАМСКОЙ СЕТИ И U-NET

Савельев Михаил Юрьевич

студент, Муромский институт (филиал) ФГБОУ ВО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых».

Шамшин Максим Николаевич

соискатель кафедры программной инженерии, Муромский институт (филиал) ФГБОУ ВО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых».

Захаров Алексей Александрович

кандидат технических наук, доцент, Муромский институт (филиал) ФГБОУ ВО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых».

Адрес: 602264, Российская Федерация, Владимирская обл., г. Муром, ул. Орловская, д. 23.

E-mail для связи: s.michail2004@mail.ru

Аннотация: В работе представлен метод сегментации изображений наноматериалов на основе нейронных сетей с использованием малого числа обучающих примеров. Выполнен сравнительный анализ современных методов обучения с ограниченными данными, включая метаобучение, трансферное обучение, генеративные сети и вариационные автокодировщики. Описаны ключевые особенности и ограничения существующих методов. Предложена гибридная архитектура, объединяющая сиамскую сеть и U-Net, что позволяет повысить точность сегментации даже при неоднородной структуре изображений. Архитектура использует метрическое обучение для сравнения пар изображений (эталон и запрос), обеспечивая устойчивое обучение и обобщение на основе 5–50 размеченных примеров. Используется евклидово расстояние для сравнения векторных представлений. Представлены результаты экспериментов, демонстрирующие высокие значения индекса Жаккара, коэффициента Дайса, точности и полноты.

Ключевые слова: обучение на малом наборе данных, сиамская сеть, U-Net, сегментация, наноматериалы.

Введение

Искусственные нейронные сети доказали свою эффективность при анализе и распознавании изображений наноматериалов и успешно применяются для обработки лабораторных данных. Благодаря своим уникальным физическим свойствам наноматериалы всё чаще используются в таких областях, как биология, энергетика, электроника и здравоохранение [1, 2].

Традиционные методы машинного обучения обычно требуют больших размеченных наборов данных. В то же время получение высококачественных изображений наноматериалов является дорогостоящим и технически сложным процессом [3]. Это приводит к необходимости разработки систем, способных эффективно обучаться на основе ограниченных наборов данных. Такое решение значительно сократит трудозатраты, необходимые для анализа наноматериалов.

Обзор методов

Подходы к обучению с использованием небольшого количества примеров (few-shot learning) позволяют избавиться от необходимости применения больших объёмов данных [4]. В отличие от классических алгоритмов, которым требуются сотни и тысячи размеченных образцов, эти подходы обеспечивают высокую точность обучения на 5–50 примерах [5].

Метаобучение

Многие методы few-shot learning основаны на метаобучении, при котором модели обучаются «умению учиться». Например, цель алгоритма MAML (Model-Agnostic Meta-Learning) — найти такие начальные параметры, которые легко адаптируются к новой задаче. Это позволяет модели обобщать полученные знания на новые классы объектов [6]. Схема алгоритма MAML представлена на рис. 1.

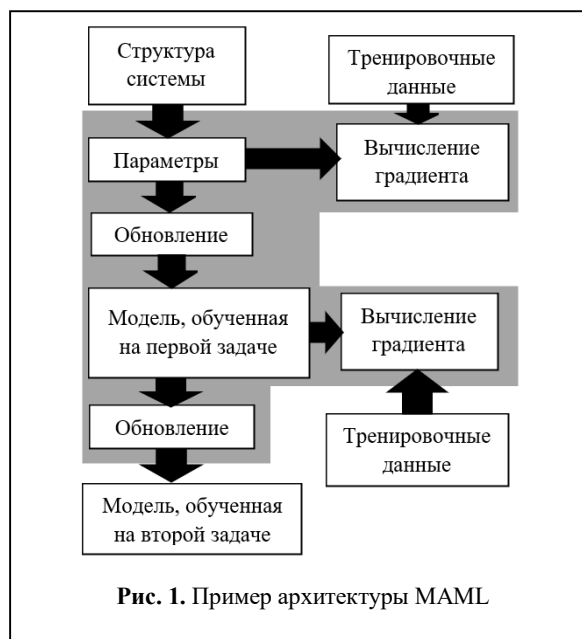
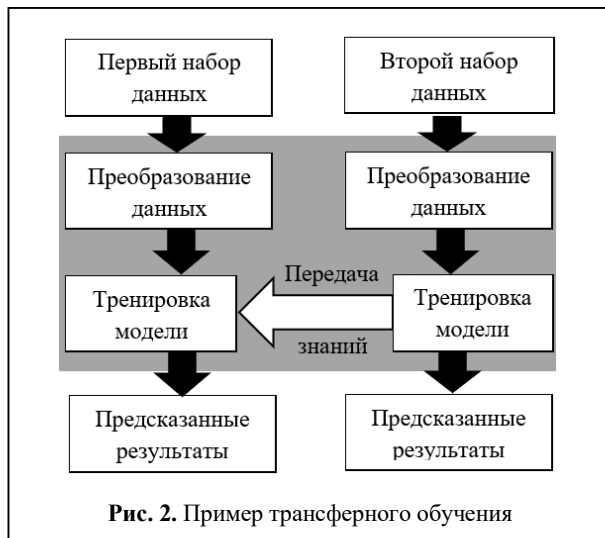


Рис. 1. Пример архитектуры MAML



Несмотря на свою эффективность, данный метод отличается вычислительной сложностью и рисками, связанными с переобучением.

Трансферное обучение

Трансферное обучение (Transfer Learning) — это стратегия машинного обучения, которая использует знания из одной задачи для выполнения другой задачи. Это достигается за счёт использования моделей, которые были предварительно предобучены на больших наборах данных [7]. Схема трансферного обучения представлена на рис. 2.

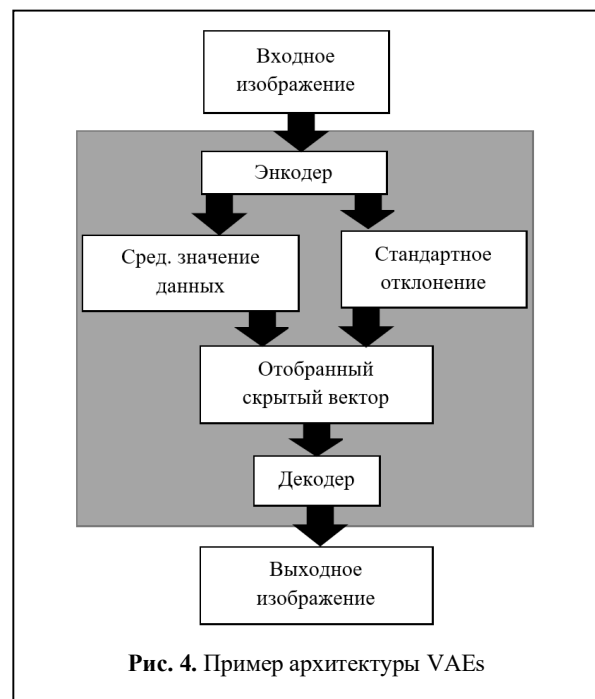
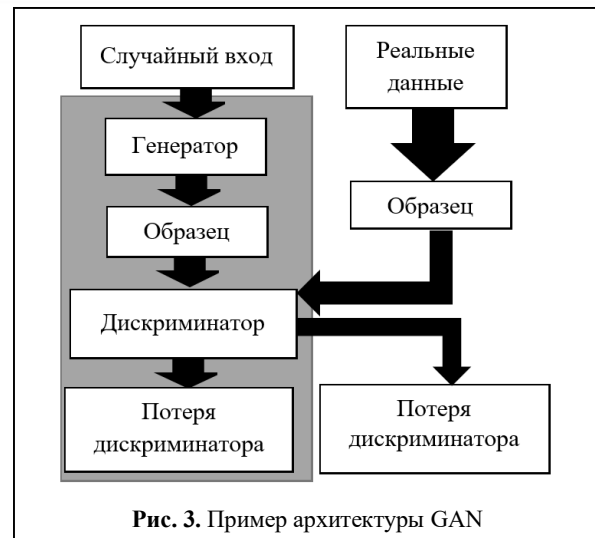
Такие модели не обучаются с нуля. Вместо этого они дообучаются (fine-tuning) на целевом наборе данных, чтобы адаптировать свои знания. Однако существенной проблемой, связанной с этим подходом, является потенциальная возможность «негативного переноса», когда предварительно усвоенные знания мешают, а не помогают решению новой задачи.

Генеративные методы

Одним из возможных решений проблемы ограниченности обучающих данных является синтез и генерация дополнительных примеров. Этот подход использует передовые методы глубокого обучения для искусственного расширения набора данных. Примерами таких методов являются генеративно-состязательные сети (GAN) [8] и вариационные автоэнкодеры (VAEs) [9]. Основным ограничением подобных подходов является присущая генеративным моделям нестабильность при обучении, что может привести к созданию нереалистичных или некорректных выборок. Схемы генеративных подходов приведены на рис. 3 и рис. 4.

Метрическое обучение

Метрическое обучение (metric learning) использует функции расстояния в пространстве признаков для оценки сходства объектов [10]. Сиамская сеть — это архитектура с двумя входными данными (эталон и запрос) и одним выходом. Все признаки из-



влекаются с помощью свёрточной нейронной сети (CNN), а сравнение происходит по расстоянию между векторными представлениями. Архитектуры CNN и сиамской сети приведены на рис. 5.

В данной работе исследуется метрический подход к обучению с ограниченными данными, в котором свёрточная сеть U-Net интегрирована в архитектуру сиамской сети.

Научная новизна исследования заключается в разработке гибридной модели, которая сочетает в себе сильные стороны U-Net, обеспечивающей точную семантическую сегментацию изображений, с возможностями сиамской сети для эффективного обучения метрикам в условиях дефицита данных.

Разработка модели и обучение

В ходе исследования был разработан алгоритм сегментации изображений наноматериалов, сочетаю-

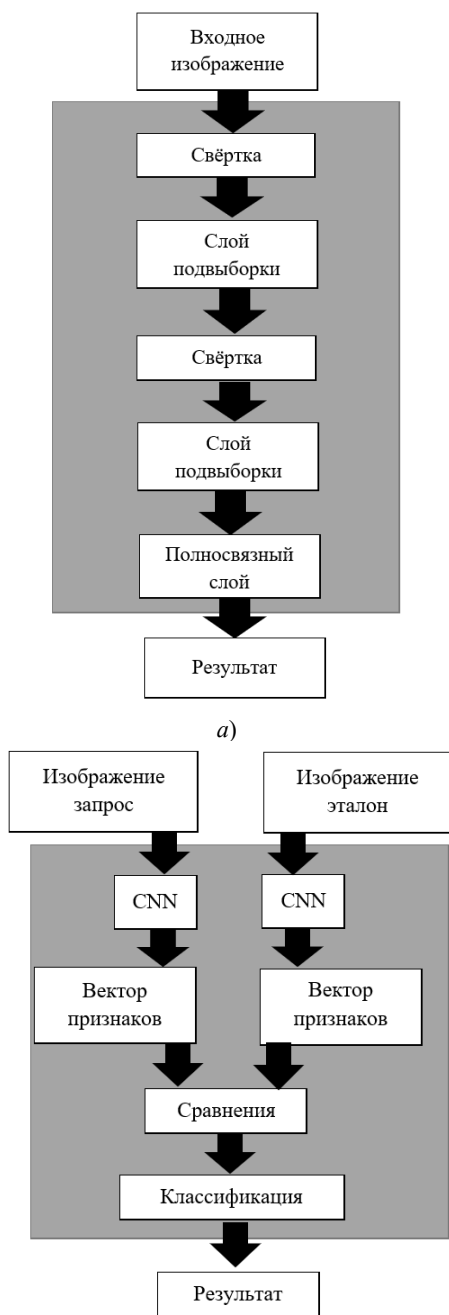


Рис. 5. а) — пример архитектуры CNN (свёрточной нейронной сети); б) — пример архитектуры сиамской сети

щий архитектуру сиамской сети и U-Net. Модель сочетает в себе способность сиамской сети сравнивать пары изображений (эталон и запрос) с высокой точностью восстановления границ объектов, обеспечиваемой U-Net [11].

Архитектура модели

Архитектура U-Net представлена на рис. 6. Основными блоками архитектуры являются:

- кодировщик (Encoder) — это последовательность свёрточных слоев с ядрами 3x3 и операциями под-

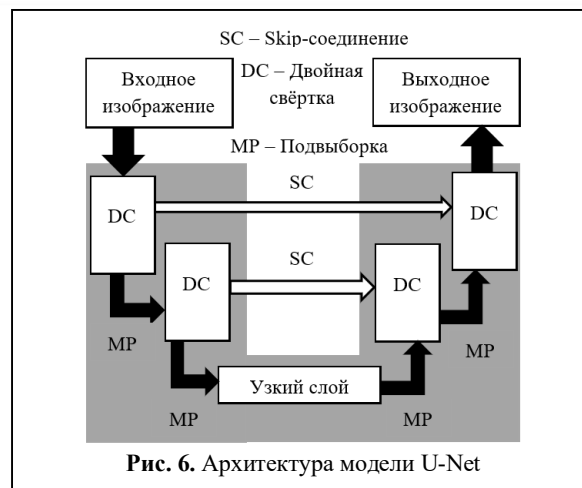


Рис. 6. Архитектура модели U-Net

выборки (MaxPooling), которые уменьшают размерность и извлекают признаки;

- узкий слой (BottleNeck) — это заключительный блок кодировщика, содержащий два свёрточных слоя с увеличенным числом фильтров;

- декодер (Decoder) выполняет обратную последовательность операций транспонированной свёртки и конкатенацию признаков из соответствующих уровней кодировщика для восстановления пространственного разрешения и уточнения границ сегментации;

- Skip-соединения (Skip connections): обеспечивают передачу детальных признаков с кодировщика в декодер через конкатенацию, сохраняя пространственную точность и улучшая восстановление границ объектов.

Полная архитектура гибридной сети показана на рис. 7. Архитектура U-Net была расширена за счёт операций, принимающих векторные представления признаков изображений эталона и запроса. Затем вычисляется евклидово расстояние между вектор-

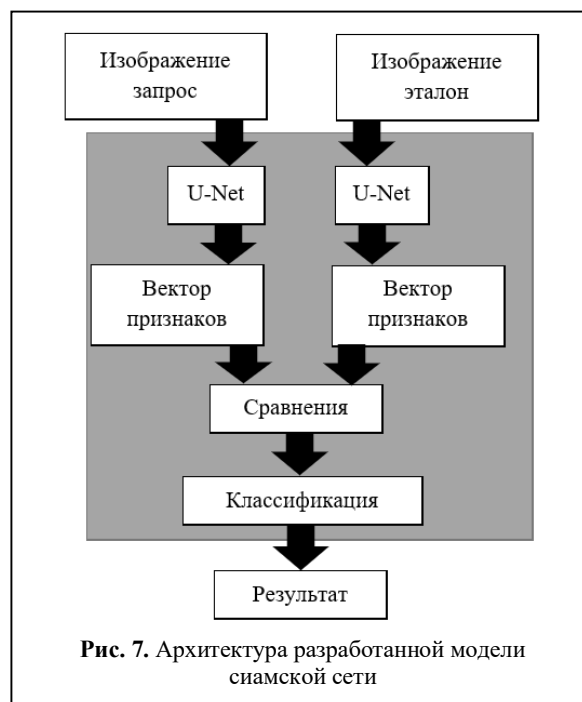


Рис. 7. Архитектура разработанной модели сиамской сети

ными представлениями $d(p_1, p_1')$ (1) [12], где p_1 — признак эталонного изображения; а p_1' — признак изображения запроса.

$$d(p, p') = \sqrt{(p_1 - p_1')^2 + (p_2 - p_2')^2 \dots (p_n - p_n')^2}. \quad (1)$$

Обучение модели

Процесс обучения модели описывается следующим образом.

На каждой итерации обучения формируется пара изображений: эталонное изображение, которое включает в себя эталонную маску сегментации, и изображение запроса, для которого должна быть сгенерирована маска.

Как эталонное изображение, так и изображение запроса обрабатываются общим кодировщиком U-Net, что формирует плотные карты признаков высокого уровня.

Признаковые карты дополнительно обрабатываются в BottleNeck-слое, который служит компактным представлением изображений.

Декодер U-Net, используя признаки изображения запроса, восстанавливает его пространственное разрешение и формирует на выходе маску сегментации.

Для обучения модели используется комбинированная функция потерь (2), объединяющая двоичную перекрёстную энтропию (3) и коэффициент Дайса (4).

Комбинированная функция потерь:

$$Loss = BCE - \ln(DICE), \quad (2)$$

где BCE — двоичная перекрёстная энтропия; $DICE$ — коэффициент Дайса.

Двоичная перекрёстная энтропия:

$$BCE = -\sum_i (y_i \ln(p_i) + (1 - y_i) \ln(1 - p_i)), \quad (3)$$

где y — истинная метка (0 или 1); p — прогнозируемая вероятность принадлежности экземпляра к классу 1.

Коэффициент Дайса:

$$DICE = 2 \frac{\sum_i y_i p_i}{\sum_i y_i + \sum_i p_i}. \quad (4)$$

Компонент бинарной кросс-энтропии обеспечивает точную классификацию на уровне пикселей. Потери по коэффициенту Дайса улучшают пространственное соответствие между предсказанной и эталонной масками, что является преимуществом для устранения дисбаланса классов [13].

Обучение длится до 50 эпох. Сохраняется лучшая модель (по валидационным метрикам) для последующего тестирования.

Результаты исследования

Для оценки качества модели использовались следующие метрики.

IoU (Intersection over Union или индекс Жаккара) (5) — мера совпадения между предсказанной и истинной масками сегментации [14]:

$$IoU = \frac{TP}{TP + FN + FP}, \quad (5)$$

где TP — истинно положительные предсказания; FN — ложно отрицательные предсказания; FP — ложно положительные предсказания.

Коэффициент Дайса (Dice) — показатель сходства между истинной и предсказанной маской, чувствительный к перекрытию объектов (4).

Точность (precision) [15]:

$$precision = \frac{|M_p \cap M_g|}{|M_p|}, \quad (6)$$

где M_p — предсказанная маска; M_g — истинная маска (ground truth).

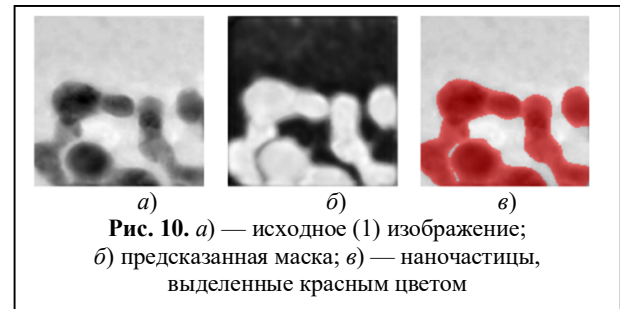
Полнота (recall) [15]:

$$recall = \frac{|M_p \cap M_g|}{|M_g|}. \quad (7)$$

Лучшая версия модели показала следующие средние результаты: $IoU = 0,85$; $Dice = 0,9$; $precision = 0,93$; $recall = 0,94$. Результаты работы модели представлены на рис. 8–15.

Метод сегментации снимков наноматериалов с помощью гибридной архитектуры сиамской сети и U-Net позволяет использовать ограниченные наборы данных для обучения модели.

Высокие значения валидационных метрик под-



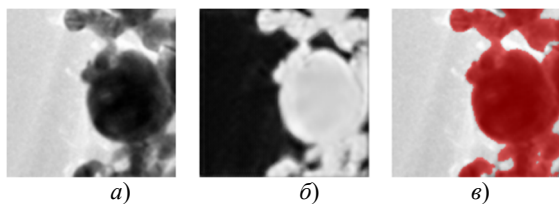


Рис. 11. а) — исходное (1) изображение; б) предсказанная маска; в) — наночастицы, выделенные красным цветом

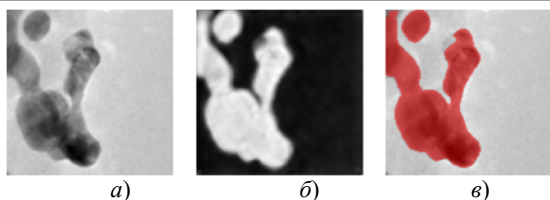


Рис. 12. а) — исходное (1) изображение; б) предсказанная маска; в) — наночастицы, выделенные красным цветом

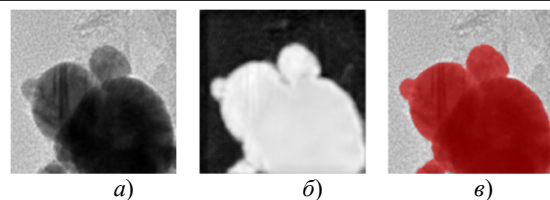


Рис. 13. а) — исходное (1) изображение; б) предсказанная маска; в) — наночастицы, выделенные красным цветом

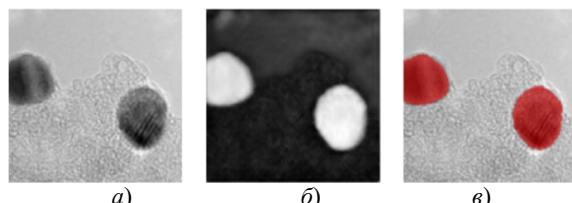


Рис. 14. а) — исходное (1) изображение; б) предсказанная маска; в) — наночастицы, выделенные красным цветом

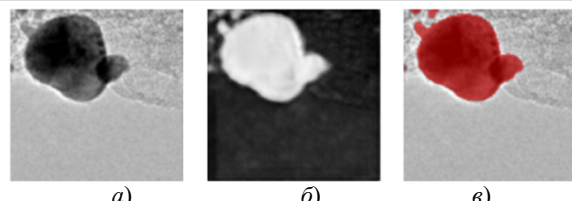


Рис. 15. а) — исходное (1) изображение; б) предсказанная маска; в) — наночастицы, выделенные красным цветом

тверждают точность и надёжность модели. Эти особенности делают метод перспективным для практического применения при анализе наноматериалов.

Заключение

Разработанная гибридная архитектура успешно решает задачу сегментации изображений наноматериалов при дефиците размеченных данных, сочетая точность сегментации на основе U-Net со способностью сиамской сети делать обобщения на основе нескольких примеров.

Перспективным направлением развития является адаптация модели для морфологического анализа, включая автоматическое измерение размера частиц, определение формы и анализ пространственного распределения. Такое расширение функциональных возможностей позволит создать комплексное решение для определения характеристик наноматериалов.

Работа выполнена в рамках государственного задания в сфере научной деятельности Министерства науки и высшего образования Российской Федерации (тема FZUN-2024-0019, госзадание ВлГУ).

Литература

1. Zhang R. Application and development prospects of nanomaterials in multiple fields // Proceedings of the 2023 International Conference on Functional Materials and Civil Engineering, 2023. Pp. 96–101.
2. Пугин Е.В., Жизняков А.Л. Алгоритмы обработки изображений для обнаружения объектов с использованием нечётких признаков // Телевизионные системы, передача и обработка изображений. 2020. № 2. С. 59–65.
3. Yao L., Chen Q. Machine learning in nanomaterial electron microscopy data analysis // Intelligent Nanotechnology: Merging Nanoscience with Artificial Intelligence. Amsterdam: Elsevier, 2023. Pp. 279–305. DOI: 10.1016/B978-0-323-85796-3.00010-X.
4. Fedele A., Guidotti R., Pedreschi D. Explaining Siamese networks in few-shot learning // Machine Learning. 2024. Pp. 7723–7760. DOI: 10.1007/s10994-024-06529-8.
5. Han G., Huang S., Ma J., He Y. Meta Faster R-CNN: Towards Accurate Few-Shot Object Detection with Attentive Feature Alignment // Proceedings of AAAI Conference on Artificial Intelligence. 2022. Vol. 36. Pp. 780–789. DOI: 10.1609/aaai.v36i1.19959.
6. Pang S., Zhao W.S., Wang S., Zhang L. PermuteMAML: exploring industrial surface defect detection algorithms for few-shot learning // Complex & Intelligent Systems. 2023. Vol. 10. Pp. 1–10. DOI: 10.1007/s40747-023-01219-9.
7. Qayyumi S.W., Park L.A.F., Obst O. Few-Shot and Transfer Learning with Manifold Distributed Datasets // Data Science and Machine Learning: Communications in Computer and Information Science. 2023. Pp. 137–149. DOI: 10.1007/978-981-99-8696-5_10.
8. Saxena D., Cao J. Generative Adversarial Networks (GANs): Challenges, Solutions, and Future Directions // ACM Computing Surveys. 2021. Vol. 54. Issue 3. No. 63. Pp. 1–42. DOI: 10.1145/3446374.
9. Abbasnejad M.E., Dick A., Hengel A. Infinite Variational Autoencoder for Semi-Supervised Learning // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. Pp. 5888–5897. DOI: 10.1109/CVPR.2017.90.

10. Chicco D. Siamese neural networks: an overview // *Methods in Molecular Biology*. 2021. Vol. 2190. Pp. 73–94. DOI: 10.1007/978-1-0716-0826-5_3.
11. Siddique N., Paheding S., Elkin C.P., Devabhaktuni V. U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications // *IEEE Access*. 2021. Vol. 9. Pp. 82031–82057. DOI: 10.1109/ACCESS.2021.3086020.
12. Wang J., Tan Y. Efficient Euclidean distance transform using perpendicular bisector segmentation // *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2011. Pp. 1625–1632. DOI: 10.1109/CVPR.2011.5995644.
13. Zhang Y., Guindon B. Application of the Dice coefficient to accuracy assessment of object-based image classification // *Canadian Journal of Remote Sensing*. 2016. Vol. 43. Pp. 48–61. DOI: 10.1080/07038992.2017.1259557.
14. Yan J., Wang H., Yan M., Wenhui D. IoU-adaptive deformable R-CNN: make full use of IoU for multi-class object detection in remote sensing imagery // *Remote Sensing*. 2019. Vol. 11. Art. no. 286. DOI: 10.3390/rs11030286.
15. Borji A., Cheng M.-M., Jiang H., Li J. Salient object detection: a benchmark // *IEEE Transactions on Image Processing*. 2015. Vol. 24. No. 12. Pp. 5706–5722. DOI: 10.1109/TIP.2015.2487833.

Поступила 21 января 2026 г.

English

SEGMENTATION OF NANOMATERIAL IMAGES USING THE HYBRID ARCHITECTURE OF SIAMESE NETWORK AND U-NET

Mikhail Yurievich Savelyev — Student, Murom Institute (branch) of the “Vladimir State University named after Alexander and Nikolay Stoletovs”.

Maxim Nikolaevich Shamshin — Applicant for a Degree, Department of Software Engineering, Murom Institute (branch) of the “Vladimir State University named after Alexander and Nikolay Stoletovs”.

Alexey Alexandrovich Zakharov — PhD in Engineering, Associate Professor, Department of Radio Engineering, Murom Institute (branch) of the “Vladimir State University named after Alexander and Nikolay Stoletovs”.

Address: 602264, Russian Federation, Vladimir region, Murom, Orlovskaya str., 23.

E-mail: s.michail2004@mail.ru

Abstract: The method of image segmentation of nanomaterials based on neural networks using a small number of training examples is presented in the paper. A comparative analysis of modern learning methods with limited data, including meta-learning, transfer learning, generative networks, and variational autoencoders has been performed. The key features and limitations of the existing methods are described. A hybrid architecture combining Siamese network and U-Net has been proposed, which makes it possible to increase segmentation accuracy even with a heterogeneous image structure. The architecture uses metric learning to compare pairs of images (reference and query), providing sustained learning and generalization based on 5–50 labeled examples. The Euclidean distance is used to compare vector representations. The experimental results demonstrating high values of the Jacquard index, Dice coefficient, precision and recall are presented.

Keywords: small data set learning, Siamese network, U-Net, segmentation, nanomaterials.

The research was carried out within the state assignment in the field of scientific activity of the Ministry of Science and Higher Education of the Russian Federation (theme FZUN-2024-0019, state assignment of the VLSU).

References

1. Zhang R. Application and development prospects of nanomaterials in multiple fields. *Proceedings of the 2023 International Conference on Functional Materials and Civil Engineering*, 2023. Pp. 96–101.
2. Pugin E.V., Zhiznyakov A.L. Image processing algorithms for detecting objects using fuzzy features. *Television systems, image transmission and processing*. 2020. No. 2. Pp. 59–65.
3. Yao L., Chen Q. Machine learning in nanomaterial electron microscopy data analysis. *Intelligent Nanotechnology: Merging Nanoscience with Artificial Intelligence*. Amsterdam: Elsevier, 2023. Pp. 279–305. DOI: 10.1016/B978-0-323-85796-3.00010-X.
4. Fedele A., Guidotti R., Pedreschi D. Explaining Siamese networks in few-shot learning. *Machine Learning*. 2024. Pp. 7723–7760. DOI: 10.1007/s10994-024-06529-8.
5. Han G., Huang S., Ma J., He Y. Meta Faster R-CNN: Towards Accurate Few-Shot Object Detection with Attentive Feature Alignment. *Proceedings of AAAI Conference on Artificial Intelligence*. 2022. Vol. 36. Pp. 780–789. DOI: 10.1609/aaai.v36i1.19959.
6. Pang S., Zhao W.S., Wang S., Zhang L. Permute-MAML: exploring industrial surface defect detection algorithms for few-shot learning. *Complex & Intelligent Systems*. 2023. Vol. 10. Pp. 1–10. DOI: 10.1007/s40747-023-01219-9.
7. Qayyumi S.W., Park L.A.F., Obst O. Few-Shot and Transfer Learning with Manifold Distributed Datasets. *Data Science and Machine Learning: Communications in Computer and Information Science*. 2023. Pp. 137–149. DOI: 10.1007/978-981-99-8696-5_10.
8. Saxena D., Cao J. Generative Adversarial Networks (GANs): Challenges, Solutions, and Future Directions. *ACM Computing Surveys*. 2021. Vol. 54. Issue 3. No. 63. Pp. 1–42. DOI: 10.1145/3446374.
9. Abbasnejad M.E., Dick A., Hengel A. Infinite Variational Autoencoder for Semi-Supervised Learning. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017. Pp. 5888–5897. DOI: 10.1109/CVPR.2017.90.

10. *Chicco D.* Siamese neural networks: an overview. *Methods in Molecular Biology*. 2021. Vol. 2190. Pp. 73–94. DOI: 10.1007/978-1-0716-0826-5_3.
11. *Siddique N., Paheding S., Elkin C.P., Devabhaktuni V.* U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications. *IEEE Access*. 2021. Vol. 9. Pp. 82031–82057. DOI: 10.1109/ACCESS.2021.3086020.
12. *Wang J., Tan Y.* Efficient Euclidean distance transform using perpendicular bisector segmentation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2011. Pp. 1625–1632. DOI: 10.1109/CVPR.2011.5995644.
13. *Zhang Y., Guindon B.* Application of the Dice coefficient to accuracy assessment of object-based image classification. *Canadian Journal of Remote Sensing*. 2016. Vol. 43. Pp. 48–61. DOI: 10.1080/07038992.2017.1259557.
14. *Yan J., Wang H., Yan M., Wenhui D.* IoU-adaptive deformable R-CNN: make full use of IoU for multi-class object detection in remote sensing imagery. *Remote Sensing*. 2019. Vol. 11. Art. No. 286. DOI: 10.3390/rs11030286.
15. *Borji A., Cheng M.-M., Jiang H., Li J.* Salient object detection: a benchmark. *IEEE Transactions on Image Processing*. 2015. Vol. 24. No. 12. Pp. 5706–5722. DOI: 10.1109/TIP.2015.2487833.