ПРИМЕНЕНИЕ НЕЙРОННОЙ СЕТИ АРХИТЕКТУРЫ U-Net ДЛЯ СЕГМЕНТАЦИИ СТМ-ИЗОБРАЖЕНИЙ

ШЕЛКОВНИКОВ Е. Ю., ШЛЯХТИН К. А., ШЕЛКОВНИКОВА Т. Е., ЕГОРОВ С. Ф.

Удмуртский федеральный исследовательский центр Уральского отделения РАН,
426067, г. Ижевск, ул. Т. Барамзиной, 34

АННОТАЦИЯ. В статье проведено исследование эффективности применения нейронной сети U-Net для сегментации СТМ-изображений. Предложена методика моделирования СТМ-изображения с наночастицами на подложке путем случайной расстановки эллипсоидов по поверхности, сгенерированной методом Diamond-Square. Разработан алгоритм генерации маски для модельного изображения с выделением границ пересекающихся эллипсоидов, позволяющий выполнять обучение нейросети для разделения соприкасающихся наночастиц. Проведено сравнение результатов сегментации модельных СТМ-изображений полученной сети с различными методами сегментации.

КЛЮЧЕВЫЕ СЛОВА: сканирующий туннельный микроскоп, сверточная нейронная сеть, архитектура U-Net, моделирование СТМ-изображения, сегментация.

ВВЕДЕНИЕ

За последние два десятилетия сканирующая зондовая микроскопия стала мощным и гибким инструментом для визуализации поверхностей с атомным разрешением [1 – 3]. Сканирующий туннельный микроскоп (СТМ) является важным звеном в этой области и широко применяется для изучения структуры поверхности, получения информации о наночастицах исследуемого образца, определяющих химические и физические свойства исследуемого материала [4]. Поэтому актуальной является задача сегментации СТМ-изображения для отделения наночастиц друг от друга на изображении и нахождения их количества. С целью повышения эффективности сегментации СТМ-изображения рассматривается возможность использования сверточной нейронной сети с архитектурой U-Net. U-Net – это архитектура сверточной нейронной сети, которая была разработана в 2015 году Олафом Роннебергером, Филиппом Фишером, Томасом Броксом для сегментации клеток на микроскопических снимках [5]. Она хорошо зарекомендовала себя и широко используется при решении задач сегментации изображений [5 – 6]. Данная архитектура содержит две части: сужающуюся (энкодер) и расширяющуюся (декодер), которые образуют U-образную структуру [7]. В сужающейся (входной) части структуры изображение, поступающее на вход, проходит серию слоёв: слои свертки с функцией активации и слои субдискретизации (пулинг). В качестве инструментов компьютерного зрения сети применяются открытые библиотеки OpenCV и Keras совместно с TensorFlow на базе языка программирования Python.

СЕГМЕНТАЦИЯ СТМ-ИЗОБРАЖЕНИЙ С ПРИМЕНЕНИЕМ НЕЙРОННОЙ СЕТИ АРХИТЕКТУРЫ U-Net

Схема сверточной нейросети архитектуры U-Net представлена на рис. 1. В обычной нейронной сети каждый нейрон связан со всеми нейронами предыдущего слоя, и каждая связь имеет свой весовой коэффициент. В сверточной нейронной сети в операциях свертки используется матрица весов небольшого размера, которую «двигают» по всему обрабатываемому слою (на входе сети – непосредственно по входному изображению). Слой свертки производит суммирование результатов поэлементного произведения каждого фрагмента изображения на матрицу - ядро свертки. Весовые коэффициенты ядра свертки неизвестны и устанавливаются в процессе обучения [8].

Рис. 1. Схема сверточной нейросети архитектуры U-Net для входного изображения с разрешением 256×256

Оператор свертки вычисляется по формуле:

$$(I * K)_{xy} = \sum_{i=1}^h \sum_{j=1}^w K_{i,j} \times I_{x+i-1,y+j-1}, \quad (1)$$

где I – двумерное изображение размером $x \times y$; K – матрица размерности $h \times w$ (так называемое ядро свертки).

Операция активации представляет собой нелинейную функцию и определяет выходной сигнал нейрона, при этом наиболее часто используемой в современных нейросетях и применяемой в данной архитектуре является функция с названием «выпрямитель» (по аналогии с однополупериодным выпрямителем в электротехнике). Нейроны с данной функцией активации называются ReLU (rectified linear unit) и имеют следующую формулу:

$$f(x) = \max(0, x). \quad (2)$$

Данная функция активации существенно ускоряет процесс обучения за счет простоты вычислений [9]. Слой субдискретизации осуществляет уменьшение размера изображения в n раз по каждой оси путем объединения соседних пикселей в областях размером $n \times n$ в один пиксель. В качестве значения полученного пикселя может использоваться среднее значение объединяемых пикселей или их максимум. В U-Net при выполнении пулинга применяется максимизация по области 2×2 пикселя. После прохождения нескольких серий сверточных и субдискретизирующих слоев входное изображение преобразуется от конкретной сетки пикселей с высоким разрешением к более абстрактным картам признаков.

Далее в декодирующей части сети используются серии слоев, аналогичных со слоями кодирующей части, только вместо субдискретизирующего слоя, уменьшающего размер изображения, применяется слой обратной свертки, который увеличивает размер изображения в n раз по обеим осям. Кроме того, в декодирующей части выполняется конкатенация карт признаков из кодирующей части. На последнем уровне свертка 1×1 используется для формирования выходного сегментированного изображения [5].

Для обучения нейросети необходимо подготовить большое количество изображений (порядка нескольких сотен). Подготовка базы снимков с СТМ является трудоемким

процессом и требует больших временных затрат, поэтому реализована модель СТМ-изображения, отражающая «холмистость» поверхности подложки и разброс наночастиц по ней. Для этого процесс обработки снимков должен осуществлять следующие операции.

Во-первых, необходимо произвести моделирование данной поверхности. Реализация модели в виде полиномиальной функции не отражает случайный характер поверхности на реальном изображении. К тому же такой фон может быть устранен современными методами обработки изображений. Поэтому для получения модели, приближенной к реальным условиям, был использован способ генерации фрактальной поверхности Diamond-Square (рис. 2, а) [10], при котором на каждой итерации работы алгоритма добавляется шум, не допускающий самоподобия. Карта высот, полученная в результате работы данного алгоритма, создает «скалистую местность», и такой «шум» достаточно сложно отфильтровать классическими методами компьютерной фильтрации.

Рис. 2. Моделирование поверхности с наночастицами:

а) – трехмерное изображение подложки модельного СТМ-изображения;
б) – двумерное модельное СТМ-изображение, используемое для обучения сети U-Net

Во-вторых, необходимо осуществить моделирование наночастиц. На полученной поверхности подложки распределяются по равномерному закону эллипсоиды, отличающиеся друг от друга размерами и высотой расположения, которые накладываются друг на друга и, тем самым, усложняют процесс сегментации. Сформированное модельное СТМ-изображение показано на рис. 2, б.

Поскольку сеть должна выделять границы и разделять пересекающиеся наночастицы, разработан алгоритм, который генерирует из модельного СТМ-изображения маску с выделенными границами частиц. Полученная маска используется для обучения U-Net.

В алгоритме (рис. 3) вычисляется приближённое значение градиента яркости изображения с помощью оператора Собеля с ядрами G_x и G_y :

$$G_x = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}; G_y = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix}, \quad (3)$$

где G_x и G_y – две матрицы, свертка с которыми дает приближенные производные по осям x и y . По значениям G_x и G_y вычисляется модуль градиента:

$$|G| = \sqrt{(G_x)^2 + (G_y)^2} \quad (4)$$

и направление градиента:

$$\theta = \arctg\left(\frac{G_y}{G_x}\right). \quad (5)$$

На рис. 4 приведены модельное СТМ-изображение и результаты работы алгоритма (где рис. 4, б иллюстрирует тот факт, что значения в полученном массиве показывают количество частиц, пересекающихся в данной точке). Это необходимо для создания бинарной маски, которая применяется для обучения сети. В маске определен фон и границы между частицами как «0», а области с частицами имеют значение «1».

Рис. 3. Блок-схема алгоритма генерации маски изображения

Рис. 4. Моделирование поверхности с наночастицами: а) – модельное СТМ-изображение; б) – результат работы алгоритма, значения в полученном массиве маски М

Таким образом, на вход сети подается маска с двумя классами для сегментации. Сеть предсказывает принадлежность пикселя к классу «частица» или «фон». Структурная схема полученной нейросети для сегментации СТМ-изображения приведена на рис. 5. Генератор модельных изображений использует данный алгоритм и генерирует пару: модельное СТМ-изображение и его маска (которые подаются на вход модуля обучения нейросети в качестве тренировочного набора и на вход модуля тестирования в качестве тестовой выборки). Тренировочный набор — это последовательность данных, которыми оперирует нейронная сеть. Далее сеть итерационно осуществляет проходы с тренировочными наборами. Одним из главных показателей при инициализации нейронной сети является не количество наборов, а так называемая эпоха. Эпоха увеличивается каждый раз, когда сеть проходит весь набор тренировочных наборов. Чем больше эпоха, тем лучше натренирована сеть и, соответственно, ее результат. Далее в модуле тестирования после окончания эпохи вычисляется ошибка уже для тестовых пар изображений, отражающая расхождение между ожидаемым и полученным ответами. Ошибка — это процентная величина, она формируется каждую эпоху и должна спадать. Как только на выходе сети вычисляемая ошибка достигнет необходимого значения (в данной сети значение 0,03), сеть прекращает обучение. Модули обучения, тестирования и сама нейросеть выполнены на языке Python с помощью библиотек Keras и TensorFlow.

Рис. 5. Структурная схема нейросети для сегментации СТМ-изображения

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Сеть обучалась методом стохастического градиентного спуска на основе входных изображений и соответствующих им масок. На вход сети подавалось 512 тренировочных изображений 256×256 пикселей. В общей сложности сеть обучалась около 2-х часов на видеокарте GeForce GTX 1070. Для оценки качества сегментации использовалась метрика IoU (Intersection over Union), которая вычислялась по формуле:

$$IoU = \frac{TP}{(TF + FP + FN)}, \quad (6)$$

где TP, TF, FP, FN – количество пикселей, правильно отнесенных к классу «частица»; правильно отнесенных к классу «фон»; неправильно отнесенных к классу «частица» и неправильно отнесенных к классу «фон», соответственно. На рис. 5 приведены результаты работы нейронной сети.

Рис. 5. Результат сегментации модельных СТМ-изображений: а) – модельное СТМ-изображение; б) – маска исходного изображения; в) – сегментированное изображение, полученное на выходе сети U-Net

Для проверки качества работы нейронной сети проведено сравнение результатов «сетевой» сегментации модельных СТМ-изображений с результатами, полученными различными методами:

- методом Оцу (с применением алгоритма, который позволяет вычислить порог бинаризации для изображения и разделить пиксели двух классов на «полезные» и «фоновые» [11]);
- детектором «Сфера» (методом детектирования поверхности, позволяющим определять области положительной кривизны на изображении, а также проводить оценку радиуса кривизны [12]);
- методом Саувола (который является одним из методов адаптивной бинаризации, анализирующим обрабатываемое изображение и вычисляющим порог бинаризации для каждого пикселя [13]).

В таблице представлены результаты сравнения сегментации на основе этих методов и сети U-Net 100 модельных изображений с количеством частиц, равным 50.

Таблица

Результаты экспериментального исследования методов сегментирования модельных СТМ-изображений

Методы	IoU		Количество частиц	
	МО	СКО	МО	СКО
Метод Оцу	0,6862	0,0485	39,8025	8,0464
Детектор Сфера	0,6543	0,0512	45,2971	5,6891
Метод Саувола	0,5586	0,1991	55,5800	9,4933
Сеть U-Net	0,9046	0,01125	50,1100	5,3674

На рис. 6 приведено СТМ-изображение, сегментированное разными методами. Анализ экспериментальных данных показал, что лучшие результаты сегментации получены с помощью нейросети U-Net.

Рис. 6. Результат сегментации реального СТМ-изображения:

а) – исходное СТМ-изображение; б) – сегментация СТМ-изображения методом Оцу;
в) – методом Саувола; г) – нейросетью U-Net; д) – с применением детектора «Сфера»

В заключение можно отметить следующее. В данной работе показана возможность применения нейронной сети архитектуры U-Net для сегментации наночастиц на СТМ-изображении. Проведенное исследование выявило высокую эффективность её применения в задаче отделения частиц от фона на изображении, при этом разработанная нейронная сеть дает возможность разделения соприкасающихся наночастиц. В работе предложена методика моделирования СТМ-изображения наночастиц на подложке путем случайной расстановки эллипсоидов на поверхности, сгенерированной алгоритмом Diamond-Square. Разработан алгоритм генерации маски для модельного изображения с выделением границ пересекающихся эллипсоидов, позволивший выполнять обучение нейросети для разделения соприкасающихся наночастиц.

СПИСОК ЛИТЕРАТУРЫ

1. Fukuma T., Onishi K., Kobayashi N., Matsuki A. and Asakawa H. Atomic-resolution imaging in liquid by frequency modulation atomic force microscopy using small cantilevers with megahertz-order resonance frequencies // *Nanotechnology*, 2012, vol. 23, no. 13, pp. 135706.
2. Gerber C. and Lang H. P. How the doors to the nanoworld were opened // *Nature Nanotechnology*, 2006, vol. 1, no. 1, pp. 3-5.
3. Шелковников Е. Ю., Тюриков А. В., Гуляев П. В., Суворов А. С., Гафаров М. Р., Липанов С. И. Гибридный туннельный микроскоп с интегрированным бипотенциостатом для изучения наночастиц // *Химическая физика и мезоскопия*. 2011. Т. 13, № 3. С. 448-451.
4. Bui K., Fauman J., Kes D., Mandiola L. T., Ciomaga A., Salazar R., Bertozzi A. L., Gilles J., Goronzy D. P., Guttentag A. I., Weiss P. S. Segmentation of scanning tunneling microscopy images using variational methods and empirical wavelets // *Pattern Analysis and Applications. Online First Articles*, 2019, pp. 1-27.
5. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation // *International Conference on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. Lecture Notes in Computer Science*, 2015, vol. 9351, pp. 234-241.
6. Krizhevsky A., Sutskever I., Hinton G. E. Imagenet classification with deep convolutional neural networks // *Advances in neural information processing systems*, 2012, pp. 1097-1105.
7. Badrinarayanan V., Kendall A., Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation // *IEEE transactions on pattern analysis and machine intelligence*, 2017, vol. 39, no. 12, pp. 2481-2495.
8. Николенко С., Кадуринов А., Архангельская Е. Глубокое обучение. Погружение в мир нейронных сетей. СПб.: Издательский дом «Питер», 2018. 482 с.
9. Nair V. and Hinton G. E. Rectified linear units improve restricted boltzmann machines // *Proceedings of the 27th International Conference on International Conference on Machine Learning (ICML'10)*, 2010, pp. 807-814.
10. Fournier A., Fussell D., Carpenter L. Computer rendering of stochastic models // *Communications of the ACM*, 1982, vol. 25, iss. 6, pp. 371-384.
11. Otsu N. A threshold selection method from gray-level histograms // *IEEE Transactions on systems, man, and cybernetics*, 1979, vol. 9, iss. 1, pp. 62-66.
12. Шелковников Е. Ю., Гафаров М. Р., Гуляев П. В., Тюриков А. В., Суворов А. С. Сегментация и идентификация СТМ-изображений модифицируемой поверхности электрохимическим туннельным микроскопом // *Химическая физика и мезоскопия*. 2012. Т. 14, № 1. С. 143-146.
13. Ибрафиллов Х. С. Исследование методов бинаризации изображений // *Вестник науки и образования*. 2017. Т. 2, № 6(30). С. 43-50.

APPLICATION OF NEURAL NETWORK OF U-Net ARCHITECTURE FOR SEGMENTATION OF NANOPARTICLES ON CTM-PROBES

Shelkovnikov E. Yu., Shlyakhtin K. A., Shelkovnikova T. E., Egorov S. F.

Udmurt Federal Research Center, Ural Branch of the Russian Academy of Sciences, Izhevsk, Russia

SUMMARY. The article deals with the use of a network with the U-Net architecture for the segmentation of STM-images. The classes "particle" and "background" were chosen as classes for segmentation, the boundaries between the particles were correlated to the "background" class. To create the network, the libraries OpenCV and Keras were used together with TensorFlow based on the Python programming language. It was applied at the network input 512 training images of 256x256 pixels. For the preparation of training images, a technique for modeling an STM image with nanoparticles on a substrate was created by random placement of ellipsoids on the surface generated by the Diamond-Square method. An algorithm for generating a mask for a model image with the selection of the boundaries of intersecting ellipsoids has been developed, allowing the neural network to be trained to separate touching nanoparticles. Created a structural diagram of the resulting neural network. A comparative analysis of segmentation methods with the results of the U-Net network operation on 100 model images has been carried out. To compare the work of the algorithms, the IOU metric was used and the number of particles found in the images was counted. The study revealed the high efficiency of its use in the task of separating particles from the background in the image, while the developed neural network makes it possible to separate the contacting nanoparticles.

KEYWORDS: scanning tunneling microscope, convolutional neural network, U-Net, modeling.

REFERENCES

1. Fukuma T., Onishi K., Kobayashi N., Matsuki A. and Asakawa H. Atomic-resolution imaging in liquid by frequency modulation atomic force microscopy using small cantilevers with megahertz-order resonance frequencies. *Nanotechnology*, 2012, vol. 23, no. 13, pp. 135706. <http://dx.doi.org/10.1088/0957-4484/23/13/135706>
2. Gerber C. and Lang H. P. How the doors to the nanoworld were opened. *Nature Nanotechnology*, 2006, vol. 1, no. 1, pp. 3-5. <https://doi.org/10.1038/nnano.2006.70>
3. Shelkovnikov E. Yu., Tyurikov A. V., Gulyaev P. V., Suvorov A. S., Gafarov M. R., Lipanov S. I. Gibrinnyy tunnel'nyy mikroskop s integrirovannym bipotentsiostatom dlya izucheniya nanochastits [Chemical Physics and Mesoscopy]. *Khimicheskaya fizika i mezoskopiya* [Chemical Physics and Mesoscopy], 2011, vol. 13, no. 3, pp. 448-451.
4. Bui K., Fauman J., Kes D., Mandiola L. T., Ciomaga A., Salazar R., Bertozzi A. L., Gilles J., Goronzy D. P., Guttentag A. I., Weiss P. S. Segmentation of scanning tunneling microscopy images using variational methods and empirical wavelets. *Pattern Analysis and Applications. Online First Articles*, 2019, pp. 1-27. <https://doi.org/10.1007/s10044-019-00824-0>
5. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. Lecture Notes in Computer Science*, 2015, vol. 9351, pp. 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
6. Krizhevsky A., Sutskever I., Hinton G. E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 2012, pp. 1097-1105. <http://dx.doi.org/10.1145/3065386>
7. Badrinarayanan V., Kendall A., Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 2017, vol. 39, no. 12, pp. 2481-2495. <https://doi.org/10.1109/TPAMI.2016.2644615>
8. Nikolenko S., Kadurin A., Arkhangel'skaya E. Glubokoe obuchenie. Pogruzhenie v mir neyronnykh setey [Deep learning. Immersion in the world of neural networks]. St. Petersburg: Izdatel'skiy dom «Piter» Publ., 2018. 482 p.
9. Nair V. and Hinton G. E. Rectified linear units improve restricted boltzmann machines. *Proceedings of the 27th International Conference on International Conference on Machine Learning (ICML'10)*, 2010, pp. 807-814. <https://dblp.uni-trier.de/rec/html/conf/icml/NairH10>
10. Fournier A., Fussell D., Carpenter L. Computer rendering of stochastic models. *Communications of the ACM*, 1982, vol. 25, iss. 6, pp. 371-384. <https://doi.org/10.1145/358523.358553>
11. Otsu N. A threshold selection method from gray-level histograms. *IEEE Transactions on systems, man, and cybernetics*, 1979, vol. 9, iss.1, pp. 62-66. <https://doi.org/10.1109/TSMC.1979.4310076>
12. Shelkovnikov E. Yu., Gafarov M. R., Gulyaev P. V., Tyurikov A. V., Suvorov A. S. Segmentatsiya i identifikatsiya STM-izobrazheniy modifitsiruemoy poverkhnosti elektrokhimicheskim tunnel'nyy mikroskopom [Segmentation and identification of the STM images the surface being modified by an electrochemical tunnel microscope . *Khimicheskaya fizika i mezoskopiya* [Chemical Physics and Mesoscopy], 2012, vol. 14, no. 1, pp. 143-146.
13. Israfilov Kh. S. Issledovanie metodov binarizatsii izobrazheniy [Study of image binarization methods]. *Vestnik nauki i obrazovaniya* [Bulletin of science and education], 2017. vol. 2, no. 6(30), pp. 43-50.

Шелковников Евгений Юрьевич, доктор технических наук, профессор, зав. лабораторией, Институт механики УдмФИЦ УрО РАН, профессор кафедры «Вычислительная техника» ИжГТУ имени М.Т. Калашникова, e-mail: evshelk@mail.ru

Шляхтин Константин Алексеевич, магистр, младший научный сотрудник, Институт механики УдмФИЦ УрО РАН

Шелковникова Татьяна Евгеньевна, магистр, младший научный сотрудник, Институт механики УдмФИЦ УрО РАН, e-mail: tanya301195@gmail.com

Егоров Станислав Феликсович, кандидат технических наук, доцент, старший научный сотрудник, Институт механики УдмФИЦ УрО РАН, доцент кафедры ВТ ИжГТУ имени М.Т. Калашникова, e-mail: stos.mitm@mail.ru