

Применение генеративно-сопоставительных нейросетей для формирования баз данных в сканирующей туннельной микроскопии

Т.Е. Шелковникова¹, С.Ф. Егоров¹, П.В. Гуляев¹

¹ Институт механики, Удмуртский федеральный исследовательский центр, Российская академия наук, 426067, Россия, г. Ижевск, ул. Татьяны Барамзиной, д. 34

Аннотация

Получение изображения со сканирующего туннельного микроскопа является трудоемким процессом, а формирование большой базы данных для обучения нейросетей может занимать сотни часов. Статья посвящена созданию методики автоматической генерации изображений сканирующей туннельной микроскопии, позволяющей существенно сокращать время подготовки таких баз данных. В статье выполнен анализ современных методов и средств автоматической обработки изображений, полученных с зондовых и электронных микроскопов. Предложено использовать генеративно-сопоставительные сети для генерации изображений, полученных с применением сканирующего туннельного микроскопа, с целью формирования обучающих баз данных изображений. Описан процесс обучения и сравнения архитектур DCGAN с применением открытых библиотек OpenCV и Keras совместно с TensorFlow, выявлена наилучшая из них посредством вычисления метрик IS, FID, KID. С использованием метода дообучения генеративно-сопоставительной нейросети SRGAN и билинейной интерполяции на базе языка программирования Python осуществлено масштабирование изображений, полученных с DCGAN. Анализ вычисленных количественных значений метрик показал, что наилучшие результаты генерации изображений получены с применением DCGAN96 и сети SRGAN. Таким образом, разработана методика автоматической генерации базы данных изображений сканирующего туннельного микроскопа, основанная на поэтапном применении DCGAN96 и SRGAN. Показаны результаты сегментации изображений со сканирующего туннельного микроскопа с применением сети U-Net, обученной на сгенерированной базе данных. Все вычисления производились на видеокарте GTX GeForce 1070.

Ключевые слова: СТМ-изображение, генеративно-сопоставительная нейросеть, методика автоматической генерации, база данных, сегментация.

Цитирование: Шелковникова, Т.Е. Применение генеративно-сопоставительных нейросетей для формирования баз данных в сканирующей туннельной микроскопии / Т.Е. Шелковникова, С.Ф. Егоров, П.В. Гуляев // Компьютерная оптика. – 2023. – Т. 47, № 2. – С. 314-322. – DOI: 10.18287/2412-6179-CO-1144.

Citation: Shelkovnikova TE, Egorov SF, Gulyaev PV. Application of generative adversarial neural networks for the formation of databases in scanning tunneling microscopy. Computer Optics 2023; 47(2): 314-322. DOI: 10.18287/2412-6179-CO-1144.

Введение

В настоящее время в сканирующей зондовой микроскопии все более широкое применение при автоматизированной обработке изображений находят методы машинного обучения на основе нейронных сетей, в частности, сверточных нейронных сетей (СНС). Высокую эффективность СНС продемонстрировали в задачах сегментации различных частиц на микроскопических изображениях [1–7], задачах микроструктурного анализа различных материалов [8–11], в задачах сегментации различных макрообъектов [12–13]. Так, в работе [1] производилась сегментация изображений, полученных с применением сканирующего зондового микроскопа (СЗМ), металлических частиц на подложке из высокоориентированного пиролитического графита. Обработка осуществлялась в два этапа. На первом этапе использовалась СНС Cascade Mask-RCNN, и с её помощью определялись

положения частиц и их контуры. Далее производилась корректировка контуров наночастиц с помощью разработанного авторами метода (показано, что достигнутая точность превышает точность работы сети). В [2] описана система искусственного интеллекта DeepSPM, основанная на машинном обучении и предназначенная для мониторинга состояния и улучшения качества работы СЗМ. DeepSPM включает в себя алгоритмический поиск информационно-значимых областей (без резких перепадов и загрязнений на образце), сверточную нейронную сеть для оценки качества полученных изображений и сеть глубокого обучения для определения состояния зонда. DeepSPM может непрерывно собирать и классифицировать данные, контролируя состояние зонда в зависимости от изменяющихся условий эксперимента. Работа [3] посвящена новому методу сегментации с использованием СНС U-Net. Показано, что он не уступает по точности существующим методам сег-

ментации, кластеризации (метод Оцу, метод К-средних, сдвиг среднего значения) наноструктурированных систем, а на некоторых изображениях и превосходит их. В [4] предложена система глубокого обучения с применением СНС, способная определять виды атомов, дефекты в присутствии загрязнений, ступеней и шума.

Для обучения нейронных сетей с целью обработки изображений, полученных с применением сканирующего туннельного микроскопа (СТМ), важной задачей является формирование баз данных достоверных СТМ-изображений (которые сходны с реальными изображениями по характеру топографии поверхности и размерами частиц на ней). Получение СТМ-изображения является трудоемким процессом, включающим следующие основные операции: подготовка и установка образца и зондирующей иглы, подвод иглы к образцу и сканирование образца (обычно в режиме постоянного туннельного тока). В процессе формирования большой базы данных (состоящей из нескольких сотен или тысяч СТМ-изображений) требуется многократно повторять эти операции. В результате этот процесс может занимать много сотен часов. Следует также отметить, что изображения в СТМ подвержены различным искажениям (возникающим, например, из-за влияния помех аппаратуры, температурного дрейфа образца, несовершенной формы зондирующего острия и др.), в результате чего при формировании базы данных изображений, пригодных для обучения нейросети, некоторые из них будут исключены. В то же время исследуемые образцы могут сильно отличаться друг от друга, поэтому для дообучения сети или даже её переобучения необходимо пересканировать каждый новый образец. Эти факторы также значительно увеличивают время подготовки базы данных. Таким образом, разработка методики, которая позволяет из десятков изображений генерировать правдоподобную базу данных (сотни или тысячи изображений) и тем самым значительно уменьшать время для подготовки базы данных для обучения нейросетей, является актуальной задачей.

В настоящее время для генерации изображений различных объектов наиболее перспективным является применение генеративно-сопоставительных сетей

(ГСН). Так, в [14] исследовался подход для обучения СНС U-Net сегментации изображений индуцированных плюрипотентных стволовых клеток на небольшом наборе размеченных данных. Для этого сначала обучали классическую ГСН на исходных изображениях с клетками. Затем генератор ГСН использовался в качестве энкодера U-Net, и она обучалась с небольшим набором аннотированных вручную изображений. Точность работы такой нейросети, обученной на небольшом наборе данных, асимптотически сходилась с точностью работы классической сети U-Net, обученной на большой базе данных. В рамках исследований в [15] решалась задача сегментации несферических, перекрывающихся частиц и агломератов частиц на изображениях сканирующей электронной микроскопии с применением различных типов ГСН (WGAN, CycleGAN) и СНС типа MultiRes UNet. Весьма полезными для нас оказались исследования работы [16], в которой предложен метод повышения разрешения изображений, использующий архитектуру ГСН SRGAN. В соответствии с предложенным в [16] подходом в настоящей работе предлагается обучать ГСН восстанавливать СТМ-изображения с принудительно заниженным разрешением. При этом главной целью данной работы является разработка методики автоматической генерации СТМ-изображений для обучения нейронных сетей для сегментации изображений наночастиц.

1. Генерация базы данных СТМ-изображений с применением ГСН

1.1. Генерация СТМ-изображений с применением DCGAN

Структура генеративно-сопоставительных нейросетей предложена в 2014 году [17]. ГСН представляет собой алгоритм машинного обучения без учителя, его структурная схема представлена на рис. 1.

Согласно рис. 1 с помощью сопоставительного процесса одновременно обучаются две модели: генеративная модель (G) и дискриминативная модель (D). На вход генеративной модели подается случайный шум z , из которого генератор будет стремиться получить изображения, похожие на обучающие данные. Генератор и дискриминатор – это многослойные перцептроны.

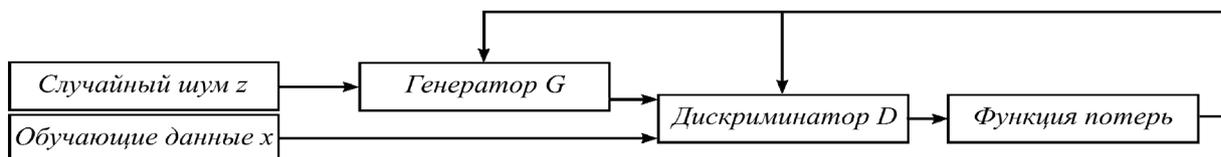


Рис. 1. Структурная схема ГСН

Обучение ГСН происходит путем одновременного обновления дискриминирующего распределения D и распределения генератора G так, чтобы дискриминатор мог различать объекты из обучающей выборки $x \sim p_{data}(x)$ и сгенерированной выборки $G(z)$ (где

$z \sim p_z(z)$). В процессе обучения обновляются пошагово веса многослойного перцептрона генератора θ_g (при фиксированных весах дискриминатора θ_d) и затем обновляются веса θ_d (при фиксированных весах генератора θ_g). После нескольких шагов обучения ГСН

должна прийти в состояние, когда $D(x) = 1/2$ (то есть дискриминатор не различает $x \sim p_{data}(x)$ и $x \sim p_g(z)$), соответственно, сеть не может больше улучшиться.

В работе для генератора и дискриминатора использовалась архитектура DCGAN [18], которая показала хорошие результаты в генерации изображений. Исходная архитектура генерировала изображения с разрешением 64×64 . Для неё предложено использовать в качестве функции активации нейронов функцию с названием Leaky ReLU (leaky rectified linear unit, LReLU):

$$f(x) = \begin{cases} x & x > 0, \\ 0,01x, & x \leq 0. \end{cases}$$

Функция Leaky ReLU существенно ускоряет процесс обучения за счет простоты вычислений [19].

Согласно рекомендациям [19] использована операция пакетной нормализации (batch normalization, BN), которая нормализует данные, подаваемые на вход слоя таким образом, чтобы математическое ожидание стало равно 0, а дисперсия – 1. Это позволяет уменьшить ковариационный сдвиг, образующийся вследствие того, что распределения тестовой и обучающей выборок различаются. Генератор DCGAN представляет собой классический декодер СНС, в данной сети состоит из серий из трех слоев: транспонированного сверточного слоя, который увеличивает размер изображения в несколько раз по обеим осям, слоя BN и LReLU. Структура применяемой сети представлена на рис. 2. На выходе генератора применялась функция активации гиперболического тангенса (tanh), которая ускоряет обучение сети и принимает значения $[-1; 1]$.

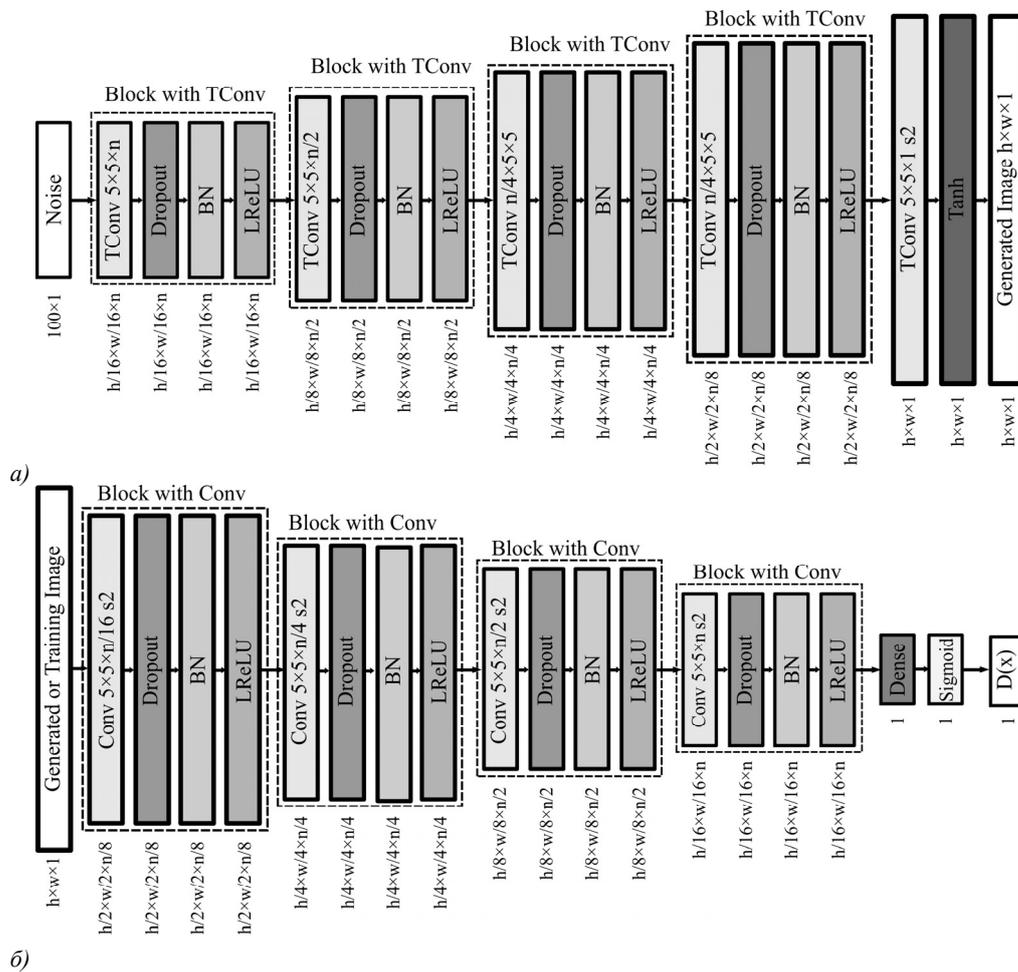


Рис. 2. Схема архитектуры генеративно-сопоставительной нейросети DCGAN для входного изображения с разрешением $h \times w$ (где h – высота изображения; w – ширина изображения; n – количество фильтров для слоев транспонированной и обычной свертки): а) генератор; б) дискриминатор

В качестве дискриминатора использовалась архитектура энкодера СНС. Структура обучаемой нами сети также состояла из серии слоев BN, LReLU, но вместо транспонированной свертки используется операция классической свертки (Conv). Таким образом, изображение, подаваемое на дискриминатор, преобразуется в карту признаков. На выходе дискримина-

тора после серии слоев со свертками (Block with Conv) использовался полносвязный слой (Dense) с функцией активации сигмоиды (Sigmoid). На выходе данной функции формируется скалярное значение $[0;1]$, которое показывает вероятность принадлежности поданного на дискриминатор изображения к обучающим данным. Для генерации СТМ-изображений классическая

DCGAN была дополнена слоями Dropout [20], которые исключают часть нейронов из слоев случайным образом, уменьшая возможность переобучения, которое возникает при обучении на относительно небольшом наборе данных. В качестве обучающей выборки использовалась база данных изображений (512×512) наночастиц [21]. База неоднородна и состоит из изображений с разным размером частиц и зашумленностью. Исходная база была увеличена до 2000 изображений с разрешением 256×256 с помощью библиотеки TensorFlow и операций, позволяющих случайно выбрать область на изображении или повернуть изображение на случайный угол.

Для генерации СТМ-изображений размером 256×256 классическая модель DCGAN (которая могла генерировать изображения 64×64) была отмасштабирована путем введения параметров h , w и n (где h и w – высота и ширина изображения; n – количество фильтров для слоев свертки и транспонированной свертки; параметр n выбирается максимально большим и ограничивается памятью видеокарты). Таким образом, было получено несколько архитектур DCGAN, способных генерировать изображения разных размеров: 32×32, 48×48, 64×64, 96×96, 128×128. С увеличением размеров изображения и соответственным увеличением количества нейронов обучение модели стало очень медленным, а качество сгенерированных изображений

было неудовлетворительным. Проблемы, возникающие при попытке генерации больших изображений на GAN, описаны в [22]. Поэтому для их решения в настоящей работе применялся подход, предложенный в работе [23], который предполагает для создания изображений сначала использовать GAN, генерирующую изображение с низким разрешением, а затем увеличивать разрешение полученных изображений с помощью дополнительной СНС.

1.2. Увеличение сгенерированных СТМ-изображений с применением SRGAN

В работе [24] представлена SRGAN – генеративно-сопоставительная нейросеть для получения высококачественных изображений с коэффициентом 4-кратного масштабирования из изображений с разрешением 64×64 в 256×256. В настоящей работе применялся подход, предложенный авторами [24], в котором для повышения разрешения использовалась SRGAN (рис. 3). В качестве генератора для этой сети использовалась такая же СНС, как и в классическом подходе, но для оценки сгенерированных данных вместо СКО применялся дискриминатор, предсказывающий, является изображение сгенерированным или реальным. Архитектура SRGAN показала значительно лучшее качество сгенерированных изображений высокого разрешения, чем СНС для улучшения разрешения изображения REDNet [25].

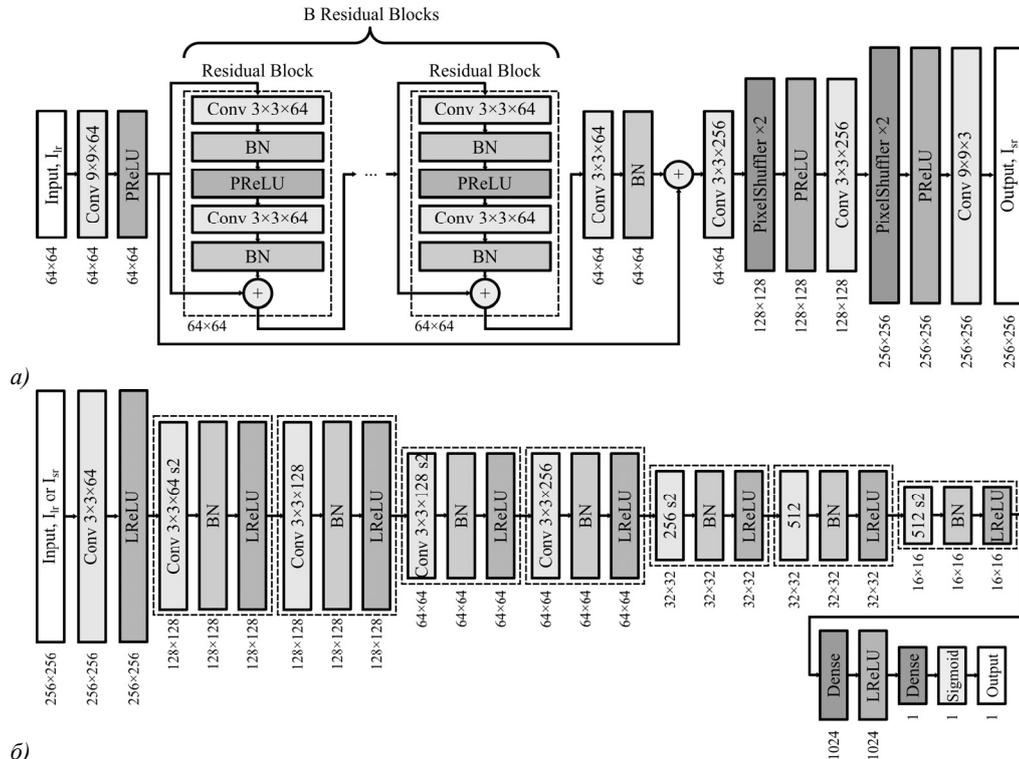


Рис. 3. Схема архитектуры генеративно-сопоставительной нейросети SRGAN: а) генератор; б) дискриминатор

Таким образом, предлагается для автоматической генерации базы данных СТМ-изображений выполнять генерацию изображений с низким разрешением с помощью DCGAN, а затем улучшать полученные изоб-

ражения с помощью SRGAN. Для повышения разрешения СТМ-изображений в работе применялась предобученная сеть SRGAN. Архитектура генератора SRGAN построена на основе СНС ResNet [26]. В

настоящей работе к входному изображению применялась свертка с ядром 9×9 , 64 каналами и активацией ParametricReLU (PReLU). Функция PReLU схожа с LReLU, но имеет дополнительный обучаемый параметр – коэффициент утечки, вместо фиксированного значения (0,01). Применение этой функции увеличивает точность сети при незначительном повышении вычислительных затрат [27]. Повышение разрешения в 2 раза в этой сети выполняется с помощью слоев субпиксельной свертки PixelShuffler, предложенной в работе [28]. На вход дискриминатора в работе подавалось изображение 256×256 с выхода генератора или реальное изображение из базы данных. Дискриминатор обучался таким образом, чтобы на его выходе была 1, если изображение реальное, и 0, если изображение создано генератором.

Полученные с SRGAN выходные изображения были размыты. Поэтому сеть была дообучена на наборе реальных СТМ-изображений. Дообучение – это процесс обучения сети, при котором на вход уже обученной сети подаются новые данные. Такой подход позволяет ускорить обучение сети и обновить веса нейронных связей для корректного увеличения изображений.

Таким образом, созданная методика подразумевает трехэтапную генерацию изображений.

1. На первом этапе осуществляется искусственное увеличение базы СТМ-изображений. Для исходных изображений с высоким разрешением итеративно (порядка 100 раз) выбираются случайные области с меньшим размером и поворачиваются на случайный угол.

2. Второй этап предполагает выполнение генерации изображений с низким разрешением и применением сети DCGAN и полученной на первом этапе базы.

3. На третьем этапе выполняется обратная операция увеличения полученных изображений с выхода DCGAN с использованием дообученной на СТМ-изображениях сети SRGAN.

2. Результаты и их обсуждение

Для оценки качества сгенерированных изображений применялись классические метрики, используемые для оценки работы ГСН: Inception score (IS), Frechet inception distance (FID), Kernel Inception Distance (KID). Метрика IS, предложенная в работе [17] в качестве альтернативы человеческой оценке качества сгенерированных изображений, показала хорошую согласованность с ней. Однако у IS есть ограничение, связанное с использованием предобученной сети классификатора InceptionV3. Изображения, подлежащие оценке, должны соответствовать требованиям разрешения входного слоя сети 299×299 , а также они должны содержать объекты классов, которые может различать сеть (предварительно обученная сеть для оценки взята из Keras, она обучена на наборе данных ILSVRC 2012 с 1000 классов) [29]. Поэтому для расчета IS сгенерированные

СТМ-изображения масштабировались, однако проблема с классами так и оставалась. Также в статье [23] подробно рассмотрены другие принципиальные ограничения данной метрики.

Более универсальной метрикой оценки качества ГСН является FID. Она предложена в [30] и не требует четкой классификации сгенерированных изображений. FID оценивает схожесть реальных и сгенерированных изображений путем вычисления расстояния Фреше между выходом предпоследнего слоя сверточной нейронной сети InceptionV3 для сгенерированных изображений и для обучающих данных. В отличие от IS у этой метрики нет ограничений, связанных с тем, что оцениваемые изображения не относятся к базе, на которой обучалась InceptionV3, так как не используется последний классифицирующий слой сети. В процессе вычисления метрики на вход предобученной сети InceptionV3 сначала подаются реальные изображения, для них рассчитываются математическое ожидание и ковариация выходов нейронов предпоследнего слоя. Отсюда FID определяется по формуле:

$$d^2(X, Y) = \|m_x - m_y\|_2^2 + Tr(C_x + C_y - 2(C_x C_y)^{1/2}),$$

где X и Y – реальное и сгенерированное представление с предпоследнего слоя InceptionV3, которые считаются двумя многомерными нормальными распределениями; m_x , C_x и m_y , C_y – математические ожидания и ковариация распределений X и Y ; Tr – операция, отображающая пространство квадратных матриц в поле (след матрицы). Численное значение данной метрики в идеальном случае равно 0. В ходе обучения сетей данная метрика убывает по мере генерации визуально лучших изображений.

В работе [31] представлена метрика Kernel Inception Distance (KID). Она (как и FID) оценивает расстояние между двумя нормальными распределениями и равна квадрату максимального среднего расхождения между распределением выхода предпоследнего слоя сети Inception на изображениях, генерируемых ГСН, и на исходных изображениях. Метрика KID рассчитывается по формуле:

$$KID(X, Y) = \frac{1}{m(m-1)} \sum_{i \neq j}^m k(x_i, x_j) + \frac{1}{m(m-1)} \sum_{i \neq j}^m k(y_i, y_j) - \frac{2}{m^2} \sum_{i=1}^m \sum_{j=1}^m k(x_i, y_j),$$

где $X = \{x_i\}_{i=1}^m$ и $Y = \{y_j\}_{j=1}^m$ – выходы предпоследнего слоя сети Inception размерностью m на сгенерированных и исходных изображениях соответственно; k – полиномиальное ядро. Авторы использовали полиномиальное ядро третьей степени:

$$k(x, y) = \left(\frac{1}{m} x^T y + 1 \right)^3.$$

В отличие от FID, метрика KID не имеет смещения и позволяет сравнивать выборки с разным количеством изображений.

Для формирования обучающей базы данных над исходным набором из N СТМ-изображений выполнялись операции, позволяющие увеличить базу данных: случайный выбор области на изображении зеркального отражения. Из каждого исходного изображения 512×512 вырезалось M случайных областей 256×256 . Операции отражения позволяли увеличить базу еще в K раз. Таким образом, исходный набор изображений увеличивался до $N \times M \times K$ изображений. Данный процесс занимает порядка нескольких секунд. Для обучения DCGAN изображения с разрешением 256×256 масштабировались до разрешений 32×32 , 48×48 , 64×64 , 96×96 , 128×128 с целью захвата большего количества признаков агломераций частиц. Далее генерировались изображения с помощью DCGAN размером 32×32 , 48×48 , 64×64 , 96×96 , 128×128 ($n=256$ для $h \leq 64$, $w \leq 64$ и $n=128$ для $h > 64$, $w > 64$) и одним цветовым каналом. Каждая нейросеть обучалась соответственно 20, 26, 30, 34, 40 часов на видеокарте GTX GeForce 1070. Далее полученные изображения масштабировались до разрешения 256×256 с применением билинейной интерполяции и дообученной сети SRGAN. Метрики рассчитывались для 100 сгенерированных разными методами изображений и для 100 исходных изображений, вычислялось их среднее значение

и SKO. Полученные значения метрик представлены в табл. 1, а сгенерированные изображения приведены на рис. 4.

Анализ табл. 1 показал, что результаты генерации СТМ-изображений, полученных с применением DCGAN96 и сети SRGAN, являются наиболее близкими к реальным. Также значения метрик FID и KID показывают, что метод SRGAN точнее билинейной интерполяции во всех случаях, кроме DCGAN32. Это связано с тем, что предобученная сеть SRGAN восстанавливает исходное СТМ-изображение согласно особенностям данных, на которых было произведено дообучение сети. Билинейная интерполяция является универсальным методом, которому не важны особенности восстанавливаемого изображения. Поэтому для размеров изображения 32×32 , когда сгенерированное DCGAN изображение практически не походит на обучающие данные, билинейная интерполяция работает лучше. Численные значения оценки IS составляли ~ 1 и практически не изменялись в ходе обучения, так как наименьшее значение оценки – это 1 (наибольшее значение метрики соответствует максимальному количеству классов, которые может определять сеть классификатор). Безусловно, возможно исправление этого недостатка путем обучения сети InceptionV3 на своей базе данных с необходимым количеством классов, однако четко классифицировать СТМ-изображения для неисследованных образцов довольно затруднительно.

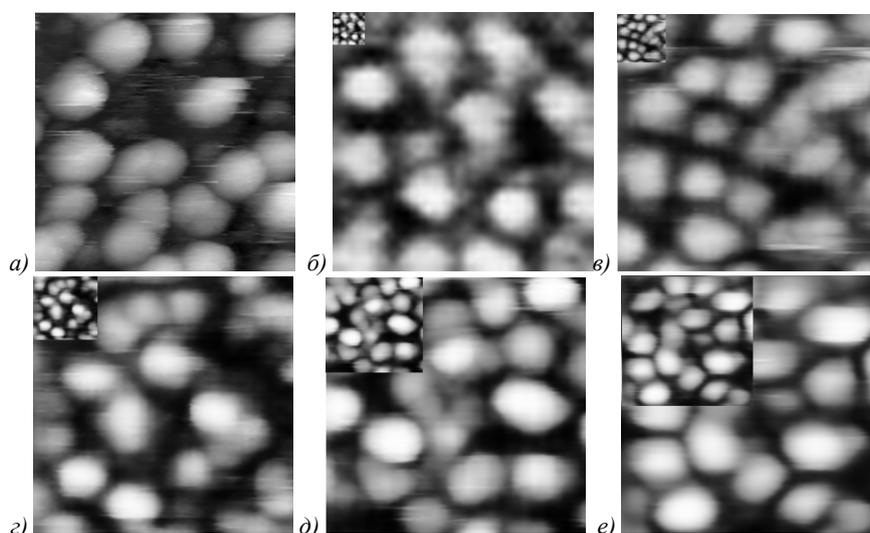


Рис. 4. СТМ-изображения, полученные с применением GCH DCGAN: а) исходное СТМ-изображение; б) изображение с разрешением 32×32 , полученное с DCGAN, и увеличенное до 256×256 его представление с применением SRGAN; в) изображение с разрешением 48×48 , полученное с DCGAN, и его увеличенное до 256×256 представление с применением SRGAN; г) изображение с разрешением 64×64 , полученное с DCGAN, и его увеличенное до 256×256 представление с применением SRGAN; д) изображение с разрешением 96×96 , полученное с DCGAN, и его увеличенное до 256×256 представление с применением SRGAN; е) изображение с разрешением 128×128 , полученное с DCGAN, и его увеличенное до 256×256 представление с применением SRGAN

Рельеф исходного СТМ-изображения, сгенерированного СТМ-изображения, полученного на DCGAN96 и увеличенного до 256×256 с применением SRGAN, представлен на рис. 5. Видно, что на полученном и исходном рельефе подложки наночастицы имеют схожую форму. В целом, сгенерированное

изображение получилось более сглаженным, при этом размеры и агломерации наночастиц соотносятся с данными из обучающей базы.

В качестве примера применения разработанной методики сформирована база данных СТМ-изображений для обучения нейронной сети сегмента-

ции наночастиц. На рис. 6 приведены результаты сегментации СТМ-изображения с помощью сети U-

Net [5], обученной на полученной базе данных и размеченной с помощью программы Labelme [32].

Табл. 1. Сравнение архитектур GCH с применением метрик IS, FID, KID

Нейросеть	IS		FID		KID	
	Mean	STD	Mean	STD	Mean	STD
DCGAN32+SRGAN	1,53	0,07	329,32	22,91	0,43	0,12
DCGAN32+Bilinear interpolation	1,52	0,06	319,36	18,46	0,39	0,07
DCGAN32+SRGAN	1,54	0,05	351,86	13,73	0,41	0,10
DCGAN48+Bilinear interpolation	1,43	0,12	346,43	14,29	0,40	0,10
DCGAN48+SRGAN	1,50	0,07	315,95	26,03	0,31	0,13
DCGAN64+Bilinear interpolation	1,49	0,10	287,94	22,45	0,32	0,05
DCGAN64+SRGAN	1,53	0,05	266,85	24,56	0,23	0,10
DCGAN96+Bilinear interpolation	1,53	0,08	335,39	17,53	0,44	0,13
DCGAN96+SRGAN	1,53	0,06	234,26	23,61	0,20	0,08
DCGAN128+Bilinear interpolation	1,51	0,07	303,51	24,00	0,27	0,18
DCGAN128+SRGAN	1,53	0,08	265,30	18,85	0,22	0,09

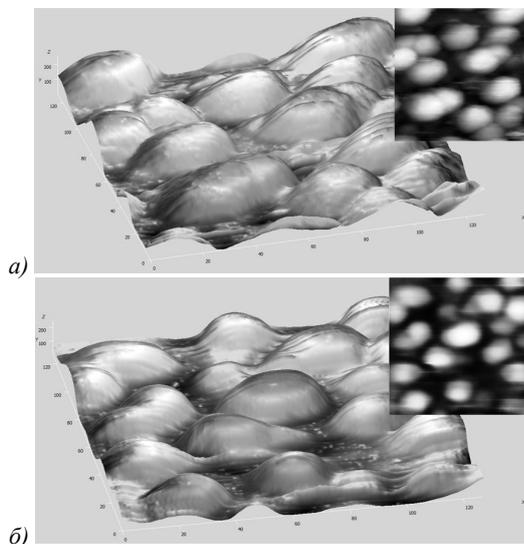


Рис. 5. СТМ-изображения, полученные с применением GCH DCGAN: а) исходное СТМ-изображение; б) изображение с разрешением, полученное с DCGAN96 и с применением SRGAN

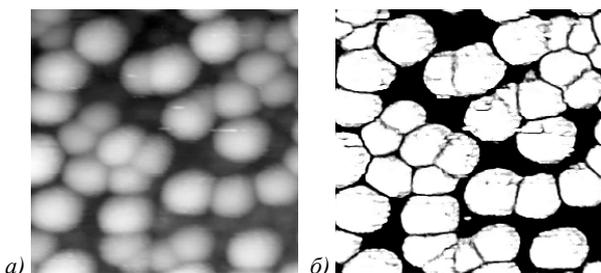


Рис. 6. а) Исходное СТМ-изображение; б) изображение, сегментированное с применением сети U-Net

Заключение

Выполнен анализ, который показал возможности применения известных сетей для формирования баз данных СТМ-изображений. Для генерации базы данных СТМ-изображений была применена генеративно-сопоставительная нейросеть архитектуры DCGAN для разных размеров входных изображений. Посредством применения GCH SRGAN сформирована база СТМ-

изображений. Разработана методика для автоматической генерации СТМ-изображений с применением методов машинного обучения. Выявлено, что наилучшие результаты получены с применением сети DCGAN96 и сети SRGAN, посредством вычисления метрик IS, FID, KID. Можно сделать вывод, что разработанная методика позволяет автоматически формировать большую базу данных СТМ-изображений (до сотен или даже тысяч) из небольшого количества входных данных (порядка нескольких десятков), что существенно уменьшает время подготовки базы данных для обучения искусственных нейронных сетей.

References

1. Okunев AG, Mashukov MYu, Nartova AV, Matveev AV. Nanoparticle recognition on scanning probe microscopy images using computer vision and deep learning. *Nanomaterials* 2020; 10(7): 1285. DOI: 10.3390/nano10071285.
2. Krull A, Hirsch P, Rother C, Schiffrin A, Krull C. Artificial-intelligence-driven scanning probe microscopy. *Commun Phys* 2020; 3: 54. DOI: 10.1038/s42005-020-0317-3.
3. Farley S, Hodgkinson JEA, Gordon OM, et al. Improving the segmentation of scanning probe microscope images using convolutional neural networks. *Mach Learn: Sci Technol* 2020; 2(1): 015015. DOI: 10.1088/2632-2153/abc81c.
4. Ziatdinov M, Fuchs U, Owen J, Randall J, Kalinin S. Robust multi-scale multi-feature deep learning for atomic and defect identification in Scanning Tunneling Microscopy on H-Si(100) 2x1 surface. *arXiv Preprint*. 2020. Source: (https://arxiv.org/abs/2002.04716). DOI: 10.48550/arXiv.2002.04716.
5. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In Book: Navab N, Hornegger J, Wells WM, Frangi AF, eds. *Medical image computing and computer-assisted intervention – MICCAI 2015*. Cham: Springer; 2015: 234-241. DOI: 10.1007/978-3-319-24574-4_28.
6. Shelkovnikov E, Shlyakhtin K, Shelkovnikova T, Egorov S. Application of neural network of U-Net architecture for segmentation of nanoparticles on STM-probes. *HFIM* 2019; 21(2): 330-336. DOI: 10.15350/17270529.2019.2.36.
7. Egorov S, Arhipov I, Tatyana S. Information system for segmentation of nanoparticles in STM-images. *CEUR Workshop Proc* 2020; 2665: 130-134.

8. Lubbers N, Lookman T, Barros K. Inferring low-dimensional microstructure representations using convolutional neural networks. *Phys Rev E* 2017; 96: 052111. DOI: 10.1103/PhysRevE.96.052111.
9. Belianinov A, Vasudevan R, Strelcov E, et al. Big data and deep data in scanning and electron microscopies: deriving functionality from multidimensional data sets. *Adv Struct Chem Imag* 2015; 1: 6. DOI: 10.1186/s40679-015-0006-6.
10. Chowdhury A, Kautz E, Yener B, Lewis D. Image driven machine learning methods for microstructure recognition. *Comput Mater Sci* 2016; 123: 176-187. DOI: 10.1016/j.commatsci.2016.05.034.
11. Li W, Field KG, Morgan D. Automated defect analysis in electron microscopic images. *Npj Comput Mater* 2018; 4: 36. DOI: 10.1038/s41524-018-0093-8.
12. Gavrilo DA. Investigation of the applicability of the convolutional neural network U-Net to a problem of segmentation of aircraft images. *Computer Optics* 2021; 45(4): 575-579. DOI: 10.18287/2412-6179-CO-804.
13. Gorbachev VA, Krivorotov IA, Markelov AO, Kotlyarova EV. Semantic segmentation of satellite images of airports using convolutional neural networks. *Computer Optics* 2020; 44(4): 636-645. DOI: 10.18287/2412-6179-CO-636.
14. Majurski M, Manescu P, Padi S, et al. Cell image segmentation using generative adversarial networks, transfer learning, and augmentations. 2019 IEEE/CVF Conf on Computer Vision and Pattern Recognition Workshops (CVPRW) 2019: 1114-1122. DOI: 10.1109/CVPRW.2019.00145.
15. Rühle B, Krumrey JF, Hodoroaba V-D. Workflow towards automated segmentation of agglomerated, non-spherical particles from electron microscopy images using artificial neural networks. *Sci Rep* 2021; 11: 4942. DOI: 10.1038/s41598-021-84287-6.
16. Zhang H, Fang C, Xie X, Yang Y, Jin D. High-throughput, high-resolution registration-free generative adversarial network microscopy. *Biomed Opt Express* 2019; 10(3): 1044-1063. DOI: 10.1364/BOE.10.001044.
17. Goodfellow IJ, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. *Adv Neural Inf Process Syst* 2014; 27.
18. Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv Preprint*. 2016. Source: <https://arxiv.org/abs/1511.06434>. DOI: 10.48550/arXiv.1511.06434.
19. Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. *Int Conf on Machine Learning (ICML)* 2010: 807-814.
20. Salimans T, Goodfellow I, Zaremba W, Cheung V, Radford A, Chen X. Improved techniques for training GANs. *Adv Neural Inf Process Syst* 2016: 2234-2242. DOI: 10.48550/arXiv.1606.03498.
21. Publications and Data. Source: <http://particlesnn.nsu.ru/text/publications>.
22. Hinton GE, Srivastava N, Krizhevsky A, Sutskever I, Salakhutdinov RR. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv Preprint*. 2012. Source: <https://arxiv.org/pdf/1207.0580.pdf>.
23. Mohsen M, Moustafa M. Generating large scale images using GANs. In: Jiang X, Hwang J-N, eds. Eleventh international conference on digital image processing (ICDIP 2019), Guangzhou, China: SPIE; 2019: 195. DOI: 10.1117/12.2540489.
24. Ledig C, Theis L, Huszar F, et al. Photo-realistic single image super-resolution using a generative adversarial network. *arXiv Preprint*. 2017. Source: <https://arxiv.org/abs/1609.04802>. DOI: 10.1109/CVPR.2017.19.
25. Mao X-J, Shen C, Yang Y-B. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. *Proc Advances in Neural Information Processing Systems* 2016: 2802-2810. DOI: 10.48550/arXiv.1603.09056.
26. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *IEEE Conf on Computer Vision and Pattern Recognition (CVPR)* 2016: 770-778. DOI: 10.1109/CVPR.2016.90.
27. He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. *IEEE Int Conf on Computer Vision (ICCV)* 2015: 1026-1034. DOI: 10.1109/ICCV.2015.123.
28. Shi W, Caballero J, Huszar F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. *IEEE Conf on Computer Vision and Pattern Recognition (CVPR)* 2016: 1874-1883. DOI: 10.1109/CVPR.2016.207.
- Barratt S, Sharma R. A note on the inception score. *arXiv Preprint*. 2018. Source: <https://arxiv.org/pdf/1801.01973.pdf>.
29. Heusel M, Ramsauer H, Unterthiner T, Nessler B, Hochreiter S. GANs trained by a two time-scale update rule converge to a local nash equilibrium. *Proc 31st Int Conf on Neural Information Processing Systems (NIPS'17)* 2017: 6629-6640.
30. Bińkowski M, Sutherland DJ, Arbel M, Gretton A. Demystifying MMD GANs. *ICLR 2018: Int Conf on Learning Representations* 2018. Source: <https://openreview.net/pdf?id=r1IUOzWCW>.
31. wkentaro/labelme. 2016. Source: <https://github.com/wkentaro/labelme>.

Сведения об авторах

Шелковникова Татьяна Евгеньевна, 1995 года рождения, в 2019 году окончила магистратуру ИжГТУ имени М.Т. Калашникова по программе «Информационно-измерительные системы». В том же году поступила в очную аспирантуру Удмуртского федерального исследовательского центра УрО РАН. Работает младшим научным сотрудником в лаборатории информационно-измерительных систем Удмуртского федерального исследовательского центра УрО РАН. Область научных интересов: обработка изображений, сканирующая туннельная микроскопия, методы машинного обучения. E-mail: shelktan@udman.ru.

Егоров Станислав Феликсович, 1967 года рождения, в 1989 году окончил факультет информатики и вычислительной техники Ижевского государственного технического университета, кандидат технических наук, доцент, работает старшим научным сотрудником в лаборатории информационно-измерительных систем Уд-

муртского федерального исследовательского центра УрО РАН. Область научных интересов: компьютерное моделирование, системы распознавания образов. E-mail: stos.mitm@mail.ru.

Гуляев Павел Валентинович, 1973 года рождения, в 1995 году окончил факультет информатики и вычислительной техники Ижевского государственного технического университета; кандидат технических наук, работает старшим научным сотрудником в лаборатории информационно-измерительных систем Удмуртского федерального исследовательского центра УрО РАН. Область научных интересов: обработка изображений, высокоточные перемещения, зондовая микроскопия. E-mail: lucac@inbox.ru.

ГРНТИ: 50.53.17

Поступила в редакцию 6 апреля 2022 г. Окончательный вариант – 8 сентября 2022 г.

Application of generative adversarial neural networks for the formation of databases in scanning tunneling microscopy

T.E. Shelkownikova¹, S.F. Egorov¹, P.V. Gulyaev¹

¹ Institute of Mechanics, Udmurt Federal Research Center, Ural Branch, Russian Academy of Sciences, 426067, Izhevsk, Russia, ul. Baramzinoi 34

Abstract

We discuss the development of a technique for automatic generation of databases of images obtained with a scanning tunneling microscope. An analysis of state-of-the-art methods and means of automatic processing of images obtained from probe and electron microscopes is carried out. We proposed using generative-adversarial networks for generating images taken with a scanning tunneling microscope to form training databases of images. A process of training and comparison of deep convolutional generative adversarial network (DCGAN) architectures using the OpenCV and Keras libraries together with TensorFlow is described, with the best of them identified by computing the metrics IS, FID, and KID. The scaling of images obtained from DCGAN is performed using a method of fine tuning of a super-resolution generative adversarial neural network (SRGAN) and bilinear interpolation based on the Python programming language. An analysis of calculated quantitative metrics values shows that the best results of image generation are obtained using DCGAN96 and SRGAN. It is found that FID and KID metric values for SRGAN method are better than values for bilinear interpolation in all cases except for DCGAN32. All calculations are performed on a GTX GeForce 1070 video card. A method for automatic generation of a scanning tunneling microscope image database based on the stepwise application of DCGAN and SRGAN is developed. Results of generation and comparison of the original image, the one obtained with DCGAN96 and the enlarged image with SRGAN are presented.

Keywords: STM-image, generative adversarial neural networks, automatic generation method, database, convolution.

Citation: Shelkownikova TE, Egorov SF, Gulyaev PV. Application of generative adversarial neural networks for the formation of databases in scanning tunneling microscopy. *Computer Optics* 2023; 47(2): 314-322. DOI: 10.18287/2412-6179-CO-1144.

Authors' information

Tatyana Evgenievna Shelkownikova (b.1995) graduated (2019) from Izhevsk State Technical University. Currently she is junior researcher of Laboratory of Information and Measurement System in Udmurt Federal Research Center, Ural Branch of the Russian Academy of Sciences. Research interests are development and improvement of methods and means of scanning tunneling microscopy. E-mail: shelktan@udman.ru.

Stanislav Felixovich Egorov (b.1967) graduated (1989) from Izhevsk State Technical University, majoring in Informatics and Computer Sciences. Candidate in Engineering Science (1998). Currently he is senior researcher of Laboratory of Information and Measurement System in Udmurt Federal Research Center, Ural Branch of the Russian Academy of Sciences. Research interests: computer modeling, pattern recognition systems. E-mail: stos.mitm@mail.ru.

Pavel Valentinovich Gulyaev (b. 1973) graduated (1995) from Izhevsk State Technical University, majoring in Informatics and Computer Sciences. Candidate in Engineering Science (2004). Currently he is senior researcher of Laboratory of Information and Measurement System in Udmurt Federal Research Center, Ural Branch of the Russian Academy of Sciences. Research interests are image processing, high-precision positioning, scanning probe microscopy. E-mail: lucac@inbox.ru.

Received April 6, 2022. The final version – September 8, 2022.
