

¹ И.А. ЕЛЬКИН

канд. физ.-мат. наук, научный сотрудник

¹ А.И. ЧУКАВИН

канд. физ.-мат. наук, старший научный сотрудник

^{1,2} О.Б. БАРЫШЕВ

аспирант

¹ Удмуртский Федеральный Исследовательский Центр УрО РАН² АО «Электонд»

СЕГМЕНТАЦИЯ ДЕФЕКТОВ ОКСИДНОГО СЛОЯ ТАНТАЛОВЫХ КОНДЕНСАТОРОВ С ПОМОЩЬЮ СВЁРТОЧНОЙ НЕЙРОННОЙ СЕТИ

Подготовлен набор данных и получена свёрточная нейронная сеть, которая с высокой точностью позволяет сегментировать дефекты диэлектрического слоя танталовых металло-оксидных конденсаторов после электрохимического окисления или после снятия катодных слоёв. Показано, что использование модели в режиме скользящего окна и варьирование шага сканирования позволяет повысить точность обнаружения дефектов и анализировать изображения произвольного размера.

Ключевые слова: танталовые конденсаторы, дефекты, глубокое обучение, семантическая сегментация.

¹ I.A. ELKIN

Cand. of Phys.-Math. Sciences, Researcher

¹ A.I. CHUKAVIN

Cand. of Phys.-Math. Sciences, Senior Researcher

^{1,2} O.B. BARYSHEV

postgraduate student

¹ Udmurt Federal Research Center Ural Branch of Russian Academy of Sciences² JSC Elecond

SEGMENTATION OF OXIDE LAYER DEFECTS IN TANTALUM CAPACITORS WITH CONVOLUTIONAL NEURAL NETWORK

The dataset has been prepared and convolutional neural network has been trained. The resulting model makes it possible to segment defects in the dielectric layer of tantalum metal oxide capacitors after electrochemical oxidation or after the removal of the cathode layers with high accuracy. It was shown that if we use the model in the sliding window mode and vary the scanning step, we can improve the quality of defects segmentation and analyze large images.

Keywords: tantalum capacitors, defects, deep learning, semantic segmentation.

DOI: 10.25791/pfim.04.2023.1271

Введение

Наиболее надёжными и пригодными для жёстких условий эксплуатации в современной электронной аппаратуре являются танталовые оксидно-электролитические конденсаторы. Конструктивно конденсатор содержит пористый анод (тантал), покрытый слоем диэлектрика (плёнка аморфного оксида тантала) и катодную обкладку, включающую электролит. Внутри и на границе диэлектрика возможно наличие дефектов: кристаллические участки

и утонения оксидного слоя, включения и примеси, а также сколы и трещины. Природа, концентрация, электрическая активность дефектов анодной плёнки определяется составом и структурой поверхностного слоя тантала, условиями формирования оксидного слоя и технологическими режимами формирования катодной обкладки.

Наличие дефектов может приводить к формированию областей с повышенной электрической проводимостью (кристаллические участки, включения), в которых происходит локальное увеличение

температуры и рост дефектной области. Кроме этого, дефекты могут вызывать электрическую деструкцию оксидного слоя, которая протекает в виде разрыва диэлектрика или в виде роста кристаллических оксидов металла [1], что может приводить к короткому замыканию и воспламенению конденсатора [2]. Повышение отказоустойчивости чаще всего достигается за счёт снижения примесей в исходном порошке и подбором технологических режимов позволяющих увеличить толщину диэлектрика и снизить его дефектность [3].

Для оценки качества диэлектрика применяется высоковольтный стресс-тест (измерение напряжения пробоя), в котором анод со сформированным диэлектриком включают в электрическую цепь и плавно повышают напряжение до отказа работы (пробоя диэлектрика). Чем выше напряжение пробоя, тем выше качество диэлектрика и, следовательно, выше надёжность и время безотказной работы конденсатора [4]. При проведении данного тестирования могут не учитываться краткосрочные пробои диэлектрика, появившиеся на ранних стадиях испытания перед «основным» пробоем, который привёл к отказу работы и невозможности дальнейшего повышения напряжения («микро пробои»). Дополнительную сложность выявления дефектов вносит рост оксидной плёнки во время испытаний. Кроме напряжения пробоя об особенностях оксидного слоя говорит морфология и площадь дефектов, образовавшихся как после испытаний, так и в процессе эксплуатации. Данная информация может быть полезна для выявления корреляций между условиями получения анодов и их качеством, а также служить для улучшения производственных процессов и повышения качества выпускаемых конденсаторов, в том числе при изменении качества материалов и технологических режимов.

Подход к обнаружению дефектов на основе визуальной информации играет важную роль в современном производстве, чтобы гарантировать высокий уровень качества и избежать ненужных потерь [5]. Этот процесс поддаётся автоматизации с помощью различных методов компьютерного зрения: классических или глубокого обучения. К классическим, например, относятся методы опорных векторов [6] и деревьев решений [7], однако они требуют ручного выделения признаков из изображений и если признаки неверны, то это значительно ухудшит точность работы модели. Методы глубокого обучения, такие как сверточные нейронные сети не требуют ручного создания признаков, они могут выполнять сквозной процесс обучения, и в результате определять участки изображений, отвечающие дефектным областям [8]. Компьютерное зрение хорошо себя зарекомендовало для решения ряда инженерных задач, например, для определения поверхностных дефектов в постоянных магнитах электродвигателей [9], стали [10], тепловых предохранителях [11], для обнаружения

аномалий в электродах литий-ионных аккумуляторов [12], а также в аддитивном производстве [13]. Для решения задач предиктивной аналитики, например, при прогнозировании шероховатости деталей [14], для классификации типов печатных плат и дефектов паяных соединений в них [15], а также для нахождения неисправных интегрированных пассивных компонентов [16].

В настоящее время поиск дефектов диэлектрического слоя танталовых металло-оксидных конденсаторов осуществляется экспертом в ручном режиме по электронно-микроскопическим изображениям. Однако это достаточно медленно, может приводить к нестабильности результатов и финансово затратно, поскольку для получения достоверной информации нужно проанализировать большое количество изображений от разных образцов. В связи с этим, целью работы являлась разработка модели глубокого обучения для автоматического обнаружения дефектов, образовавшихся в аморфном оксидном слое и определения их площади.

Материалы и методики

В работе исследовались танталовые аноды различных производителей, полученные из агломерированного порошка плотность и пористость которых была на уровне 5500 кг/м^3 и 67%, соответственно. Процесс производства анодов включал несколько этапов, в начале проводились прессование и высокотемпературное спекание. После этого выполнялось электрохимическое окисление (формовка) в проточном электролите 0,1 % ортофосфорной кислоты (сопротивление $250 \text{ Ом}\cdot\text{см}$) с использованием танталового катода. Затем образцы в течение 2 ч промывали в дистиллированной воде (сопротивление $5 \text{ МОм}\cdot\text{см}$) при температуре $70 \pm 5^\circ\text{C}$ с проточностью не менее 0,5 л/мин и сушили 15 мин при 150°C . На завершающей стадии в течение 60 минут проводили имитацию температурного технологического воздействия при 250°C в температурной камере на воздухе. После приготовления образцов получали изображения поверхности с помощью электронного микроскопа Thermo Fisher Quattro S (рис. 1), из которых формировался набор данных (датасет).

Для этого брались изображения образцов с шириной горизонтального поля (HFW) от 30 до 300 мкм при ширине изображения 1536 пикселей от разных моделей конденсаторов. Из них вырезались фрагменты таким образом, что при подаче на вход нейросети, размером 256×256 , размер пикселя был равен 98 нм. Например, из изображений с HFW 30, 150 и 300 мкм вырезались фрагменты 1280×1280 , 256×256 и 128×128 пикселей, соответственно, размер которых в дальнейшем приводился к размеру входа нейросети (фрагменты большего размера сжимались, а меньшего растягивались). Эта операция выполнялась для

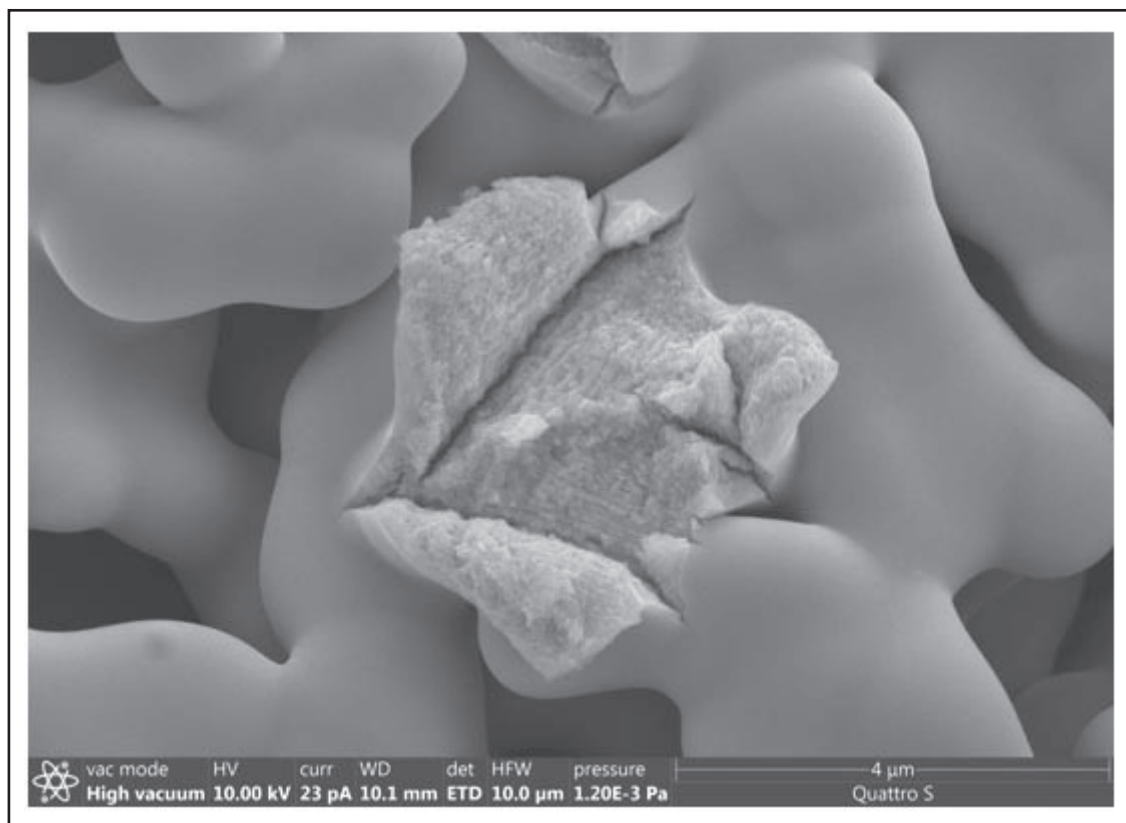


Рис. 1. Дефект оксидного слоя, образовавшийся в результате роста кристаллической структуры

повышения обобщающей способности модели, поскольку на её вход подавались изображения с одинаковым увеличением, но различным разрешением, а также для экономии памяти видеопроцессора (Nvidia RTX 3060Ti). Кроме этого, данный подход позволит повысить стабильность результатов обработки экспериментальных данных, поскольку изображения могут быть проанализированы в едином масштабе.

На изображениях могли присутствовать дефекты двух видов: составные, образовавшиеся в результате разрыва оксидного слоя, а также сплошные, образовавшиеся из-за нарушения его структуры (рис. 2). Для каждого изображения с помощью LabelStudio готовили три сегментационные карты: первая содержала кристаллизовавшиеся участки аморфного слоя (оболочка дефекта), вторая – металлические участки, оголившиеся во время разрыва диэлектрика (ядро), а третья – области, не относящиеся к дефектам (фон). В результате каждому пикселю на исходном изображении соответствовал пиксель на одном из сегментов.

Изображения объединялись в массив, который делился на обучающую, валидационную и тестовую выборки: обучающая использовалась для определения параметров модели, валидационная для определения точности модели в процессе обучения, а тестовая для финальной оценки качества. Изображения из обучающего набора дополнительно

подвергались аугментации (повороты на 90, 180 и 270 градусов) для увеличения объёма выборки и улучшения пространственной инвариантности. Для повышения масштабной инвариантности в обучающую выборку дополнительно вводились экспериментально полученные изображения, у которых размер пикселя после масштабирования отличался от остальных (98 нм) как в большую, так и меньшую сторону. После всех преобразований обучающая выборка содержала 1325 (до аугментации 574) изображений, а валидационная и тестовая по 192 изображения.

Результаты и обсуждение

С помощью фреймворка TensorFlow [17] и подготовленного датасета обучалась нейросеть с архитектурой U-Net [18], в которую дополнительно были добавлены слои BatchNormalization и Dropout [19,20], позволяющие повысить скорость обучения и снизить тенденцию к переобучению (рис. 3), суммарное количество обучаемых параметров было на уровне $7,8 \cdot 10^6$. Во время обучения нейросети изображение дефекта подаётся на вход, обрабатывается набором свёрток, преобразуется с помощью нелинейной функции ReLU [21] (выражение 1), в результате получается тензор первичных признаков (например, это могут быть линии разного наклона).

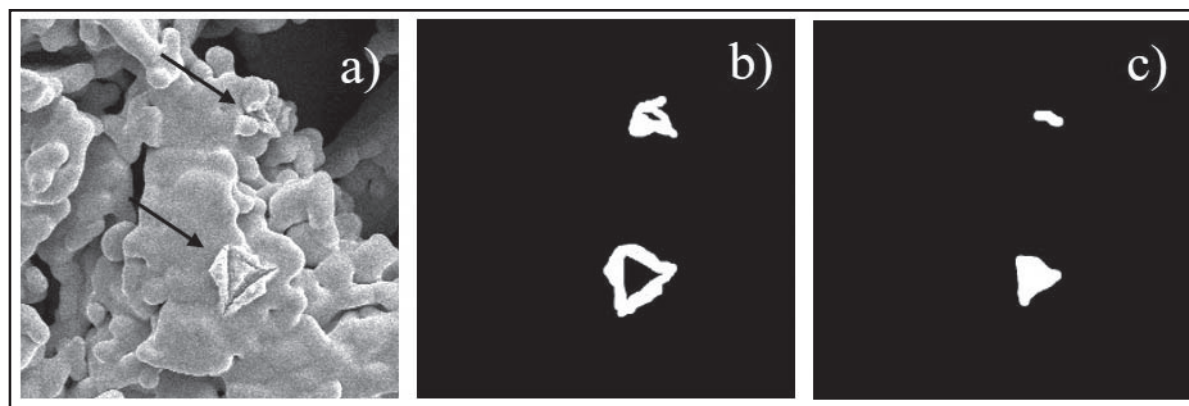


Рис. 2. Изображение дефектов (a), сегменты, содержащие оболочки (b) и ядра (c)

$$f(x) = \max(0, x). \quad (1)$$

После этого, понижается длина и ширина тензора выделенных признаков, снова обрабатывается набором свёрток при этом его глубина и сложность признаков увеличиваются (могут определяться не только наклонные линии, но и границы объектов). В результате повторения этих операций (снижение длины и ширины, повышение глубины) формируется тензор признаков высокого порядка, из которого также с помощью свёрток и повышения разрешения восстанавливается изображение сегментов. Характерной особенностью U-Net является добавление тензоров, полученных в процессе снижения разрешения (левая часть нейросети), к тензорам, образовавшимся в процессе его повышения (правая часть). Эта процедура

нужна для того, чтобы модель могла улавливать мелкие детали в структуре дефектов и как следствие точнее декодировать изображение.

На последнем шаге формируется тензор с параметрами длины и ширины как у входного изображения, но содержащий 3 канала в глубину по одному на каждый тип дефектов (оболочка, ядро, фон). Далее, его значения с помощью многопеременной логистической функции (Softmax) пересчитываются в вероятность образования каждого типа дефекта (выражение 2), где x_i – значение, полученное после свёртки и активации для каждого из C каналов, \hat{y}_i – предсказание модели.

$$\hat{y}_i = \frac{e^{x_i}}{\sum_{i=1}^C e^{x_i}}. \quad (2)$$

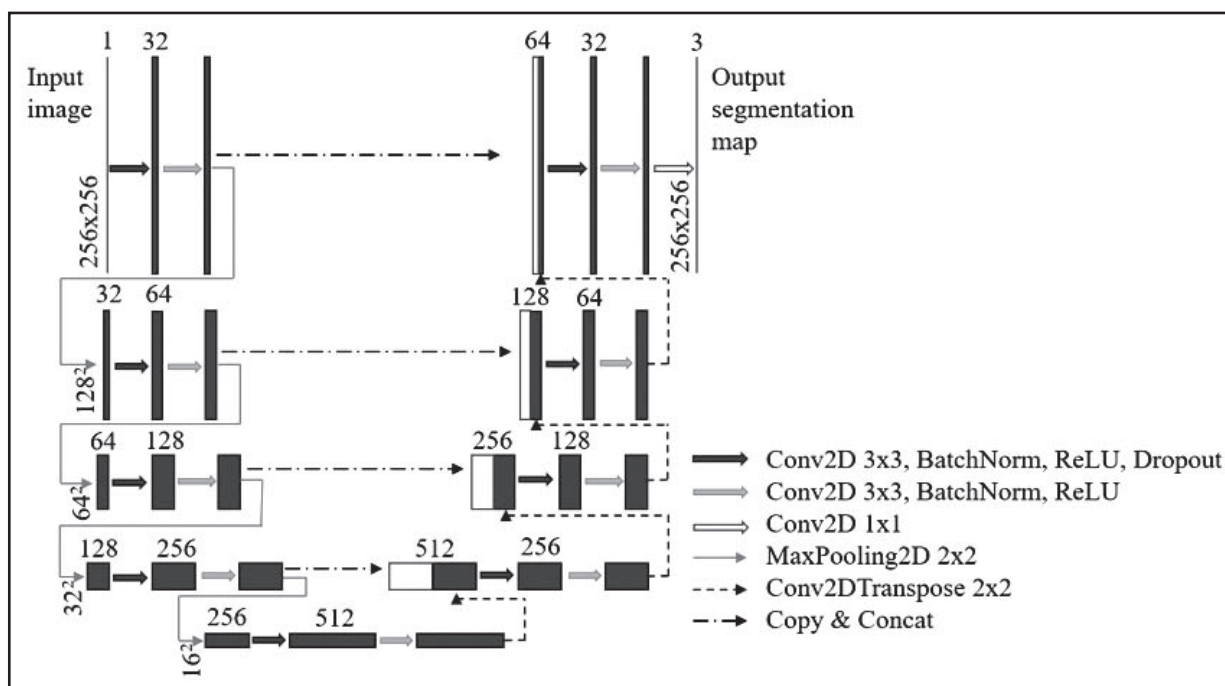


Рис. 3. Схема нейросети и основные операции

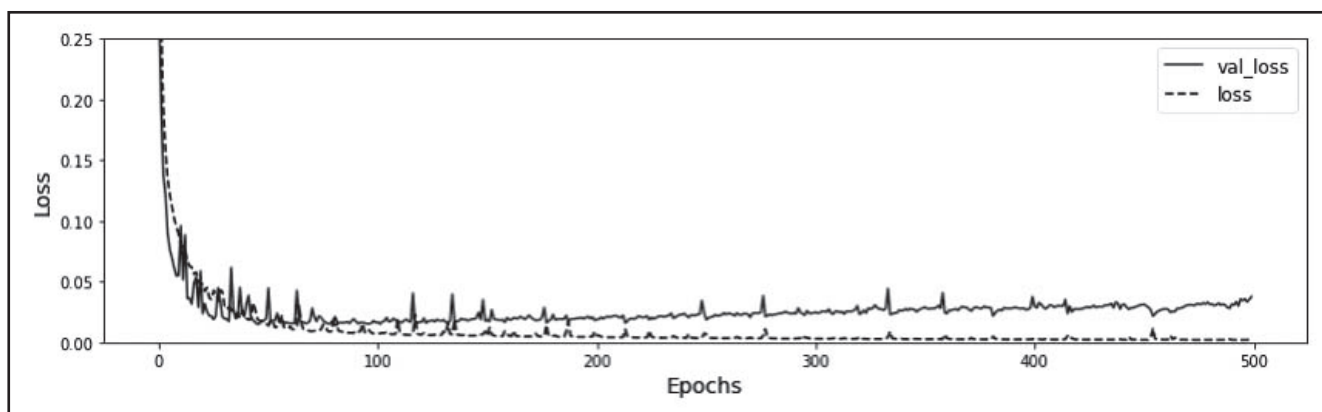


Рис. 4. Изменение величины функции потерь в процессе обучения для тренировочной (loss) и валидационной (val_loss) выборок

Данный тензор сравнивается с сегментационной маской и вычисляется значение функции потерь, которое итерационно минимизируется в результате подгонки параметров сети (коэффициенты свёрток) методом градиентного спуска. Таким образом нейросеть обучается находить дефекты оксидного слоя.

В процессе обучения (500 эпох) минимизировалась кросс-энтропийная функция потерь (выражение 3), где y_i и \hat{y}_i – истинное и предсказанное значения, N – количество пикселей в канале выходного тензора, C – количество каналов.

$$L(y, \hat{y}) = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C y_{i,j} \log \hat{y}_{i,j}. \quad (3)$$

Процесс осуществлялся с помощью модифицированного градиентного спуска Adam [22], учитывающего значения градиентов, вычисленных на предыдущих итерациях, скорость обучения равнялась 0,0001, а размер пакета – 4. Были использованы малые значения пакета и скорости обучения для того, чтобы с одной стороны снизить влияние дисбаланса классов, а с другой избежать предельно быстрого роста величины градиентов (взрыв градиентов).

На рис. 4 можно видеть, что модель имеет небольшую склонность к переобучению (снижение величины функции потерь на тренировочной выборке и рост на валидационной), эта особенность может быть связана с относительно небольшим размером обучающей выборки (574 изображений до аугментации) и значительным преобладанием фона над дефектами. По этой причине, в качестве итоговых параметров модели выбирались, соответствующие минимуму на валидационной выборке (в районе 220 эпохи), т.е. до начала переобучения.

После обучения точность работы модели проверялась на тестовой выборке. Были посчитаны значения метрики, равной отношению области пересечения и области объединения истинных и предсказанных дефектов: $IoU = \text{true} \cap \text{pred} / \text{true} \cup \text{pred}$. Значение IoU для оболочек, ядер дефектов и фона было 0,80, 0,79, 0,83, соответственно. В среднем значение этой метрики для всех трех типов дефектов на тестовой выборке соответствовало 0,81, что является хорошим результатом. На рис. 5 представлено предсказание обученной сети на изображении из тестового набора данных. Можно видеть, что дефекты,

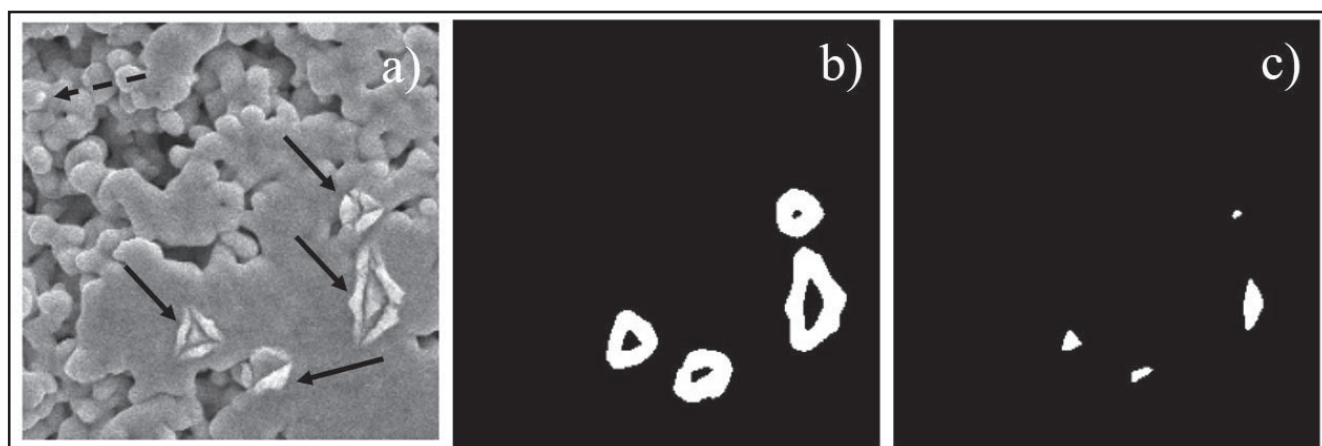


Рис. 5. Изображение дефектов (a), предсказанные моделью оболочки (b) и ядра (c). Сплошными стрелками указаны корректно обнаруженные дефекты, штриховой – ложноотрицательное срабатывание

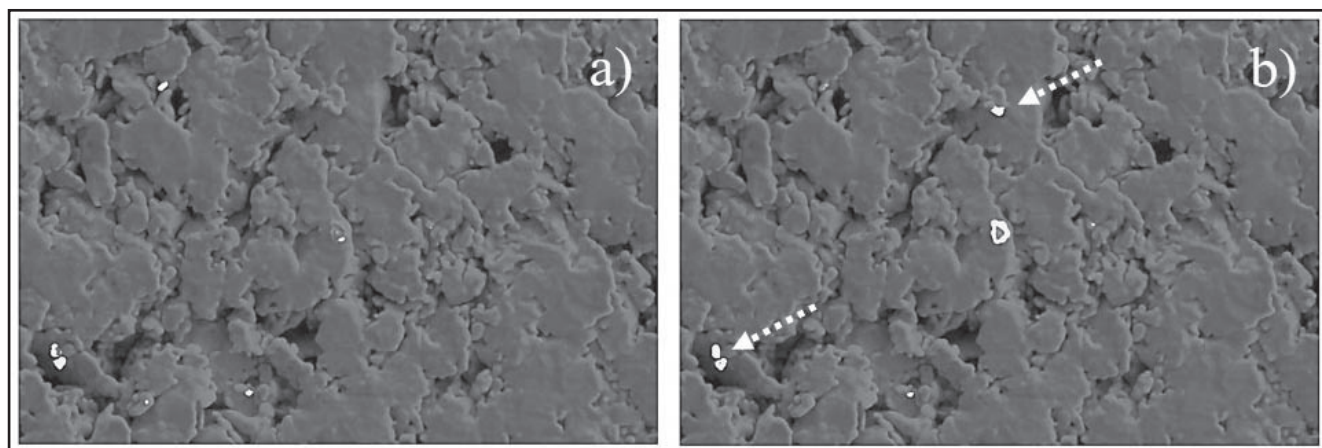


Рис. 6. Обнаружение дефекта в образце с шагом сканирования 256 (a) и 32 (b) пикселя. Пунктирными стрелками указаны ложноположительные срабатывания

отмеченные сплошными стрелками, определены довольно точно, выделены их оболочки и ядра.

Изображения, полученные после испытаний конденсаторов, имели размер, который в несколько раз превышал размер входа обученной нейросети: 1024×1536 и 256×256 , соответственно. По этой причине они анализировались в режиме скользящего окна, соответствующего размеру входа модели. Для этого происходило накопление сигнала для каждого из сегментов изображения, определение сегмента с максимальным значением, присвоение ему единицы и нулей оставшимся двум. Таким образом пиксель на изображении соотносился с типом сегмента. На рис. 6 можно видеть, как изменяется качество обнаружения при уменьшении шага сканирования. Видно, что уменьшение шага приводит к более точному определению дефекта. Это связано с тем, что при использовании большого шага, в поле действия модели может попасть только небольшая часть дефекта, которая окажется проигнорирована. При уменьшении шага

сканирования увеличивается вероятность захвата дефекта целиком (или большей его части), что приводит к более точному определению. Кроме этого, многократное обнаружение дефекта приводит к накоплению высокого уровня сигнала, что повышает уверенность отнесения его к правильному классу. На рис. 7 можно видеть работу модели (с шагом сканирования 32 пикселя) на изображении анода, где были найдены практически все дефекты, определена их суммарная площадь в пикселях: оболочка – 9300, ядро – 303.

Отдельно нужно отметить наличие на рис. 5-7 ложноположительных (пунктирные стрелки) и ложноотрицательных (штриховые стрелки) срабатываний. Ложноположительным является выделение в качестве дефекта области фона (не содержащей дефект), а ложноотрицательным – наоборот выделение в качестве фона дефектной области. К ложноположительным были отнесены крупные сгустки анодированного материала на границах порошковых агломератов. По-видимому, эти области воспринимаются

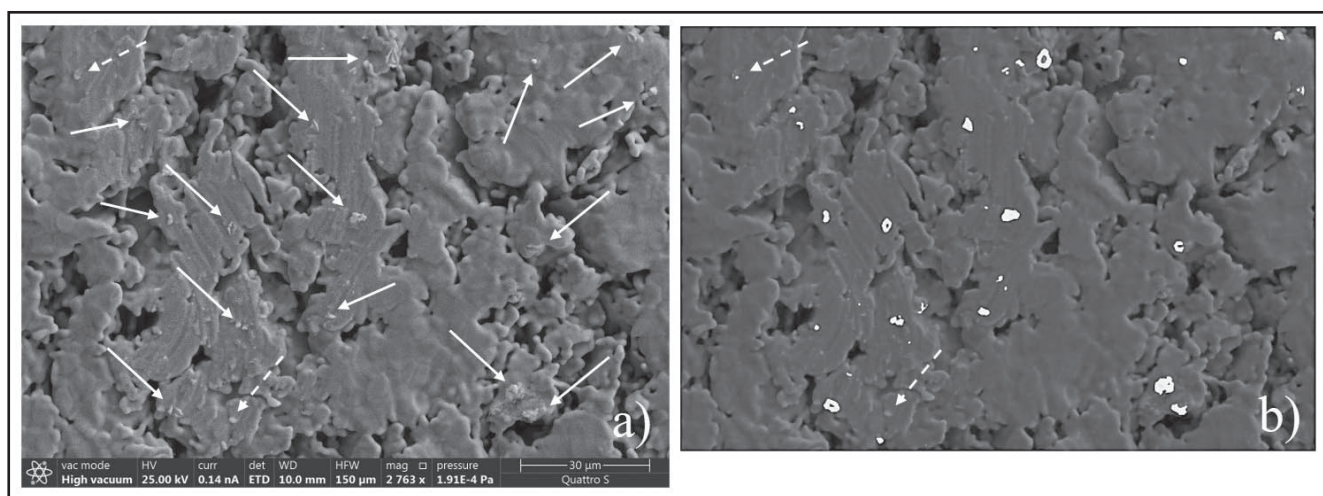


Рис. 7. Изображение поверхности анода (a) и сегментационной карты дефектов (b). Сплошными стрелками указаны корректно обнаруженные дефекты, штриховыми – ложноотрицательные срабатывания

моделью как дефекты из-за схожей с дефектами выпуклой формы и повышенной яркости, которая повышается при отклонении от среднего уровня поверхности. В качестве ложноотрицательных срабатываний были выделены каплевидные дефекты малого размера, которые могут восприниматься как края анодированных агломератов. Можно сказать, что с низкой вероятностью происходит ошибочное разделение выступающих участков на границах и дефектов, форма которых схожа с подобными участками, из-за слабой представленности данных примеров в обучающем наборе.

Заключение

Дефекты диэлектрического слоя танталовых оксидно-электролитических конденсаторов играют критическую роль при их производстве и эксплуатации. Автоматическое определение дефектов является актуальной задачей и позволит ускорить процесс проведения испытаний, а также может быть применено в серийном производстве для оценки влияния изменения технологических режимов и материалов. В работе подготовлен набор данных и обучена свёрточная нейронная сеть с архитектурой U-Net, которая позволяет определять дефекты оксидного слоя со средним значением метрики IoU равным 81%. Режим скользящего окна позволил анализировать изображения, размер которых гораздо больше тех, что содержатся в обучающем наборе данных. Точность работы модели в таком случае может быть повышена за счёт уменьшения шага сканирования и накопления интенсивности для каждого из сегментов изображения. Показано, что существует вероятность некорректного определения выступающих участков на границах анодированных агломератов и дефектов, форма которых схожа с подобными участками.

Работа выполнена в рамках темы государственного задания Минобрнауки РФ (проект № ББ_2021_121030100003-7). При выполнении работы использовалось оборудование ЦКП “Центр физических и физико-химических методов анализа, исследования свойств и характеристик поверхности, наноструктур, материалов и изделий” Уд-мФИЦ УрО РАН.

Список литературы (References)

1. Khanin S.D. Structure inhomogeneities of the oxide dielectric and the properties of tantalum capacitors // *Materials Science Forum*. 1995. Vol. 185. Pp. 573–580.
2. Freeman Y. Tantalum and niobium-based capacitors // *Springer International Publishing AG*, 2018. 120 p. <https://doi.org/10.1007/978-3-319-67870-2>
3. Horacek I., Zednicek T., Zednicek S. et al. High CV tantalum capacitors—Challenges and limitations // *AVX Czech Republic sro*, 2008. Pp. 11.
4. Freeman Y., Hahn R., Lessner P. et al. Reliability and critical applications of tantalum capacitors // *CARTS-Conference. – Components Technology Institute INC*. 2007. Vol. 27. Pp. 193–204.
5. Wang J., Zhao S., Xu C. et al. Brain-inspired interpretable network pruning for smart vision-based defect detection equipment // *IEEE Transactions on Industrial Informatics*. 2022. Vol. 19(2). Pp. 1666–1673. <https://doi.org/10.1109/TII.2022.3188349>
6. Wang Y., Sun Y., Lv P. et al. Detection of line weld defects based on multiple thresholds and support vector machine // *Ndt & E International*. 2008. Vol. 41(7). Pp. 517–524. <https://doi.org/10.1016/j.ndteint.2008.05.004>
7. Wang Y., Xia H., Yuan X. et al. Distributed defect recognition on steel surfaces using an improved random forest algorithm with optimal multi-feature-set fusion // *Multimedia Tools and Applications*. 2018. Vol. 77. Pp. 16741–16770. <https://doi.org/10.1007/s11042-017-5238-0>
8. Wang T., Chen Y., Qiao M. et al. A fast and robust convolutional neural network-based defect detection model in product quality control // *The International Journal of Advanced Manufacturing Technology*. 2018. Vol. 94. Pp. 3465–3471. <https://doi.org/10.1007/s00170-017-0882-0>
9. Huang Y., Qiu C., Yuan K. Surface defect saliency of magnetic tile // *The Visual Computer*. 2020. Vol. 36. Pp. 85–96. <https://doi.org/10.1007/s00371-018-1588-5>
10. Hao R., Lu B., Cheng Y. et al. A steel surface defect inspection approach towards smart industrial monitoring // *Journal of Intelligent Manufacturing*. 2021. Vol. 32. Pp. 1833–1843. <https://doi.org/10.1007/s10845-020-01670-2>
11. Sun TH., Tien FC., Tien FC. et al. Automated thermal fuse inspection using machine vision and artificial neural networks // *Journal of Intelligent Manufacturing*. 2016. Vol. 27. Pp. 639–651. <https://doi.org/10.1007/s10845-014-0902-y>
12. Badmos O., Kopp A., Bernthaler T. et al. Image-based defect detection in lithium-ion battery electrode using convolutional neural networks // *Journal of Intelligent Manufacturing*. 2020. Vol. 31. Pp. 885–897. <https://doi.org/10.1007/s10845-019-01484-x>
13. Davtalab O., Kazemian A., Yuan X. et al. Automated inspection in robotic additive manufacturing using deep learning for layer deformation detection // *Journal of Intelligent Manufacturing*. 2022. Vol. 33(3). Pp. 771–784. <https://doi.org/10.1007/s10845-020-01684-w>

14. Grzenda M., Bustillo A. Semi-supervised roughness prediction with partly unlabeled vibration data streams // *Journal of Intelligent Manufacturing*. 2019. Vol. 30(2). Pp. 933–945. <https://doi.org/10.1007/s10845-018-1413-z>
15. An K., Zhang Y. LPViT: A transformer-based model for PCB image classification and defect detection // *IEEE Access*. 2022. Vol. 10. Pp. 42542–42553. <https://ieeexplore.ieee.org/document/9760408>
16. Chuang C.H., Hou K.W., Wu C.W. et al. A deep learning-based screening method for improving the quality and reliability of integrated passive devices // *IEEE International Test Conference (ITC)*, 2020. Pp. 1–9. <https://ieeexplore.ieee.org/document/9325221>
17. Abadi M., Barham P., Chen J. et al. Tensorflow: a system for large-scale machine learning // *OSDI*. 2016. Vol. 16. Pp. 265–283.
18. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // *MICCAI*, 2015. Pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
19. Ioffe S., Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift // *International conference on machine learning*, PMLR, 2015. Pp. 448–456.
20. Srivastava N., Hinton G, Krizhevsky A. et al. Dropout: a simple way to prevent neural networks from overfitting // *The journal of machine learning research*. 2014. Vol. 15(1). Pp. 1929–1958.
21. Agarap A.F. Deep learning using rectified linear units (relu) // *arXiv preprint arXiv:1803.08375*, 2018.
22. Kingma D.P., Ba J. Adam: A method for stochastic optimization // *arXiv preprint arXiv:1412.6980*, 2014.

Сведения об авторах

¹ Елькин Иван Александрович, канд. физ.-мат. наук, научный сотрудник

¹ Чукавин Андрей Игоревич, канд. физ.-мат. наук, старший научный сотрудник

^{1,2} Барышев Олег Борисович, аспирант

¹ Удмуртский Федеральный Исследовательский Центр УрО РАН
426067, Российская Федерация, Ижевск, ул. Т. Барамзиной, 34

² АО «Электонд»

427968, Российская Федерация, Сарапул, ул. Калинина, 3

Information about authors

¹ Elkin Ivan Aleksandrovich, Cand. of Phys.-Math. Sciences, Researcher

¹ Chukavin Andrey Igorevich, Cand. of Phys.-Math. Sciences, Senior Researcher

^{1,2} Baryshev Oleg Borisovich, postgraduate student

¹ Udmurt Federal Research Center Ural Branch of Russian Academy of Sciences
426067, Russian Federation, Izhevsk, st. T. Baramzina, 34

² JSC Elecond

427968, Russian Federation, Sarapul, Kalinina st., 3