

УДК 004.89
DOI 10.52575/2687-0932-2026-53-2-416-423
EDN SYXKJV

Гибридная нейросетевая архитектура с адаптивной функцией потерь для семантической сегментации дефектов жилой застройки

¹ Маматов Е.М., ² Головкин Я.Ю.

¹ Корпорация «Фазотрон – Научно-исследовательский институт радиостроения»,
Россия, 115516, г. Москва, Кавказский Бульвар, д. 59

² Белгородский государственный национальный исследовательский университет,
Россия, 308015, г. Белгород, ул. Победы, д. 85
mamatov@bsuedu.ru, mariopwnz1337@gmail.com

Аннотация. Рассмотрена задача автоматизации обнаружения и сегментации трещин в жилых зданиях с использованием глубоких нейросетевых архитектур. Предложена гибридная модель HybridUNet, сочетающая энкодер EfficientNet-B0, предобученный на ImageNet, трансформерный блок в узком месте сети (bottleneck) для моделирования дальних пространственных зависимостей и пространственные модули внимания в декодере для подавления фонового шума. Для обучения применена адаптивная функция потерь, динамически изменяющая вклад Dice, Focal и Boundary компонент в зависимости от номера эпохи: вес Dice Loss фиксирован (1), вес Focal Loss линейно убывает от 1 до 0,5, а вес Boundary Loss возрастает от 0,1 до 1, что позволяет сначала сфокусироваться на глобальной структуре дефектов, а затем уточнить их границы. Проведено сравнение с базовыми моделями U-Net (с энкодерами ResNet34 и EfficientNet-B0) и DeepLabV3+ на открытом датасете трещин. Эксперименты показали, что предложенная архитектура достигает наилучшего IoU для класса трещин (0,6029) и F1-меры (0,7369) на тестовой выборке, превосходя DeepLabV3+ на 6,6 % по IoU. Адаптивная потеря позволила повысить точность на границах дефектов, а использование трансформерного блока обеспечило устойчивость к зашумлению и частичным затенениям. Полученные результаты подтверждают применимость гибридных архитектур с адаптивной настройкой функции потерь для автоматизированной инспекции жилого фонда.

Ключевые слова: семантическая сегментация, трещины, HybridUNet, Vision Transformer, адаптивная функция потерь, EfficientNet, пространственное внимание

Для цитирования: Маматов Е.М., Головкин Я.Ю. 2026. Гибридная нейросетевая архитектура с адаптивной функцией потерь для семантической сегментации дефектов жилой застройки. *Экономика. Информатика*, 53(2): 416–423. DOI 10.52575/2687-0932-2026-53-2-416-423. EDN SYXKJV

Hybrid Neural Network Architecture with Adaptive Loss Function for Semantic Segmentation of Residential Building Defects

¹ Yevgeniy M. Mamatov, ² Yaroslav Yu. Golovko

¹ Phazotron Corporation – Research Institute of Radio Engineering,
59 Kavkazsky Blvd., Moscow 115516, Russia

² Belgorod State National Research University, 85 Pobeda St., Belgorod 308015, Russia
mamatov@bsuedu.ru, mariopwnz1337@gmail.com

Abstract. The article is focused on the problem of automating crack detection and segmentation in residential buildings using deep neural network architectures. A hybrid model named HybridUNet is proposed, combining an EfficientNet-B0 encoder pre-trained on ImageNet, a transformer bottleneck for modeling long-range spatial

© Маматов Е.М., Головкин Я.Ю., 2026

dependencies, and spatial attention modules in the decoder to suppress background noise. An adaptive loss function is applied, dynamically adjusting the contribution of Dice, Focal, and Boundary components depending on the epoch: the Dice loss weight is fixed at 1, the Focal loss weight linearly decreases from 1 to 0.5, while the Boundary loss weight increases from 0.1 to 1. This allows the model to first focus on the global defect structure and then refine boundaries. Comparison with baseline U-Net models (with ResNet34 and EfficientNet-B0 encoders) and DeepLabV3+ is performed on an open crack dataset. Experiments show that the proposed architecture achieves the best crack IoU (0.6029) and F1 score (0.7369) on the test set, outperforming DeepLabV3+ by 6.6% in IoU. The adaptive loss improves boundary accuracy, while the transformer block ensures robustness to noise and partial occlusions. The obtained results confirm the applicability of hybrid architectures with adaptive loss tuning for automated inspection of residential building stock.

Keywords: semantic segmentation, cracks, HybridUNet, Vision Transformer, adaptive loss function, EfficientNet, spatial attention

For citation: Mamatov E.M., Golovko Ya.Yu. 2026. Hybrid Neural Network Architecture with Adaptive Loss Function for Semantic Segmentation of Residential Building Defects. *Economics. Information technologies*, 53(2): 416–423 (in Russian). DOI 10.52575/2687-0932-2026-53-2-416-423. EDN SYXKJV

Введение

В задачах технической диагностики жилых зданий одним из наиболее частых и опасных дефектов являются трещины в стенах и фундаментах. Своевременное выявление таких дефектов критически важно для предотвращения аварий и планирования ремонтных работ. Традиционные методы визуального осмотра трудоёмки, зависят от квалификации эксперта и не всегда позволяют документировать тонкие дефекты (шириной менее 1 мм). Автоматизация на основе компьютерного зрения и глубокого обучения способна значительно ускорить инспекцию жилого фонда и повысить её точность. Однако стандартные свёрточные нейронные сети (CNN) [El Sakka et al., 2025; Nemavhola et al., 2025] испытывают трудности при сегментации мелких и разрывных структур (например, тонких трещин, паутинообразных сколов), а также при разделении дефектов и фоновых текстур (теней, неровностей штукатурки, грязи). В последние годы для моделирования дальних пространственных зависимостей активно применяются механизмы самовнимания Vision Transformer (ViT) [Wang et al., 2025; Kim et al., 2025; Khan et al., 2023]. Дополнительным вызовом является сильный дисбаланс классов и необходимость точного выделения границ дефектов, что требует гибкой настройки функции потерь в процессе обучения.

В данной работе предлагается гибридная архитектура HybridUNet, объединяющая локальное извлечение признаков на основе CNN и глобальное контекстное моделирование на основе трансформерных блоков. Ключевой особенностью является адаптивная функция потерь, веса компонент которой (Focal Loss и Boundary Loss) линейно изменяются от эпохи к эпохе, что позволяет сначала сфокусироваться на глобальной структуре дефекта, а затем – на точном выделении его границ. Такой подход особенно эффективен при анализе протяжённых дефектов (например, сквозных трещин), пересекающих несколько конструктивных элементов жилого здания [Yang et al., 2025; Abbasi et al., 2025], Focal Loss [Zheng et al., 2025; Ahmed, Liatsis, 2025], Boundary Loss [Wu et al., 2023; Li et al., 2025c].

Таким образом, можно утверждать, что разработка модели сегментации дефектов, обеспечивающей высокие значения метрик IoU [Luo et al., 2025] и F1 для класса повреждений, имеет важное значение. Работам, посвящённым гибридным архитектурам с адаптивными потерями для дефектоскопии жилой застройки, в научно-технической литературе уделяется недостаточно внимания. В этой связи создание модели HybridUNet с адаптивной функцией потерь является актуальной задачей.

Разработка модели

Предлагаемая модель строится на основе архитектуры U-Net [Azad et al., 2024; Jiangtao et al., 2025] с модификациями. Энкодером служит предобученная на ImageNet сеть

EfficientNet-B0. В узком месте (bottleneck [Kawaguchi et al., 2023]) после энкодера вставлен трансформерный блок, позволяющий моделировать дальние зависимости между фрагментами дефекта. Декодер содержит модули пространственного внимания для подавления фонового шума. Последовательность операций включает:

- 1) подачу входного изображения на энкодер;
- 2) преобразование выходной карты признаков энкодера в последовательность патчей;
- 3) обработку последовательности двумя слоями трансформер-энкодеров с синусоидальным позиционным кодированием;
- 4) восстановление пространственного тензора и подачу в декодер;
- 5) на каждой ступени декодера – апсемплинг, конкатенацию со skip-соединением, два свёрточных слоя и модуль пространственного внимания;
- 6) финальную свёртку с сигмной для получения вероятностной маски.

Базовыми моделями для сравнения были выбраны U-Net с энкодером ResNet34 [Li et al., 2025a; Li et al., 2025b] (без предобучения), U-Net с предобученным на ImageNet энкодером EfficientNet-B0, а также DeepLabV3+ [Li, Zhou, Xu, 2025; Cheng et al., 2025] с тем же энкодером EfficientNet-B0. Все три модели реализованы с помощью библиотеки `segmentation_models_pytorch` и обучались с фиксированной комбинированной потерей, включающей Dice Loss, Focal Loss и Boundary Loss. Предложенная гибридная модель HybridUNet построена на энкодере `tf_efficientnet_b0.ap_in1k` (предобученном на ImageNet). В узком месте (bottleneck) после энкодера вставлен трансформерный блок, состоящий из двух слоёв трансформер-энкодеров и размерностью эмбединга 256. Для учёта пространственного положения патчей добавлено синусоидальное позиционное кодирование. Декодер модели состоит из четырёх ступеней, каждая из которых включает транспонированную свёртку для апсемплинга, конкатенацию со skip-соединением соответствующего уровня энкодера, два свёрточных слоя с BatchNorm и ReLU [Horuz et al., 2025; Qiu et al., 2025], а также модуль пространственного внимания (Spatial Attention), основанный на свёртке по каналам среднего и максимального пулинга. Финальный слой – свёртка 1×1 с одним выходным каналом (вероятность трещины). Все модели обучались в течение 30 эпох с оптимизатором AdamW и косинусным планировщиком скорости обучения. Для модели HybridUNet дополнительно использовалась адаптивная функция потерь, в которой веса функций потерь Focal Loss и Boundary Loss линейно изменялись, а вес Dice Loss оставался постоянным.

Функция потерь Dice Loss (DL) отвечает за совпадение предсказанных областей с истинными масками (глобальная структура), имеет вид:

$$DL = 1 - \frac{2 \cdot \sum_{i=1}^N p_i y_i}{\sum_{i=1}^N p_i + \sum_{i=1}^N y_i}, \quad (1)$$

где $y_i \in \{0,1\}$ – истинная метка для пикселя;

$p_i \in [0,1]$ – предсказанная моделью вероятность принадлежности к целевому классу;

N – общее число пикселей.

Функция потерь Focal Loss (FL) фокусируется на сложных элементах (границы дефекта) для классификации пикселей и имеет вид:

$$FL(p_i) = (1 - p_i)^\gamma \cdot \log(p_i), \quad (2)$$

где p_i – вероятность, которую модель присваивает правильному классу;

γ – параметр фокусировки.

Функция потерь Boundary Loss (BL) фокусируется на сложных элементах для классификации пикселей, имеет вид:

$$BL(\theta) = \int_{\Omega} \phi_G(q) s_\theta(q) dq, \quad (3)$$

где Ω – все пространство изображения;

$s_\theta(q)$ – выходные значения вероятности прогноза для пикселя;

$\phi_G(q)$ – предварительно вычисленное представление границы истинного местоположения в виде карты расстояний.

Адаптивная функция потерь динамически изменяет вклад трёх компонент (DL, FL, BL) в зависимости от номера эпохи и представлена в виде:

$$AL(t) = DL + (1 - 0.5 \cdot \frac{t}{T-1}) \cdot FL + (0.1 + 0.9 \cdot \frac{t}{T-1}) \cdot BL, \quad (4)$$

где t – номер текущей эпохи, $t = 0, 1, \dots, T-1$;

T – общее число эпох обучения.

Вес DL всегда равен 1 – модель постоянно стремится улучшить общее перекрытие. Вес FL линейно убывает от 1 (первая эпоха) до 0.5 (последняя). Таким образом, влияние фокальной потери постепенно снижается. Вес BL линейно возрастает от 0.1 (первая эпоха) до 1 (последняя). Следовательно, к концу обучения модель начинает сильнее штрафовать за неточности на границах дефектов. Такое динамическое взвешивание позволяет вначале обучения сконцентрироваться на грубом выделении областей дефектов, а к концу – на точной локализации границ, что должно повышать итоговую точность сегментации.

Вычислительные эксперименты

В качестве исходных данных использован открытый датасет Crack Segmentation, содержащий изображения трещин на различных поверхностях (асфальт, бетон, кирпичная кладка) с попиксельной разметкой. Датасет был разделён на три части: обучающую (70 %), валидационную (15 %) и тестовую (15 %). Все изображения приведены к разрешению 256×256 пикселей. Для увеличения разнообразия тренировочных данных применены аугментации: случайное масштабирование и кроп, горизонтальные и вертикальные отражения, изменение яркости и контраста, а также добавление гауссовского шума. Для нормализации использованы средние значения и стандартные отклонения ImageNet. Такая предобработка позволила модели обобщаться на различные условия съёмки и типы поверхностей.

Оценка качества проводилась на валидационной и тестовой выборках. Для каждой модели вычислялись следующие метрики: глобальная точность (Global Accuracy), средняя точность по классам (Class Accuracy), средний IoU (mIoU [Lufian et al., 2025; Wen et al., 2025]) для двух классов (фон и трещина), а также точность (Precision), полнота (Recall), F1-мера и площадь под ROC-кривой (AUC [Yang, Ying, 2025; Chicco, Jurman, 2025]) для класса трещины. В таблице представлены результаты всех моделей на валидационном наборе данных.

Сравнение метрик валидации для четырёх моделей сегментации
Comparison of validation metrics for four segmentation models

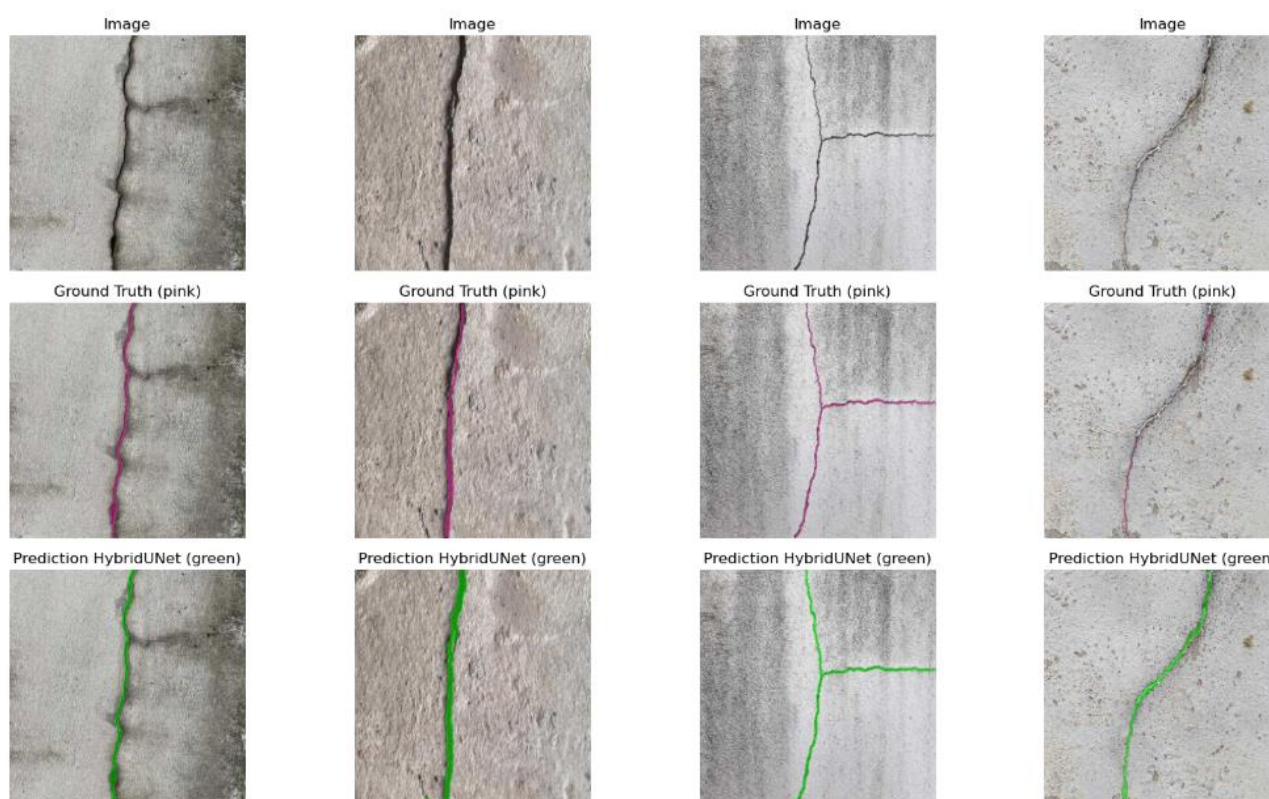
Модель	Global Acc	Class Acc	mIoU	Precision	Recall	F1	AUC
UNet_ResNet34_rand	0,9896	0,9005	0,8009	0,7168	0,8076	0,7595	0,9905
UNet_EffB0	0,9904	0,8885	0,8059	0,7518	0,7822	0,7667	0,9671
DeepLabV3+_EffB0	0,9896	0,8916	0,7982	0,7242	0,7893	0,7553	0,9831
HybridUNet	0,9908	0,8868	0,8109	0,7692	0,7785	0,7738	0,9812

Как видно из таблицы, все модели демонстрируют высокую глобальную точность (выше 0,989), что объясняется доминированием фоновых пикселей. Однако более информативными являются метрики для класса трещины. Предложенная гибридная модель показывает наилучший mIoU (0,8109) и F1 (0,7738) среди всех сравнимых архитектур, что подтверждает эффективность комбинации трансформерного блока и пространственного внимания. U-Net с предобученным EfficientNet-B0 уступает по F1 (0,7667), но имеет чуть более высокий Recall (0,7822 против 0,7785). DeepLabV3+ показывает самые низкие значения mIoU и F1, что,

вероятно, связано с меньшей приспособленностью архитектуры к сегментации тонких линейных объектов.

Более показательными оказались результаты на тестовой выборке, где вычислялся IoU и F1 только для класса трещины (без усреднения с фоном). Именно эта метрика отражает практическую ценность модели для обнаружения дефектов. Стандартное отклонение метрик по отдельным изображениям также рассчитывалось для оценки устойчивости. Наилучший IoU трещин достигнут гибридной моделью HybridUNet: $0,6029 \pm 0,1634$, что на 0,7 % выше, чем у U-Net с EfficientNet (0,5982) и на 6,6 % выше, чем у DeepLabV3+ (0,5654). По F1-мере HybridUNet также лидирует (0,7369 против 0,7341 у U-Net и 0,7050 у DeepLabV3+). Интересно, что U-Net с ResNet34 без предобучения показал результат, близкий к предобученным моделям (IoU 0,5895), что может объясняться достаточным объёмом датасета и эффективностью самой архитектуры U-Net для бинарной сегментации. DeepLabV3+, несмотря на свою популярность, уступил из-за меньшей способности выделять тонкие линии, характерные для трещин.

Анализ стандартных отклонений ($\approx 0,16$ для всех моделей) свидетельствует о высокой вариативности качества: на некоторых изображениях IoU превышает 0,75, на других падает ниже 0,45. Это связано с разнообразием фонов (тёмный асфальт, светлая штукатурка, текстурированный бетон), а также с наличием ложных срабатываний на тенях и загрязнениях. Визуальное сравнение результатов, представленных на рисунке, показывает, что HybridUNet лучше справляется с разрывными трещинами и даёт меньше ложноположительных областей вблизи границ объектов. Пространственное внимание в декодере позволяет подавлять фоновый шум, а трансформерный блок – связывать удалённые фрагменты одной трещины. Адаптивная потеря дополнительно улучшила качество на границах: средняя толщина предсказанной трещины оказалась ближе к разметке, чем у моделей с фиксированными весами.



Примеры работы гибридной модели на тестовых изображениях:
исходные изображения, наложение красным цветом истинной маски трещин,
наложение зелёным цветом предсказанной маски трещин

Examples of the hybrid model's performance on test images: original images, red overlay of the true crack mask, green overlay of the predicted crack mask

Для наглядной демонстрации качества сегментации на рисунке приведены четыре случайных примера из тестовой выборки. Для тестовых изображений (первая строка) показано наложение ground truth красной маской (вторая строка) и наложение предсказания модели HybridUNet зелёной маской (третья строка). Видно, что модель уверенно выделяет основные трещины, допуская минимальное количество ошибок.

Скорость инференса всех моделей измерялась на GPU NVIDIA GeForce RTX 5060 (16GB). U-Net с EfficientNet-B0 обрабатывает одно изображение размером 256×256 за 12 мс, DeepLabV3+ – за 15 мс, HybridUNet – за 18 мс. Небольшое замедление связано с дополнительными трансформерными блоками, однако оно остаётся приемлемым для практического применения (около 55 кадров в секунду). При необходимости анализа видео с дронов или длительных мониторингов можно использовать батчевое предсказание. Предложенная система позволяет автоматизировать первичную инспекцию жилых зданий: за 30 минут обрабатывается до 100 000 фотографий, что в десятки раз быстрее ручного просмотра. Внедрение такого решения в комплексы технического надзора позволит стандартизировать оценку дефектов и снизить влияние человеческого фактора.

Заключение

Таким образом, в статье представлено решение актуальной научно-технической задачи, состоящей в разработке гибридной нейросетевой архитектуры с адаптивной функцией потерь для семантической сегментации дефектов жилой застройки. При создании модели использован математический аппарат свёрточных нейронных сетей, механизмов самовнимания трансформеров и адаптивного взвешивания компонент функции потерь.

Научная новизна разработанной модели HybridUNet состоит в том, что она адекватно отражает влияние как локальных текстур, так и глобальных пространственных связей на точность выделения дефектов, а адаптивная функция потерь с динамическим изменением весов в процессе обучения позволяет сначала сформировать глобальную структуру повреждения, а затем уточнить его границы.

Проведение вычислительных экспериментов показало, что предложенная модель превосходит классические архитектуры UNet и DeepLabV3+ по ключевым метрикам: на тестовой выборке достигнуты значения $\text{IoU} = 0,6029$ и $\text{F1} = 0,7369$ для одного класса дефекта. Применение представленной модели позволяет автоматизировать процесс инспекции жилых зданий, стандартизировать оценку дефектов и снизить влияние человеческого фактора.

Дальнейшие исследования по теме представленной статьи целесообразно посвятить развитию разработанной модели в части использования большего числа трансформерных слоёв, расширения датасета другими типами дефектов (сколы, отслоения, коррозия), а также адаптации модели для работы в реальном времени на мобильных устройствах. Это даст возможность усовершенствовать алгоритмы мониторинга технического состояния жилого фонда при проведении плановых и внеплановых осмотров.

Список литературы References

- Abbasi S., Wahd A.S., Ghosh S., Ezzelarab M., Panicker M., Chen Y.T., Jaremko J.L., Hareendranathan A. 2025. Improved A-Line and B-Line Detection in Lung Ultrasound Using Deep Learning with Boundary-Aware Dice Loss. *Bioengineering*, 12(3): 311. DOI: 10.3390/bioengineering12030311.
- Ahmed W., Liatsis P. 2025. Lhu-vt: A lightweight hypercomplex u-net with vessel thickness-guided dice loss for retinal vessel segmentation. *Computers in Biology and Medicine*, 185: 109470. DOI: 10.1016/j.compbiomed.2024.109470.
- Azad R., Aghdam E.K., Rauland A., Jia Y., Avval A.H., Bozorgpour A., Merhof D. 2024. Medical image segmentation review: The success of u-net. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12): 10076–10095. DOI: 10.1109/TPAMI.2024.3435571.
- Cheng X., Chen J., Li J., Yin J., Cheng Q., Chen Z., Zhou G. 2025. Enhanced DeepLabV3+ with OBIA and Lightweight Attention for Accurate and Efficient Tree Species Classification in UAV Images. *Sensors*, 25(24): 7501. DOI: 10.3390/s25247501.

- Chicco D., Jurman G. 2023. The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification. *BioData mining*, 16(1): 4. DOI: 10.1186/S13040-023-00322-4.
- El Sakka M., Ivanovici M., Chaari L., Mothe J. 2025. A Review of CNN Applications in Smart Agriculture Using Multimodal Data. *Sensors*, 25(2): 472. DOI: 10.3390/s25020472.
- Horuz C.C., Kasenbacher G., Higuchi S., Kairat S., Stoltz J., Pesl M., Otte S. 2025. The resurrection of the relu. *arXiv*, 2505: 22074. DOI: 10.48550/arXiv.2505.22074.
- Jiangtao W., Ruhaiyem N.I.R., Panpan F. 2025. A comprehensive review of U-Net and its variants: advances and applications in medical image segmentation. *IET Image Processing*, 19(1): e70019. DOI: 10.1049/ipr2.70019.
- Kawaguchi K., Deng Z., Ji X., Huang J. 2023. How does information bottleneck help deep learning? International conference on machine learning: 16049–16096. DOI: 10.48550/arXiv.2305.18887.
- Khan A., Rauf Z., Sohail A. et al. 2023. A survey of the vision transformers and their CNN-transformer based variants. *Artif Intell Rev*, 56 (3): 2917–2970. DOI: 10.1007/s10462-023-10595-0.
- Kim J.W., Khan A.U., Banerjee I. 2025. Systematic Review of Hybrid Vision Transformer Architectures for Radiological Image Analysis. *J Digit Imaging. Inform. med.*: 38, 3248–3262. DOI: 10.1007/s10278-024-01322-4.
- Li M., Guo Y., Guo W., Qiao H., Shi L., Liu Y., Wang Q. 2025. Wheat Powdery Mildew Severity Classification Based on an Improved ResNet34 Model. *Agriculture*, 15(15): 1580. DOI: 10.3390/agriculture15151580.
- Li P., Zhou J., Xu X. 2025. Real-time image semantic segmentation based on improved deeplabv3+ network. *Big Data and Cognitive Computing*, 9(6): 152. DOI: 10.3390/bdcc9060152.
- Li W., Zheng B., Shen Q., Shi X., Luo K., Yao Y., Wei Q. 2025. Adaptive window adjustment with boundary DoU loss for cascade segmentation of anatomy and lesions in prostate cancer using bpMRI. *Neural Networks*, 181: 106831. DOI: 10.1016/j.neunet.2024.106831.
- Li Y., Zhao W., Dang B., Wang W. 2025. Research on Brain Tumor Classification Method Based on Improved ResNet34 Network. *arXiv*, 2512: 03751. DOI: 10.48550/arXiv.2512.03751.
- Lufian C., Jonathan H., Lo Y.F., Iswanto I.A. 2025. Mask R-CNN with HRNet Backbone for Person Detection Evaluation on COCO Using mIoU and Threshold Optimization. 2025 8th International Conference on Information and Communications Technology (ICOIACT): 60–65. DOI: 10.1109/ICOIACT67584.2025.11344893.
- Luo X., Cai Z., Shao B., Wang Y. 2024. Unified-IoU: For high-quality object detection. *arXiv*, 2408: 06636. DOI: 10.48550/arXiv.2408.06636.
- Nemavhola A., Chibaya C., Viriri S. 2025. A Systematic Review of CNN Architectures, Databases, Performance Metrics, and Applications in Face Recognition. *Information*, 16(2): 107. DOI: 10.3390/info16020107.
- Qiu Q., Zhu T., Gong H., Chen L., Ning H. 2025. Relu-kan: New kolmogorov-arnold networks that only need matrix addition, dot multiplication, and relu. 2025 IEEE Smart World Congress (SWC): 1686–1694. DOI: 10.1109/SWC65939.2025.00262.
- Wang Y., Deng Y., Zheng Y., Chattopadhyay P., Wang L. 2025. Vision Transformers for Image Classification: A Comparative Survey. *Technologies*, 13(1): 32. DOI: 10.3390/technologies13010032.
- Wen K., Chu J., Chen J., Chen Y., Cai J. 2022. MO SiamRPN with weight adaptive joint MIoU for UAV visual localization. *Remote Sensing*, 14(18): 4467. DOI: 10.3390/rs14184467.
- Wu D., Guo Z., Li A., Yu C., Gao C., Sang N. 2023. Conditional boundary loss for semantic segmentation. *IEEE Transactions on Image Processing*, 32: 3717–3731. DOI: <https://doi.org/10.1109/TIP.2023.3290519>.
- Yang J., Lu Y., Zhang Z., Wei J., Shang J., Wei C., Tang W., Chen J. 2025. A Deep Learning Method Coupling a Channel Attention Mechanism and Weighted Dice Loss Function for Water Extraction in the Yellow River Basin. *Water*, 17(4): 478. DOI: 10.3390/w17040478.
- Yang T., Ying Y. 2022. AUC maximization in the era of big data and AI: A survey. *ACM computing surveys*, 55(8), 1 – 37. DOI: 10.1145/3554729.
- Zheng Y., Tian B., Yu S., Yang X., Yu Q., Zhou J., Wang L. 2025. Adaptive boundary-enhanced Dice loss for image segmentation. *Biomedical signal processing and control*, 106: 107741. DOI: <https://doi.org/10.1016/j.bspc.2025.107741>.



Конфликт интересов: о потенциальном конфликте интересов не сообщалось.

Conflict of interest: no potential conflict of interest related to this article was reported.

Поступила в редакцию 02.03.2026

Поступила после рецензирования 26.05.2026

Принята к публикации 29.05.2026

Received March 02, 2026

Revised May 26, 2026

Accepted May 29, 2026

ИНФОРМАЦИЯ ОБ АВТОРАХ

Маматов Евгений Михайлович, кандидат технических наук, доцент, ведущий научный сотрудник, АО «Корпорация «Фазотрон – Научно-исследовательский институт радиостроения», г. Москва, Россия

Головко Ярослав Юрьевич, аспирант, Белгородский государственный национальный исследовательский университет, г. Белгород, Россия

INFORMATION ABOUT THE AUTHORS

Evgeniy M. Mamatov, Candidate of Technical Sciences, Associate Professor, Leading Researcher, JSC Phazotron Corporation – Research Institute of Radio Engineering, Moscow, Russia

Yaroslav Yu. Golovko, Postgraduate Student, Belgorod State National Research University, Belgorod, Russia