



А.Е. Аязбай¹, А.М. Узбекбаев¹, Ғ.Қ. Сергазин²

¹Energo University, Алматы, Казахстан

²Mukhametzhan Tynyshbayev ALT University, Алматы, Казахстан

E-mail: niipntkz@gmail.com

СБОР И ПРЕДОБРАБОТКА ДАННЫХ С SWIR-КАМЕР ДЛЯ ОБУЧЕНИЯ НЕЙРОННОЙ СЕТИ ДЛЯ РАСПОЗНАВАНИЯ ОБЪЕКТОВ

Аннотация. В статье рассматривается задача сбора и предварительной обработки данных, получаемых с SWIR-камер, для последующего обучения нейронной сети распознаванию объектов в условиях ограниченной видимости. Актуальность исследования обусловлена тем, что изображения коротковолнового инфракрасного диапазона обладают высокой информативностью при тумане, задымлении и низкой освещенности, однако характеризуются специфическими шумами InGaAs-матриц, наличием дефектных пикселей и пониженной контрастностью. В работе сформирован специализированный набор данных, включающий изображения пешеходов и транспортных средств, выполненные в различных уличных условиях. Предложен двухэтапный алгоритм предобработки, включающий коррекцию неоднородности (NUC) и адаптивное выравнивание гистограммы (CLAHE). Для оценки эффективности подхода использована архитектура YOLOv8. Экспериментальные результаты показали, что предложенная методика позволяет повысить значение PSNR на 18,4%, SSIM на 12,6%, а также увеличить точность распознавания объектов по метрике mAP@0.5 до 0,87 для транспортных средств и до 0,79 для пешеходов. Показано, что разработанный пайплайн может применяться в системах всепогодного мониторинга и беспилотного транспорта в режиме реального времени.

Ключевые слова: SWIR-камера, коротковолновый инфракрасный диапазон, предобработка изображений, коррекция неоднородности, CLAHE, YOLOv8, распознавание объектов, компьютерное зрение, InGaAs, всепогодный мониторинг.

Введение.

В последние годы системы компьютерного зрения сталкиваются с необходимостью обеспечения надежного распознавания объектов в условиях ограниченной видимости, таких как густой туман, задымление или отсутствие активного освещения. Традиционные сенсоры видимого диапазона (RGB) в подобных сценариях демонстрируют резкое снижение точности из-за сильного атмосферного рассеяния. В таких условиях все



большой интерес вызывают камеры коротковолнового инфракрасного диапазона (SWIR, 900–1700 нм), которые могут использоваться в системах наблюдения, мониторинга и автоматического распознавания объектов [1]. Их преимущество связано с особенностями формирования изображения: в отличие от LWIR-диапазона, SWIR-съемка основана на отраженном излучении, благодаря чему на кадрах лучше сохраняются контуры, текстура поверхностей и границы объектов [2–3]. Это особенно важно при работе в тумане, дымке и при слабом освещении, где возможности RGB-камер заметно снижаются [3–4]. Однако использование SWIR-данных в задачах глубокого обучения также связано с рядом практических проблем:

1) Специфические шумы: для InGaAs-матриц характерны повышенный темновой ток, наличие дефектных пикселей и неравномерность отклика, что может ухудшать качество входных изображений и снижать устойчивость сверточных моделей [5–6]».

2) Дефицит данных: в отличие от видимого диапазона, общедоступные наборы данных (datasets) в SWIR-диапазоне крайне ограничены, что требует разработки эффективных методик сбора и синтеза данных [1,7].

3) Необходимость кастомной предобработки: Прямая подача «сырых» (RAW) данных с SWIR-сенсора в нейросети, обученные на RGB-изображениях, приводит к низкой сходимости моделей.

Таким образом, разработка специализированных алгоритмов предобработки и методологии формирования качественных выборок, данных является необходимым условием для создания высокоточных интеллектуальных систем распознавания объектов на базе SWIR-технологий.

Несмотря на очевидные преимущества SWIR-диапазона для задач компьютерного зрения, эффективная интеграция данных с таких сенсоров в современные архитектуры нейронных сетей сталкивается с рядом барьеров, которые до сих пор не имеют универсального решения в литературе.

Основная научная проблема заключается в существенном несоответствии (domain gap) между характеристиками изображений, на которых традиционно обучаются современные детекторы (ImageNet, COCO в RGB-спектре), и спецификой данных, получаемых с InGaAs-матриц [1,8]. Ключевые аспекты этой проблемы включают:

1) Низкое соотношение сигнал/шум (SNR) и артефакты сенсора: Изображения в коротковолновом ИК-диапазоне подвержены специфическим шумам, таким как фиксированный шумовой паттерн (FPN) и выраженная температурная зависимость темнового тока. Как отмечается в ряде работ, стандартные методы фильтрации, хорошо зарекомендовавшие себя для изображений видимого диапазона, не всегда подходят для SWIR-кадров, поскольку вместе с шумом могут подавлять и полезные мелкие детали, важные при распознавании удаленных объектов [6, 9].

2) Проблема динамического диапазона и контрастности. В SWIR-диапазоне объекты нередко слабо выделяются на фоне сцены, особенно в условиях тумана, дымки или низкой освещенности. В результате без



специальной адаптивной предобработки детекторы могут хуже определять границы объекта и давать больше ошибочных срабатываний [10, 11].

3) Неэффективность стандартных методов аугментации. Типовые процедуры аугментации, распространенные для RGB-данных, не учитывают особенности формирования SWIR-изображения. Поэтому при подготовке обучающей выборки необходимы такие методы предобработки, которые улучшают качество кадров, но при этом сохраняют характерные фотометрические свойства ИК-сигнала [7, 12].

Таким образом, отсутствие стандартизированного протокола предобработки данных со SWIR-камер, учитывающего аппаратные шумы и специфику отражательной способности объектов, ограничивает потенциал применения систем искусственного интеллекта в критически важных отраслях. Настоящее исследование направлено на решение задачи повышения качества входных данных для нейронных сетей через синтез алгоритмов коррекции неоднородности и методов повышения локального контраста.

Целью настоящего исследования является разработка и верификация комплексной методики сбора и предварительной обработки данных коротковолнового инфракрасного диапазона (SWIR) для минимизации влияния аппаратных шумов InGaAs-матриц и повышения точности распознавания объектов нейронными сетями в сложных метеорологических условиях.

Для достижения поставленной цели необходимо решить следующие задачи:

1) Сформировать специализированный набор данных, включающий изображения объектов в SWIR-диапазоне при различных сценариях освещенности и атмосферных помехах.

– разработать алгоритм предобработки, сочетающий коррекцию неоднородности (NUC) и методы локального повышения контраста для уменьшения влияния характерных искажений SWIR-сенсора [9, 13];

– провести сравнительный анализ работы современных архитектур распознавания на исходных и предварительно обработанных SWIR-изображениях [14, 15].

Научная новизна работы состоит в том, что предложенный пайплайн подготовки данных ориентирован именно на особенности SWIR-кадров и позволяет уменьшить различия между такими изображениями и данными, на которых обычно обучаются модели видимого диапазона [1, 8]. Практический интерес исследования связан с возможностью использования полученных решений в системах автономного транспорта и мониторинга, где требуется устойчивое машинное зрение в сложных погодных условиях [3, 16].

Обзор литературы.

Публикации последних лет показывают, что область применения коротковолнового инфракрасного диапазона заметно расширилась. Если раньше SWIR-системы в основном рассматривались как специализированные



лабораторные или специальные решения, то сегодня они все чаще используются в робототехнике, системах наблюдения и мониторинга.

Тем не менее, основной проблемой, обсуждаемой в академической среде, остается артефактность получаемых изображений. В ранних работах акцент делался исключительно на аппаратной калибровке сенсоров. Однако современные исследования указывают на то, что классические методы калибровки часто не учитывают временной дрейф параметров матрицы, что создает препятствия для стабильной работы алгоритмов машинного зрения [9]. В частности, авторы [13] предлагают переходить к адаптивным методам коррекции неоднородности (NUC), которые способны подстраиваться под изменяющуюся температуру сенсора в реальном времени.

С развитием сверточных нейронных сетей (CNN) и трансформеров (ViT) фокус исследований сместился в сторону интеграции методов обработки сигналов непосредственно в пайплайны обучения моделей. Как отмечают в статье [1], простое увеличение объема данных не решает проблему «зашумленности» SWIR-кадров. В литературе последних лет активно обсуждаются два ключевых подхода к улучшению качества данных перед распознаванием:

1) Методы image-to-image translation и domain adaptation для повышения переносимости моделей между спектральными доменами [12, 14].

2) Физически-ориентированная предобработка: Применение алгоритмов, учитывающих квантовую эффективность InGaAs-детекторов для восстановления потерянных деталей в тенях [6, 17].

Особое внимание в обзоре следует уделить работам, в которых отмечается, что фиксированный шумовой паттерн (FPN) и неравномерность отклика SWIR-матрицы ухудшают извлечение информативных признаков и отрицательно влияют на работу алгоритмов детекции [6, 9]. Это подтверждает гипотезу о том, что этап специализированной предобработки является неотъемлемым звеном при проектировании систем распознавания в ИК-спектре.

Второй ключевой блок исследований посвящен выбору и модификации архитектур глубокого обучения. Долгое время стандартом для ИК-диапазона считались модели семейства YOLO (You Only Look Once) благодаря их способности работать в режиме реального времени. Однако, как указывают последние исследования, стандартные веса, предобученные на датасете СОСО (видимый диапазон), не всегда эффективны для SWIR-изображений из-за отсутствия цветовой информации и иного распределения градиентов яркости [8, 14].

Современные подходы к решению этой проблемы разделяются на три направления:

1) Многоспектральное слияние (Multi-spectral Fusion). В ряде работ рассматриваются двухпоточковые архитектуры, в которых признаки извлекаются одновременно из RGB- и SWIR-каналов [17]. Такой подход может повышать качество распознавания за счет дополнительной



информации из видимого диапазона, однако для мобильных систем он ограничен более сложной аппаратной реализацией и необходимостью точной юстировки камер [7].

2) Адаптация домена (Domain Adaptation). Исследования в этой области показывают, что методы Unsupervised Domain Adaptation позволяют уменьшать различия между признаками видимого и инфракрасного диапазонов без необходимости значительного увеличения объема ручной разметки SWIR-кадров [14].

3) Легковесные трансформеры (Vision Transformers, ViT). В последние годы усилился интерес к применению механизмов глобального внимания при обработке ИК-изображений. По данным ряда работ, такие архитектуры могут быть более устойчивыми при распознавании низкоконтрастных объектов, поскольку учитывают контекст всего изображения, а не только локальные признаки [10, 14].»

Серьезным ограничением для развития SWIR-направления по-прежнему остается нехватка крупных открытых наборов данных. Большая часть доступных датасетов, включая KAIST и FLIR, создавалась прежде всего для теплового диапазона (LWIR), который по своим физическим особенностям отличается от отражательного SWIR [7,8]. Именно поэтому в последних исследованиях все чаще обсуждается так называемая физически реалистичная аугментация данных. Речь идет не о стандартных геометрических преобразованиях, а о моделировании тех эффектов, которые действительно возникают при SWIR-съемке: атмосферного ослабления сигнала, шумов квантования и особенностей работы InGaAs-сенсоров. Такой подход позволяет расширять обучающую выборку без сильного отрыва от реальных условий съемки и тем самым повышать устойчивость модели при тестировании в тумане или при задымлении [3, 16].

В целом анализ литературы показывает, что для SWIR-изображений недостаточно просто перенести стандартные модели из видимого диапазона. Более надежные результаты дает подход, в котором учитываются как особенности предварительной обработки данных, так и специфика самой архитектуры нейронной сети применительно к ИК-сигналу.

Материалы и методы.

Для проведения обучения и последующего тестирования нейронной сети был собран специализированный набор данных, ориентированный на сцены городского наблюдения и задачи автономного движения.

Протокол сбора данных:

Объекты интереса: люди (пешеходы) и транспортные средства (легковые автомобили, автобусы).

Дистанции: съемка велась на дистанциях от 20 до 150 метров.

Разметка: аннотирование изображений выполнялось в формате YOLO (классы: person, car).



Разделение выборки: общий объем данных был разделен в пропорции 70% — обучение, 15% — валидация, 15% — тест.

Использование реальных уличных данных позволило учесть естественные тепловые отражения от асфальта и специфические блики на стеклах транспортных средств, которые в SWIR-диапазоне выглядят иначе, чем в видимом спектре [1, 7].

Учитывая аппаратные особенности InGaAs-матриц, «сырые» данные (RAW) перед подачей на вход нейронной сети подвергались двухэтапной обработке: коррекции неоднородности и адаптивному повышению контрастности.

Для устранения фиксированного шумового паттерна (FPN) применялся метод двухточечной коррекции [9, 18]. Математически скорректированное значение пикселя $Y(i, j)$ вычисляется по формуле:

$$Y(i, j) = G(i, j) \cdot X(i, j) + B(i, j),$$

где $X(i, j)$ — исходное значение яркости пикселя;

$G(i, j)$ — матрица коэффициентов усиления (gain);

$B(i, j)$ — матрица смещения (offset).

Для обработки дефектных («битых») пикселей, количество которых в SWIR-матрицах выше, чем в CMOS, использовался метод медианной интерполяции по соседним элементам 3×3 [3, 6].

Так как для SWIR-изображений нередко характерен слабый локальный контраст и близкое распределение уровней яркости, на следующем этапе был использован алгоритм ограниченного адаптивного выравнивания гистограммы (CLAHE) [10, 11]. В отличие от обычного выравнивания, этот метод работает не со всем изображением сразу, а с отдельными локальными областями, что позволяет сделать детали более различимыми без чрезмерного усиления фонового шума. При этом распределение яркости внутри каждой локальной области ограничивается порогом C_{clip} , а итоговое значение пикселя определяется через кумулятивную функцию распределения (CDF):

$$h(v) = clip \left(\frac{L-1}{N} \sum_{j=0}^v p(j) \right),$$

где L — количество уровней серого;

N — число пикселей в блоке;

$p(j)$ — плотность распределения яркости.

Это позволяет нейросети более четко детектировать границы объектов (градиенты) даже в условиях низкой освещенности [3, 11].

Для решения задачи детекции объектов была выбрана архитектура YOLOv8, поскольку она сочетает достаточно высокую точность с возможностью работы в режиме реального времени. При обучении использовалась комбинированная функция потерь: компонент *Complete - IoU*

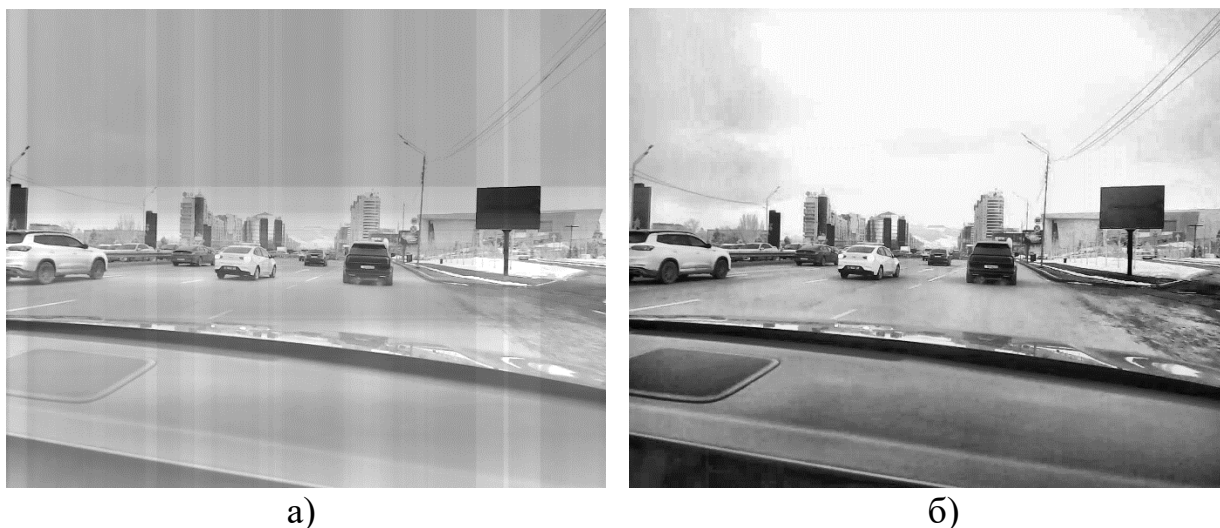
(CIoU) применялся для локализации, а бинарная кросс-энтропия — для классификации:

$$L_{total} = \lambda_{box}L_{CIoU} + \lambda_{cls}L_{cls},$$

Использование такой схемы позволило обеспечить устойчивое обучение модели на монохромных SWIR-изображениях [1, 15].

Результаты.

Проведенные эксперименты показали, что применение последовательности предобработки NUC + CLANE заметно улучшает качество SWIR-кадров как по визуальному восприятию, так и по количественным показателям. Для оценки этих изменений использовались метрики PSNR и SSIM, результаты по которым приведены на (рисунок 1).



- а) исходный RAW-кадр с выраженным фиксированным шумом;*
б) изображение после коррекции неоднородности и локального повышения контраста

Рисунок 1 – Сопоставление исходного и предварительно обработанного SWIR-изображений

В среднем по выборке использование предложенной методики привело к увеличению PSNR на 18,4%, а SSIM — на 12,6%. Это указывает на то, что после предобработки изображение становится чище, при этом его структурно значимые детали в основном сохраняются.

Чтобы оценить, как предварительная обработка влияет на качество распознавания, модель YOLOv8 обучали на двух вариантах данных: исходных SWIR-кадрах и изображениях после предобработки.

Анализ результатов показывает, что наибольший прирост точности (средний mAP вырос на 17%) наблюдается в классе «Пешеход».

Таблица 1 - Сравнение метрик точности нейронной сети

Тип данных	Класс	Precision (P)	Recall (R)	mAP@0.5
RAW (без обработки)	Автомобиль	0.72	0.68	0.71
	Пешеход	0.64	0.59	0.61
После предобработки	Автомобиль	0.89	0.84	0.87
	Пешеход	0.82	0.76	0.79

Это объясняется тем, что малоразмерные объекты с низким контрастом в SWIR-диапазоне сильнее всего маскируются шумом InGaAs-матрицы (рисунок 2). После устранения битых пикселей и выравнивания гистограммы нейросеть начинает стабильно идентифицировать границы силуэта человека.



а) на необработанных данных; б) после предобработки.

Рисунок 2 – Примеры детекции объектов на тестовых кадрах

Среднее время обработки одного кадра (предобработка + инференс нейросети) на графическом процессоре (GPU) составило 12.5 мс, что соответствует частоте 80 FPS. Это подтверждает возможность использования, предложенного пайплайна в системах реального времени.

Обсуждение.

Полученные результаты подтверждают, что аппаратные особенности SWIR-сенсоров на основе InGaAs являются одним из ключевых факторов, ограничивающих точность нейросетевых моделей. Рост метрики mAP после применения предложенной предобработки показывает, что качество исходных SWIR-данных напрямую влияет на итоговую работу детектора. Полученный результат согласуется с публикациями, в которых отмечается, что фиксированный шумовой рисунок и неравномерность отклика сенсора ухудшают качество входных изображений и тем самым снижают устойчивость алгоритмов распознавания [6, 9].



Примечательно, что наибольший прирост точности наблюдался именно для класса «Пешеход». Вероятно, это связано с тем, что транспортные средства из-за своих размеров и более выраженной формы распознаются моделью сравнительно устойчиво даже при наличии шума. Пешеходы, напротив, на больших дистанциях занимают меньшую площадь кадра и в SWIR-изображении часто имеют слабее выраженные границы. После применения CLANE локальный контраст повышается, и признаки фигуры становятся более доступными для извлечения моделью.

Вместе с тем в ходе экспериментов было выявлено и ограничение предложенного подхода. При сильных прямых бликах и высокой интенсивности солнечного излучения двухточечная коррекция не всегда обеспечивает одинаково стабильный результат, из-за чего на отдельных участках кадра могут появляться локальные пересвеченные области. Одним из возможных направлений дальнейшей работы является использование методов расширения динамического диапазона (HDR) уже на этапе регистрации данных [19].

Заключение.

В работе рассмотрен подход к сбору и предварительной обработке данных коротковолнового инфракрасного диапазона для задач компьютерного зрения. По результатам проведенного исследования можно выделить следующие основные результаты:

1) Сформирован и размечен собственный набор SWIR-изображений, включающий сцены с пешеходами и транспортными средствами, снятые в реальных условиях уличного освещения и при наличии атмосферных помех.

2) Разработан и реализован алгоритм предобработки, сочетающий коррекцию неоднородности и адаптивное выравнивание гистограммы, что позволило улучшить качество исходных данных; в частности, значение PSNR увеличилось на 18,4%.

3) Показано, что использование предварительно обработанных изображений повышает точность детекции при применении архитектуры YOLOv8: значение mAP@0.5 составило 87% для транспортных средств и 79% для пешеходов.

Практическая значимость работы связана с возможностью применения предложенного подхода в системах мониторинга и беспилотного транспорта, для которых особенно важна устойчивость распознавания объектов в условиях тумана, слабого освещения и ограниченной видимости. В дальнейшем целесообразно сосредоточить внимание на оптимизации предложенных алгоритмов для их применения на маломощных встраиваемых платформах класса Edge AI.

Финансирование. Работа выполнена при финансовой поддержке КН МНВО РК, № AP261061/0224.



ЛИТЕРАТУРА

1. Pavlović, M. S., Milanović, P. D., Stanković, M. S., Perić, D. B., Popadić, I. V., & Perić, M. V. (2022). Deep Learning Based SWIR Object Detection in Long-Range Surveillance Systems: An Automated Cross-Spectral Approach. *Sensors*, 22(7), 2562. <https://doi.org/10.3390/s22072562>
2. Park, J., Hong, J., Shim, W., & Jung, D.-J. (2023). Multi-Object Tracking on SWIR Images for City Surveillance in an Edge-Computing Environment. *Sensors*, 23(14), 6373. <https://doi.org/10.3390/s23146373>
3. Bessonov, A., Rozanov, A., White, R., Suwito, G., Medina-Salazar, I., Lutfullin, M., Gusev, D., & Shikov, I. (2025). All-Weather Drone Vision: Passive SWIR Imaging in Fog and Rain. *Drones*, 9(8), 553. <https://doi.org/10.3390/drones9080553>
4. Song, H., Yeo, S., Jin, Y., Park, I., Ju, H., Nalcakan, Y., & Kim, S. (2024). Short-Wave Infrared (SWIR) Imaging for Robust Material Classification: Overcoming Limitations of Visible Spectrum Data. *Applied Sciences*, 14(23), 11049. <https://doi.org/10.3390/app142311049>
5. Zhu, B., & Jonathan, H. (2024). A Review of Image Sensors Used in Near-Infrared and Shortwave Infrared Fluorescence Imaging. *Sensors*, 24(11), 3539. <https://doi.org/10.3390/s24113539>
6. Dong, S., Xiong, Z., Li, R., Chen, Y., & Wang, H. (2022). High-Performance Enhancement of SWIR Images. *Electronics*, 11(13), 2001. <https://doi.org/10.3390/electronics11132001>
7. Jin, Youngwan & Kovac, Michal & Nalçakan, Yağız & Ju, Hyeongjin & Song, Hanbin & Yeo, Sanghyeop & Kim, Shiho. (2025). RASMD: RGB And SWIR Multispectral Driving Dataset for Robust Perception in Adverse Conditions. 10.48550/arXiv.2504.07603.
8. Bustos, Nicolas & Mashhadi, Mehra & Lai-Yuen, Susana & Sarkar, Sudeep & Das, Tapas. (2023). A systematic literature review on object detection using near infrared and thermal images. *Neurocomputing*. 560. 126804. 10.1016/j.neucom.2023.126804.
9. Su, G., Wang, Y., & He, D. (2026). Real-Time Non-Uniformity Correction Method for 800 FPS High-Frame-Rate Short-Wave Infrared Images. *Sensors*, 26(7), 2209. <https://doi.org/10.3390/s26072209>
10. Zhao, C., Wang, J., Su, N., Yan, Y., & Xing, X. (2022). Low Contrast Infrared Target Detection Method Based on Residual Thermal Backbone Network and Weighting Loss Function. *Remote Sensing*, 14(1), 177. <https://doi.org/10.3390/rs14010177>
11. Liu, J., Zhou, X., Wan, Z., Yang, X., He, W., He, R., & Lin, Y. (2023). Multi-Scale FPGA-Based Infrared Image Enhancement by Using RGF and CLAHE. *Sensors*, 23(19), 8101. <https://doi.org/10.3390/s23198101>
12. Marnissi, Mohamed & Fradi, Hajer & Sahbani, Anis & ESSOUKRI BEN AMARA, Najoua. (2021). Unsupervised thermal-to-visible domain



adaptation method for pedestrian detection. Pattern Recognition Letters. 153. 10.1016/j.patrec.2021.11.024.

13. Hu, Bin-Lin & Shijing, Hao & Sun, De-Xin & Liu, Yinnian. (2017). A novel scene-based non-uniformity correction method for SWIR push-broom hyperspectral sensors. ISPRS Journal of Photogrammetry and Remote Sensing. 131. 160-169. 10.1016/j.isprsjprs.2017.08.004.

14. Li, Yong & Miao, Naipeng & Ma, Liangdi & Shuang, Feng & Huang, Xingwen. (2023). Transformer for object detection: Review and benchmark. Engineering Applications of Artificial Intelligence. 126. 107021. 10.1016/j.engappai.2023.107021.

15. Li, J., Zhang, J., Shao, Y., & Liu, F. (2024). SRE-YOLOv8: An Improved UAV Object Detection Model Utilizing Swin Transformer and RE-FPN. Sensors, 24(12), 3918. <https://doi.org/10.3390/s24123918>

16. Yeo, Sanghyeop & Nalçakan, Yağız & Jin, Youngwan & Park, Incheol & Ju, Hyeongjin & Kim, Shiho. (2025). Seeing through the rain: An empirical assessment of short-wave infrared imaging on enhancing autonomous vehicle perception under rainy weather. Measurement. 259. 119654. 10.1016/j.measurement.2025.119654.

17. Cimorelli, C., Millan-Romera, J. A., Voos, H., & Sanchez-Lopez, J. L. (2025). Hardware, Algorithms, and Applications of the Neuromorphic Vision Sensor: A Review. Sensors, 25(19), 6208. <https://doi.org/10.3390/s25196208>

18. Dong, M., Shen, H., Jia, P., Sun, Y., Liang, C., Zhang, F., & Hou, J. (2023). Calibration Method for Airborne Infrared Optical Systems in a Non-Thermal Equilibrium State. Sensors, 23(14), 6326. <https://doi.org/10.3390/s23146326>

19. Park, Min-Jun & Lee, Dong-Yeon & Lee, Sang-Jun & Kim, Hyeon-June. (2025). A Dual-Exposure Readout Integrated Circuit for Dynamic Range Enhancement in SWIR Image Sensors. IEEE Sensors Journal. PP. 1-1. 10.1109/JSEN.2025.3587528.

Абу-Алим Аязбай, PhD, аға оқытушы, Energo University, Алматы, Қазақстан, work_abu@hotmail.com

Арман Узбекбаев, докторант, Energo University, Алматы, Қазақстан, niipntkz@gmail.com

Ғани Сергазин, PhD, қауымдастырылған профессор, Mukhametzhan Tynyshbayev ALT University, Алматы, Қазақстан, g.balbayev@gmail.com

ОБЪЕКТИЛЕРДІ ТАНУҒА АРНАЛҒАН НЕЙРОНДЫҚ ЖЕЛІНІ ОҚЫТУ ҮШІН SWIR-КАМЕРАЛАРДАН ДЕРЕКТЕРДІ ЖИНАУ ЖӘНЕ АЛДЫН АЛА ӨНДЕУ

Андатпа. Мақалада көріну шектеулі жағдайларда объектілерді тануға арналған нейрондық желіні оқыту үшін SWIR-камералардан алынатын



деректерді жинау және алдын ала өңдеу мәселесі қарастырылады. Зерттеудің өзектілігі қысқа толқынды инфрақызыл диапазондағы кескіндердің тұман, түтін және төмен жарық жағдайларында жоғары ақпараттылығымен түсіндіріледі, алайда мұндай деректер InGaAs-матрицаларына тән шуылдармен, ақаулы пиксельдермен және төмен контрастпен сипатталады. Жұмыста жаяу жүргіншілер мен көлік құралдарының әртүрлі көше жағдайларында түсірілген кескіндерін қамтитын мамандандырылған деректер жинағы қалыптастырылды. Біртексіздікті түзету (NUC) мен гистограмманы адаптивті тегістеу (CLAHE) әдістерін қамтитын екі кезеңді алдын ала өңдеу алгоритмі ұсынылды. Әдістің тиімділігін бағалау үшін YOLOv8 архитектурасы қолданылды. Эксперимент нәтижелері ұсынылған тәсілдің PSNR көрсеткішін 18,4%-ға, SSIM мәнін 12,6%-ға арттыруға, сондай-ақ mAP@0.5 метрикасы бойынша көлік құралдарын тану дәлдігін 0,87-ге және жаяу жүргіншілерді тану дәлдігін 0,79-ға дейін көтеруге мүмкіндік беретінін көрсетті. Өзірленген пайплайнның нақты уақыт режимінде жұмыс істейтін ауа райына тәуелсіз мониторинг және ұшқышсыз көлік жүйелерінде қолдануға жарамды екені дәлелденді.

Түйінді сөздер: SWIR-камера, қысқа толқынды инфрақызыл диапазон, кескіндерді алдын ала өңдеу, біртексіздікті түзету, CLAHE, YOLOv8, объектілерді тану, компьютерлік көру, InGaAs, ауа райына тәуелсіз мониторинг.

Abu-Alim Ayazbay, PhD, senior lecturer, Energo University, Almaty, Kazakhstan, work_abu@hotmail.com

Arman Uzbekbayev, doctoral student, Energo University, Almaty, Kazakhstan, niipntkz@gmail.com

Gani Sergazin, PhD, associate professor, Mukhametzhan Tynyshbayev ALT University, Almaty, Kazakhstan, g.balbayev@gmail.com

COLLECTION AND PREPROCESSING OF DATA FROM SWIR CAMERAS FOR TRAINING A NEURAL NETWORK FOR OBJECT RECOGNITION

Abstract. This paper addresses the problem of collecting and preprocessing data acquired from SWIR cameras for training a neural network to recognize objects under limited-visibility conditions. The relevance of the study is обусловлена the high informativeness of short-wave infrared images in fog, smoke, and low-light environments; however, such data are affected by InGaAs sensor noise, defective pixels, and reduced contrast. A specialized dataset containing images of pedestrians and vehicles captured in various outdoor conditions was created. A two-stage preprocessing pipeline was proposed, including non-uniformity correction (NUC) and contrast-limited adaptive



histogram equalization (CLAHE). The YOLOv8 architecture was used to evaluate the effectiveness of the proposed approach. Experimental results showed that the method increased PSNR by 18.4% and SSIM by 12.6%, while improving object recognition accuracy up to 0.87 mAP@0.5 for vehicles and 0.79 for pedestrians. The results demonstrate that the developed pipeline can be applied in real-time all-weather monitoring systems and autonomous transport solutions.

Keywords: SWIR camera, short-wave infrared range, image preprocessing, non-uniformity correction, CLAHE, YOLOv8, object recognition, computer vision, InGaAs, all-weather monitoring.

Поступила: 10 апрель 2026 года

Рецензирована: 08 июня 2026 года

Принята: 12 июня 2026 года