

УДК 004.93

СОВРЕМЕННЫЕ ПАРАДИГМЫ РАСПОЗНАВАНИЯ ОБРАЗОВ: ТЕОРИЯ, МЕТРИКИ И ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ

Бийбосунова С. К., Ли Гуанчже, Ши Сянтао

КГУ им. И. Арабаева

В статье представлен системный анализ эволюции методов распознавания образов – от классических подходов, основанных на ручном проектировании признаков, до современных архитектур глубокого обучения. Рассматриваются теоретические основы представления объектов в виде признаковых пространств. Результаты исследования демонстрируют полный жизненный цикл разработки интеллектуальных систем и обосновывают выбор конкретных методов в зависимости от условий эксплуатации и доступных вычислительных ресурсов.

Ключевые слова: Распознавание образов, классификация, метрики качества, нейронные сети.

ОБРАЗДАРДЫ ТААНУУНУН ЗАМАНБАП ПАРАДИГМАЛАРЫ: ТЕОРИЯ, МЕТРИКАЛАР ЖАНА ПРАКТИКАДА КОЛДОНУЛУШУ

Бийбосунова С. К., Ли Гуанчже, Ши Сянтао

И. Арабаев атындагы Кыргыз мамлекеттик университети

Макалада образдарды таануу методдорунун эволюциясына – белгилерди кол менен долбоорлоорго негизделген классикалык ыкмалардан баштап, терең окутуунун заманбап архитектураларына чейинки системалуу талдоо берилген. Объекттерди белгилер мейкиндиги түрүндө көрсөтүүнүн теориялык негиздери каралат. Изилдөөнүн жыйынтыктары интеллектуалдык системаларды иштеп чыгуунун толук жашоо циклин көрсөтөт жана эксплуатациялоо шарттарына, ошондой эле жеткиликтүү эсептөө ресурстарына жараша конкреттүү методдорду тандоону негиздейт.

Баштапкы сөздөр: Образдарды таануу, классификациялоо, сапат метрикалары, нейрондук тармактар.

CURRENT PARADIGMS IN IMAGE RECOGNITION: THEORY, METRICS AND PRACTICAL APPLICATIONS.

Biibosunova S. K., Li Guangzhe, Shi Xiangtao

Kyrgyz State University named after I. Arbaeva

This article presents a systematic analysis of the evolution of image recognition methods – from classical approaches based on manually designed features to modern deep learning architectures. The theoretical foundations of representing objects as feature spaces are examined. The research results demonstrate the full life cycle of intelligent system development and justify the selection of specific methods based on operating conditions and available computational resources.

Keywords: Pattern recognition, classification, quality metrics, neural networks.

Введение

Теория распознавания образов представляет собой одно из ключевых направлений современной информатики, искусственного интеллекта и прикладной математики. Ее задача состоит в том, чтобы по наблюдаемым данным выявить свойства объекта, отнести его к определенному классу, обнаружить структуру или принять решение в условиях неопределенности. В прикладном смысле это означает автоматическое понимание изображений, текстов, аудиосигналов, биометрических характеристик и множества других типов данных, которые ранее анализировались человеком вручную.

Актуальность темы объясняется сразу несколькими факторами. Во-первых, цифровые данные генерируются в огромных объемах, и ручная обработка становится экономически нецелесообразной. Во-вторых, развитие вычислительной техники и графических ускорителей сделало возможным обучение сложных моделей на больших выборках. В-третьих, распознавание образов уже стало основой для интеллектуальных сервисов в промышленности, медицине, транспорте, безопасности, банковской сфере и пользовательских приложениях [1].

Современный этап развития распознавания образов характеризуется переходом от жестко заданных правил и ручного

выделения признаков к данным-ориентированным методам, где система самостоятельно извлекает информативные представления из примеров. При этом классические методы не утратили своей значимости: они сохраняют ценность там, где объем данных ограничен, требуется высокая интерпретируемость или существуют жесткие вычислительные ограничения.

Цель данного текста заключается в том, чтобы системно рассмотреть современные методы теории распознавания образов, показать их место в общей структуре интеллектуальных систем и обосновать практическую разработку сервиса, демонстрирующего сравнение классического и нейросетевого подходов на задаче анализа изображений [1, 2].

Теоретические основы распознавания образов

Под образом в теории распознавания понимается совокупность признаков, описывающих некоторый объект или явление. Объектом может быть фотография изделия, дорожный знак, отпечаток пальца, голосовая команда, фрагмент текста или результат медицинского исследования. Для того чтобы объект мог быть обработан машиной, его необходимо представить в числовой форме. Такое представление называют вектором признаков или пространством признаков описаний [3].

Центральная идея распознавания образов состоит в том, что объекты одного класса обладают общими характеристиками, а объекты разных классов различаются по значимым признакам [4, 5, 6].

В реальных приложениях важен не только выбор алгоритма, но и качество данных, корректность разметки, согласованность признакового пространства, устойчивость к шуму и способность модели обобщать знания на новые примеры.

Основные этапы решения задачи распознавания

Независимо от конкретного алгоритма практическая система распознавания образов обычно строится по этапам. Сначала происходит сбор данных: формируется корпус изображений, сигналов или иных объектов, на которых будет обучаться и проверяться модель. На этом шаге критически важны полнота выборки, баланс классов и соответствие данных реальным условиям эксплуатации.

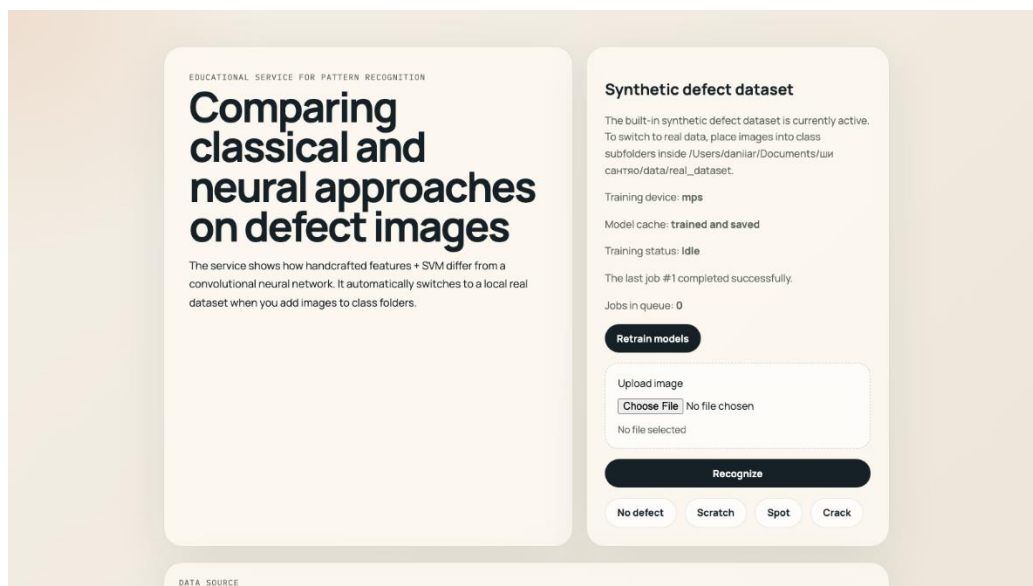


Рис.1. Интерфейс учебного веб-сервиса для сравнительного анализа классического и нейросетевого подходов

Далее выполняется предварительная обработка. Для изображений это может быть изменение размера, перевод в оттенки серого, нормализация яркости, фильтрация шумов, выравнивание контраста, выделение интересующей области. Цель этапа состоит в том, чтобы уменьшить влияние случайных факторов и привести объекты к единому формату, удобному для дальнейшего анализа.

Следующий этап связан с извлечением признаков. В классических системах признаки проектируются человеком: текстурные характеристики, градиенты, гистограммы интенсивности, контурные дескрипторы, геометрические моменты. В современных нейросетевых подходах признаки формируются самой моделью автоматически в процессе обучения. Это принципиальное отличие определило скачок в качестве систем компьютерного зрения и распознавания речи.

После этого происходит обучение модели и ее валидация. Набор данных делят на обучающую, валидационную и тестовую части. На обучающей выборке модель подбирает параметры, на валидационной настраиваются гиперпараметры, а тестовая часть используется для независимой оценки качества. Завершающий этап связан с внедрением: модель интегрируется в сервис, получает новые данные от пользователя и формирует прогноз с указанием класса, вероятности и дополнительных диагностических сведений.

Классические методы распознавания образов

Классические методы сформировали фундамент распознавания образов задолго до распространения глубокого обучения. Их сильная

сторона заключается в прозрачной логике работы и сравнительно невысоких требованиях к вычислительным ресурсам. Наиболее простым методом является k-ближайших соседей, при котором решение для нового объекта принимается на основе сходства с уже известными примерами. Метод интуитивно понятен, однако чувствителен к выбору метрики и масштабу признаков [4].

Широкое распространение получили методы линейной и нелинейной классификации. Логистическая регрессия подходит для случаев, когда классы можно разделить с помощью почти линейной границы, в то время как SVM (машина опорных векторов) может обучаться более сложным границам принятия решений с помощью ядерных функций. SVM часто используются при работе с многомерным пространством признаков и относительно небольшим количеством образцов. Раньше типичный рабочий процесс в компьютерном зрении включал извлечение признаков, их нормализацию и классификацию с помощью SVM.

Особый интерес представляют также деревья решений, случайные леса и градиентный бустинг. Они очень эффективны при работе с табличными данными, позволяют ранжировать важность признаков и часто используются в контексте гибридного подхода, при котором извлекаются признаки изображения и применяется метод ансамблевой классификации. Для анализа текстуры часто используются признаки, основанные на дескрипторах HOG, LBP, признаках типа Haar, локальной статистике и моментах изображения.

Классические методы имеют серьезный недостаток, заключающийся в том, что качество модели напрямую зависит от качества признаков, выбранных инженером.

Если признаки недостаточно выразительны, алгоритм не сможет компенсировать это автоматически. Именно поэтому с ростом сложности данных и требований к точности все большую роль стали играть методы глубокого обучения, которые строят представление объектов из самих данных [6].

Современные методы и глубокое обучение

Современный этап развития распознавания образов связан прежде всего с искусственными нейронными сетями и глубоким обучением [7]. Их главное преимущество заключается в способности автоматически выделять иерархию признаков: от простых локальных контуров и текстур до сложных абстрактных структур. В результате модель самостоятельно формирует внутреннее представление объектов, которое оказывается

более устойчивым и информативным, чем заранее заданные вручную дескрипторы.

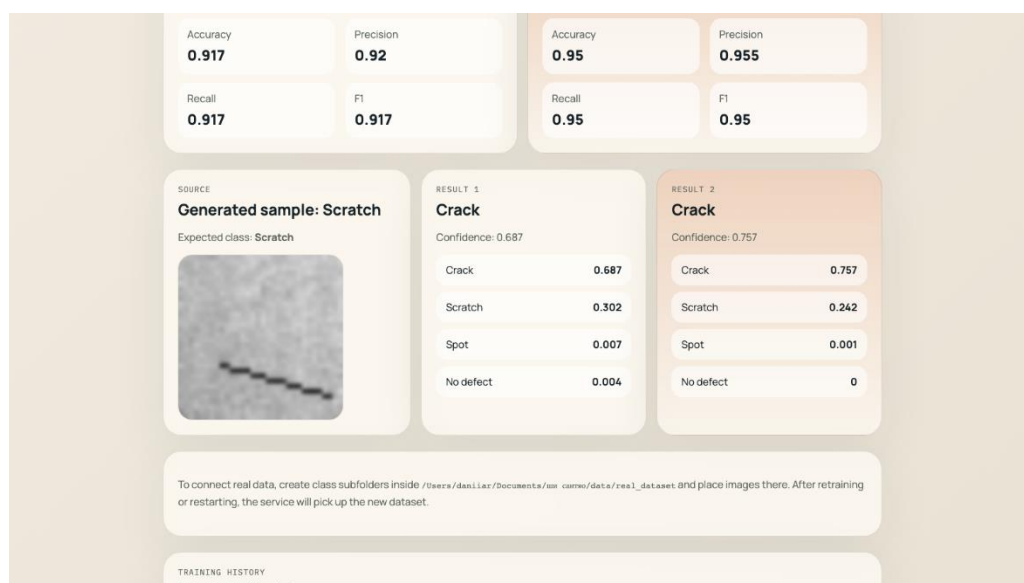


Рис.2. Результаты инференса моделей и сравнительный анализ метрик качества классификации.

Для обработки изображений особенно важны сверточные нейронные сети. Они используют локальные фильтры, общие веса и операции подвыборки, благодаря чему эффективно выявляют пространственные закономерности. Классические архитектуры семейства LeNet, AlexNet, VGG, ResNet, EfficientNet и их производные стали основой большинства практических решений в компьютерном зрении. В задачах классификации сеть определяет класс изображения, в задачах детектирования она дополнительно локализует объект, а в задачах сегментации выделяет пиксели интересующей области [8].

Большое значение имеет transfer learning, то есть дообучение уже подготовленной модели на новой предметной области.

При этом окончательное решение в ответственных сценариях остается за специалистом, а сама модель выступает в роли инструмента поддержки принятия решений [9, 10].

В транспортных и городских системах распознавание образов применяется для обнаружения дорожных знаков, контроля полос движения, распознавания номеров, мониторинга трафика и предотвращения инцидентов. В сфере безопасности используются методы идентификации лиц, анализа поведения, обнаружения запрещенных предметов и поиска аномалий на видеопотоке. В

пользовательских сервисах распознавание образов лежит в основе поиска по изображению, умных камер, фильтров, биометрической аутентификации и систем рекомендаций.

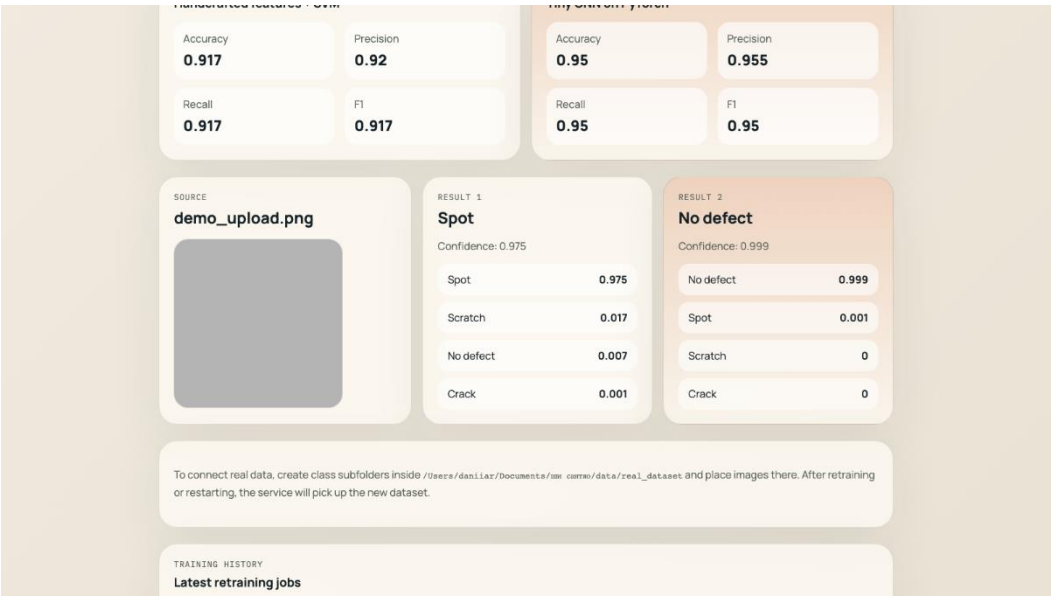


Рис.3. Интерфейс загрузки пользовательского изображения и сопоставление результатов классификации двумя контурами

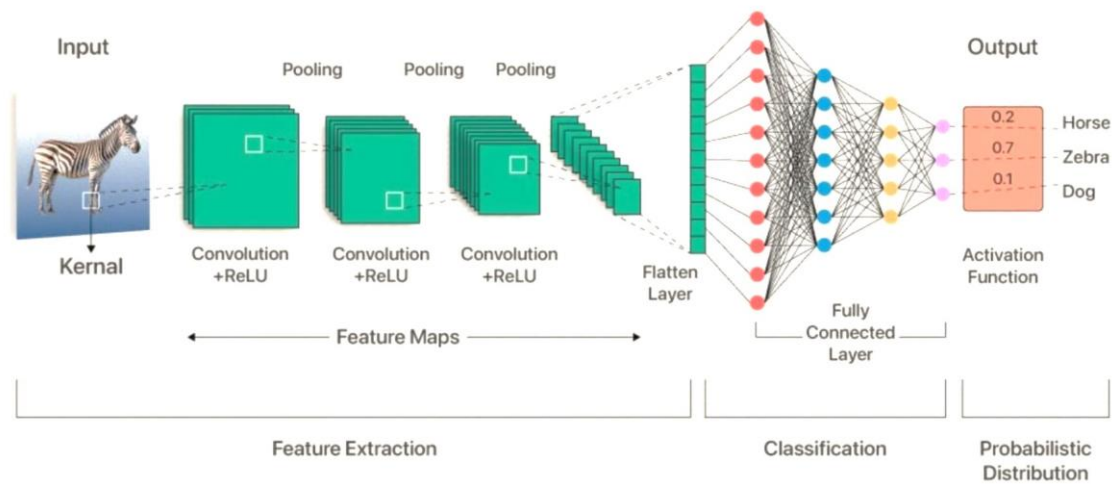


Рис.4. Типовая архитектура сверточной нейронной сети (CNN) и этапы обработки визуальных данных

Таким образом, практическая ценность распознавания образов определяется не только высокой точностью алгоритма, но и способностью встроиться в реальный бизнес-процесс. Хорошая система должна быть

понятной, устойчивой к шуму, измеримой по качеству, достаточно быстрой и экономически оправданной.

Концепция разрабатываемого сервиса

В качестве практической разработки целесообразно создать учебный веб-сервис, демонстрирующий сравнение двух подходов к распознаванию изображений: классического признакового алгоритма и нейросетевой модели. Такой сервис удобен как для изучения теории, так и для последующего расширения под реальную прикладную задачу, например контроль дефектов на производстве. Пользователь получает не абстрактное описание алгоритмов, а наглядный инструмент, в котором можно загрузить изображение, увидеть прогнозы двух моделей и сравнить их поведение.

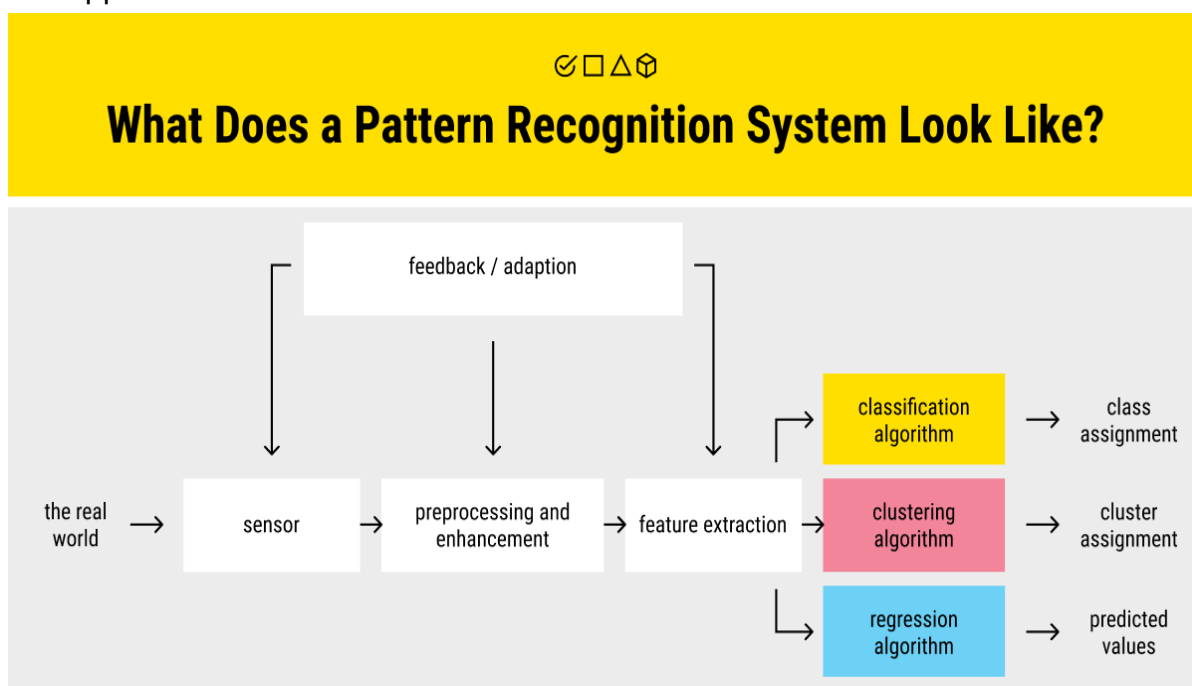


Рис.5. Обобщенная функциональная схема системы распознавания образов и основные этапы обработки данных

Архитектурно сервис должен включать модуль подготовки данных, модуль обучения, API для получения изображений от пользователя, блок вывода результатов и интерфейс с метриками качества. В учебном MVP допустимо использовать синтетический датасет изображений дефектов. Такой подход позволяет развернуть демонстрационную систему без внешних зависимостей и одновременно показать весь pipeline: генерацию данных, выделение признаков, обучение, валидацию и инференс.

Классический контур может быть построен на комбинации заранее вычисленных признаков и классификатора SVM. Конструкция нейронной сети должна представлять собой многослойный перцептрон или сверточную сеть, в зависимости от доступных ресурсов и библиотек. Она должна предоставлять показатели производительности моделей, вероятности классов для классификации, объяснение результатов, а также набор тестовых примеров, позволяющих пользователю проверить их.

Практическая значимость рассматриваемого сервиса заключается в двух аспектах. Во-первых, он служит обучающим набором для экспериментов по классификации изображений на основе глубокого обучения.

Заключение

Распознавание изображений находится на стыке математики, машинного обучения и инженерных дисциплин. Традиционно классическая математика использовалась для формального описания элементов и характеристик изображений, а также для поддержки процессов принятия решений. Однако математика нейронных сетей выводит эту область в новое измерение: компьютеры могут видеть, слышать и извлекать информацию из обширных баз данных изображений, демонстрируя выдающуюся эффективность в компьютерном зрении, обработке речи и мультимодальной обработке.

Распознавание изображений используется во многих областях: промышленной автоматизации, медицине, беспилотных автомобилях, системах безопасности, мобильной и бытовой электронике, а также в автоматизации дома и офиса. Поэтому выбор подходящего алгоритма распознавания изображений для конкретного приложения обычно сводится к сопоставлению требований приложения с характеристиками доступных алгоритмов. Критерии, которые следует учитывать в этом контексте, включают качество данных для обучения и оценки, скорость работы приложения, необходимость объяснения результатов и, наконец, уровень сложности приложения. Предлагая широкий спектр проверенных и современных алгоритмов, OpenCV облегчает выбор подходящего метода для каждого приложения.

ЛИТЕРАТУРА

1. Бишоп, К. Распознавание образов и машинное обучение = Pattern Recognition and Machine Learning [Текст] / К. Бишоп. – Springer, 2006. – 738 с.

2. Гудфеллоу, Я. Глубокое обучение = Deep Learning [Текст] / Я. Гудфеллоу, И. Бенджио, А. Курвилль. – М. : ДМК Пресс, 2018 (ориг. изд. MIT Press). – 652 с.
3. Гонсалес, Р. Цифровая обработка изображений = Digital Image Processing [Текст] / Р. Гонсалес, Р. Вудс. – М. : Техносфера, 2012 (ориг. изд. Pearson). – 1104 с.
4. Хасты, Т. Элементы статистического обучения = The Elements of Statistical Learning [Текст] / Т. Хасты, Р. Тибширани, Дж. Фридман. – Springer, 2009. – 745 с.
5. Szeliski, R. Computer Vision: Algorithms and Applications [Text] / R. Szeliski. – Springer, 2010. – 812 p.
6. LeCun, Y. Deep learning [Text] / Y. LeCun, Y. Bengio, G. Hinton // Nature. – 2015. – Vol. 521. – P. 436–444.
7. Scikit-learn: Machine Learning in Python [Electronic resource] : official documentation. – URL: <https://scikit-learn.org> (date of access: 18.03.2026).
8. FastAPI framework [Electronic resource] : interactive API documentation. – URL: <https://fastapi.tiangolo.com> (date of access: 18.03.2026).
9. NumPy [Electronic resource] : fundamental scientific computing with Python. – URL: <https://numpy.org> (date of access: 18.03.2026).
10. Pillow (PIL Fork) [Electronic resource] : Python Imaging Library. – URL: <https://python-pillow.org> (date of access: 18.03.2026).