

DOI: 10.12731/3033-5965-2025-15-4-407

EDN: FIZCUW

УДК 004.85:004.932.2:681.5.015



Научная статья | Искусственный интеллект в промышленных системах

АНАЛИЗ АРХИТЕКТУР СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ СИСТЕМ ТЕХНИЧЕСКОГО ЗРЕНИЯ

А.А. Акулов, Д.С. Талдыкин, А.В. Гришкина, Н.М. Ганжа

Аннотация

Обоснование. Современные системы машинного зрения в значительной степени опираются на сверточные нейронные сети, которые демонстрируют высокую результативность при анализе изображений за счет способности извлекать иерархические представления признаков. Повышение качества распознавания достигается увеличением глубины архитектур и числа параметров, но данный подход сопровождается ростом вычислительных затрат, повышенной чувствительностью к шуму и снижением устойчивости к вариациям масштаба и освещенности. Актуализируется задача поиска архитектурных решений, обеспечивающих более эффективное извлечение признаков без существенного увеличения вычислительной сложности. Перспективным направлением является использование многомасштабных сверточных операций, позволяющих учитывать локальный и глобальный контекст изображения, а также внедрение механизмов внимания, обеспечивающих адаптивную концентрацию вычислительных ресурсов на наиболее информативных областях и признаках.

Цель – формирование архитектурного подхода к построению сверточных нейронных сетей, основанного на совмещении многомасштабной сверточной обработки и механизмов внимания, ориентированного на повышение эффективности алгоритмов машинного зрения при сохранении приемлемого уровня вычислительных затрат.

Материалы и методы. В статье рассмотрены три архитектуры сверточных нейронных сетей, включающие базовую CNN, многомасштабную CNN и расширенную модель с механизмом внимания. Проанализированы операции двумерной свертки, пакетной нормализации и нелинейной активации, а также способы агрегации признаков, полученных на различных пространственных масштабах. Выполнена аналитическая оценка вычислительной сложности архитектур с учетом числа параметров, глубины сети и объема операций прямого прохода.

Результаты. Применение многомасштабной сверточной обработки расширяет область восприятия сети и повышает устойчивость к вариациям размеров объектов за счет одновременного анализа признаков на разных уровнях детализации. Внедрение механизмов внимания обеспечивает адаптивное перераспределение весов между признаками, снижает влияние шумовых и нерелевантных компонентов и повышает селективность анализа изображений.

Ключевые слова: машинное зрение; нейронные сети; сверточная нейронная сеть; механизм внимания; многомасштабные свертки; устойчивость алгоритмов; диагностика и мониторинг

Для цитирования. Акулов, А. А., Талдыкин, Д. С., Гришкина, А. В., & Ганжа, Н. М. (2025). Анализ архитектур сверточных нейронных сетей для систем технического зрения. *Transportation and Information Technologies in Russia / Транспорт и информационные технологии*, 15(4), 105–128. <https://doi.org/10.12731/3033-5965-2025-15-4-407>

Original article | Artificial Intelligence in Industrial Systems

ANALYSIS OF CONVOLUTIONAL NEURAL NETWORK ARCHITECTURES FOR MACHINE VISION SYSTEMS

A.A. Akulov, D.S. Taldykin, A.V. Grishkina, N.M. Ganzha

Abstract

Background. Modern machine vision systems largely rely on convolutional neural networks, which demonstrate high performance in image

analysis due to their ability to extract hierarchical feature representations. Improvement in recognition quality is commonly achieved by increasing network depth and the number of parameters; however, this approach is associated with higher computational costs, increased sensitivity to noise, and reduced robustness to variations in scale and illumination. This circumstance highlights the need for architectural solutions that enable more efficient feature extraction without a significant increase in computational complexity. A promising direction involves the use of multiscale convolutional operations that capture both local and global image context, as well as the integration of attention mechanisms that provide adaptive concentration of computational resources on the most informative regions and features.

Purpose. To formulate an architectural approach to the design of convolutional neural networks based on the combination of multiscale convolutional processing and attention mechanisms, aimed at improving the efficiency of machine vision algorithms while maintaining an acceptable level of computational cost.

Materials and methods. The article examines three convolutional neural network architectures, including a baseline CNN, a multiscale CNN, and an extended model incorporating an attention mechanism. Operations of two dimensional convolution, batch normalization, and nonlinear activation are analyzed, along with methods for aggregating features obtained at different spatial scales. An analytical assessment of computational complexity is performed with consideration of the number of parameters, network depth, and the volume of operations during the forward pass.

Results. The application of multiscale convolutional processing expands the receptive field of the network and improves robustness to variations in object size through simultaneous analysis of features at different levels of detail. The integration of attention mechanisms provides adaptive redistribution of feature weights, reduces the influence of noise and irrelevant components, and enhances the selectivity of image analysis.

Keywords: machine vision; neural networks; convolutional neural network; attention mechanism; multiscale convolutions; algorithm robustness; diagnostics and monitoring

For citation. Akulov, A. A., Taldykin, D. S., Grishkina, A. V., & Gan-zha, N. M. (2025). Analysis of convolutional neural network architectures for machine vision systems. *Transportation and Information Technologies in Russia*, 15(4), 105–128. <https://doi.org/10.12731/3033-5965-2025-15-4-407>

Введение

Современные методы машинного зрения развиваются преимущественно в направлении повышения точности извлечения признаков при одновременном снижении вычислительных затрат. Применение сверточных нейронных сетей показало высокую эффективность в задачах классификации, детектирования и сегментации изображений, что обусловлено их способностью формировать иерархические представления признаков различного уровня абстракции. Однако рост глубины и числа параметров моделей нередко приводит к избыточным вычислениям и ухудшению обобщающей способности. Вследствие этого актуальной задачей остается поиск архитектурных решений, обеспечивающих баланс между точностью и скоростью обработки данных.

Одним из направлений повышения эффективности является использование многомасштабных сверточных блоков, позволяющих учитывать контекст изображений на разных уровнях детализации [14]. Подобный подход обеспечивает более полное описание локальных и глобальных структур, что особенно важно при анализе сложных сцен, где объекты различаются по размеру, форме и текстуре. Дополнительный прирост качества достигается при внедрении механизмов внимания, которые распределяют вычислительные ресурсы между наиболее информативными областями изображения. Совмещение многомасштабных операций и внимания формирует гибридные архитектуры, способные адаптивно выделять значимые признаки и игнорировать шумовые компоненты.

В рамках данной статьи рассмотрены три архитектуры сверточных нейронных сетей: базовая CNN, многомасштабная CNN-MS и расширенная CNN-MS-Attn, включающая блок внимания [2]. Проведен сравнительный анализ их структурных особенностей, опи-

саны формальные зависимости, определяющие операции свертки, нормализации и активации, а также предложены схемы для оценки вычислительной сложности. Рассмотрены подходы к проектированию архитектур с точки зрения оптимального соотношения числа параметров, глубины сети и объема операций. Приведены теоретические обоснования выбора многомасштабных фильтров и механизмов внимания, позволяющих повысить устойчивость сети к вариациям масштаба и контраста входных изображений.

Материалы и методы

Технологию машинного зрения можно представить как двухуровневую систему, объединяющую аппаратные и алгоритмические компоненты. Аппаратный уровень отвечает за сбор визуальных данных, в то время как алгоритмический уровень выполняет обработку данных, сегментацию и классификацию [5].

Были рассмотрены три архитектурных подхода к построению сверточных нейронных сетей: стандартная CNN, многомасштабная CNN-MS и модифицированная CNN-MS-Attn с механизмом внимания. Каждая архитектура отличается способом обработки входных данных, принципом агрегирования признаков и вычислительной сложностью [3].

Стандартная CNN. Базовая сверточная нейронная сеть CNN выполняет двумерную свертку входного изображения с ядром фиксированного размера. Входное изображение обозначено как

$$I \in RH \times W \times C, \tag{1}$$

где H и W – пространственные размеры, C – число каналов.

Свертка с ядром K размером $kh \times kw$ вычисляется по формуле

$$S(i,j) = (I \cdot K)(i,j) = \sum_{m=0}^{k_h-1} \sum_{n=0}^{k_w-1} I(i+m, j+n) \cdot K(m,n), \tag{2}$$

i, j – координаты позиции в выходной карте признаков;

m, n – координаты элементов ядра свертки;

k_p, k_w – высота и ширина ядра свертки;

$K(m,n)$ – вес (коэффициент) ядра свертки в позиции (m,n) ;

Результирующая карта признаков формируется последовательным применением таких операций с последующей пакетной нормализацией и активацией:

$$F^{(l)} = ReLU(BN(S^{(l)})), \quad (3)$$

Каждый слой сети выполняет переход от локальных структур к более абстрактным признакам [1]. Глобальное усреднение по пространству и полносвязный слой завершают архитектуру, обеспечивая классификацию признаков. Базовая CNN характеризуется устойчивостью и простотой реализации, однако ограничена фиксированным масштабом восприятия, что снижает ее эффективность при анализе объектов разных размеров [11].

CNN-MS: многомасштабная свертка. Для повышения устойчивости к вариациям масштаба используется многомасштабный подход. В данной архитектуре входная карта признаков обрабатывается несколькими сверточными ядрами различных размеров $K(s)$, где $s=1..S$. Для каждого масштаба выполняется свертка:

$$S^{(s)}(i, j) = m = \sum_{m=0}^{k_h^{(s)}-1} \sum_{n=0}^{k_w^{(s)}-1} I(i+m, j+n) \cdot K^{(s)}(m, n), \quad (4)$$

Такое объединение расширяет количество каналов и обеспечивает одновременное восприятие контекстной информации на разных уровнях детализации [6]. После агрегации выполняется нормализация и свертка 1×1 , приводящая число каналов к требуемому размеру:

$$F^{(l)} = ReLU(BN(Conv_{1 \times 1}(S_{MS}))), \quad (5)$$

Расширенная модель CNN-MS-Attn включает механизм внимания, который регулирует вклад признаков, полученных с разных масштабов. Для каждого масштаба s вводится карта весов внимания $A(s)(i, j)$. Выходной сигнал многомасштабного блока с вниманием определяется выражением:

$$s_{MS-Attn}(i, j) = \sum_{s=1}^S A^{(s)}(i, j) \cdot S^{(s)}(i, j), \quad (6)$$

где :

$$A^{(s)}(i, j) = \frac{\exp(\varphi(S^{(s)}(i, j)))}{\sum_{r=1}^S \exp(\varphi(S^{(r)}(i, j)))}, \quad (7)$$

где φ – функция генерации весов, реализуемая в виде линейного преобразования или небольшой сверточной сети.

Такое представление позволяет архитектуре динамически определять, какой масштаб содержит наиболее значимую информацию в конкретной пространственной позиции. При этом суммарная сумма весов по всем масштабам равна единице, что обеспечивает корректную нормировку распределения внимания [17]. Введение механизма внимания позволяет улучшить селективность восприятия признаков и повысить устойчивость сети к шуму и нерелевантным фрагментам изображения. На рисунках 1, 2 и 3 отображены основные элементы представленных в работе архитектур.



Рис. 1. Архитектура CNN

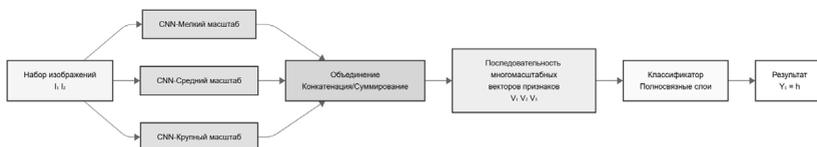


Рис. 2. Архитектура CNN-MS

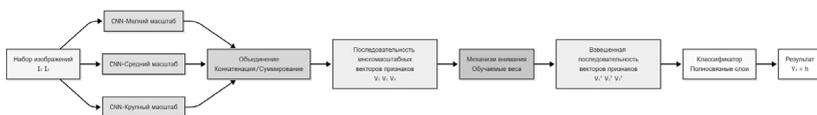


Рис. 3. Архитектура CNN-MS-Attn

Для стабилизации обучения используется пакетная нормализация, обеспечивающая приведение статистик активаций к нулевому среднему и единичной дисперсии. В качестве функции активации применяется $ReLU(x) = \max(0, x)$, что предотвращает насыщение градиентов и ускоряет сходимость [8].

Предварительная обработка включала нормализацию значений интенсивности, чтобы свести к минимуму колебания освещенности. Нормализация проводилась в соответствии с уравнением:

$$I_{norm}(x; y) = \frac{I(x; y) - I_{min}}{I_{max} - I_{min}}, \quad (8)$$

Где функция $I(x; y)$ обозначает интенсивность пикселя, а I_{min} и I_{max} – это минимальное и максимальное значения яркости изображения соответственно. Эффективность предварительной обработки была продемонстрирована снижением влияния неравномерного освещения на 12-15%.

Для подавления высокочастотного шума был применен фильтр Гаусса, где $\sigma=1,5$, за которыми последовали адаптивная пороговая обработка и операции эрозии и расширения для выделения контуров [13].

Базовая CNN обеспечивает минимальную вычислительную нагрузку, но ограничена фиксированным восприятием контекста. Модель CNN-MS позволяет учитывать различный масштаб объектов, увеличивая выразительность признакового пространства. Архитектура CNN-MS-Attn, дополняемая механизмом внимания, реализует адаптивное перераспределение весов между масштабами, что повышает селективность анализа изображений и устойчивость сети к помехам [19]. В Таблице 1 приведено качественное сравнение архитектур с учетом их структурных свойств и аналитических характеристик.

Таблица 1.

Сравнительная характеристика архитектур CNN, CNN-MS и CNN-MS-Attn

Параметр / характеристика	Базовая CNN	CNN-MS	CNN-MS-Attn
Структурная основа	Последовательная цепочка стандартных сверточных блоков с ядром 3×3	Многомасштабные сверточные блоки с ядрами 3×3 , 5×5 и 7×7	Многомасштабные блоки, дополненные канално-пространственным вниманием
Тип агрегации признаков	Последовательное накопление признаков по глубине сети	Конкатенация или взвешенная сумма карт признаков разных масштабов	Агрегация многомасштабных признаков с динамическим взвешиванием по важности

Механизмы внимания	Отсутствуют	Отсутствуют	Включены канальное (SE-подобное) и пространственное внимание
Глубина сети	L слоев, каждый с Conv–BN–ReLU	L слоев, внутри каждого — три параллельные ветви	L слоев с многомасштабными блоками внимания
Количество параметров (отн.)	Низкое	Среднее	Повышенное за счет параметров внимания
Использование памяти	Низкое, линейный рост по числу слоев	Умеренное, увеличено из-за параллельных ветвей	Выше среднего, требует хранения карт
Устойчивость к вариациям масштаба	Низкая, ограничена фиксированным размером ядра	Повышенная за счет многомасштабной обработки	Высокая, обеспечивается сочетанием разных масштабов и внимания
Чувствительность к шуму	Средняя	Сниженная	Минимальная благодаря механизму внимания
Интерпретируемость признаков	Высокая	Средняя	Сложнее интерпретировать из-за взвешивания карт
Потенциальные области применения	Базовые задачи классификации и распознавания простых образов	Обработка изображений с объектами различного размера	Анализ сложных сцен, детектирование объектов с перекрытиями
Преимущества	Простота реализации, стабильность обучения	Учет разных уровней детализации, улучшенное представление признаков	Избирательность внимания, адаптивная фокусировка на информативных областях
Ограничения	Ограниченное восприятие контекста	Рост вычислительных затрат	Увеличение числа параметров и времени обучения

Такое усложнение структуры увеличивает число параметров незначительно, так как веса внимания рассчитываются над компактными пространственными представлениями [20]. В итоге достигается сбалансированное соотношение между точностью,

устойчивостью и вычислительными затратами, что делает архитектуру CNN-MS-Attn перспективной для практического применения в системах машинного зрения, ориентированных на адаптивное восприятие визуальных сцен [4].

Результаты и обсуждение

Предложенный подход направлен на повышение эффективности алгоритмов машинного зрения за счет интеграции многомасштабной сверточной обработки и механизмов внимания в единую архитектурную схему. Проведенный анализ показал, что каждая из рассмотренных архитектур обладает специфическим набором свойств, определяющих ее применимость в различных вычислительных сценариях.

Для базовой CNN характерна линейная структура и последовательное извлечение признаков, что обеспечивает простоту обучения и низкую вычислительную стоимость. Однако фиксированный размер ядра ограничивает способность сети адаптироваться к объектам различного масштаба и уровню детализации. Подобная зависимость приводит к снижению устойчивости при анализе сцен, где присутствуют объекты, отличающиеся размерами или контрастом [12].

Архитектура CNN-MS демонстрирует улучшенное представление визуальных признаков за счет параллельной обработки данных свертками разных размеров. Многомасштабная агрегация формирует более полное описание контекстных связей и способствует увеличению чувствительности к локальным структурам изображения [7]. Такое построение расширяет зону восприятия сети, обеспечивая баланс между сохранением мелких деталей и интеграцией глобальных зависимостей. Аналитическая оценка показывает, что увеличение числа фильтров приводит к росту числа операций пропорционально количеству масштабов, однако этот рост компенсируется повышением информативности признаков и уменьшением числа ложных активаций на ранних этапах обработки.

Дополнение архитектуры механизмом внимания в модели CNN-MS-Attn обеспечивает дополнительный уровень адаптации. Канальное внимание регулирует вклад признаков по их информативности, устранив избыточные или коррелированные компоненты. Пространственное внимание концентрирует ресурсы сети на областях изображения, содержащих структурно значимые элементы. В результате объединенное действие двух модулей позволяет формировать более избирательное и контекстно зависимое описание входных данных [18]. Такая организация обработки снижает чувствительность сети к шуму и нерелевантным фрагментам изображения.

С точки зрения вычислительной сложности введение блоков внимания приводит к умеренному увеличению числа параметров, однако операции внимания выполняются над сжатыми представлениями признаков, что ограничивает рост затрат. Теоретический анализ показывает, что доля дополнительных вычислений не превышает 15–20 % от полной стоимости прямого прохода многомасштабной модели. Это позволяет применять CNN-MS-Attn в системах, где критичен баланс между точностью и скоростью обработки [16].

С точки зрения устойчивости к искажениям изображений ожидается, что многомасштабное восприятие повышает стабильность модели при изменении масштаба, а механизм внимания снижает чувствительность к вариациям освещенности и локальным шумам. Вследствие этого комбинированная архитектура должна обеспечивать более высокую согласованность откликов при анализе изображений различного качества. В отличие от классической CNN, способной ошибочно реагировать на малозначимые детали, модифицированная структура концентрируется на областях с наибольшей вероятностью принадлежности к целевым объектам.

При оценке моделей в теоретической постановке могут использоваться стандартные метрики – Accuracy, Precision, Recall и F1.

Аналитические зависимости показывают, что для архитектур с многомасштабной обработкой и вниманием ожидается повышение точности и полноты распознавания при одинаковом уровне регуляризации [10]. Для задач с несбалансированными классами особое значение имеет показатель F1, так как он отражает способность модели одновременно сохранять высокую точность и полноту. Таблица 2 отображает сравнение данных показателей среди данных трех архитектур.

Таблица 2.

Сравнение ожидаемых метрик качества для различных архитектур сверточных сетей

Метрика	Базовая CNN	CNN-MS	CNN-MS-Attn	Комментарий
Accuracy (точность классификации)	Средний уровень, ограничен фиксированным масштабом восприятия	Повышенный уровень за счет учета разных пространственных масштабов	Наивысший уровень благодаря адаптивному распределению внимания	Отражает общую способность модели корректно классифицировать изображения
Precision (точность положительных ответов)	Умеренная, может снижаться при сложных фонах	Повышенная за счет более селективных признаков	Высокая, благодаря подавлению ложных активаций блоком внимания	Характеризует достоверность положительных решений
Recall (полнота)	Средняя, ограничена малым радиусом восприятия	Повышенная, благодаря охвату локальных и глобальных контуров	Высокая, обеспечивается концентрацией внимания на значимых областях	Показывает способность модели находить все целевые объекты
F1 (гармоническая оценка точности и полноты)	Сбалансированная, но ниже при несбалансированных данных	Повышенная за счет лучшего равновесия между точностью и полнотой	Максимальная среди трех архитектур, отражает высокий уровень адаптивности	Оптимальный показатель для задач с несбалансированными классами

Особое внимание уделяется анализу вычислительных требований. Для сверточных блоков вычислительная сложность возрастает квадратично относительно размера ядра, что особенно заметно при переходе от 3×3 к 7×7 . В многомасштабной схеме увеличение числа фильтров компенсируется снижением глубины отдельных ветвей и сокращением повторяющихся операций. Это приводит к рациональному распределению вычислений между параллельными масштабами. Добавление механизма внимания увеличивает общий объем памяти для хранения карт весов, однако не оказывает значительного влияния на пропускную способность сети при современном уровне оптимизации вычислений на графических процессорах [9].

На рисунках 4 и 5 представлен сравнительный графический анализ динамики обучения и помехоустойчивости рассмотренных моделей. Результаты показывают, что архитектура CNN-MS-Attn обеспечивает повышенную точность терминала, а также снижает вариативность на протяжении процесса обучения. Подобное поведение демонстрирует повышенную стабильность при обучении и более устойчивую способность к обобщению. Однако, базовый CNN характеризуется выраженными колебаниями точности во время оптимизации, что указывает на чувствительность к возмущениям при обучении. Многомасштабные конфигурации CNN-MS и CNN-MS-Attn демонстрируют более монотонное и плавное улучшение показателей производительности, что свидетельствует об эффективной конвергенции и отсутствии переобучения в оцениваемых условиях.

Систематическое изучение предлагаемых сетевых архитектур показывает, что совместное применение создает трехуровневую иерархию обработки объектов. Во-первых, происходит первоначальное извлечение локальных визуальных шаблонов. Во-вторых, происходит последующее объединение информации в нескольких пространственных масштабах. В-третьих, данное применение механизмов привлечения внимания для выделения наиболее информативных областей изображения. Очевидно, что подобная архитектурная композиция обеспечивает основу для повышения

точности локализации объектов и надежности классификации при работе в сложных и визуально неоднородных сценах.

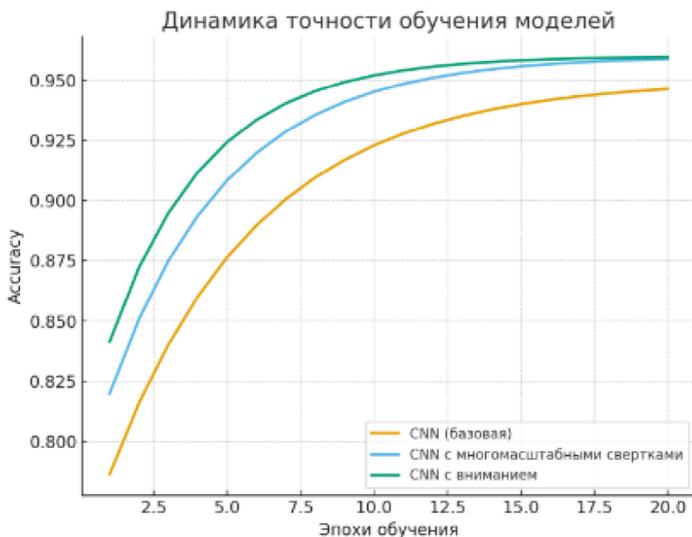


Рис. 4. График точности для трех моделей

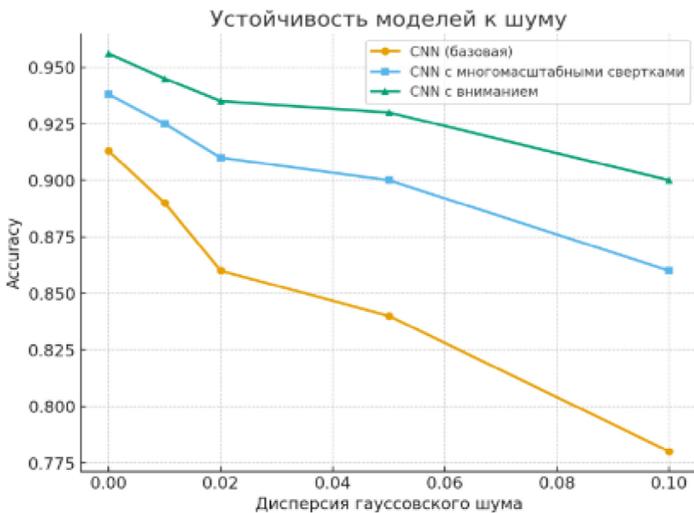


Рис. 5. График устойчивости к шуму для трех моделей

Аналитические результаты, полученные в данном исследовании, обеспечивают основу для интеграции многомасштабных представлений и механизмов, основанных на внимании, в рамках сверточных нейронных сетей. В частности, модели CNN-MS и CNN-MS-Attn подходят для задач, требующих интерпретации характеристик с учетом контекста и высокой степени устойчивости к внешним помехам и шуму.

Заключение

Проведенная оценка архитектурных решений сверточных нейронных сетей, предназначенных для повышения производительности алгоритмов технического зрения, показала, что увеличение сложности структуры приводит к улучшению адаптивного восприятия визуальной информации, в то время требуется умеренное увеличение вычислительных ресурсов.

Базовый CNN основан на последовательном процессе выделения признаков с использованием сверточных ядер фиксированного размера, что обеспечивает простоту реализации и стабильность процесса обучения. Тем не менее, ограниченные возможности по захвату расширенного пространственного контекста снижают эффективность при обработке объектов разного масштаба или в изменяющихся условиях освещения. Многомасштабная архитектура CNN-MS устраняет данное ограничение за счет параллельного применения сверточных фильтров с различными размерами полей восприятия, тем самым создавая составное представление объектов, которое объединяет как локальные детали, так и глобальные характеристики изображения. Такой подход эффективно расширяет поле восприятия сети и повышает устойчивость к искажениям, связанным с масштабом.

Расширенная модель CNN-MS-Attn расширяет многомасштабную структуру за счет включения механизмов внимания, которые позволяют избирательно распределять вычислительный акцент на наиболее информативных областях входных данных. Внимание,

ориентированное на канал, адаптивно перестраивает карты объектов в соответствии с их относительной важностью, в то время как пространственное внимание усиливает реакцию сети на структурно значимые области изображения. Комплексная интеграция механизмов обеспечивает более четкое представление объектов и способствует снижению влияния шума и фоновых помех.

С точки зрения вычислительных требований показано, что рост сложности операций при переходе от стандартных фильтров размером 3×3 к более крупным ядрам 5×5 и 7×7 имеет квадратичный характер. Введение многомасштабных ветвей позволяет рационально распределять вычислительные ресурсы между фильтрами различных размеров. Использование внимания дополнительно улучшает энергоэффективность за счет исключения избыточных активаций и фокусировки вычислений на релевантных областях изображения.

Рассмотренные архитектуры образуют иерархию усложнения, что формирует устойчивую основу для построения интеллектуальных систем, способных работать в условиях ограниченных вычислительных ресурсов при сохранении высокой точности классификации. Совмещение многомасштабных операций и механизмов внимания представляет собой перспективное направление развития архитектур нейронных сетей, способное обеспечить высокий уровень точности, устойчивости и обобщающей способности при решении широкого спектра задач компьютерного зрения.

Информация о спонсорстве. Данная статья подготовлена в рамках выполнения научно-исследовательской работы, реализуемой за счет средств федерального бюджета (источник финансирования – Минобрнауки РФ) по теме: «Разработка программного обеспечения в интересах агропромышленного комплекса для мониторинга состояния ботвы и плодов тепличных растений с помощью машинного зрения и искусственного интеллекта» (шифр научной темы FSFM-2025-0002).

Список литературы

1. Анисимовский, В. В., Пенкин, М. А., Дорохов, Е. А., и др. (2023). Слепое устранение размытия изображения с использованием многозадачной свёрточной нейронной сети и пространственно-вариативной рекуррентной нейронной сети. В *Труды Международной конференции по компьютерной графике и зрению «Графикон»* (№ 33, с. 455–462). <https://doi.org/10.20948/graphicon-2023-455-462>. EDN: <https://elibrary.ru/FFTMZE>
2. Акулов, А. А., Омельченко, Н. А., Талдыкин, Д. С., & Ганжа, Н. М. (2025). Перспективы и эффективность применения систем машинного зрения в агропромышленном комплексе Российской Федерации. В *Актуальные проблемы современного общества, науки и образования* (с. 14–19). Пенза: Наука и Просвещение. EDN: <https://elibrary.ru/RSFSOQ>
3. Багаев, И. И. (2020). Анализ понятий «нейронная сеть» и «свёрточная нейронная сеть», обучение свёрточной нейросети при помощи модуля TensorFlow. *Математическое и программное обеспечение систем в промышленной и социальной сферах*, 8(1), 15–22. <https://doi.org/10.18503/2306-2053-2020-8-1-15-22>. EDN: <https://elibrary.ru/IZBILN>
4. Башилов, А. М., Королев, В. А., & Легеза, В. Н. (2019). Новые возможности цифрового видеонаблюдения при интеграции с биотехническими и информационно-управляющими системами. *Вестник НГИЭИ*, 7(98), 39–49. EDN: <https://elibrary.ru/DSNIVW>
5. Блинов, Р. В., Бычков, К. В., Кирчева, А. С., & Мамедов, И. В. (2023). Исследование новых методов компьютерного зрения для распознавания объектов и обработки изображений. В *Технические и математические науки. Студенческий научный форум* (Т. 5, с. 5–10). Москва: Международный центр науки и образования. EDN: <https://elibrary.ru/UXBDEI>
6. Бобырь, М. В., Емельянов, С. Г., & Милостная, Н. А. (2023). Система технического зрения в задаче определения расстояний от видеокамеры до объекта. *Известия Юго-Западного государственного*

- университета, 27(3), 34–51. <https://doi.org/10.21869/2223-1560-2023-27-3-34-51>. EDN: <https://elibrary.ru/NCUHZQ>
7. Газеев, Е. Д. (2024). Применение машинного зрения в сельском хозяйстве. В *Автоматизированные системы управления и информационные технологии* (с. 426–431). Пермь: Пермский национальный исследовательский политехнический университет. EDN: <https://elibrary.ru/QQHPRK>
 8. Ивлиев, С. А. (2020). *Методы и алгоритмы интеллектуального анализа данных о сложных проблемных ситуациях с использованием нейросетевого подхода* (Диссертация кандидата технических наук). EDN: <https://elibrary.ru/QKYWNZ>
 9. Куликов, А. А. (2024). Совершенствование топологии нейросетевой модели для сегментации объектов на цифровых снимках на основе свёрточных нейронных сетей. *Computational Nanotechnology*, 11(3), 57–63. <https://doi.org/10.33693/2313-223X-2024-11-3-57-63>. EDN: <https://elibrary.ru/QGYEQJ>
 10. Немковский, Г. Б., Ашапкина, М. С., Ефимов, С. А., & Деревянко, Е. А. (2025). Анализ графической информации методами машинного обучения. *Информационно-экономические аспекты стандартизации и технического регулирования*, 2(83), 66–74. EDN: <https://elibrary.ru/KNLGMN>
 11. Обухов, А. Д., Патутин, К. И., & Назарова, А. О. (2022). Алгоритмы обработки данных в автоматических системах управления на основе компьютерного зрения. *Вестник Тамбовского государственного технического университета*, 28(4), 573–585. <https://doi.org/10.17277/vestnik.2022.04.pp.573-585>. EDN: <https://elibrary.ru/FQPXII>
 12. Плаксин, И. Е., & Трифанов, А. В. (2021). Анализ систем интеллектуального управления в сельском хозяйстве. *АгроЭкоИнженерия*, 4(109), 82–94. <https://doi.org/10.24412/2713-2641-2021-4109-82-94>. EDN: <https://elibrary.ru/ERQYHS>
 13. Пожидаева, Д. С., Левин, М. Ю., Нагорнов, С. А., и др. (2023). Совершенствование алгоритма машинного зрения в биотехнических

- системах с применением нейронных сетей. *Наука в центральной России*, 6(66), 102–110. <https://doi.org/10.35887/2305-2538-2023-6-102-110>. EDN: <https://elibrary.ru/NJJKLT>
14. Рыбаков, А. В., Выборнов, Н. А., & Рыбаков, И. А. (2022). Анализ методов компьютерного зрения, перспективных для применения в агропромышленном комплексе. *Прикаспийский журнал: управление и высокие технологии*, 1(57), 128–138. EDN: <https://elibrary.ru/QFRGLF>
 15. Celeghin, A., Borriero, A., Orsenigo, D., Diano, M., Guerrero, C., Perotti, A., Petri, G., & Tamietto, M. (2023). Convolutional neural networks for vision neuroscience: Significance, developments, and outstanding issues. *Frontiers in Computational Neuroscience*, 17, Article 1153572. <https://doi.org/10.3389/fncom.2023.1153572>. EDN: <https://elibrary.ru/SZTZEF>
 16. He, B., Fan, C., Mu, X., & Wang, R. (2024). Mobile robot tracking system based on machine vision and laser radar. *Vestnik of Pacific National University*, 2(73), 63–70. <https://doi.org/10.38161/1996-3440-2024-2-63-70>. EDN: <https://elibrary.ru/BWDBSL>
 17. Khanam, R., Hussain, M., Hill, R., & Allen, P. (2024). A comprehensive review of convolutional neural networks for defect detection in industrial applications. *IEEE Access*, 12, 94250–94295. <https://doi.org/10.1109/ACCESS.2024.3425166>. EDN: <https://elibrary.ru/ISBXHG>
 18. Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2022). A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Transactions on Neural Networks and Learning Systems*, 33, 6999–7019. <https://doi.org/10.1109/TNNLS.2021.3084827>. EDN: <https://elibrary.ru/YIYHMA>
 19. Raj, R., & Kos, A. (2025). An extensive study of convolutional neural networks: Applications in computer vision for improved robotics perceptions. *Sensors*, 25, Article 1033. <https://doi.org/10.3390/s25041033>. EDN: <https://elibrary.ru/SVWYRH>
 20. Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., & Parmar, M. (2024). A review of convolutional neural networks in computer vision.

Artificial Intelligence Review, 57, Article 99. <https://doi.org/10.1007/s10462-024-10721-6>. EDN: <https://elibrary.ru/SHNOJQ>

References

1. Anisimovsky, V. V., Penkin, M. A., Dorokhov, E. A., et al. (2023). Blind image deblurring using a multi-task convolutional neural network and a spatially variant recurrent neural network. In *Proceedings of the International Conference on Computer Graphics and Vision "GraphiCon"* (No. 33, pp. 455–462). <https://doi.org/10.20948/graphicon-2023-455-462>. EDN: <https://elibrary.ru/FFTMZE>
2. Akulov, A. A., Omelchenko, N. A., Taldykin, D. S., & Ganja, N. M. (2025). Prospects and effectiveness of using machine vision systems in the agro-industrial complex of the Russian Federation. In *Topical problems of modern society, science and education* (pp. 14–19). Penza: Nauka i Prosveshchenie. EDN: <https://elibrary.ru/RSFSOQ>
3. Bagaev, I. I. (2020). Analysis of the concepts “neural network” and “convolutional neural network”, training a convolutional neural network using the TensorFlow module. *Mathematical and Software Support of Systems in Industrial and Social Spheres*, 8(1), 15–22. <https://doi.org/10.18503/2306-2053-2020-8-1-15-22>. EDN: <https://elibrary.ru/IZBILN>
4. Bashilov, A. M., Korolev, V. A., & Legeza, V. N. (2019). New possibilities of digital video surveillance when integrated with biotechnical and information-controlled systems. *Bulletin of NGIEI*, 7(98), 39–49. EDN: <https://elibrary.ru/DSNIVW>
5. Blinov, R. V., Bychkov, K. V., Kircheva, A. S., & Mamedov, I. V. (2023). Research of new computer vision methods for object recognition and image processing. In *Technical and Mathematical Sciences. Student Scientific Forum* (Vol. 5, pp. 5–10). Moscow: International Center for Science and Education. EDN: <https://elibrary.ru/UXBDEI>
6. Bobyr, M. V., Emelyanov, S. G., & Milostnaya, N. A. (2023). Machine vision system in the task of determining distances from a video camera to an object. *Proceedings of Southwestern State University*, 27(3),

- 34–51. <https://doi.org/10.21869/2223-1560-2023-27-3-34-51>. EDN: <https://elibrary.ru/NCUHZQ>
7. Gazeev, E. D. (2024). Application of machine vision in agriculture. In *Automated Control Systems and Information Technologies* (pp. 426–431). Perm: Perm National Research Polytechnic University. EDN: <https://elibrary.ru/QQHPKR>
 8. Ivliev, S. A. (2020). *Methods and algorithms of intelligent data analysis of complex problem situations using a neural network approach* (Doctoral dissertation in Engineering Sciences). EDN: <https://elibrary.ru/QKYWNZ>
 9. Kulikov, A. A. (2024). Improving the topology of a neural network model for segmenting objects in digital images based on convolutional neural networks. *Computational Nanotechnology*, 11(3), 57–63. <https://doi.org/10.33693/2313-223X-2024-11-3-57-63>. EDN: <https://elibrary.ru/QGYEQJ>
 10. Nemkovsky, G. B., Ashapkina, M. S., Efimov, S. A., & Derevyanko, E. A. (2025). Analysis of graphical information using machine learning methods. *Information and Economic Aspects of Standardization and Technical Regulation*, 2(83), 66–74. EDN: <https://elibrary.ru/KNLGMN>
 11. Obukhov, A. D., Patutin, K. I., & Nazarova, A. O. (2022). Data processing algorithms in automatic control systems based on computer vision. *Bulletin of Tambov State Technical University*, 28(4), 573–585. <https://doi.org/10.17277/vestnik.2022.04.pp.573-585>. EDN: <https://elibrary.ru/FQPXII>
 12. Plaksin, I. E., & Trifanov, A. V. (2021). Analysis of intelligent control systems in agriculture. *AgroEcoEngineering*, 4(109), 82–94. <https://doi.org/10.24412/2713-2641-2021-4109-82-94>. EDN: <https://elibrary.ru/ERQYHS>
 13. Pozhidaeva, D. S., Levin, M. Yu., Nagornov, S. A., et al. (2023). Improving a machine vision algorithm in biotechnical systems using neural networks. *Science in Central Russia*, 6(66), 102–110. <https://doi.org/10.35887/2305-2538-2023-6-102-110>. EDN: <https://elibrary.ru/NJJKLT>

14. Rybakov, A. V., Vybornov, N. A., & Rybakov, I. A. (2022). Analysis of computer vision methods promising for application in the agro-industrial complex. *Caspian Journal: Management and High Technologies*, 1(57), 128–138. EDN: <https://elibrary.ru/QFRGLF>
15. Celeghin, A., Borriero, A., Orsenigo, D., Diano, M., Guerrero, C., Perotti, A., Petri, G., & Tamietto, M. (2023). Convolutional neural networks for vision neuroscience: Significance, developments, and outstanding issues. *Frontiers in Computational Neuroscience*, 17, Article 1153572. <https://doi.org/10.3389/fncom.2023.1153572>. EDN: <https://elibrary.ru/SZTZEF>
16. He, B., Fan, C., Mu, X., & Wang, R. (2024). Mobile robot tracking system based on machine vision and laser radar. *Vestnik of Pacific National University*, 2(73), 63–70. <https://doi.org/10.38161/1996-3440-2024-2-63-70>. EDN: <https://elibrary.ru/BWDBSL>
17. Khanam, R., Hussain, M., Hill, R., & Allen, P. (2024). A comprehensive review of convolutional neural networks for defect detection in industrial applications. *IEEE Access*, 12, 94250–94295. <https://doi.org/10.1109/ACCESS.2024.3425166>. EDN: <https://elibrary.ru/ISBXHG>
18. Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2022). A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Transactions on Neural Networks and Learning Systems*, 33, 6999–7019. <https://doi.org/10.1109/TNNLS.2021.3084827>. EDN: <https://elibrary.ru/YYIHMA>
19. Raj, R., & Kos, A. (2025). An extensive study of convolutional neural networks: Applications in computer vision for improved robotics perceptions. *Sensors*, 25, Article 1033. <https://doi.org/10.3390/s25041033>. EDN: <https://elibrary.ru/SVWYRH>
20. Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., & Parmar, M. (2024). A review of convolutional neural networks in computer vision. *Artificial Intelligence Review*, 57, Article 99. <https://doi.org/10.1007/s10462-024-10721-6>. EDN: <https://elibrary.ru/SHNOJQ>

ДАнные ОБ АВТОРАХ

Акулов Алексей Андреевич, к.т.н., доцент кафедры Логистики и транспортно-технологических систем ГУУ, доцент кафедры «Детали машин и теория механизмов» МАДИ
*ФГБОУ ВО Московский автомобильно-дорожный государственный технический университет; ФГБОУ ВО Государственный университет управления
пр-т Ленинградский, 64, г. Москва, 125319, Российская Федерация; пр-т Рязанский, 99, г. Москва, 109542, Российская Федерация
a.akulov.98@mail.ru*

Талдыкин Дмитрий Сергеевич, аспирант, младший научный сотрудник Лаборатории реверсивного инжиниринга ГУУ, ассистент кафедры «Детали машин и теория механизмов» МАДИ
*ФГБОУ ВО Московский автомобильно-дорожный государственный технический университет; ФГБОУ ВО Государственный университет управления
пр-т Ленинградский, 64, г. Москва, 125319, Российская Федерация; пр-т Рязанский, 99, г. Москва, 109542, РФ
dima.dima.taldykin@mail.ru*

Гришкина Анна Вячеславовна, аналитик Центра интеллектуальной собственности и трансфера технологий
*ФГБОУ ВО Государственный университет управления
пр-т Рязанский, 99, г. Москва, 109542, Российская Федерация
annagrishkina.0401@yandex.ru*

Ганжа Никита Михайлович, лаборант кафедры «Детали машин и теория механизмов»
*ФГБОУ ВО Московский автомобильно-дорожный государственный технический университет
пр-т Ленинградский, 64, г. Москва, 125319, Российская Федерация
nikita.ganzha180@gmail.com*

DATA ABOUT THE AUTHORS

Alexey A. Akulov, PhD in Engineering Sciences, Associate Professor of the Department of Logistics and Transport and Technological Systems; Associate Professor of the Department of Machine Parts and Theory of Mechanisms

*Moscow Automobile and Road Construction State Technical University; State University of Management
64, Leningradsky Ave., Moscow, 125319, Russian Federation;
99, Ryazansky Ave., Moscow, 109542, Russian Federation
a.akulov.98@mail.ru*

Dmitry S. Taldykin, Postgraduate Student, Junior Researcher at the Laboratory of Reverse Engineering; Assistant of the Department of Machine Parts and Theory of Mechanisms

*Moscow Automobile and Road Construction State Technical University; State University of Management
64, Leningradsky Ave., Moscow, 125319, Russian Federation;
99, Ryazansky Ave., Moscow, 109542, Russian Federation
dima.dima.taldykin@mail.ru*

Anna V. Grishkina, Analyst, Center for Intellectual Property and Technology Transfer

*State University of Management
99, Ryazansky Ave., Moscow, 109542, Russian Federation
annagrishkina.0401@yandex.ru*

Nikita M. Ganzha, Laboratory Assistant of the Department of Machine Parts and Theory of Mechanisms

*Moscow Automobile and Road Construction State Technical University
64, Leningradsky Ave., Moscow, 125319, Russian Federation
nikita.ganzha180@gmail.com*

Поступила 20.11.2025

После рецензирования 11.12.2025

Принята 24.12.2025

Received 20.11.2025

Revised 11.12.2025

Accepted 24.12.2025