

УДК 004.896

**ДВУХЭТАПНОЕ РАСПОЗНАВАНИЕ  
ГРАФИЧЕСКИХ ПРИМИТИВОВ В ВИДЕОПОТОКЕ****Долженкова М.Л., Прозорова Т.Г., Чистяков Г.А., Эсаулов О.К.***ФГБОУ ВО «Вятский государственный университет», Киров,**e-mail: gennadiychistyakov@gmail.com*

В статье предлагается двухэтапный подход к распознаванию графических примитивов в видеопотоке, суть которого состоит в использовании независимых детекторов и классификаторов, предназначенных для работы при различных состояниях внешней среды. С помощью данного подхода решается одна из важных проблем машинного зрения – проблема идентификации сигналов светофора. Актуальность исследования обосновывается отсутствием доступных средств, пригодных для применения в реальных условиях. Отличительной чертой является также применение на этапе детектирования нейросетевых моделей глубокого обучения, использующих в своей архитектуре семейство сверточных слоев. При этом предлагается использовать предобученные модели с измененными весовыми коэффициентами заключительных полносвязных слоев. В работе описываются особенности и недостатки современных методов решения задач данного класса, приводятся основные концепции и положения предлагаемого подхода, а также рассматриваются основные этапы проектирования структуры системы машинного зрения, основанной на данном подходе. Приведенная структура может являться основой для создания более гибкой и адаптированной к реальным условиям системы детектирования и классификации светофоров. Кроме того, предложенный подход в целом может быть обобщен для эффективного решения задач, имеющих аналогичную формальную постановку.

**Ключевые слова:** компьютерное зрение, машинное обучение, детектирование объектов, распознавание объектов, обнаружение и распознавание светофоров

**TWO-STAGE RECOGNITION OF GRAPHIC PRIMITIVES IN VIDEO FLOW****Dolzhenkova M.L., Prozorova T.G., Chistyakov G.A., Esaulov O.K.***Federal Budget Educational Institution of Higher Education Vyatka State University,**Kirov, e-mail: gennadiychistyakov@gmail.com*

The article proposes a two-step approach to recognizing graphic primitives in a video stream, the essence of which is the use of independent detectors and classifiers designed to work in various states of the external environment. One of the important problems of machine vision is solved with this approach – the problem of identifying traffic lights. The relevance of the study is justified by the lack of available tools that are suitable for use in real conditions. A distinctive feature is also the use of deep learning neural network models that use a family of convolutional layers in their architecture. It is proposed to use pre-trained models with modified weight coefficients of the final fully connected layers. The paper describes the features and disadvantages of modern methods for solving problems of this class, describes the main concepts and provisions of the proposed approach, and also discusses the main stages of designing the structure of the computer vision system based on this approach. This structure can be the basis for creating a more flexible and adapted to the real conditions system of detection and classification of traffic lights. In addition, the proposed approach as a whole can be generalized to effectively solve problems that have a similar formulation.

**Keywords:** computer vision, machine learning, object detection, object recognition, detection and recognition of traffic lights

В настоящее время теория искусственного интеллекта является передовым математическим аппаратом, повсеместно внедряемым во многие отрасли человеческой деятельности. Машинное обучение, как одно из направлений искусственного интеллекта, активно развивается и на сегодняшний день применяется как эффективный инструмент решения широкого спектра задач, отличающихся низкой степенью формализуемости и требующих обработки больших объемов неструктурированных или слабоструктурированных данных.

С развитием машинного обучения совершенствуются и другие технологии, основанные на нем. В частности, в последние годы значительно возросли возможности компьютерного зрения, которое позволя-

ет автоматизировать процессы, связанные с извлечением информации [1]. Так, например, становится возможным создание интеллектуальных систем автоматического и вспомогательного управления транспортным средством, ориентирующихся в пространстве и адекватно распознающих происходящие во внешней среде события [2].

Одной из частных проблем, которую необходимо решить при разработке подобного рода систем, является проблема обнаружения в видеопотоке специфических графических примитивов [3]. На основе проведенного анализа был сделан вывод, что на данный момент применяемые для распознавания светофоров подходы и методы дают приемлемое для практического использования качество только в условиях, близких

к идеальным [2–4]. В связи с этим был сделан вывод об актуальности разработки системы, способной решать рассмотренную задачу вне зависимости от состояния внешней среды.

*Современное состояние проблемы.* Существующие методы выделения графических примитивов в видеопотоке способны определить, имеются ли на изображении интересующие пользователя элементы, и указать, где именно находится каждый объект и какова занимаемая им область [5]. При этом, как правило, выполняется сканирование входного изображения окном и классификация подученных результатов («интересующий объект» или «не интересующий объект»). Основными требованиями, предъявляемыми к данным методам, являются: высокое качество распознавания, работа в режиме реального времени и устойчивость к факторам внешней среды. Максимально эффективно с задачей классификации справляются методы машинного обучения. Однако, на практике один и тот же объект, который необходимо обнаружить или распознать, представляется абсолютно различным вектором пикселей даже при малейших воздействиях внешних факторов, что затрудняет корректное обучение классификатора и дальнейшее его использование [6]. Данная проблема решается с помощью нахождения характерных признаков объекта (цвет, форма, признаки Хаара, локальные бинарные шаблоны, направленные градиенты), которые определяют его уникальные черты и особенности.

Ведущие мировые автомобилестроительные концерны оснащают свои автомобили автономными или полуавтономными системами управления, использующими компьютерное зрение. Так, компания Tesla предложила систему *Navigate on Autopilot*, которая позволяет распознавать и реагировать на сигналы светофора и знаки «Движение без остановки запрещено», самостоятельно проезжать перекрестки с круговым движением. Audi Group разработала и презентовала интеллектуальную систему, информирующую автомобилистов о работе светофоров в городе. Электронный помощник под названием *Green Light Optimized Speed Advisory* позволяет распознавать расстояние до следующего перекрестка, зеленый сигнал светофора, а также оптимизировать скорость движения.

К сожалению, особенности реализации подобных систем не находятся в открытом доступе. Однако анализ свободных решений, не сопоставимых с коммерческими по качеству, позволяет сделать вывод о том, что используемый в большинстве работ подход

применения единой универсальной модели в качестве детектора и классификатора объектов в реальной жизни крайне затруднителен, поскольку является слишком обобщенным и не может учитывать множество различных факторов, условий и ситуаций (время суток, состояние внешней среды, локальные особенности местности применения системы и так далее) [7].

Кроме того, следует отметить, что базовые методы детектирования, основанные на классических подходах, не обеспечивают высокой точности обнаружения, поэтому в настоящее время более перспективными являются методы на основе глубокого обучения.

Из вышерассмотренного следует также, что рациональным является построение системы на основе композиции двух независимых подсистем: детектирования и классификации, – каждая из которых будет функционировать с применением своего математического аппарата [5].

#### **Материалы и методы исследования**

В ходе работы были детально исследованы несколько возможных вариантов организации системы [8], в результате чего были сделаны следующие выводы.

1. При применении для детектирования методов глубокого обучения целесообразно использовать стороннюю предобученную модель. Переориентация модели на поиск примитивов нужного типа может быть осуществлена за счет удаления верхних слоев сети с сохранением сверточной структуры, а также последующим дообучением нужного числа полносвязных слоев с определением соответствующих весовых коэффициентов посредством дообучения на наборе заранее подготовленных примеров.

2. Эффективное решение задачи классификации возможно с применением классических методов машинного обучения.

3. Для достижения высокого качества работы системы при различных условиях предлагается использование нескольких детекторов и классификаторов, предназначенных для конкретных состояний внешней среды и выбираемых динамически на основе сведений извне.

4. Приемлемые с точки зрения качества результаты достижимы при последовательной обработке видеопотока с определенной частотностью и дальнейшей агрегацией сведений, полученных на смежных кадрах.

5. Совершенствование работы подсистем может быть достигнуто путем использования групп эвристик, предназначенных для устранения дефектов (например, вероятностная реализация отсека ложных срабатываний).

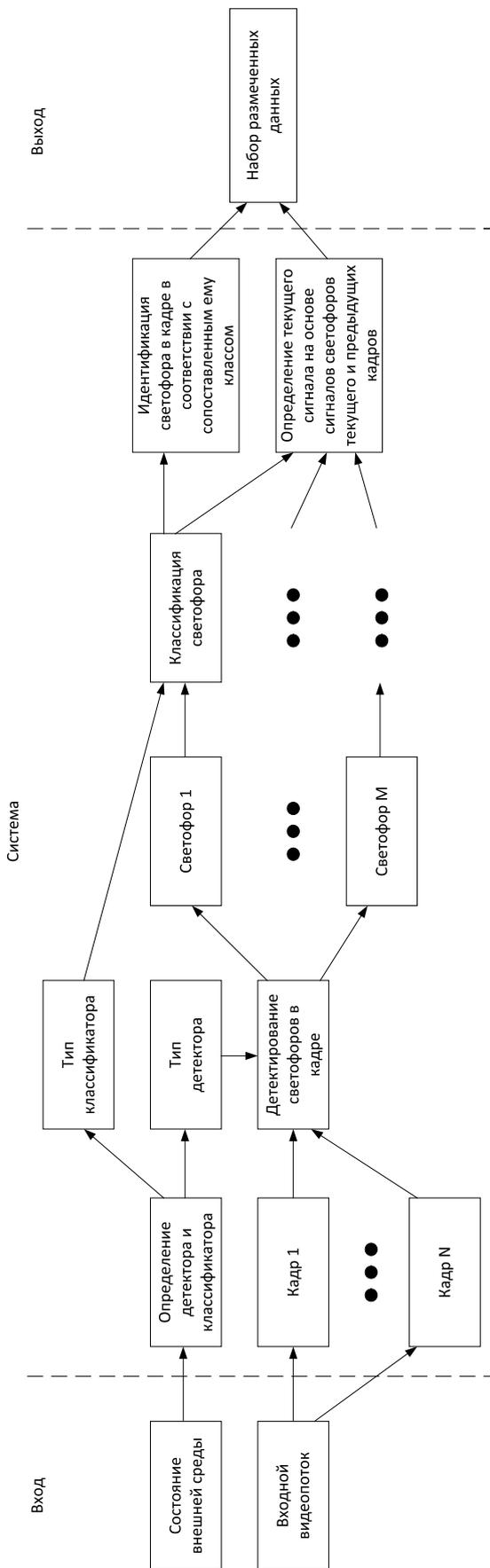


Рис. 1. Общая схема работы системы

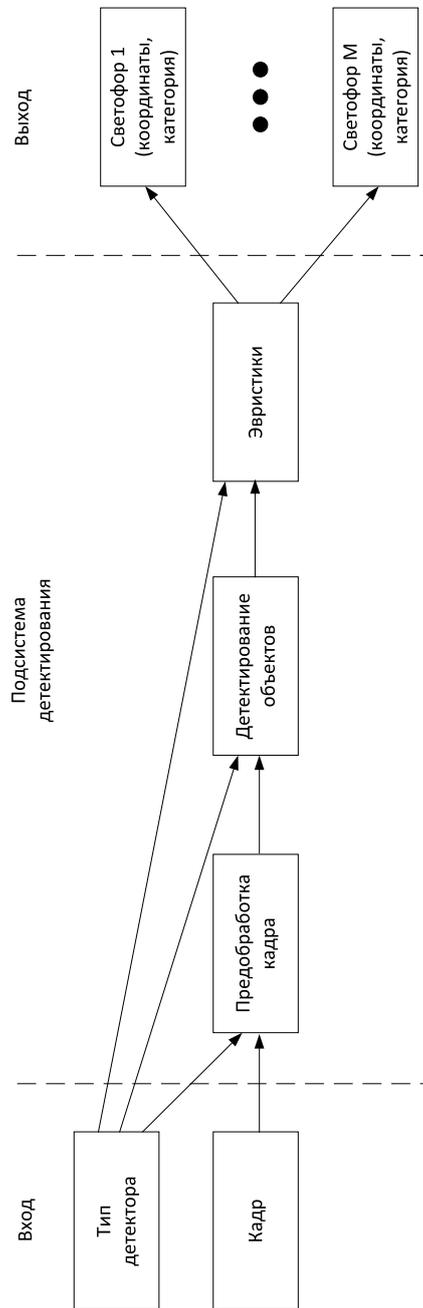


Рис. 2. Общая схема работы подсистемы детектирования

Определенная в соответствии с данными тезисами общая схема работы системы представлена на рис. 1.

Входными данными для системы является видеопоток, а также описание внешней ситуации, на основании которого осуществляется выбор подходящего детектора и классификатора. В частном случае выбор может осуществляться вручную. Поступивший на вход видеопоток разбивается на кадры, после чего осуществляется анализ полученных изображений в соответствии с определенной частотой: в зависимости от характеристик видеопотока, производительности имеющихся аппаратных ресурсов и требования выполнения обработки в режиме реального времени детектирование примитивов может происходить как на всех кадрах, так и на их подмножестве, выбранном с определенным периодом.

Задача обнаружения светофора решается подсистемой детектирования, схема работы которой представлена на рис. 2.

В первую очередь изображение проходит необходимую предварительную обработку – изменение размеров, повышение качества, устранение шумов. После этого кадр подается на вход соответствующей обученной нейросетевой сверточной модели, предназначенной для детектирования светофоров при текущей внешней ситуации [9].

Результатом работы сети является множество пар координат, ограничивающих найденные объекты, а также вероятности принадлежности объектов к определенным категориям (вертикальная ориентация, горизонтальная ориентация, трамвайный светофор и так далее). Далее к полученным данным применяется ряд эвристических алгоритмов (фильтрация по размерам, фильтрация по вероятностям принадлежности и так далее), после чего они выдаются подсистемой. Важно отметить, что совокуп-

ность применяемых эвристик определяется типом используемого детектора.

Далее найденные фрагменты необходимо классифицировать: определить текущий сигнал каждого светофора и дополнительную информацию (наличие вспомогательного сигнала-стрелки или дополнительной секции, ситуацию использования неактивной секции в качестве таймера и так далее). Задача классификации решается подсистемой, схема работы которой представлена на рис. 3.

На вход подсистемы подается изображение светофора, его категория, а также непосредственно классификатор. Следует отметить, что именно тип классификатора и категория определяют логику работы подсистемы.

На первом этапе входное изображение проходит предобработку – нормализуются яркость и контрастность, выполняется необходимая фильтрация, а также ряд аффинных преобразований. Затем из изображения выделяются характерные признаки, которые являются базой для классификации, его моделью машинного обучения. Если тип классификатора и категория светофора позволяют, то возможно одновременное использование нескольких моделей, причем как в виде ансамбля (для повышения точности распознавания сигнала), так и в виде независимого набора, выявляющего различные параметры (сигнал, присутствие на фрагменте активного таймера и так далее).

Так как наличие нескольких параметров экспоненциально повышает вероятность ложной классификации, на данном этапе используется ряд эвристических оптимизаций, направленных на исправление возможных ошибок. На выходе подсистема сообщает класс, соответствующий изображению светофора.

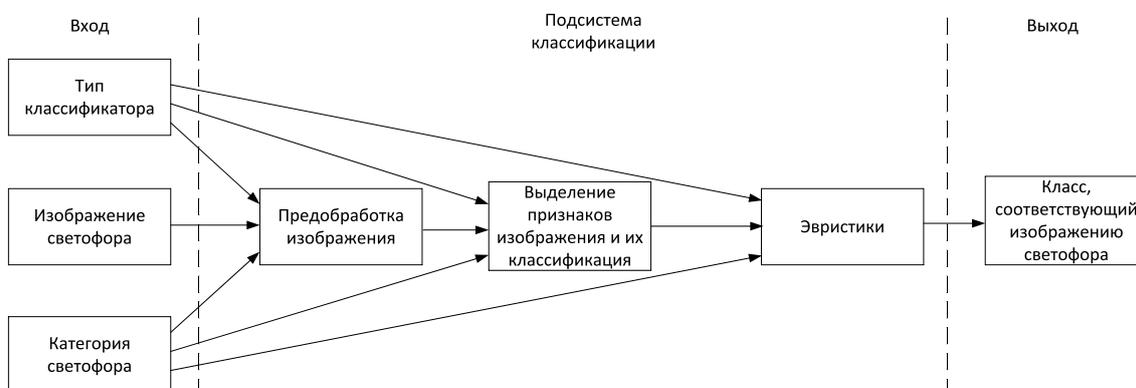


Рис. 3. Общая схема работы подсистемы классификации

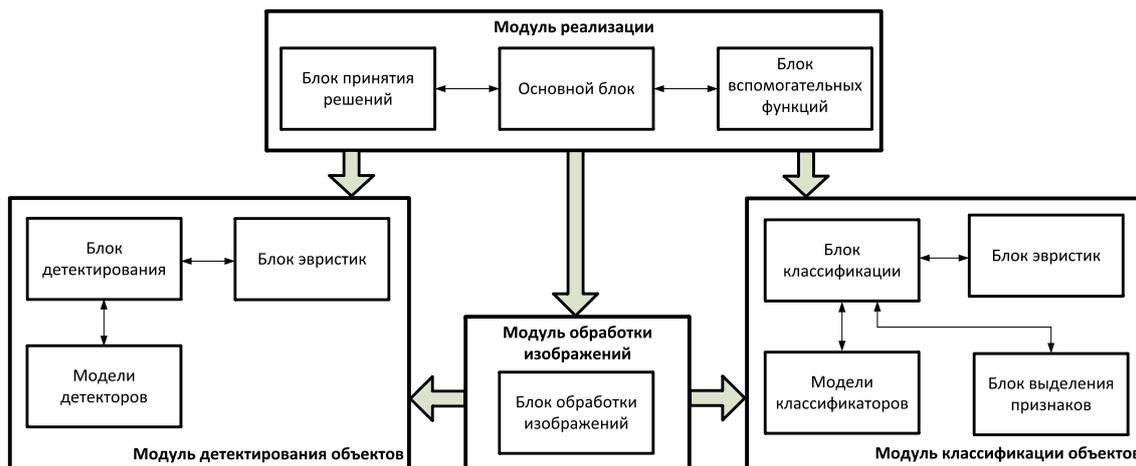


Рис. 4. Структура системы

По завершении этапов детектирования и классификации к полученным данным применяется набор комплексных эвристик, основанных на анализе цепочки сведений, полученных при обработке предшествующих кадров. Данное решение позволяет значительно повысить качество работы системы в целом, за счет выявления ложных срабатываний, вызванных быстроизменяющейся внешней средой. Примером подобной ситуации является восприятие системой нехарактерного блика в солнечную погоду за желтый сигнал светофора.

В итоге с каждым кадром сопоставляется набор размеченных данных, характеризующих все найденные объекты и включающих в себя сведения об их координатах на изображении, принадлежности к определенной категории и коэффициенту уверенности в принятом решении, основной сигнал светофора, а также дополнительную информацию о наличии стрелок, таймеров и тому подобного.

#### Результаты исследования и их обсуждение

На основе представленных схем была предложена структура системы распознавания светофоров в видеопотоке, изображенная на рис. 4 и включающая в себя четыре основных модуля: реализации, детектирования, классификации и обработки.

Центральным является модуль реализации, обеспечивающий синхронную работу прочих компонентов системы. В основном блоке реализуется рассмотренный выше алгоритм детектирования и классификации. Блок вспомогательных функций отвечает за дополнительный функционал: выделение найденных примитивов в кадре, вывод от-

ладочной информации в консоль, визуализация обработанного видеопотока в отдельном окне по запросу, сохранение данных на внешнем хранилище и так далее. Блок принятия решений содержит дополнительную логику и реализует набор комплексных эвристик, позволяющих определять основной сигнал светофора на основе цепочки проанализированных кадров.

Модуль детектирования объектов состоит из блоков детектирования и эвристики, а также хранит предобученные модели детекторов для различных внешних условий. Блок детектирования реализует соответствующий функционал для работы с моделями на основе глубокого обучения: загрузка моделей в память, создание сессии детектирования, запуск процесса обработки конкретного кадра, интерпретация результатов работы. Блок эвристик реализует вспомогательные алгоритмы для каждой модели детектора.

Модуль классификации объектов включает в себя аналогичные блоки, предназначенные для работы с моделями машинного обучения. Кроме того, в его состав входит блок выделения признаков, в котором реализуется функционал выявления существенных особенностей изображений.

Модуль обработки реализует все необходимые для работы системы функции предобработки.

#### Заключение

В ходе исследования был предложен двухэтапный подход к распознаванию графических примитивов в видеопотоке, отличающийся совместным использованием нейросетевых детекторов на основе глубокого обучения и классических класси-

каторов, что в целом позволило обеспечить приемлемое качество определения объектов вне зависимости от внешних условий. Экспериментальная апробация, проводимая на наборе размеченных вручную записей с автомобильного регистратора, с использованием самостоятельно реализованных инструментов на базе библиотеки Keras, показала успешное распознавание более 95% сигналов светофоров. Следует также отметить, что применяемый комплекс эвристик существенно снижает процент ложных срабатываний детекторов. Пример результата работы системы представлен по ссылке [10].

Перспективным направлением дальнейшей работы является адаптация предложенного подхода к решению задач распознавания других типов объектов.

#### Список литературы

1. Николенко С., Кадурин А., Архангельская Е. Глубокое обучение. Погружение в мир нейронных сетей. СПб.: «Питер», 2018. 480 с.
2. Нагайцев М.В., Сайкин А.М., Ендачев Д.В. «Беспилотные» автомобили – этапы разработки и испытаний // Журнал автомобильных инженеров. 2012. № 5 (76). С. 32–39.
3. Дьяков А.С. Анализ отечественного и зарубежного опыта создания безэкипажных наземных транспортных средств специального назначения // Труды НАМИ. 2017. № 4 (271). С. 75–90.
4. Золотых Н.Ю., Кустикова В.Д., Мееров И.Б. Обзор методов поиска и сопровождения транспортных средств на потоке видеоданных // Вестник Нижегородского университета им. Н.И. Лобачевского. 2012. № 5–2. С. 348–358.
5. Долженкова М.Л., Чистяков Г.А. Система управления автономным беспилотным транспортным средством на основе продукционной модели представления знаний // Научно-технический вестник Поволжья. 2018. № 9. С. 99–102.
6. Ковалев Д.А., Юдин Д.А. Распознавание сигналов секций светофора на основе анализа цветowych пятен на изображении // Математические методы в технике и технологиях – ММТТ. 2018. Т. 8. С. 25–29.
7. Алешин С.П., Бородина Е.А. Нейросетевое распознавание классов в режиме реального времени // Инженерный вестник Дона. 2013. № 1 (24). С. 11.
8. Косовская Т.М., Тимофеев А.В. Иерархическое описание классов и нейросетевое распознавание сложных образов // Нейрокомпьютеры: разработка, применение. 2007. № 6. С. 30–33.
9. Рашка С., Мирджалили В. Python и машинное обучение. Машинное и глубокое обучение с использованием Python, scikit-learn и TensorFlow. М.: «Вильямс», 2019. 656 с.
10. Chistyakov G., Esaulov O. Example of using of two-stage graphic primitives recognition system [Электронный ресурс]. URL: <https://zenodo.org/record/2652062> (дата обращения: 26.04.2019).