

Анализ параметров обучения каскадного классификатора для поиска элементов лица

С.С. Волкова

Аннотация. Задача определения положения элементов лица на цифровом изображении актуальна во многих областях знаний. Одним из подходов к ее решению является метод построения каскада классификаторов. Эффективность этого подхода зависит от подбора параметров модели. В данной работе предложен метод, позволяющий получить оптимальные параметры обучения, при которых достигаются лучшие значения точности детектирования при сокращении ложноположительных откликов. В результате применения метода на примере обнаружения внешнего угла правого глаза были получены практические значения параметров обучения.

Ключевые слова: каскад классификаторов, метод Виолы-Джонса, моделирование антропометрических точек лица, параметры обучения модели.

Введение

Определение положения антропометрических точек лица на изображении лица актуально во многих областях знаний. Для решения этой задачи разработан ряд методов, в том числе: вычисление параметрических характеристик и их сравнение (фильтры Габора, яркостные характеристики), нейронные сети, а также каскады классификаторов на основе признаков, к которым относится и рассматриваемый метод Виолы-Джонса [1]. Несмотря на прозрачность метода, его работоспособность во многом зависит от деталей процедуры обучения.

В данной работе мы предлагаем подход к определению параметров обучения для поиска элементов лица, позволяющий проанализировать влияние параметров на эффективность моделирования.

1. Метод Виолы-Джонса

Напомним суть предложенного Виолой и Джонсом метода. Метод состоит из трех этапов: вычисление интегрального изображения, определение возможных положений объектов и

группировка результатов по положению и размерам. Интегральное представление изображения является двумерной матрицей того же размера, что и исходное изображение. Значение каждого элемента вычисляется по формуле (1):

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (1)$$

В работе [2] авторы представили признаки Хаара, отличные от прямоугольных. В этом случае необходимо находить интегральное представление, повернутое под определенным углом к исходному изображению.

На этапе определения возможных положений объектов производится обработка каскада классификаторов для каждого скользящего окна изменяемых размеров на изображении, который основывается на технологии AdaBoost [3]. В результате применения технологии бустинга, некоторое количество слабых классификаторов могут стать одним сильным:

$$f(x) = \sum_{t=1}^T \alpha_t \cdot h_t(x), \quad (2)$$

где $h_i(x)$ - слабый классификатор, α_i - вес данного классификатора в каскаде.

В случае, когда результат сильного классификатора оказывается меньше порогового значения, объект считается обнаруженным. В качестве слабых классификаторов в методе можно использовать признаки Хаара [1, 4] и LBP [5]. Признаки Хаара определяют разницу между взвешенной интенсивностью пикселей в прямоугольных областях внутри скользящего окна, признаки LBP определяют яркостную разницу в окрестности пикселя.

Последний этап производится с целью группировки близлежащих возможных областей нахождения объекта с дальнейшим их усреднением для вывода результата работы метода. Для выполнения этого этапа могут быть использованы алгоритмы кластеризации [6].

2. Методика исследования работоспособности метода локализации

Работоспособность методов локализации принято характеризовать их эффективностью и скоростью. В работе рассматривается влияние параметров обучения классификатора на эффективность метода, которая определяется ошибками первого и второго рода [7]. Практика показывает, что при выборе параметров модели приходится соблюдать оптимальный баланс между ошибками первого и второго рода, поскольку улучшение одной из этих метрик, как правило, приводит к ухудшению другой.

Для определения влияния каждого из параметров обучения зафиксируем значение остальных и произведем процедуру обучения с изменением одного параметра. Далее произведем оценку обученных классификаторов на фиксированной базе изображений, содержащих одно лицо различных размеров, в различных ракурсах и масштабах. Для этого воспользуемся функцией распределения отклонений допустимых результатов от точного, описанной нами в [8].

Функция распределения отклонения результирующей точки локализации от расставленной человеком-экспертом на тестовой базе определяется формулой (3):

$$F_{DEV}(dev) = P(DEV \leq dev, \mathfrak{Z}), \quad (3)$$

где DEV - случайная величина, заданная на наборе изображений $\mathfrak{Z} = \{\mathfrak{Z}_1, \mathfrak{Z}_2, \dots, \mathfrak{Z}_n\}$ с мерой $P = P(DEV = dev, \mathfrak{Z})$, определяющая значение отклонения результата от точного.

Индикатор отклонения принятого решения потенциальной точки $a(x_i)$ при известной истинной точке x_i определяется как

$$I(deviation) = I(x_i, a(x_i), deviation) = \begin{cases} 1, & \frac{\|x_i, a(x_i)\|}{\|C_l, C_r\|} = deviation \\ 0, & \text{в противном случае,} \end{cases} \quad (4)$$

где $C_l, C_r \in \mathbb{R}^2$ - точки истинных положений между внешними углами глаз.

При анализе влияния параметра обучения зафиксируем объект поиска как внешний угол правого глаза. Анализ влияния параметров обучения на другие объекты представляется аналогичным образом.

Тестирование производится на базе FERET [9]. Набор содержит в себе 1187 изображений, выбранных из 14126 изображений лиц, представляющих 1199 человек. Расстояние между внешними углами глаз на фотографиях в базе составляет в среднем 80 пикселей, положение человека относительно камеры различное.

3. Численный эксперимент. Влияние различных параметров обучения на работоспособность метода

3.1. Влияние различных размеров вырезаемых изображений

Для обучения каскада классификаторов требуется два класса изображений: позитивные образцы и негативные. В негативных образцах содержатся изображения, не содержащие объект. Позитивные изображения (Рис. 1) - изображения внешнего угла глаза, извлеченные с кадров, снятых с камеры. Размер изображений формировался в долях относительно расстояния между внешними углами глаз человека. На Рис. 2 показано влияние долей вырезаемых



Рис. 1. Примеры изображений, используемых для обучения классификатора, направленного на поиск внешнего угла правого глаза

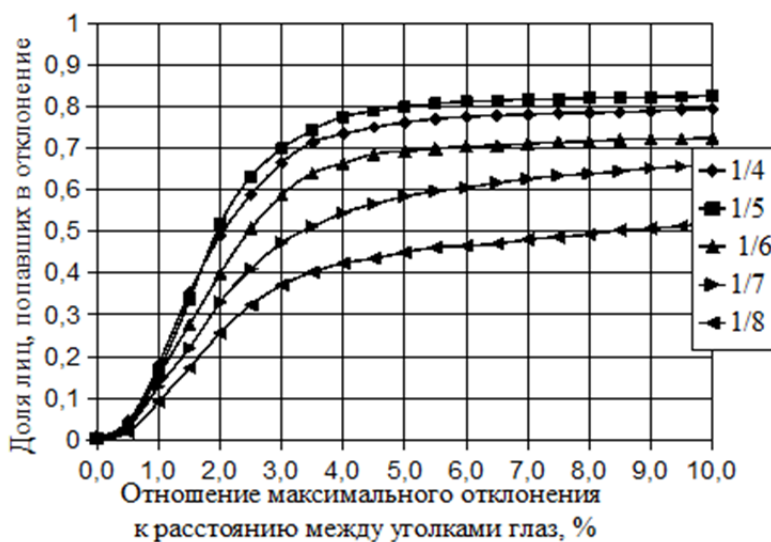


Рис. 2. Влияние доли вырезаемых изображений относительно внешних углов глаз на эффективность работы метода поиска внешнего угла правого глаза

изображений относительно расстояния между внешними углами глаз. Анализ результата показывает, что эффективной долей для обнаружения внешнего угла правого глаза является 1/5 от расстояния между внешними углами глаз. Увеличение размера вырезаемых областей не приводит к существенным улучшениям, при этом вносит ошибку, связанную с попаданием в область интереса зон, находящихся вне лица.

3.2. Влияние размера окна каскада классификаторов

При обучении каскадного классификатора одним из важных параметров обучения является размер стандартного окна, в котором будет идентифицироваться объект. От размера окна зависит количество признаков Хаара в различных наборах. Существует несколько возможных наборов признаков Хаара: базовый (детек-

торы краев и признаки, определяющие горизонтальные, вертикальные и диагональные линии), основной (все элементы базового набора и признаки, идентифицирующие области с ярко выраженным центром) и расширенный (включает в себя признаки основного набора, а также наклонные признаки). В таблице отображено количество признаков Хаара в различных наборах при различных размерах окна.

Количество признаков Хаара в различных наборах (без ограничений по площади признаков для несимметричного объекта)

Набор	Размер окна			
	12 x 12	16 x 16	20 x 20	24 x 24
базовый	10 344	32 384	78 460	162 336
основной	13 168	41 600	101 329	210 400
расширенный	15 846	50 878	125 199	261 600

На Рис. 3 показано влияние размера окна каскада классификаторов на эффективность работы метода. На Рис. 4 показана та же зависимость в увеличенном масштабе части осей.

Наиболее эффективным показывает себя размер 20x20. При больших размерах используется большое количество признаков, что вызывает эффект переобучения, при меньших значениях размера, признаков оказывается недостаточно.

3.3. Зависимость работоспособности метода от величины порога суммарного веса значимых примеров (weighttrimming)

На каждом шаге обучения AdaBoost происходит перераспределение весов всех примеров. Вес каждого примера задает значимость его на очередном шаге обучения.

Задание порога суммарного веса значимых примеров позволяет эффективно использовать обучающие данные, путем использования только

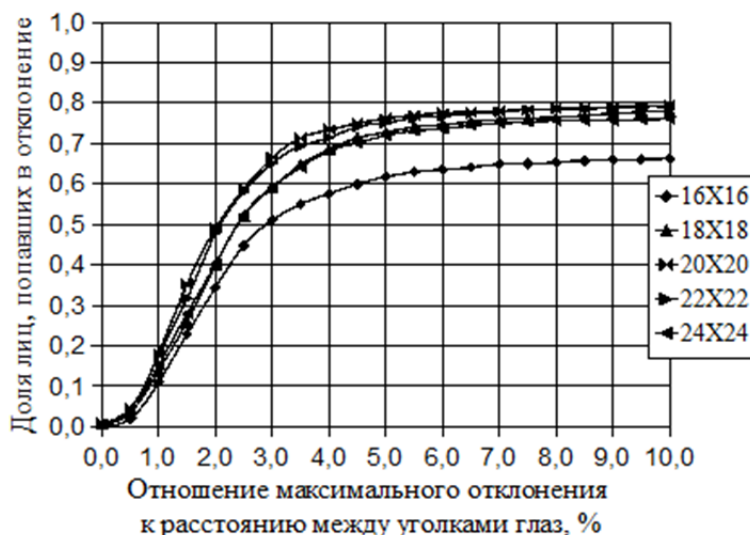


Рис. 3. Влияние размера окна каскада классификаторов на эффективность работы метода поиска внешнего угла правого глаза

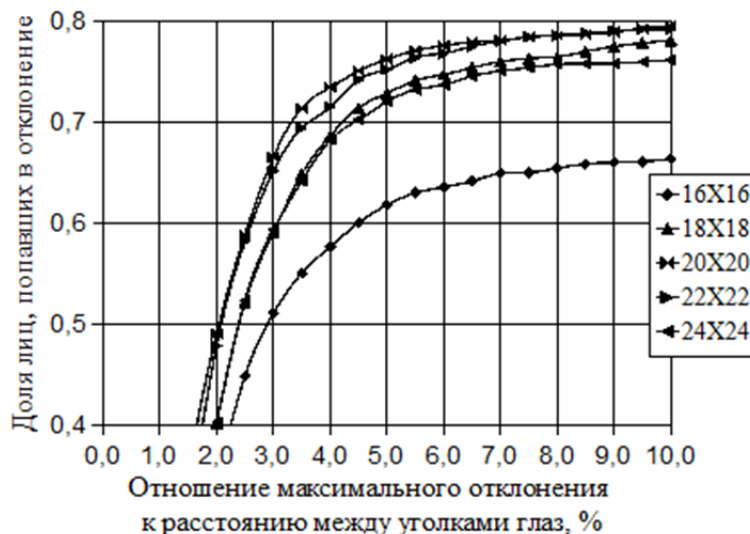


Рис. 4. Влияние размера окна каскада классификаторов на эффективность работы метода поиска внешнего угла правого глаза (увеличенный масштаб)

тех образцов, которые наиболее значимы для обучения. Пороговое значение β определяет вес примеров, участвующих в формировании классификатора на текущей стадии. На каждой итерации принимается значение $t(\beta)$ как β -тый квантиль распределения весов на обучающих данных, т. е. только те наблюдения, которые несут вес $1 - \beta$ от суммарного веса, используются для обучения. Обычно, данные, используемые для обучения, несут 90 – 99% от

суммарного веса. На Рис. 5 отображено влияние указанного параметра на эффективность работы, Рис. 6 отображает аналогичную зависимость в другом масштабе.

Наиболее эффективным для нашей задачи является значение 0,965. При уменьшении значения эффективность падает вследствие исключения значимых примеров, при увеличении значения падение обусловлено попаданием примеров со слишком малыми весами.

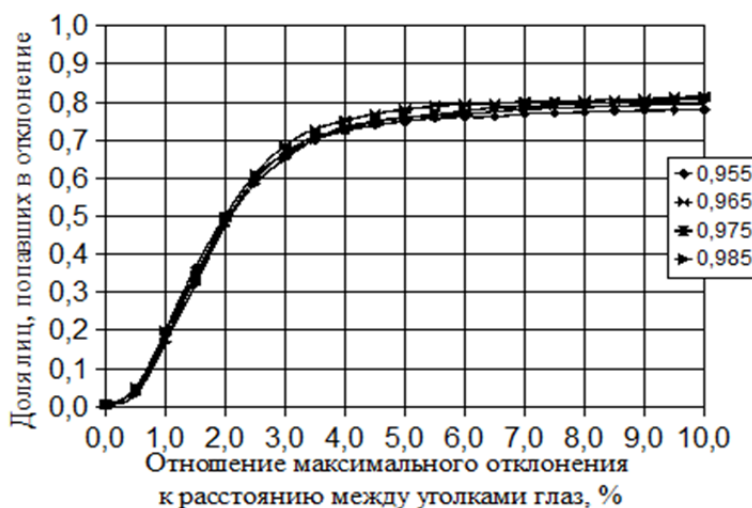


Рис. 5. Влияние величины порога суммарного веса на эффективность работы метода поиска внешнего угла правого глаза

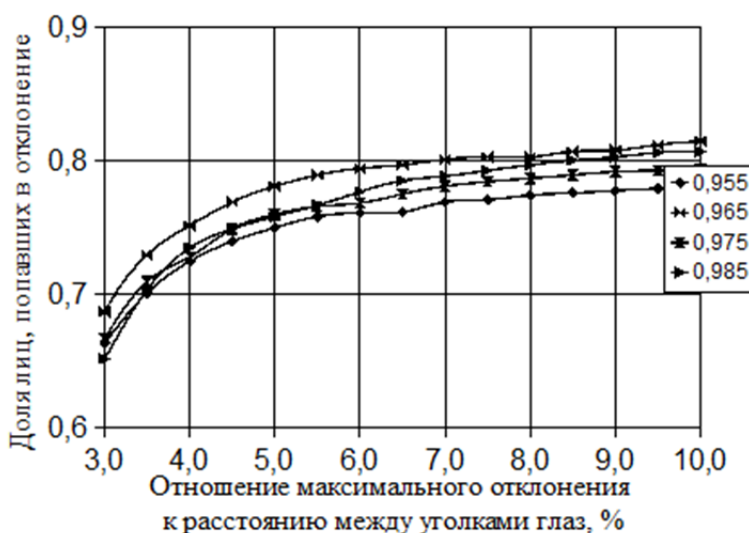


Рис. 6. Влияние величины порога суммарного веса на эффективность работы метода поиска внешнего угла правого глаза (увеличенный масштаб)

3.4. Зависимость работоспособности метода от максимального допустимого результата ложных обнаружений на каждой стадии (maxfalsealarm)

Параметр задается для указания максимального значения доли ложных обнаружений на каждой из стадий. Доля ложных обнаружений на каждой стадии (F_i) вычисляется по формуле (5) и определяется как отношение числа негативных примеров, которые были детектированы на данной стадии (f_i) к общему числу негативных примеров (N_{neg}).

$$F_i = \frac{f_i}{N_{neg}} \quad (5)$$

Суммарное значение уровня ложных тревог оценивается по формуле (6)

$$F = F_i^n, \quad (6)$$

где F_i - доля ложных обнаружений на каждой стадии, n - число стадий.

Согласно определению, чем больше значение допустимой доли ложных срабатываний, тем большее число объектов будет обнаружено. Качество найденных объектов при этом будет падать. Характер зависимости на Рис. 7 отражает этот результат. Однако, при использовании каскадов классификаторов в задаче локализации

антропометрических точек лица, важен не столько процент верных обнаружений, сколько значение максимального отклонения найденной точки от указанной человеком-экспертом. В связи с чем, в случае оценки параметра maxFalseAlarm, возникает необходимость оценивать графики отклонений не только при делении на общее число изображений, как сделано при оценке предыдущих параметров, но и при делении на число изображений, которые попали в область допустимых отклонений. Подобное действие позволяет определить, насколько точно найденные объекты обнаружены.

Исследуя график эффективности при делении на число изображений, попавших в допустимое отклонение (Рис. 8, Рис. 9), можно сделать вывод, что уменьшение максимального значения доли ложных срабатываний позволяет обнаруживать точки более точно к расставленным человеком-экспертом.

В связи с чем важным этапом при обучении является нахождение значения, которое позволит получить классификатор без большой потери доли найденных объектов, при этом также и без потерь значений отклонений.

Одним из возможных решений данной проблемы является построение нескольких классификаторов с различным значением максимальной доли ложных срабатываний и их дальнейшая комбинация.

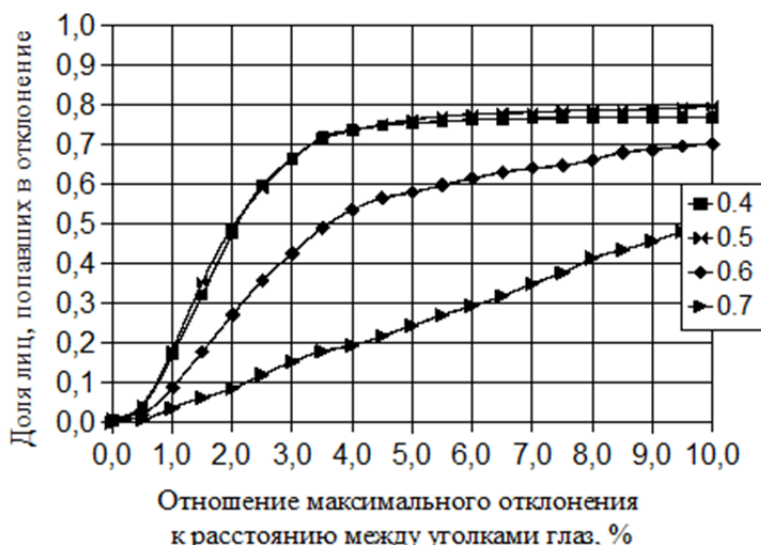


Рис. 7. Влияние величины максимального допустимого результата ложных обнаружений на каждой стадии на эффективность работы метода поиска внешнего угла правого глаза при делении на общее число изображений

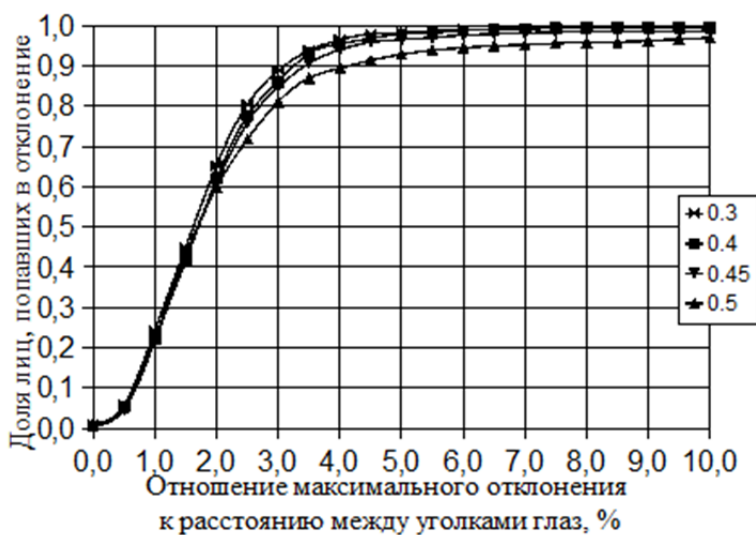


Рис. 8. Влияние величины максимального допустимого результата ложных обнаружений на каждой стадии на эффективность работы метода поиска внешнего угла правого глаза при делении на число изображений с допустимым отклонением

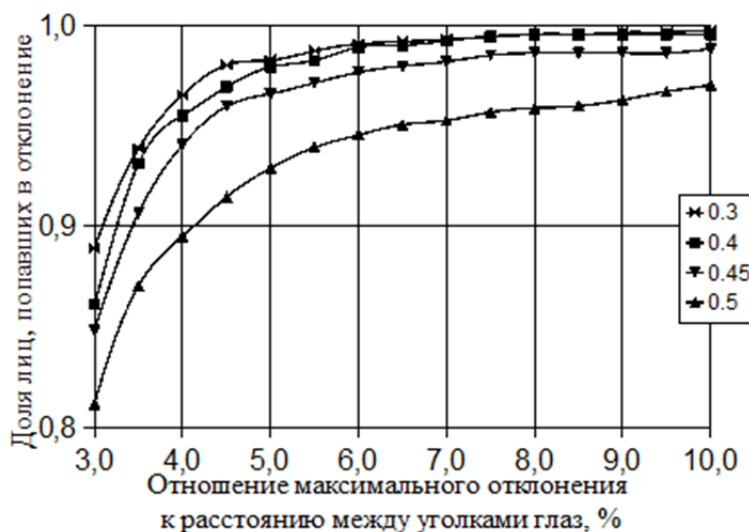


Рис. 9. Влияние величины максимального допустимого результата ложных обнаружений на каждой стадии на эффективность работы метода поиска внешнего угла правого глаза при делении на число изображений с допустимым отклонением (увеличенный масштаб)

Заключение

Активное использование методов локализации антропометрических точек лица, основанное на построении каскада классификаторов, напрямую ставит вопрос о надежности обучаемой модели и зависимости получаемых результатов от параметров обучения. Для метода Виолы-Джонса важными параметрами являются

характеристики обучающей выборки, размер каскада классификаторов, минимальные и максимальные коэффициенты верного и ложного обнаружения на каждой из стадий.

Для решения поставленной задачи разработан критерий эффективности метода, позволяющий на тестовой выборке определять зависимости и осуществлять выбор лучших параметров обучения. Критерий основывается

на анализе ложных и истинных срабатываний на эталонной базе, представленном в зависимости от размера отклонения найденной точки.

При проведении исследования влияния каждого конкретного параметра обучения остальные параметры были зафиксированы.

Поскольку все характеристики можно оценить по экспериментальным данным, предложенный критерий можно использовать для обучения модели других объектов на изображении лица.

Литература

1. Viola P., Jones M. Rapid object detection using a boosted cascade of simple features. // Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. December 2001, vol. 1, pp. 1063–6919.
2. Lienhart, R. An Extended Set of Haar-like Features for Rapid Object Detection / R. Lienhart, J. Maydt // IEEE ICIP. – 2002. – P. 900-903.
3. Kachurin A. Introduction to Boosting [Электронный ресурс] / А. Kachurin. – 2006. – Режим доступа: <http://www.cs.ucf.edu/courses/cap6411/cap6411/spring2006/Lecture10.pdf>
4. Freund Y., Robert E. Schapire A Short Introduction to Boosting// Shannon Laboratory, USA, 1999., pp. 771-780
5. Liao S., Zhu X., Zhen Lei, Lun Zhang, Stan Z. Li. Learning Multi-scale Block Local Binary Patterns for Face Recognition // International Conference on Biometrics (ICB), 2007, pp. 828-837
6. Барсегян, А.А. Анализ данных и процессов: учеб.пособие / А.А. Барсегян, М.С. Куприянов, И.И. Холод и др. – СПб.: БХВ-Петербург, 2009. – 512 с.
7. ГОСТ Р 50779.10-2000. «Статистические методы. Вероятность и основы статистики. Термины и определения.» - С. 26
8. Малышева С.С. Метод оценки качества работы алгоритмов классификации в задачах моделирования антропометрических точек лица// Вестник Саратовского государственного технического университета.- 2014. №1 (74), том 1.- С. 104-107.
9. Phillips P. J., Moon H., Rizvi S.A., Rauss P.J. The FERET evaluation methodology for face-recognition algorithms. // IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(10):1090-1104, 2000.

Волкова Светлана Сергеевна. Окончила Вологодский государственный педагогический университет в 2012 году, аспирантуру в 2015 году. Автор 13 печатных работ. Область научных интересов: обработка изображений, обнаружение объектов на изображении, системы биометрической идентификации. E-mail: malysheva.svetlana.s@gmail.com