Разработка и программная реализация системы мониторинга и анализа контента на видео и изображениях в интернет-пространстве

М. А. Руденко1, В.В. Милюков2, М.А. Крапивина3, Руденко А.В.4

ФГАУО ВО «Крымский федеральный университет им. В.И. Вернадского»

1rudenko.ma@cfuv.ru, 2milyukov.vv@cfuv.ru, 3miladakrapivina@gmail.com, 4rudenkoandre@mail.ru

***Аннотация*. Статья посвящена исследованию технологии компьютерного зрения и решению задач мониторинга и анализа контента на видео и изображениях в интернет-пространстве. С этой целью получена модель нейронной сети с высокими оценками точности, предложена модель анализа контента на основе комплекса математических метрик**

***Ключевые слова: мониторинг, контент, противоправный контент, обработка изображений, анализ метрик, сверточные нейронные сети, YOLO***

# Введение

Сегодня, в эпоху активного развития цифровых технологий, информация стала более доступной, её создание и распространение все более упрощается. Этот факт имеет как и очевидные положительные, так и отрицательные стороны. Согласно последним исследованиям в области СМИ и массовых телекоммуникаций: «Высокая доступность и широкий охват аудитории являются ключевыми факторами использования возможностей социальных медиа различными деструктивными сообществами и организациями (в частности, экстремистскими)» [1].

Современные системы контроля противоправного контента в сети Интернет в большинстве своем опираются на анализ текстового содержимого, проверке всевозможных фильтров и ручном формировании списков ресурсов экспертами. В это же время особенностью экстремистского контента является явная самоидентификация авторов контента. Как правило, экстремистские группировки явно признаются в своих преступлениях, показывают свою атрибутику. К атрибутике, в частности, относятся флаги, эмблемы, нашивки. То есть в экстремистских материалах часто есть ограниченный набор символики на иллюстрирующих изображениях и видео.

В связи с этим актуальной задачей является разработка и программная реализация системы мониторинга и анализа контента на видео и изображениях в интернет-пространстве. Эта система может применяться как самостоятельное решение, и как дополнительный инструмент.

# Обзор используемых инструментов

В настоящее время разработан целый ряд инструментов и библиотек для обработки изображений и создано множество алгоритмов распознавания изображений.

В последнее время для решения задачи поиска объектов на изображениях широкое распространение получили алгоритмы, основанные на применении региональных глубоких сверточных нейронных сетей (Regional Convolutional Neural Networks, R-CNN) и сетей детектирования YOLO (You Only Look Once), которые принципиально ориентированы на решение указанной задачи с одновременной классификацией найденных объектов. Архитектура YOLO была выбрана как наиболее оптимальное решение по совокупности точности и быстродействия [2, 3].

# Понятие противоправного контента

Многие страны и межгосударственные организации описывают понятие «противоправного контента» по-разному. Еврокомиссия о мерах по эффективной борьбе с незаконным контентом в Интернете в своих рекомендациях определяет: «незаконный контент означает любую информацию, которая не соответствует законодательству Союза или законодательству соответствующего государства-члена» [4].

Иными словами, противоправным контентом является информация, содержание которой не соответствует законодательству государства.

Особая категория противоправного контента – экстремистские материалы. Материалы, которые призывают к экстремистской деятельности, либо оправдывают ее, являются экстремистскими.

Так, согласно федеральному закону от 25 июля 2002 г. № 114-ФЗ "О противодействии экстремистской деятельности": «экстремистская деятельность (экстремизм):

- насильственное изменение основ конституционного строя и (или) нарушение территориальной целостности Российской Федерации (в том числе отчуждение части территории Российской Федерации), за исключением делимитации, демаркации, редемаркации Государственной границы Российской Федерации с сопредельными государствами;

- публичное оправдание терроризма и иная террористическая деятельность; …» и «экстремистская организация - общественное или религиозное объединение либо иная организация, в отношении которых по основаниям, предусмотренным настоящим Федеральным законом, судом принято вступившее в законную силу решение о ликвидации или запрете деятельности в связи с осуществлением экстремистской деятельности» [5].

Экстремистская деятельность, как правило, имеет четкую региональную локализацию, то есть при борьбе с распространением экстремистских материалов важно учитывать географию деятельности таких группировок.

# Модель нейронной сети

Применяемая классификация объектов основана на наиболее опасных в РФ и особо актуальных для Республики Крым, в виду граничного положения с Украиной, содержит следующие классы:

* ColSwast - свастика (ЦВ);
* GraySwast - свастика (ЧБ);
* Igil - Игил;
* RSector -Правый сектор;
* Taliban - Талибан;
* Ichkeria - Ичкерия;
* Azov - Азов.

Модель обучена на датасете, содержащем 163 изображения с достаточно равномерном распределением по классам (рис.1) и показывает хорошие результаты точности (рис.2).

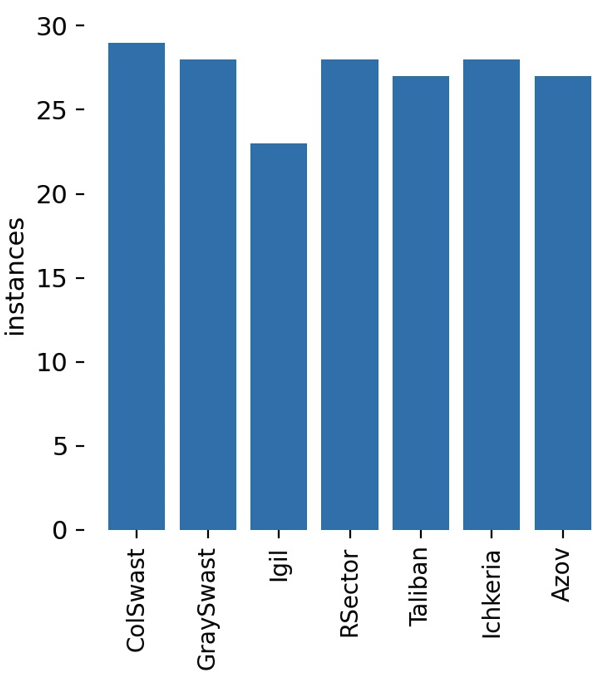
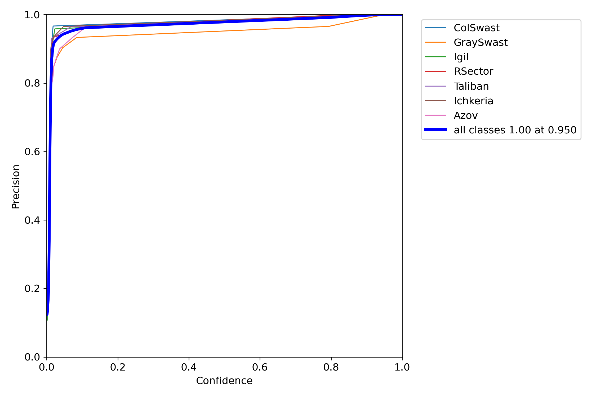


Рис. 1 Распределение классов первой версии обученной модели нейросети

  
Рис. 2 Точность обученной модели

# Модель анализа объектов на изображениях и видео

Оригинальные метрики детектирования YOLO:

Dyolo = (n, x, y, w, h, confidence), (1)

где n – номер класса;

x и y – относительные координаты bounding box объекта;

w и h – относительные ширина и высота bounding box объекта, соответственно;

с – confidence (коэффициент уверенности сети).

На основе этих метрик была разработана модель анализа объектов на изображениях и видео.

Математические метрики рассматриваемой модели: площадь объекта, центр фокусировки объекта, количество объектов, максимальный конфидент, средний конфидент, максимальная площадь, средняя площадь, площадь покрытия, дисперсия абсцисс и дисперсия ординат.

Площадь – метрика для одного отдетектированного объекта. Центр фокусировки, количество объектов, максимальный конфиденс (confidence), площадь покрытия могут считаться не только в рамках одного детектируемого объекта заданного класса, но и для всех найденных объектов одного класса (это происходит, когда в результате отдетектировано несколько изображений). Все остальные считаются в рамках одного класса.

Площадь объекта (S) – геометрическая площадь объекта, представляющее из себя произведение ширины и высоты ограничивающей рамки (bounding box) объекта на отдетектированном изображении. Величина относительна размеру изображения. Площадь объекта явно показывает какую часть от всего изображения занимает отдетектированный объект.

, (2)

где w- относительная ширина bounding box объекта,

h – относительная высота bounding box объекта.

Центр фокусировки (X, Y) – средняя точка центров bounding box найденных объектов; координаты точки – среднее арифметическое относительных координаты bounding box объектов:

, (3)

где x- относительная абсцисса объекта,

n - количество объектов класса.

, (4)

где y- относительная ордината объекта,

n - количество объектов класса.

Если координаты центра фокусировки равны (или очень близки) к координатам (0,5; 0,5) это значит, что объекты (объект) занимают центральное положение на изображении. Чем центральнее находится объект на изображении, тем большую смысловую нагрузку он несет для человеческого восприятия. Если рассматривать интерпретацию значения метрики в предметной области поиска противоправного контента, можно определить следующее: чем ближе центр фокуса отдетектированного объекта к центру фокуса всего изображения, тем более неслучайно нахождение объекта и «более противоправно» само изображение.

Количество объектов (n) – число найденных объектов одного класса или общее число найденных объектов. Эту метрику можно интерпретировать по-разному, к примеру, большое n объектов класса «флаг» может свидетельствовать, что на источнике изображено шествие или парад.

Максимальный конфиденс () – максимальное значение конфиденса для объектов одного класса или для всех найденных объектов. Конфиденс (confidence, коэффициент уверенности) — это метрика качества прогнозирования, которая показывает, на сколько сеть уверена (не уверена) в своих собственных результатах, представляет собой число от 0 до 1 эквивалентное проценту уверенности.

, (5)

где Сk = {C0, C1, … Cn} последовательность значений коэффициента уверенности объектов. Очевидно, может принимать значение от 0 до 1. Если значение максимального конфиденса близко к 1 — это значит, что с высокой точностью распознан как минимум один объект. Значение этой метрики может использоваться как граничное значение определения наличия экстремистского контента. К примеру, если Cmax < 0,3, то на изображениях скорее всего нет экстремистского контента.

Средний конфиденс () - среднее значение конфиденса для объектов одного класса; считается как среднее арифметическое и принимает значение от 0 до 1. Может интерпретироваться как уверенность нахождения класса объекта.

, (6)

где С - значение коэффициента уверенности объекта,

n - количество объектов класса.

Максимальная площадь () - максимальное значение площади для объектов одного класса. Может принимать значение от 0 до 1.

, (7)

где Sk = {S0, S0, … Sn} последовательность значений площадей объектов.

Средняя площадь () - среднее значение площади объектов одного класса; считается как среднее арифметическое, принимает значение от 0 до 1.

, (8)

где S - значение площади объекта,

n - количество объектов класса.

Площадь покрытия () – сумма площадей для объектов одного класса или для всех найденных объектов. Так как объекты на изображении могут перекрывать друг друга, возможно, что . Это свидетельствует о кучности отдетектированных объектов.

, (9)

где S - значение площади объекта,

n - количество объектов класса.

Дисперсия — это число, равное среднему квадрату отклонений значений случайной величины от ее среднего значения. Здесь считается дисперсия абсцисс и ординат. Чем больше эти значения, тем разрозненней объекты на изображении. Так как координаты - относительные величины, дисперсии также относительны.

Дисперсия абсцисс () – показывает отклонения объектов от центра фокусировки по оси абсцисс. В случае единичного распознанного изображения принимает значение 0.

, (10)

где x- относительная абсцисса объекта,

X- абсцисса центра фокусировки,

n - количество объектов класса.

Дисперсия ординат () – показывает отклонения объектов от центра фокусировки по оси абсцисс. В случае единичного распознанного изображения принимает значение 0.

, (11)

где x- относительная ордината объекта,

Y- ордината центра фокусировки,

n - количество объектов класса.

# Реализация системы мониторинга и анализа контента на видео и изображениях

Предложенные математические модели были реализованы в алгоритмах и программных модулях разработанной системы мониторинга и анализа контента на видео и изображениях (рис 3).

Модуль анализа принимает заявку в виде JSON-файла, где прописана ссылка на анализируемый ресурс. Модуль предварительной подготовки скачивает все файлы ресурса, фильтрует их по типу, конвертирует оставшиеся изображения и видео в формат, совместимый для детектирования, и запускает детектирование подготовленных ресурсов. В результате детектирования формируется директория с определенной внутренней структурой (рис.4).



Рис.3 Структура системы

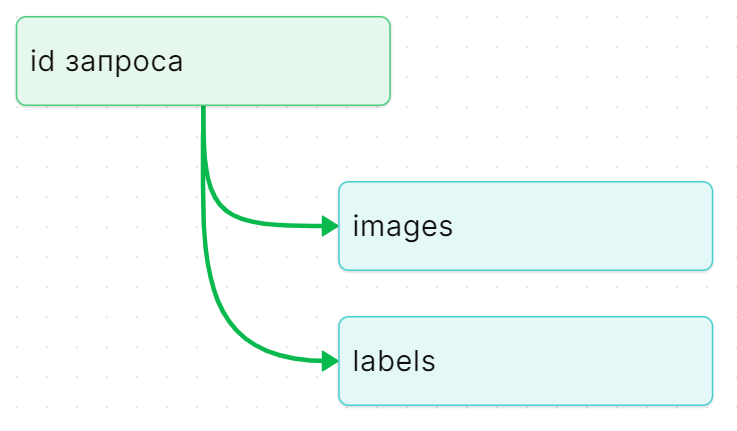


Рис.4 Структура директории

Модуль анализа метрик принимает расположение директории и в зависимости от содержимого, формирует текстовый отчет.

Рассмотрим две тестовые ситуации: распознано одно изображение (рис.5, 6,7) и распознано несколько изображений (рис.8, 9).

Одно изображение.



Рис. 5 Распознанное изображение

Метрики изображения:

3 0.544156 0.756067 0.228525 0.331402 0.943758

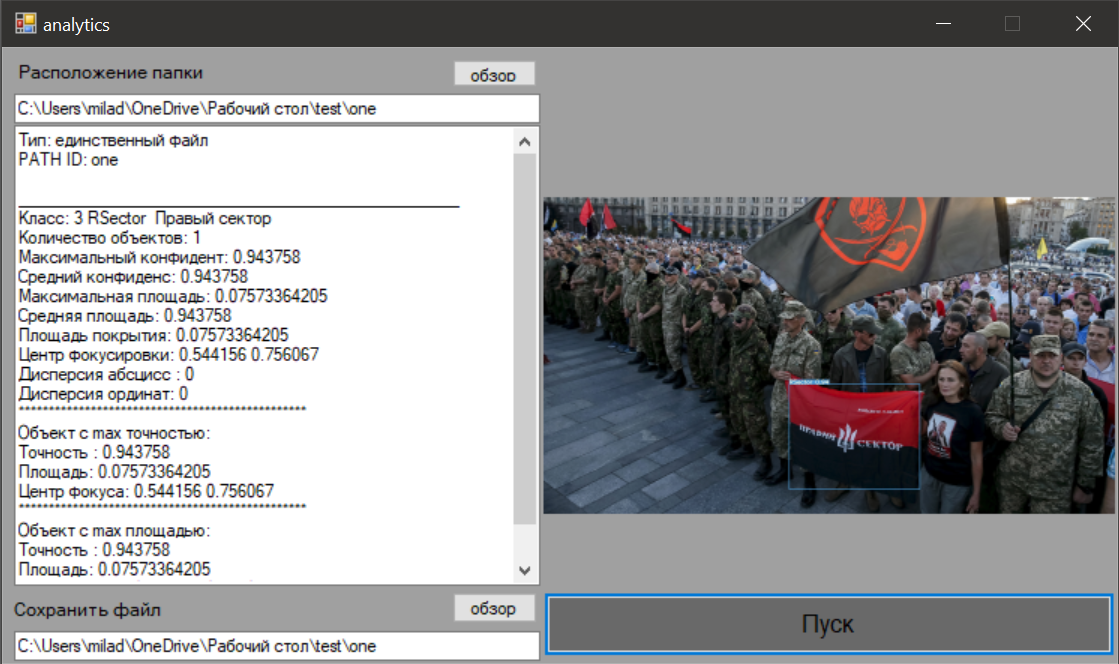


Рис. 6 Интерфейс приложения при одном изображении

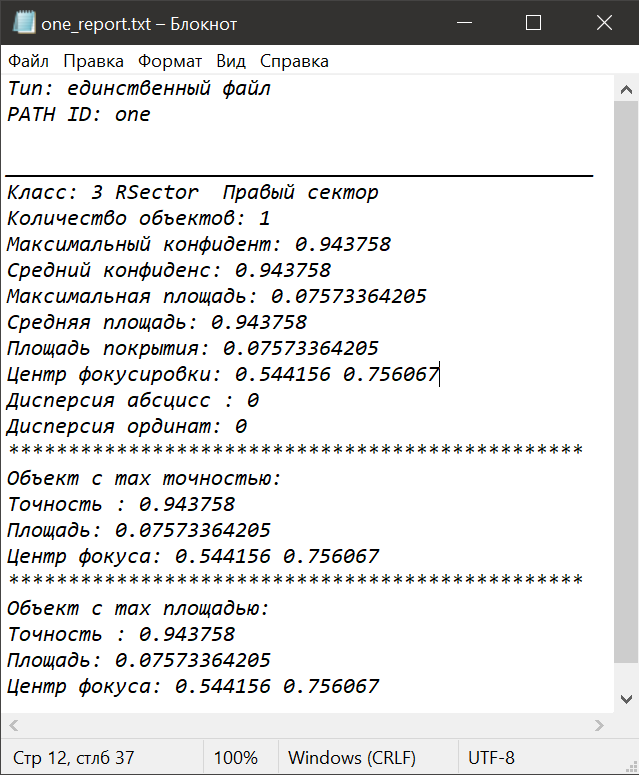


Рис. 7 Отчет по запросу с одним изображением

Несколько изображений.

Пусть распознаны следующие изображения:

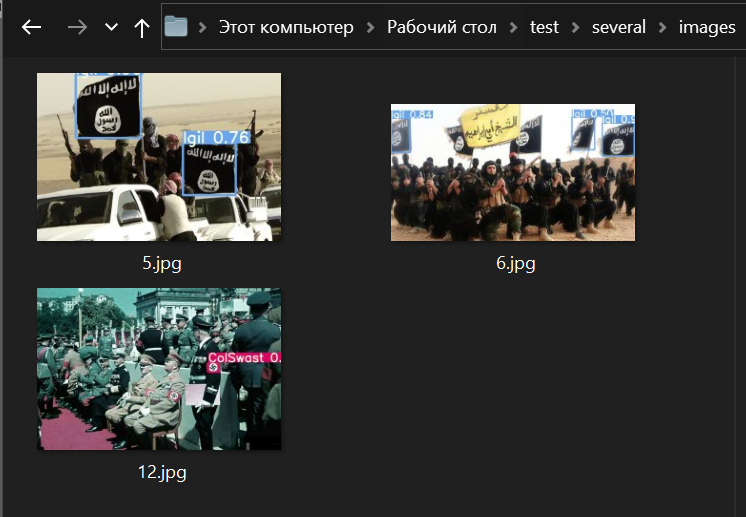


Рис. 8 Распознанные изображения

Текстовые метрики распознанных изображений имеют следующий вид:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 5.txt | 2 | 0.70744 | 0.57326 | 0.216898 | 0.307692 | 0.761032 | 0.066738 |
| 2 | 0.29256 | 0.197802 | 0.267339 | 0.377289 | 0.955227 | 0.100864 |
| 6.txt | 2 | 0.788672 | 0.216667 | 0.092969 | 0.236111 | 0.499497 | 0.021951 |
| 2 | 0.039063 | 0.230556 | 0.078125 | 0.25 | 0.843372 | 0.019531 |
| 2 | 0.932422 | 0.260417 | 0.130469 | 0.243056 | 0.912715 | 0.031711 |
| 12.txt | 0 | 0.723866 | 0.48963 | 0.04142 | 0.060741 | 0.896578 | 0.002516 |

Так как изображений несколько, стал доступен список изображений с объектами, имеющими максимальные значения конфиденса и площади.

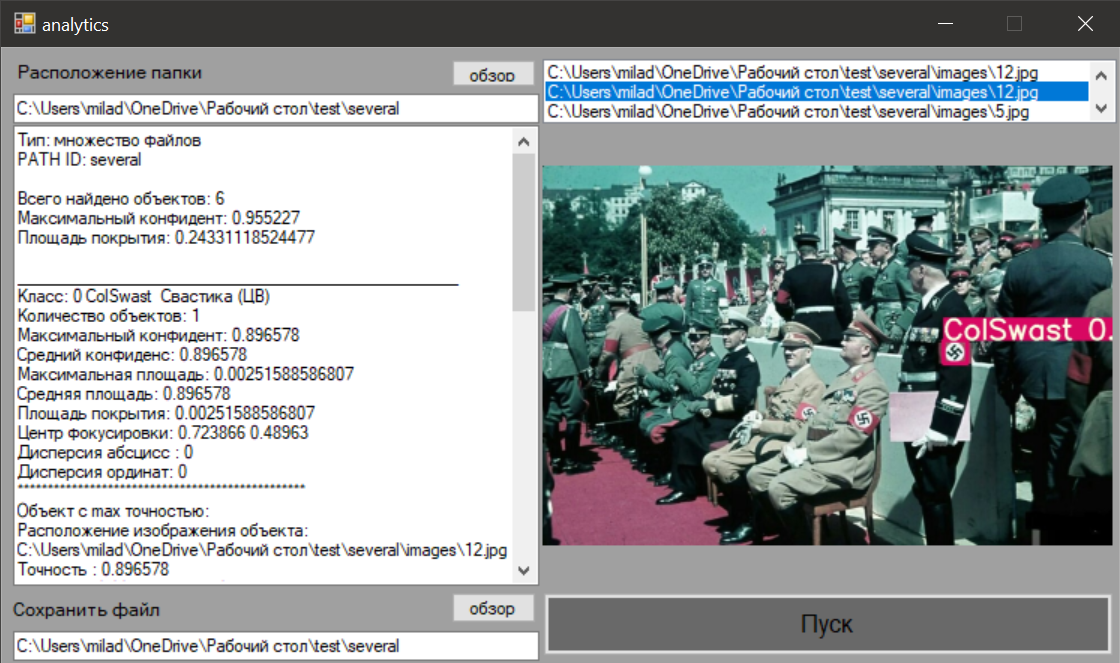


Рис. 9 Интерфейс приложения при нескольких изображениях

Текст отчета:

Тип: множество файлов

PATH ID: several

Всего найдено объектов: 6

Максимальный конфидент: 0.955227

Площадь покрытия: 0.24331118524477

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Класс: 0 ColSwast Свастика (ЦВ)

Количество объектов: 1

Максимальный конфидент: 0.896578

Средний конфиденс: 0.896578

Максимальная площадь: 0.00251588586807

Средняя площадь: 0.896578

Площадь покрытия: 0.00251588586807

Центр фокусировки: 0.723866 0.48963

Дисперсия абсцисс : 0

Дисперсия ординат: 0

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Объект с max точностью:

Расположение изображения объекта:

C:\Users\milad\OneDrive\Рабочий стол\test\several\images\12.jpg

Точность : 0.896578

Площадь: 0.00251588586807

Центр фокуса: 0.723866 0.48963

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Объект с max площадью:

Расположение изображения объекта:

C:\Users\milad\OneDrive\Рабочий стол\test\several\images\12.jpg

Точность : 0.896578

Площадь: 0.00251588586807

Центр фокуса: 0.723866 0.48963

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Класс: 2 Igil Игил

Количество объектов: 5

Максимальный конфидент: 0.955227

Средний конфиденс: 0.7943686

Максимальная площадь: 0.100864063971

Средняя площадь: 0.04815905987534

Площадь покрытия: 0.2407952993767

Центр фокусировки: 0.5520313 0.2957404

Дисперсия абсцисс : 0.11106202297516

Дисперсия ординат: 0.01967168195144

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Объект с max точностью:

Расположение изображения объекта:

C:\Users\milad\OneDrive\Рабочий стол\test\several\images\5.jpg

Точность : 0.955227

Площадь: 0.100864063971

Центр фокуса: 0.29256 0.197802

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Объект с max площадью:

Расположение изображения объекта:

C:\Users\milad\OneDrive\Рабочий стол\test\several\images\5.jpg

Точность : 0.955227

Площадь: 0.100864063971

Центр фокуса: 0.29256 0.197802

Аналитические расчеты, представленные в табличной форме на рис.10, подтверждают корректность вывода модуля анализа.



Рис. 10 Аналитические расчеты метрик отчета для нескольких изображений.

Модуль отправки заявки отправляет отчет пользователю, также в виде JSON файла.

##### Заключение

В результате исследования разработана модель анализа параметров детектированных объектов заданных классов на изображениях и видео, реализовано программное обеспечение, алгоритмически использующее данную модель. Тестирование показало высокую эффективность разработки, что позволяет использовать ее в государственных организациях, в особенности, общеобразовательных учреждениях для анализа контента на предмет наличия в первую противоправного контента.

##### Список литературы

1. Карданов, А. Р., Карданова, Д. А. Экстремистский контент в социальных медиа: анализ основных трендов и мер противодействия / А. Р. Карданов, Д. А. Карданова // Журнал прикладных исследований. — 2022. — № 11 т.2. — С. 169-172
2. Руденко М.А., Руденко А.В. Нечеткая модель классификации медицинских изображений на основе нейронных сетей // XXIV Международная конференция по мягким вычислениям и измерениям (SCM-2021). Сборник докладов. Санкт-Петербург. 26 – 28 мая 2021 г. СПб.: СПбГЭТУ «ЛЭТИ». 351 с. ISBN 978-5-7629-2864-92
3. YOLOv5: state-of-the-art модель для распознавания объектов. — Текст: электронный // neurohive.io: [сайт]. — URL: https://neurohive.io/ru/papers/yolov5-state-of-the-art-model-dlyaraspoznavaniya-obektov/ (дата обращения: 28.04.2022).
4. European Commission, Secretariat-General Commission Recommendation (EU) 2018/334 of 1 March 2018 on measures to effectively tackle illegal content online / European Commission, Secretariat-General // — 2018. — № C/2018/1177. — P. 50–61.
5. Об информации, информационных технологиях и о защите информации: Федеральный закон № 149-ФЗ: [принят Государственной Думой 8 июля 2006 г.: одобрен Советом Федерации 14 июля 2006 г.]. – Текст: непосредственный