

Санкт-Петербургский государственный университет

Математико-механический факультет
Кафедра информационно-аналитических систем

Жилкин Федор Игоревич

**Гендерная классификация лиц людей на
фотографиях в сфере компьютерной
криминалистики**

Бакалаврская работа

Научный руководитель:
к.ф.-м.н., доцент каф. ИАС СПбГУ Е. Г. Михайлова

Рецензент:
Главный специалист по данным ПАО Сбербанк Бугайченко Д. Ю.

Санкт-Петербург
2021

SAINT-PETERSBURG STATE UNIVERSITY

Software and Administration of Information Systems
Sub-Department of Analytical Information Systems

Zhilkin Fedor

Gender estimation of human faces based on photos in the forensics sphere

Graduation Thesis

Scientific supervisor:
Associate Professor, Ph.D. Mikhailova E. G. Amvrosy Vibegallo

Reviewer:
Chief Data Scientist of Sberbank PJSC Bugaychenko D. Y.

Saint-Petersburg
2021

Оглавление

Введение	5
1. Постановка задачи	7
2. Обзор предметной области	8
2.1. Криминалистические факторы	8
2.2. Обзор открытых наборов данных для гендерной классификации	9
2.2.1. VGGFace2	9
2.2.2. DiF	10
2.2.3. IMDb-WIKI	11
2.2.4. Unfiltered faces for gender and age classification . .	12
2.2.5. FairFace	13
2.3. Обзор существующих алгоритмов гендерной классификации	13
2.3.1. Gil Levi and Tal Hassner model trained on Adience	14
2.3.2. VGG-Face model fine-tuned on IMDb-Wiki	15
2.3.3. FairFace	15
2.4. Выводы из обзора предметной области	15
3. Наборы данных для обучения и тестирования алгоритма	17
3.1. Набор данных для обучения	17
3.2. Тестирующий набор данных	18
4. Алгоритм гендерной классификации	20
4.1. Архитектура EfficientNet	20
4.2. Обучение на подготовленном наборе данных	21
5. Анализ работы алгоритмов	23
5.1. Качественная характеристика работы алгоритмов	23
5.2. Скорость работы алгоритмов на CPU	25
5.3. Анализ результатов	26
5.4. Примеры работы построенного алгоритма	26

6. Интеграция в продукт BelkaSoft Evidence Center X	29
Заключение	30
Список литературы	32

Введение

Компьютерная криминалистика – это прикладная наука о раскрытии преступлений, связанных с компьютерной информацией, об исследовании цифровых доказательств, методах поиска, получения и закрепления таких доказательств. С развитием технологического прогресса компьютерная криминалистика оказывается все более востребованной, так как постоянно растет число различных цифровых носителей: персональных компьютеров, смартфонов, систем видеонаблюдения, планшетов и т.д.

Цифровые источники фотографий (камеры видеонаблюдения, веб-камеры, смартфоны, фотоаппараты) особо интересны криминалистам, так как с помощью этих изображений можно получить визуальную информацию о совершенном преступлении. В частности, с помощью городских камер видеонаблюдения можно оперативно установить местонахождение подозреваемого.

Данных с таких источников может быть очень много, и становится необходимым автоматизировать процесс обработки поступающих изображений. Выделим следующий ряд задач: автоматическое обнаружение лиц людей на фотографиях, кластеризация лиц, получение информации о человеке (гендер и возраст). При этом успешное автоматическое определение пола человека по его фотографии существенно ускоряет процесс анализа имеющихся фотографий.

Фотографии, полученные с камер видеонаблюдения, смартфонов и других цифровых источников, в контексте решения криминалистических задач характеризуются особыми факторами (т.н. криминалистическими факторами), которые значительно усложняют автоматизацию обработки и классификации фотографий. Подробно о криминалистических факторах будет рассказано в главе 2.1. Однако, большинство существующих алгоритмов гендерной классификации создавались для статичных, хорошо освещенных портретных фотографий, не обладающих остальными криминалистическими факторами. Также криминалистов интересует фактор скорости работы алгоритма гендерной клас-

сификации, т.к. время обработки фотографий или видеосъемки может быть ограничено жесткими рамками вследствие необходимости быстрого реагирования на преступление. Наконец, важным этапом является поиск аналогов данного изображения в большом объеме фотоматериалов.

В соответствии с вышеперечисленным возникает цель данного исследования – составить такой алгоритм, который быправлялся с задачей гендерной классификации в рамках сферы компьютерной криминастики.

1. Постановка задачи

Целью данной работы является создание системы автоматического анализа фотографий людей для решения задач в сфере компьютерной криминалистики. Данная система должна определять пол человека по его фотографии. Для успешного достижения данной цели были поставлены следующие задачи:

- рассмотреть существующие алгоритмы гендерной классификации и наборы данных, наиболее актуальные для сферы криминалистики;
- создание набора данных для обучения и тестирования алгоритма гендерной классификации с учетом криминалистических факторов;
- создание алгоритма машинного обучения для решения задачи гендерной классификации в сфере криминалистики;
- сравнение разработанного алгоритма с другими алгоритмами, решающими эту же задачу, но не учитывающими криминалистические факторы;
- встраивание созданного алгоритма в продукт BelkaSoft Evidence Center X (BEC X).

2. Обзор предметной области

2.1. Криминалистические факторы

Как было сказано во введении, фотографии в области криминалистики обладают определенными свойствами перечисленными ниже.

- Качество съемки ниже среднего, т.к. фотографии получены с непрофессиональных фотоаппаратов, или же лицо может быть искажено, т.к. фотография была сделана в движении – **Quality**.
- Плохое освещение вследствие того, что съемка проводилась в темное время суток или в темном помещении – **Illumination**.
- Лицо может занимать несущественную часть изображения (т.е. быть маленьким) из-за удаленного источника съемки (например, камеры видеонаблюдения в метро) – **Scale**.
- Лицо может быть чем-то прикрыто (окклюзия), например, если подозреваемый прячется за маской или шарфом – **Occlusion**.
- Лицо может выражать эмоцию, например, если человек кричит, плачет, улыбается – **Expression**.
- На лицо может быть нанесен макияж или грим – **Make-Up**.
- Лицо может быть повернуто на определенный угол относительно источника съемки – **Pose**.

Все эти факторы играют значительную роль в криминалистике, ведь почти все фотографии, полученные с камер видеонаблюдения или смартфонов, обладают как минимум одним из вышеперечисленных свойств.

Однако большинство уже существующих алгоритмов гендерной классификации предназначены для портретных фотографий хорошего качества, т.к. большинство из существующих наборов данных для обучения не содержат фотографий, обладающие вышеперечисленными криминалистическими факторами. Соответственно, необходимо разрабо-

тать алгоритм гендерной классификации, который смог бы справляться с криминалистическими фотографиями. Для начала требуется создать соответствующий набор данных. В этом обзоре будут показаны наиболее интересные для текущей задачи наборы данных, а также алгоритмы, обученные на этих наборах данных и показавшие наилучшие результаты на известных бенчмарках.

2.2. Обзор открытых наборов данных для гендерной классификации

На данный момент существует достаточно много наборов данных для решения задачи гендерной классификации, однако не все подходят для задачи гендерной классификации лиц в области компьютерной криминалистики. Отметим, что мы будем рассматривать только те наборы данных, которые находятся в открытом доступе для просмотра и использования.

2.2.1. VGGFace2

Набор данных VGGFace2 [10] состоит примерно из 3.31 миллиона изображений, разделенных на 9131 класс, каждый из которых представляет собой определенного человека – знаменитость. Набор данных разделен на две части: одна – для обучения, а другая – для тестирования. Последняя содержит около 170000 изображений, разделенных на 500 классов-личностей, в то время как все остальные изображения принадлежат к оставшимся 8631 классам, доступным для обучения. При построении набора данных авторы сосредоточили свои усилия на достижении очень низкого уровня шума меток (лейблов) и высокого разнообразия положения лица (pose) и возрастов, таким образом, сделав набор данных VGGFace2 подходящим выбором для обучения современных моделей глубокого обучения для задач, связанных с лицом человека. Однако авторы отмечают, что распределение личностей в наборе данных VGGFace2 может быть нерепрезентативным для глобальной популяции людей, т.е. набор данных не отличается разнообразием по возрасту, по-

лу, расовой принадлежности. По криминалистическим факторам набор данных так же нельзя назвать полностью подходящим, т.к. он отвечает лишь фактору Pose. Примеры изображений набора данных представлены на рис. 1.



Рис. 1: Примеры изображений лиц набора данных VGGFace2 (источник: набор данных VGGFace2)

2.2.2. DiF

Набор данных Diversity in Faces [6] был представлен компанией IBM в 2019 году. Он насчитывает около 1 млн. изображений лиц людей, размеченных по возрасту и полу. Авторы используют изображения из набора данных YFCC-100M¹, в котором данные собирались путем получения фотографий пользователей фотохостинга Flickr². Далее авторы статьи с использованием алгоритмов машинного обучения и оригинальных алгоритмов для разметки набора данных получили все необходимые метки (возраст, пол, координаты лица). Набор данных Diversity in Faces обладает большим объемом фотографий и разнообразием этих фотографий, однако он не удовлетворяет всем криминалистическим фак-

¹<http://projects.dfki.uni-kl.de/yfcc100m/>, дата обращения: 25.04.2021

²<https://www.flickr.com/>, дата обращения: 25.04.2021

торам, описанным выше. Набор данных доступен для использования после предварительной заявки-обращения авторам.

2.2.3. IMDb-WIKI

Набор данных IMDb-WIKI [11] создавался путем сбора фотографий наиболее популярных актеров с сайта IMDb³, далее эти фотографии размечались по гендеру и возрасту. Остальная часть набора данных была собрана с сайта Wikipedia⁴ таким же образом. В общей сложности было получено 460 000 изображений лиц 20 284 знаменитостей с IMDb и 62 300 из Википедии, то есть в общей сложности около 522 000 изображений лиц. Такое число уникальных персон способствует достаточно широкому разнообразию лиц в наборе данных, однако авторы не предоставляют подробной информации о расовой принадлежности этих персон. Также почти все лица в этом наборе данных представляют собой портретные фотографии в хорошем качестве и при хорошем освещении, что делает этот набор данных не совсем подходящим для задачи гендерной классификации в сфере криминалистики. Примеры изображений набора данных представлены на рис. 2.



Рис. 2: Примеры изображений лиц набора данных IMDb-WIKI (источник: набор данных IMDb-WIKI

³<https://www.imdb.com/>, дата обращения: 25.04.2021

⁴<https://www.wikipedia.org/>, дата обращения: 25.04.2021

2.2.4. Unfiltered faces for gender and age classification

Набор данных Unfiltered faces for gender and age classification (Adience) [2] содержит около 26 000 изображений лиц людей. Изображения собирались с сервиса Flickr, со смартфонов (iPhone 5 или выше) и с других мобильных устройств, что дает достаточно большое количество уникальных персон в наборе данных. Его главная особенность – это то, что все изображения были собраны с реальных фотографий лиц без дополнительной обработки, т.е. эти изображения наиболее приближены к реальным снимкам, из-за чего этот набор данных заслуживает особого внимания. Однако изображения в наборе данных Adience отвечают только немногим вышеперечисленным криминалистическим факторам: Scale, Pose, Expression. Примеры изображений набора данных представлены на рис. 3.



Рис. 3: Примеры изображений лиц набора данных Adience (источник: набор данных Adience)

2.2.5. FairFace

FairFace [4] – достаточно большой по объему набор данных – около 108 000 изображений лиц. Набор данных создавался так же, как и Diversity in Faces, однако затем был вручную сбалансирован по расовой, возрастной и гендерной принадлежности. Примеры изображений набора данных представлены на рис. 4.

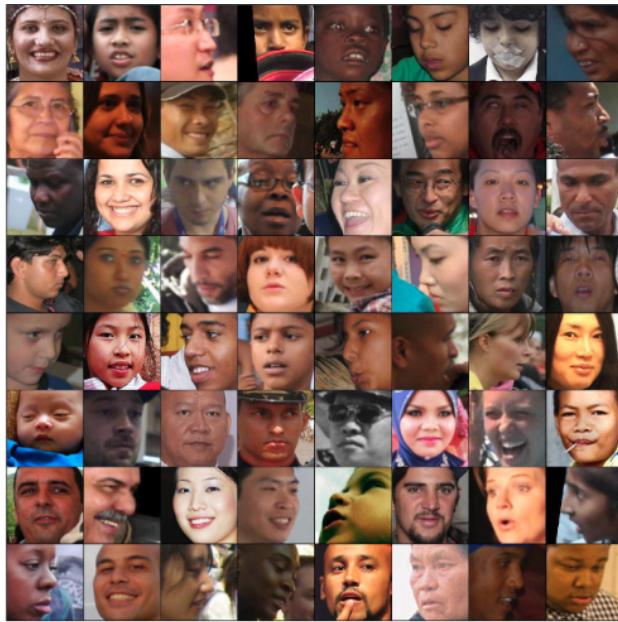


Рис. 4: Примеры изображений лиц набора данных FairFace (источник: набор данных FairFace)

2.3. Обзор существующих алгоритмов гендерной классификации

Качество работы алгоритма гендерной классификации напрямую зависит от набора данных, на котором этот алгоритм был обучен. В предыдущей главе были рассмотрены наиболее интересные для исследования существующие наборы данных. В этой главе будут рассмотрены алгоритмы, обученные на рассмотренных наборах данных и показывающие наиболее высокие результаты по качеству работы на соответствующих бенчмарках. В дальнейшем выбранные алгоритмы будут сравниваться по качеству и скорости работы с построенным алгоритмом.

Таким образом, для исследования были отобраны следующие алгоритмы гендерной классификации: Алгоритм Gil Levi and Tal Hassner, обученный на наборе данных Adience; алгоритм VGG-Face, обученный на наборе данных IMDb-WIKI; алгоритм FairFace, обученный на наборе данных FairFace.

2.3.1. Gil Levi and Tal Hassner model trained on Adience

Оригинальный алгоритм [5] был предложен Gil Levi и Tal Hassner в 2015 году. Авторы используют очень простую архитектуру сверточных нейронных сетей, аналогичную CaffeNet [8] и AlexNet [3]. Сеть использует 3 сверточных слоя, 2 полностью связанных слоя и конечный выходной слой:

- Conv1 : Первый сверточный слой имеет 96 узлов размера ядра 7;
- Conv2 : Второй уровень conv имеет 256 узлов с размером ядра 5;
- Conv3 : Третий уровень conv имеет 384 узла с размером ядра 3;
- Два полностью соединенных слоя имеют по 512 узлов каждый.

Архитектура сверточной нейронной сети представлена на рис. 5. Ав-

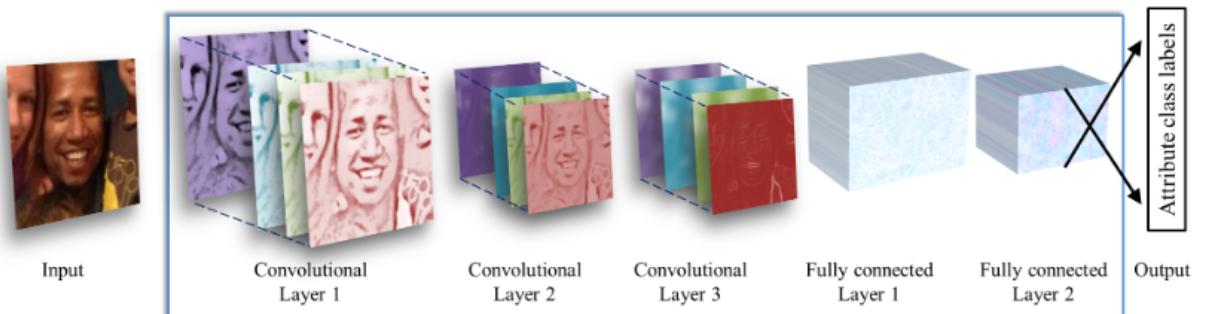


Рис. 5: Архитектура Gil Levi and Tal Hassner model

торы используют свой же набор данных Adience для обучения модели. Метрика F1 на тестовой части набора данных Adience составляет 86.8 ± 1.4 .

2.3.2. VGG-Face model fine-tuned on IMDb-Wiki

Алгоритм VGG-Face [9] был построен специально для задач, связанных с лицами. Он основан на архитектуре сверточных нейронных сетей, представленной на рис. 6. Далее алгоритм был дообучен (Transfer



Рис. 6: Архитектура VGG-Face model

learning [13]) на наборе данных IMDb-Wiki для задачи гендерной классификации. Алгоритм показывает высокую точность определения гендеря на тестовой части обучающего множества с метрикой F1 равной 95.8 ± 1.1

2.3.3. FairFace

Алгоритм FairFace [4] был предложен авторами одноименного набора данных. Алгоритм основан на архитектуре сверточных нейронных сетей ResNet-34, далее обучен на самом наборе данных FairFace. Авторы сравнивают построенный ими алгоритм с алгоритмами, обученными на наборе данных UTKFace, LWFA+, CelebA. Их алгоритм показывает наивысшие результаты с метрикой F1 равной 94.89 ± 1.1 .

2.4. Выводы из обзора предметной области

Были рассмотрены следующие наборы данных для обучения и тестирования алгоритмов гендерной классификации в сфере криминалистики: VGGFace2, DiF, IMDb-WIKI, FairFace, Adience (см. табл. 1). Колонки в табл. 1 имеют следующее значение:

- **#photos** – количество изображений в наборе данных;

- **#subjects** – количество уникальных персон в наборе данных;
- **IsWild** – фотографии, обладающие, вышеописанными криминалистическими факторами;
- **Div (Diversity)** – разнообразие набора данных, т.е. наличие фотографий людей разного пола, разного возраста, разной расовой принадлежности.

Data set	#photos	#subjects	IsWild	Diversity
VGGFace2	3.3M	9131	±	–
DiF	0.97M	?	+	+
IMDb-WIKI	523k	20284	–	±
FairFace	108k	?	+	+
Adience	26.6k	?	+	±

Таблица 1: Характеристики рассматриваемых наборов данных

Также были рассмотрены наиболее интересные для исследования алгоритмы машинного обучения, обученные на этих наборах данных: Gil Levi and Tal Hassner model, VGG-Face, FairFace. В дальнейшем эти алгоритмы будут сравниваться с построенным алгоритмом по скорости и качеству работы на тестовом наборе данных.

3. Наборы данных для обучения и тестирования алгоритма

Данные для обучения предложенного алгоритма и для тестирования всех рассматриваемых алгоритмов должны обладать вышеописанными криминалистическими факторами. Особое внимание необходимо обратить на данные для тестирования, так как они должны быть максимально приближены к реальным изображениям.

3.1. Набор данных для обучения

Для обучения алгоритма гендерной классификации будет использован набор данных WiderFace [12]. Набор данных WiderFace содержит около 32 000 изображений и 393 000 лиц. Изображения в наборе данных WiderFace обладают всеми вышеперечисленными криминалистическими факторами: размер (Scale), поворот (Pose), окклюзия (Occlusion), экспрессия (Expression), макияж (Makeup), освещение (Illumination). Примеры изображений из набора данных WiderFace представлены на рис. 7.



Рис. 7: Пример изображений из набора данных WiderFace

Авторы разметили набор данных по лицам, т.е. он содержит координаты каждого лица в этом наборе данных, что делает его отличным выбором для задачи детектирования лиц людей в реальных условиях.

ях (In the wild). Однако для данного исследования этого недостаточно, т.к. нам нужны метки гендера к каждому лицу (“Male”/”Female”). Для создания набора данных для обучения была выбрана часть изображений, лица на которых не меньше 30*30 пикселей. Таким образом мы получили 4635 изображений. Далее для каждого лица на каждом изображении вырезаем лицо, ставим ему соответствующую метку (“Male”/”Female”/”Skip”) и записываем в файл. Стоит отметить, что метки ставились вручную, поскольку существующие алгоритмы не способны правильно предсказать гендер по лицу в наборе данных WiderFace (в силу его специфиности). Метки ”Male” и ”Female” ставились при полной уверенности в гендерной принадлежности человека, чье лицо мы размечаем.

Всего было получено 12 620 размеченных по гендеру лиц, где 6289 ”Male” (0.49 от общего числа) и 6331 ”Female” (0.51 от общего числа). Остальные 4678 лиц были пропущены (метка ”Skip”). Построенный набор данных доступен по ссылке⁵.

3.2. Тестирующий набор данных

Для тестирования алгоритмов был собран набор данных, состоящий из 52 изображений, включающий в себя 309 изображений лиц, отвечающих всем криминалистическим факторам. Изображений были получены со смартфонов (iPhone 7 и более новые модели), открытых уличных камер видеонаблюдения, фотографий из Санкт-Петербургского метрополитена и из интернета. Изображения из тестирующего набора данных представлены на рис. 8.

⁵<https://www.kaggle.com/fedoszhilkin/genderwiderface>, дата обращения: 25.04.2021

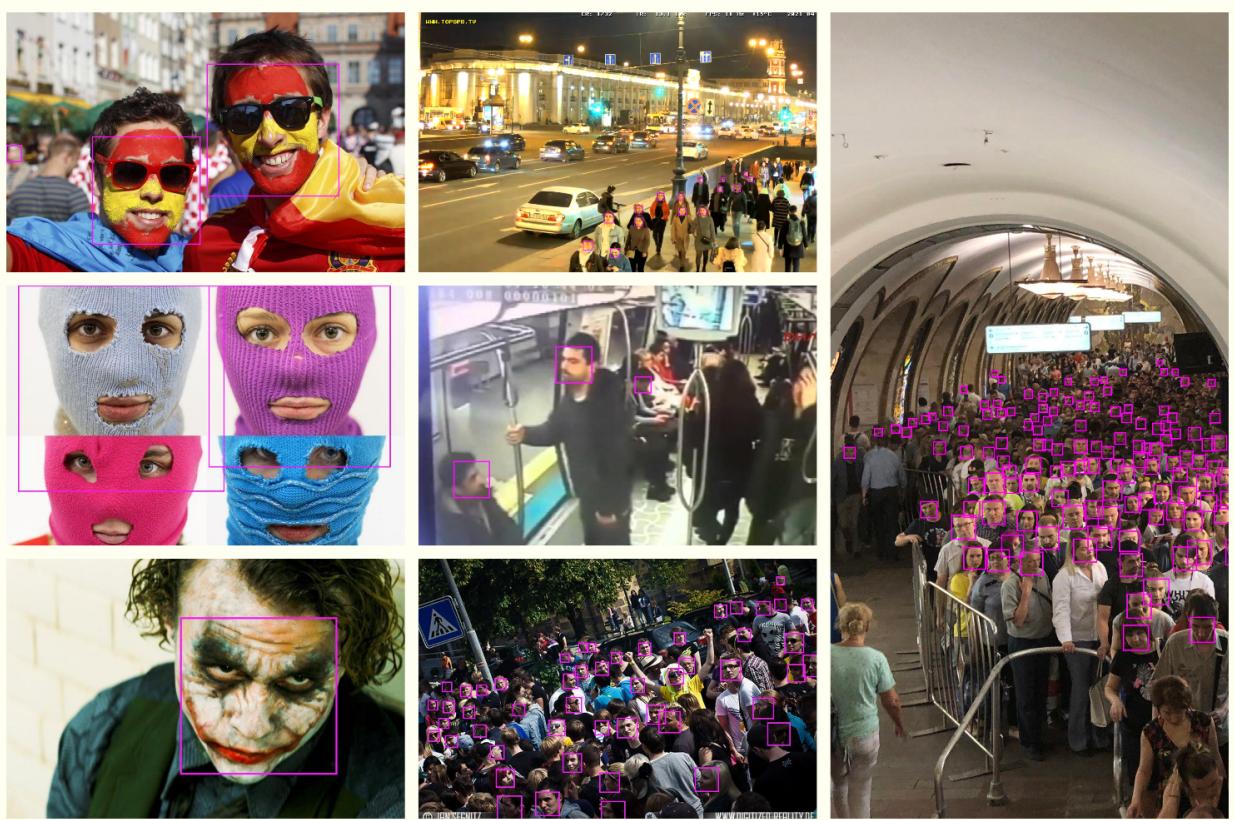


Рис. 8: Примеры тестовых данных

4. Алгоритм гендерной классификации

4.1. Архитектура EfficientNet

Задача определения гендера по фотографии является классической задачей классификации изображений. Выбор архитектуры алгоритма машинного обучения для задачи будет происходить путем сравнения наиболее современных архитектур алгоритмов классификации на общеизвестных бенчмарках. На данный момент существует множество архитектур классифицирующих алгоритмов: AlexNet, GoogLeNet, VGG, ResNet, SENet, Inception и другие. Однако в 2019 году была представлена архитектура EfficientNet [7], которая показывала наивысшие результаты на бенчмарке Imagenet⁶ в рамках задачи классификации изображений. Данная архитектура обходит state-of-the-art подходы по скорости и качеству работы. Таким образом, EfficientNet-B7 обогнала рассмотренные сверточные архитектуры и стала топ-1 по точности для задачи классификации изображений на Imagenet. EfficientNet-B1 в 7.6 раз меньше и в 5.7 раз быстрее ResNet-152. Такой прирост эффективности был достигнут благодаря масштабированию (scaling) моделей и балансирования между глубиной и шириной (количества каналов) сети, а также разрешения изображений в сети. Авторы статьи предлагают новый метод составного масштабирования (compound scaling method), который равномерно масштабирует глубину/ширину/разрешение с фиксированными пропорциями между ними. Из существующего метода под названием «Neural Architecture Search» (NAS [1]) для автоматического создания новых сетей и своего собственного метода масштабирования авторы получают новый класс моделей под названием EfficientNets.

Начальная модель EfficientNet-B0 (см. рис. 9) использует архитектуру MBConv, схожую с MobileNetV2 и MnasNet. Затем путем масштабирования параметров появились архитектуры EfficientNet-B1, EfficientNet-B2, ..., EfficientNet-B7.

Начиная с архитектуры EfficientNet-B4, прирост качества работы

⁶<http://www.image-net.org/index>, дата обращения: 25.04.2021

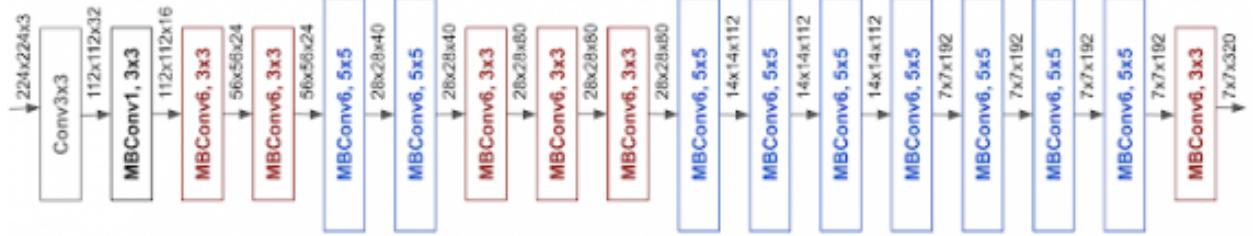


Рис. 9: Архитектура модели EfficientNet-B0

несравним с приростом количества параметров. Поскольку для нас важен фактор сокости работы алгоритма на CPU, для исследования были выбраны архитектуры EfficientNet-B0, EfficientNet-B2, EfficientNet-B4. Выбранные архитектуры будут обучены на подготовленном наборе данных и затем будут сравниваться с выбранными алгоритмами гендерной классификации.

4.2. Обучение на подготовленном наборе данных

Выбранные архитектуры EfficientNet-B0, EfficientNet-B2, EfficientNet-B4 были обучены на построенном наборе данных. Обучение длилось до тех пор, пока модель не переставала улучшаться в течение последних 20 эпох. После этапа обучения всех архитектур был построен график обучения (функция потерь на обучении/количество эпох). Результаты представлены на рис. 10.

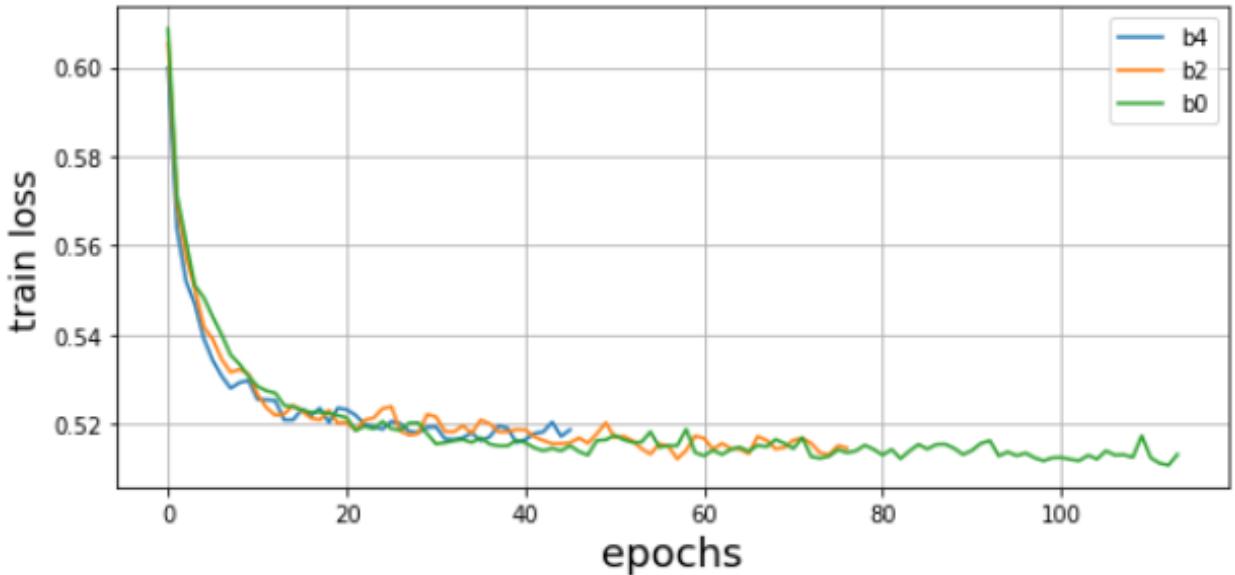


Рис. 10: Результаты обучения EfficientNet-B0, EfficientNet-B2, EfficientNet-B4

Также были получены результаты качества работы построенных алгоритмов на тестовой части обучающего множества. Каждая из моделей обучалась на подготовленном наборе данных 10 раз. В табл. 2 представлены лучшие результаты моделей. Дисперсия в таком случае была не больше, чем 3-4 знака после запятой. Для сравнения также приведены результаты моделей архитектуры ResNet-50 и Inception-V3.

Method	Precision	Recall	F1
EfficientNet-B0	0.912	0.919	0.948
EfficientNet-B2	0.978	0.959	0.969
EfficientNet-B4	0.977	0.965	0.968
ResNet-50	0.749	0.718	0.733
Inception-V3	0.789	0.749	0.768

Таблица 2: Качество работы построенных алгоритмов на тестовой части обучающего множества

5. Анализ работы алгоритмов

В данной главе будет определяться качество работы рассматриваемых алгоритмов с помощью тестирующего набора данных, а также измеряться скорость работы каждого из рассматриваемых алгоритмов.

5.1. Качественная характеристика работы алгоритмов

Изображения лиц из тестирующего набора данных подавались на вход каждому из рассматриваемых алгоритмов, включая построенный алгоритм. Далее считались интересующие нас метрики:

- Recall показывает насколько точно алгоритм способен отличить мужчину от "немужчины" (доля правильно предсказанных мужчин);
- Precision показывает насколько точно алгоритм способен отличить женщину от "неженщины" (доля правильно предсказанных женщин);
- F1 – среднее гармоническое precision и recall;
- AUC – площадь под ROC-кривой⁷.

Полученные результаты качества работы алгоритмов представлены в табл. 3. Также были построены ROC-кривые (см. рис. 11).

⁷https://en.wikipedia.org/wiki/Receiver_operating_characteristic, дата обращения: 25.04.2021

Method	Trained on	Precision	Recall	F1	AUC
EfficientNet-B0	Our	0.924	0.872	0.897	0.901
EfficientNet-B2	Our	0.911	0.887	0.899	0.894
EfficientNet-B4	Our	0.902	0.897	0.9	0.899
Gil Levi model	Adience	0.893	0.6	0.718	0.741
VGG-Face	IMDb-WIKI	0.649	0.985	0.782	0.546
FairFace	FairFace	0.829	0.897	0.862	0.798

Таблица 3: Качество работы алгоритмов на тестирующем наборе данных

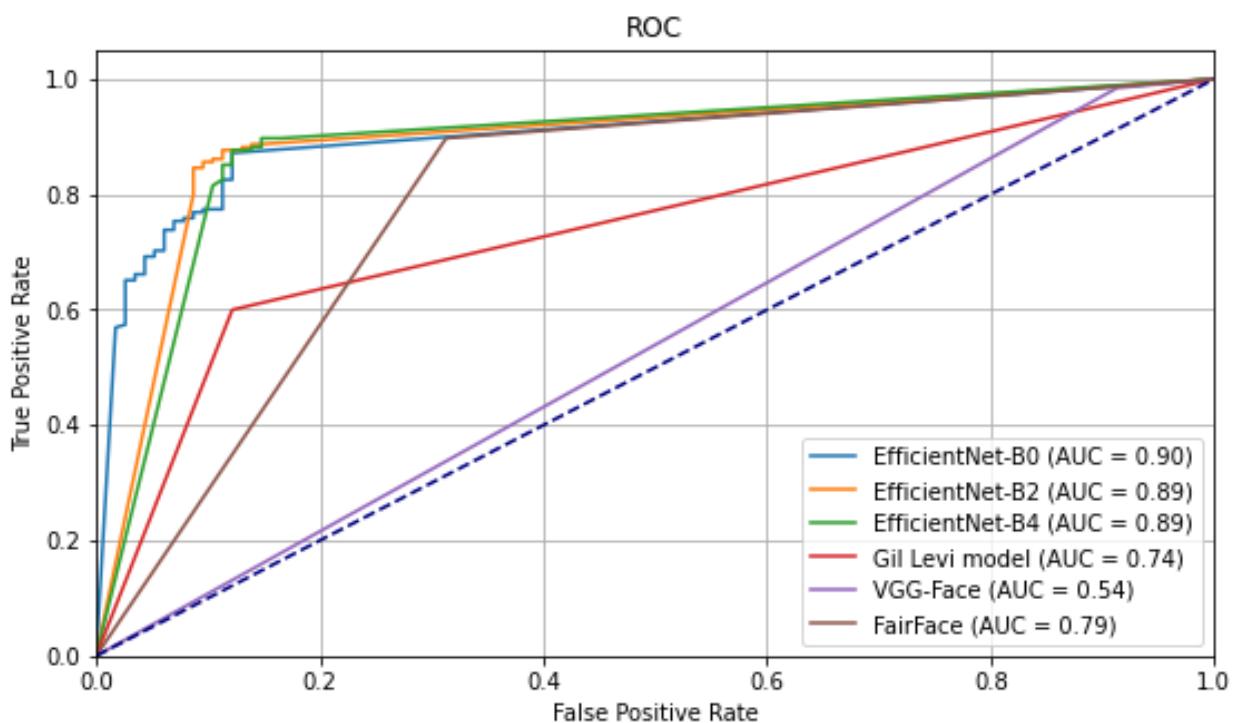


Рис. 11: ROC-кривая на тестирующем множестве

Анализируя таблицу, можно сделать следующие выводы.

- Наши построенные алгоритмы показывают наивысшие результаты по качеству работы.
- Разницу в качестве между выбранными архитектурами (B0, B2, B4) будем считать незначительной.

- FairFace показывает второе место по качеству работы, после наших алгоритмов.

5.2. Скорость работы алгоритмов на CPU

Для получения скорости работы алгоритмов возьмем изображение лица размером 84*84 и прогоним его через каждый алгоритм 1000 раз. Тестирование скорости работы будет производиться на CPU. Характеристики среды выполнения:

- CPU – Intel(R) Core(TM) i5-6200U CPU @ 2.30GHz 2.40 GHz;
- RAM – 16 Gb;
- OC – Windows 10 x64;
- ЯП – Python 3.9.1.

Результаты скорости работы алгоритмов представлены в табл. 4. и на рис. 12.

Method	Min time	Max time	Avg time
EfficientNet-B0	0.24	0.39	0.28
EfficientNet-B2	0.39	1.2	0.51
EfficientNet-B4	0.76	1.66	0.87
Gil Levi model	0.12	0.24	0.13
VGG-Face	1.57	2.87	1.75
FairFace	0.87	1.61	0.94

Таблица 4: Скорость работы алгоритмов на CPU, включая мин., макс., среднее время

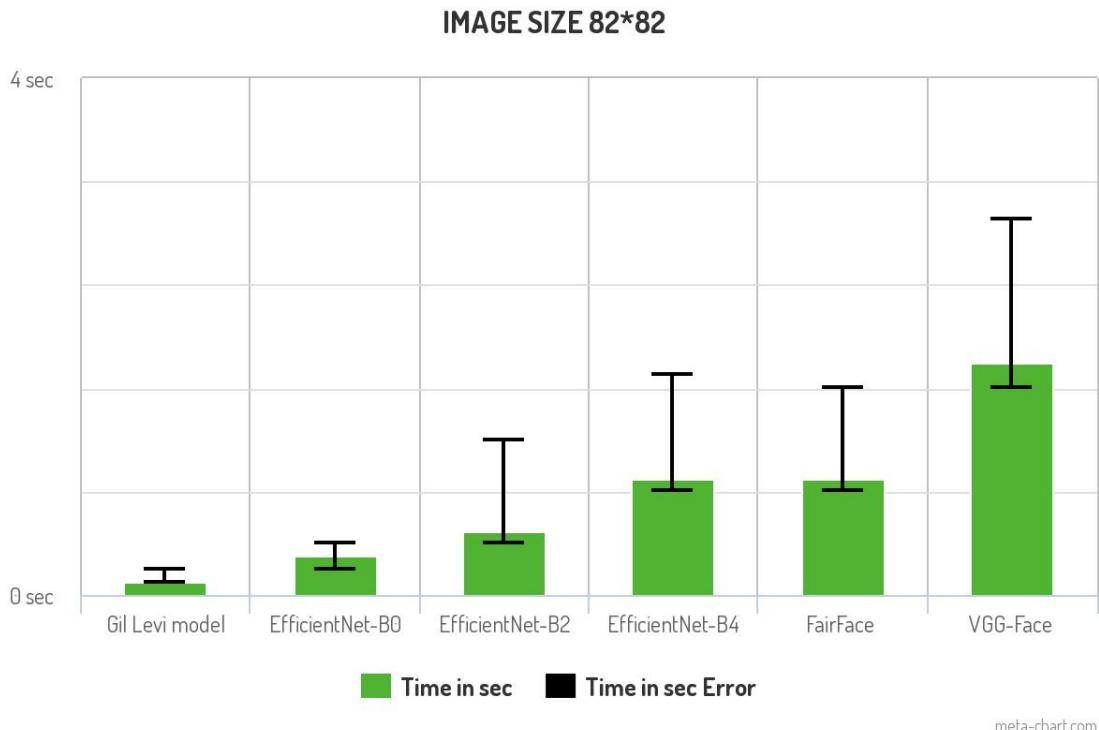


Рис. 12: Скорость работы алгоритмов на CPU, включая мин., макс., среднее время

Из полученных данных видно, что наиболее быстрым является алгоритм, построенный Gil Levi и Tal Hassner, со значением 0.13 сек. Вторым по скорости является собственный построенный алгоритм на архитектуре EfficientNet-B0 со значением 0.28 сек.

5.3. Анализ результатов

Были проведены основные тестирования алгоритмов – скорость и качество работы алгоритма. Исходя из полученных результатов, был выбран наиболее подходящий алгоритм – построенный алгоритм EfficientNet-B0. Он показывает одни из самых высоких результатов по качеству и скорости работы.

5.4. Примеры работы построенного алгоритма

Рис. 13, 14, 15 демонстрируют работу построенного алгоритма.



Рис. 13: Пример работы: камера в метро

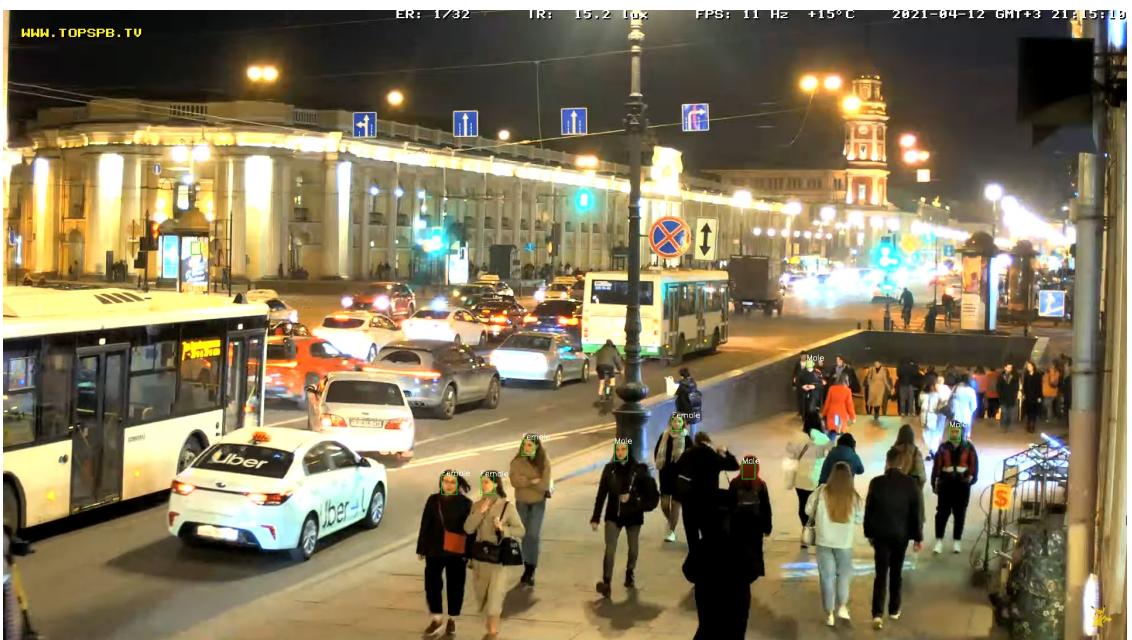


Рис. 14: Пример работы: камера видеонаблюдения

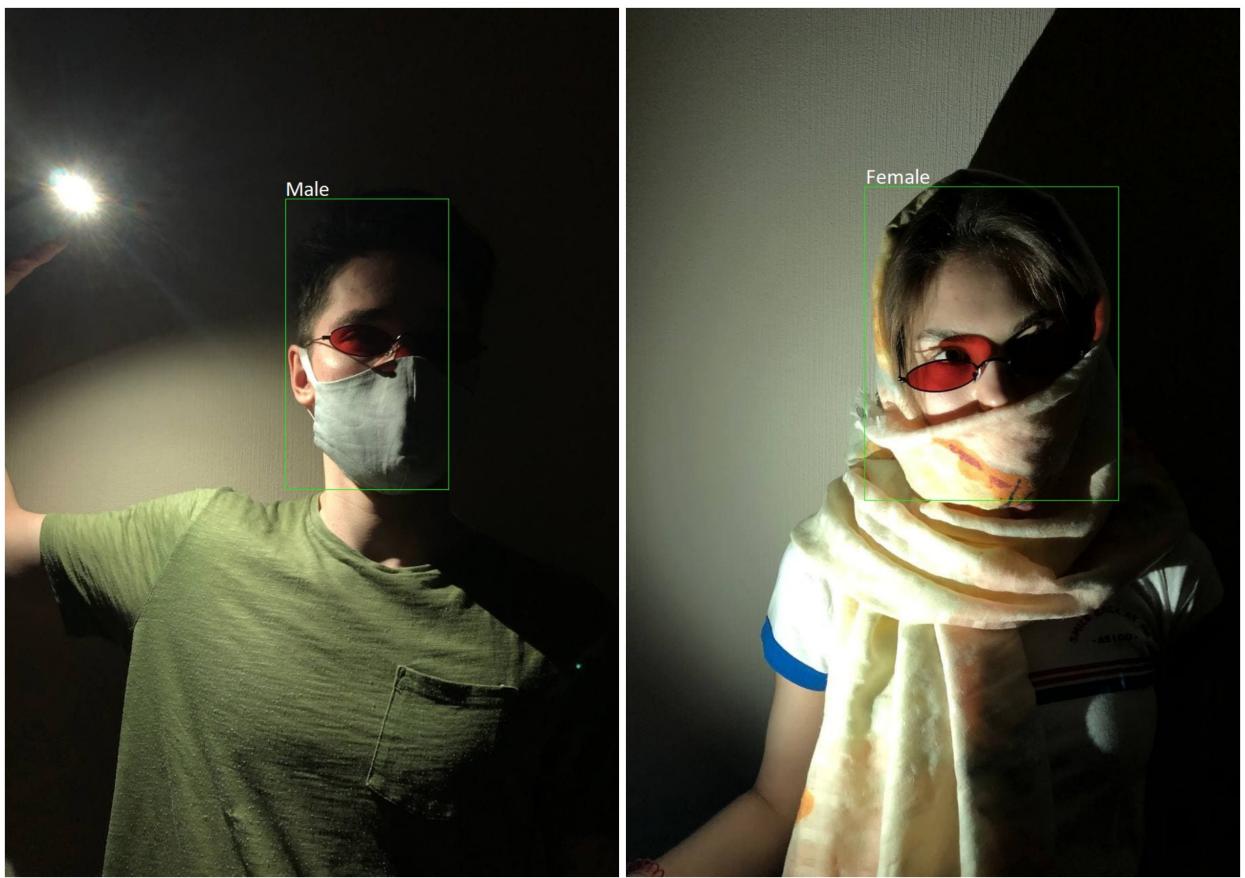


Рис. 15: Пример работы: засвет, маска, очки

6. Интеграция в продукт BelkaSoft Evidence Center X

Продукт BelkaSoft Evidence Center X (BEC X) предназначен для криминалистов, работающими с цифровыми источниками данных. С его помощью криминалист может снять данные с доступного ему источника данных (смартфон, жесткий диск, удаленный ПК). Далее полученные данные анализируются и составляется отчет по найденным уликам. Гендерная классификации поможет криминалистам автоматизировать поиск нужного человека на фотографии для дальнейшей кластеризации или идентификации личности.

Продукт работает под платформой .NET, поэтому необходимо конвертировать построенную модель в формат onnx (Open standard for machine learning interoperability)⁸ для дальнейшего использования с помощью библиотеки onnxruntime⁹. Пример использования построенного алгоритма под платформой .NET (WPF desktop application) доступен по ссылке¹⁰. Рис. 16 демонстрирует работу этого примера.

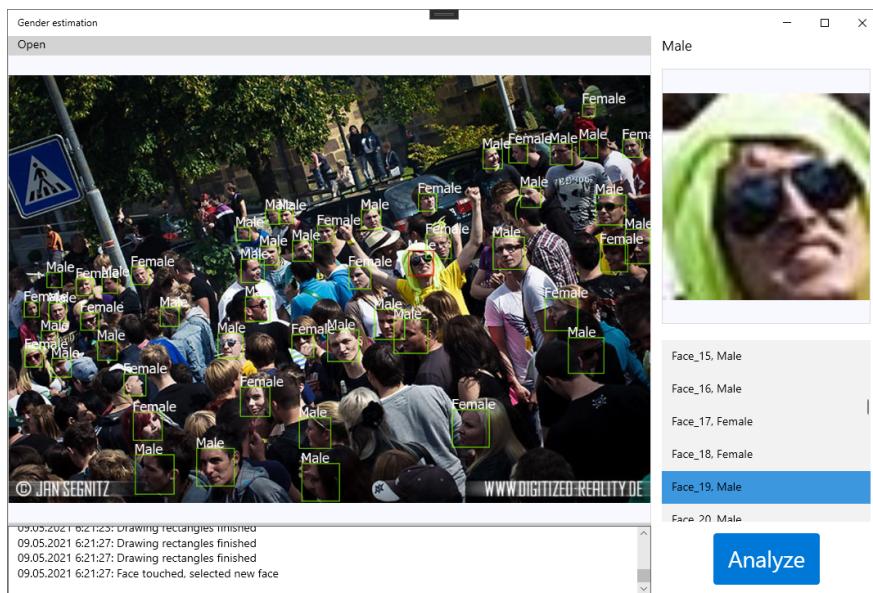


Рис. 16: Пример работы .NET приложения

⁸<https://github.com/onnx/onnx>, дата обращения: 25.04.2021

⁹<https://github.com/microsoft/onnxruntime>, дата обращения: 25.04.2021

¹⁰<https://github.com/Feodoros/ForensicGenderSex/tree/master/PredictGenderWPF>, дата обращения: 25.04.2021

Заключение

В ходе выполнения работы были получены следующие результаты.

- Рассмотрены наиболее актуальные для исследования существующие наборы данных (VGGFace2, DiF, IMDb-WIKI, Adience, Fairface) и алгоритмы (Gil Levi and Tal Hassner model, VGG-Face, FairFace).
- Создан набор данных для обучения алгоритма гендерной классификации, основанный на наборе данных WiderFaces, включающий 12620 лиц людей и удовлетворяющий приведенным криминалистическим факторам. В качестве тестирующего набора данных был собран набор фотографий с камер видеонаблюдения в метро и на улицах, фотографий со смартфонов, а также специально подготовленных фотографий людей в масках и с нанесенным макияжем.
- Создан алгоритм машинного обучения для задачи гендерной классификации на основе архитектуры EfficientNet, получены метрики (recall, precision, F1-мера) на тестовой части обучающего множества.
- Созданный алгоритм был сопоставлен по качеству и скорости работы со следующими существующими алгоритмами: Gil Levi and Tal Hassner model, VGG-Face, FairFace. Созданный алгоритм показывает наивысшие результаты по качеству работы с метрикой F1-мера 0.9, а также является одним из самых быстрых алгоритмов по скорости работы со скоростью обработки фотографии на CPU 0.28 сек.
- Разработанный алгоритм гендерной классификации интегрирован в продукт BelkaSoft Evidence Center.
- Исходный код исследования, а также построенные наборы данных доступны по ссылке¹¹.

¹¹<https://github.com/Feodoros/ForensicGenderSex>, дата обращения: 25.04.2021

- Создано Python-решение, а также десктопное приложение (платформа .NET) для тестирования построенного алгоритма. Решения доступны по ссылке¹².

¹²<https://github.com/Feodoros/ForensicGenderSex/tree/master/PredictGenderWPF>, дата обращения: 25.04.2021

Список литературы

- [1] Barret Zoph Quoc V. Le. Neural Architecture Search with Reinforcement Learning. — 2017. — Access mode: <https://arxiv.org/abs/1611.01578>.
- [2] Eran Eidinger Roee Enbar, Hassner Tal. Age and Gender Estimation of Unfiltered Faces // Transactions on Information Forensics and Security (IEEE-TIFS), special issue on Facial Biometrics in the Wild. — 2014. — Vol. 9. — P. 2170 – 2179.
- [3] ImageNet Large Scale Visual Recognition Challenge / Olga Russakovsky, Jia Deng, Hao Su et al. // International Journal of Computer Vision (IJCV). — 2015. — Vol. 115, no. 3. — P. 211–252.
- [4] Karkkainen Kimmo, Joo Jungseock. FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age for Bias Measurement and Mitigation // Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. — 2021. — P. 1548–1558.
- [5] Levi Gil, Hassner Tal. Age and Gender Classification using Convolutional Neural Networks // IEEE Workshop on Analysis and Modeling of Faces and Gestures (AMFG).
- [6] Michele Merler Nalini Ratha Rogerio S. Feris John R. Smith. Diversity in Faces // <https://arxiv.org/pdf/1901.10436.pdf> (online; accessed: 25.04.2021).
- [7] Mingxing Tan Quoc V. Le. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks // International Conference on Machine Learning, 2019. — 2019. — Access mode: <https://arxiv.org/abs/1905.11946>.
- [8] Mishkin Dmytro, Sergievskiy Nikolay, Matas Jiri. Systematic evaluation of convolution neural network advances on the Imagenet // Computer Vision and Image Understanding. — 2017. — Access mode: <https://arxiv.org/abs/1608.07603>.

mode: <http://www.sciencedirect.com/science/article/pii/S1077314217300814>.

- [9] Parkhi Omkar. Deep Face Recognition. — 2015. — Access mode: <https://www.robots.ox.ac.uk/~vgg/publications/2015/Parkhi15/parkhi15.pdf>.
- [10] Qiong Cao Li Shen Weidi Xie Omkar M. Parkhi, Zisserman Andrew. VGGFace2: A dataset for recognising faces across pose and age // Conference: 2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018). — 2018. — Access mode: <https://arxiv.org/pdf/1710.08092.pdf> (online; accessed: 25.04.2021).
- [11] Rothe Rasmus, Timofte Radu, Gool Luc Van. Deep expectation of real and apparent age from a single image without facial landmarks // International Journal of Computer Vision. — 2018. — Vol. 126, no. 2-4. — P. 144–157.
- [12] WIDER FACE: A Face Detection Benchmark / Shuo Yang, Ping Luo, Chen Change Loy, Xiaoou Tang // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2016.
- [13] Wikipedia. Transfer learning. — Access mode: https://en.wikipedia.org/wiki/Transfer_learning (online; accessed: 25.04.2021).