|  |  |
| --- | --- |
| **СОГЛАСОВАНО**  Доцент кафедры  ИАНИ ННГУ, к.ф.-м.н.  \_\_\_\_\_\_\_\_\_\_\_\_\_ Д.А. Яшунин  «\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_2019 г. | **УТВЕРЖДАЮ**  Профессор кафедры  ИАНИ ННГУ, д.т.н.  \_\_\_\_\_\_\_\_\_\_\_\_\_ Н.В. Старостин  «\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_2019 г. |

**Пояснительная записка № 5**

**«ПЗ по известным наборам изображений (ПЗ5)»**

**Этап 1. Подготовка обзоров на существующие подходы к решению   
задачи и построение технологического стека**

**НИР «Разработка и реализация программного обеспечения   
для распознавания лиц на групповых фотографиях»**

**(Шифр ПО «AFR»)**

Ответственный исполнитель

\_\_\_\_\_\_\_\_\_\_\_\_ М.М. Годовицын

«\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_2019 г.

**Н. Новгород 2019**

Оглавление

[1 Введение 3](#_Toc5030298)

[2 Датасет AgeDB 3](#_Toc5030299)

[3 Датасет UmdFaces 4](#_Toc5030300)

[4 Датасет MS-Celeb-1M 4](#_Toc5030301)

[5 Датасет VGGFace2 5](#_Toc5030302)

[6 Датасет IARPA Janus Benchmark-C 7](#_Toc5030304)

[7 Датасет Cross-Age LFW 8](#_Toc5030305)

[8 Датасет UTKFaces 9](#_Toc5030306)

[Список используемых источников 11](#_Toc5030307)

# 

# Введение

Для решения задач детекции лица человека на изображении и его идентификации необходим набор данных, на которых будет происходить обучение алгоритма, решающего эти задачи. Таких наборов существует большое множество, но все они не включают в себя лица, которые необходимо распознать с помощью ПО «AFR». Поэтому возникает необходимость создания нового набора данных. Для этого требуется провести анализ существующих решений, с целью выявления принципов построения наборов.

# Датасет AgeDB



Рисунок 1. Примеры изображений.



Рисунок 2. Распределение изображений в датасете.

Содержит 16488 изображений лиц 568 людей с метками о их возрасте. Может быть использован для распознавания лиц при наличии возрастной разницы. Также этот набор может быть использован в инвариантных возрасту экспериментах по распознаванию лиц в реальных условиях: чувствительность алгоритма распознавания лица может быть измерена при увеличении возрастного разрыва между изображениями одного и того же человека. Возрастные метки проставлены точно, так как AgeDB собран *вручную*. Более того, все люди были зафиксированы бесконтрольно, в реальных условиях, то есть имеют разные позы, эмоции на лице, могут быть даже частично загорожены. С целью проведения подобных экспериментов на основе AgeDB была создана база данных LFW: изображения разбиты согласно предопределённому значению возрастной разницы (5, 10, 20 и 30 лет)

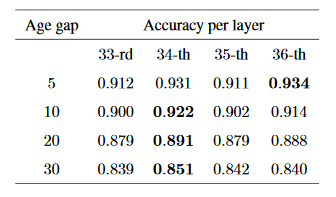


Таблица 1. Зависимость точности распознавания лиц от возраста и номера слоя глубокой сети VGG Face на датасете LFW

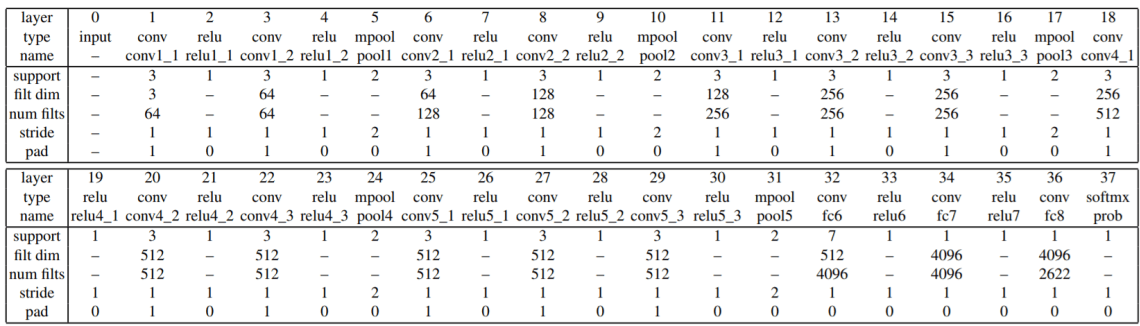


Таблица 2. Архитектура VGG Face.

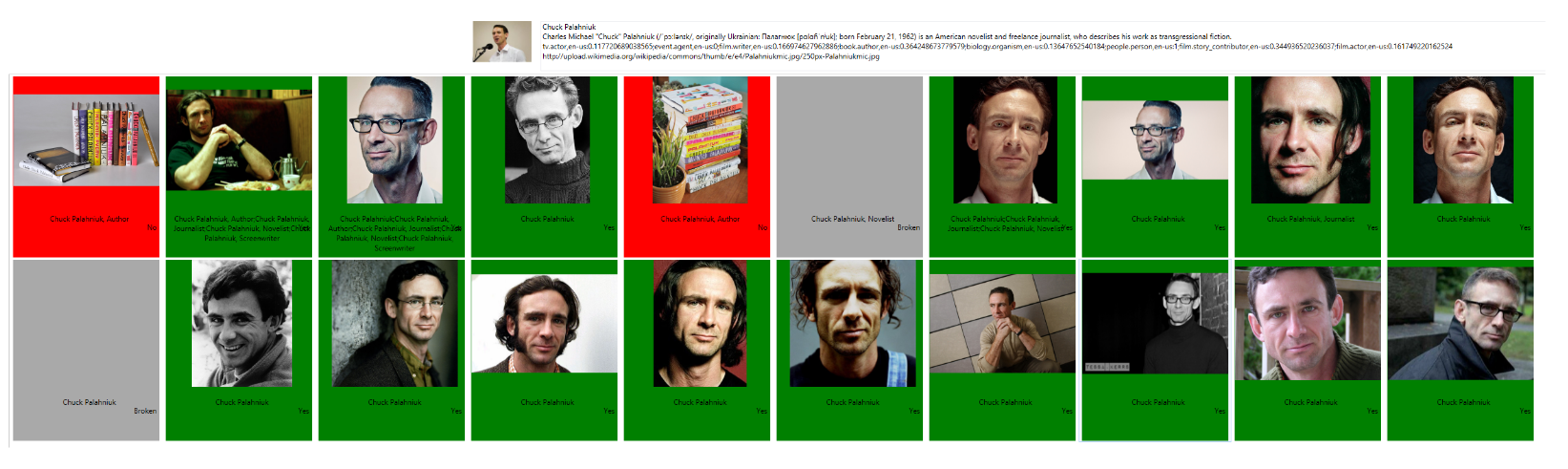
# http://umdfaces.io/header.jpgДатасет UmdFaces

Рисунок 3. Примеры изображений.

Этот набор был пропущен через свёрточную нейронную сеть и затем снабжён информацией о положении ограничительной рамки, половой принадлежности, которые затем были проверены людьми и скорректированы. Набор состоит из 367888 изображений для 8277 людей, разделённых на непересекающиеся тренировочные и тестовые части. Может быть использован для задач детекции и распознавания лиц.

# Датасет MS-Celeb-1M

Рисунок 3. Пример изображений.



Набор, состоящий из 10 миллионов изображений лиц знаменитостей. Он используется больше для оценки метрики расстояния между похожими лицами. Он намеренно включает в себя лица похожих друг на друга знаменитостей (и даже близнецов), а также очень разные изображения одного человека (разный возраст, наличие макияжа, проведённые пластические операции, в том числе по смене пола).

Выходные данные, получаемые при использовании этого набора, более близки к привычной человеческой оценке, так как распознанное лицо затем связывается с соответствующей о нём информацией в базе знаний (если распознали известную актрису, можно будет узнать не только её имя, но и также сам факт того, что она является актрисой, фильмы, где она снималась и т. д.)

Около 30000 меток были тщательно проверены и выставлены экспертами вручную.

# Датасет VGGFace2

Состоит из 3.31 миллиона изображений 9131 знаменитостей, включающий в себя широкий спектр национальностей и профессий. Может быть использован для задач детекции, распознавания. Изображения различны по позам людей, их возрасту, по освещению и фону. Также набор сбалансирован по половой принадлежности (58% мужчин). На каждого человека в среднем приходится 360 изображений ( от 80 до 843). Имеются подтверждённые людьми данные о положении ограничительных рамок и пяти опорных точек лица. Набор включает в себя тренировочных 8630 классов и 500 тестовых.

Таблица 3. Этапы формирования датасета. (M=Manual, A=Automatic)

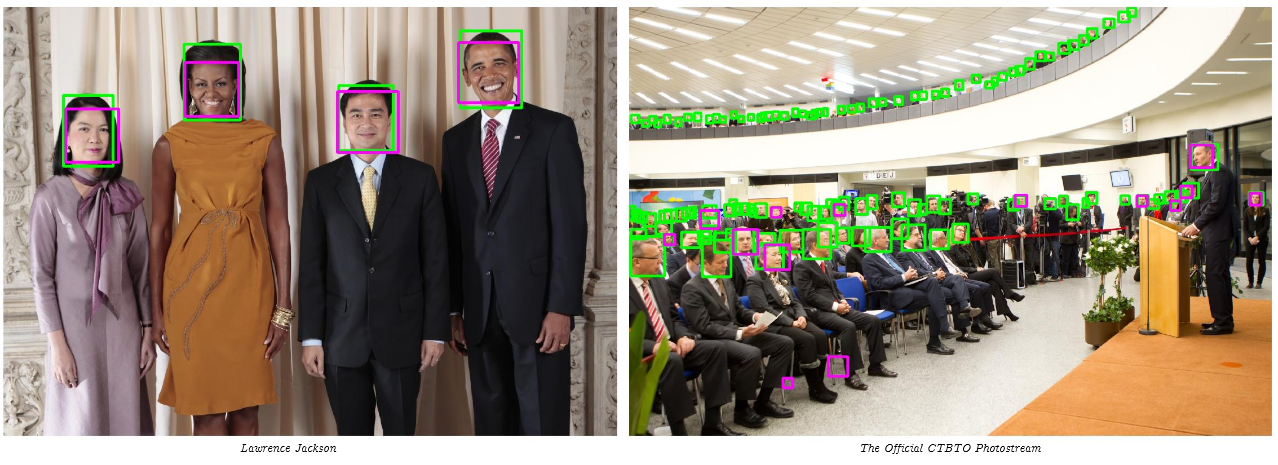
Рисунок 4. Примеры изображений.

# Датасет IARPA Janus Benchmark-C



Рисунок 5. Примеры изображений.

Рисунок 6. Зелёным цветом обозначены ограничительные рамки, информация о которых содержится в датасете. Фиолетовые рамки получены с помощью метода Виолы-Джонса.



Состоит из 11779 (117542 кадра) видеофайлов и 31334 изображений (из них 10040 изображений, на которых лица отсутствуют) 3531 людей. Снабжён информацией о положении ограничительных рамок с помощью Amazon Mechanical Turk. Может быть использован для задачи детекции и распознавания лиц.

# Датасет Cross-Age LFW

Это модифицированный набор LFW с добавлением разницы в возрасте одного и того же человека. Состоит из 13233 изображений лиц 5749 людей, содержит информацию об имени, расе, половой принадлежности. Может быть использован для задач детекции, распознавания лиц.

Процесс построения:

1. сбор необработанных изображений из интернета;
2. запуск детектора лица и ручная проверка полученных результатов в случае, если на снимке присутствуют два и более человека;
3. обрезка и масштабирование обнаруженных лиц;
4. устранение дубликатов;
5. оценка корректности меток;
6. получение опорных точек лица и выравнивание;
7. оценка возраста и формирование тренировочного и тестового наборов: выбор пар с самой большой возрастной разницей в качестве положительных и людей с тем же полом и расой в качестве отрицательных пар.

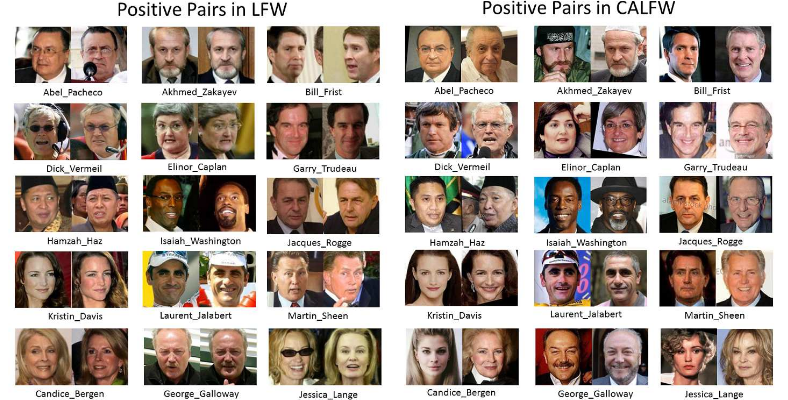
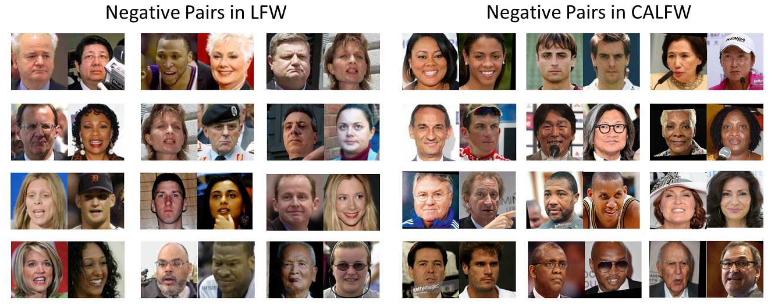


Figure 8. Негативные пары изображений в LFW и модифицированном CALFW

Рисунок 7. Позитивные пары изображений в LFW и модифицированном CALFW.

# Датасет UTKFaces

Состоит из 23708 изображений в естественной среде или с выровненными и обрезанными гранями разрешения 128x128 (присутствует только одно лицо) с 68 опорными точками и метками о половой принадлежности ( 0 или 1 для мужского пола и женского), возрасте (от 0 до 106 лет) и этнической принадлежности (от 0 до 4 для людей с белой кожей, тёмной, азиатов и индийцев). Может быть использован для задач детекции, распознавания, поиска опорных точек лица.



Figure 9. Примеры изображений.

# Список используемых источников

1. <http://openaccess.thecvf.com/content_cvpr_2017_workshops/w33/papers/Moschoglou_AgeDB_The_First_CVPR_2017_paper.pdf>
2. <https://arxiv.org/pdf/1611.01484.pdf>
3. <https://arxiv.org/pdf/1607.08221.pdf>
4. <https://arxiv.org/pdf/1710.08092.pdf>
5. <https://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Klare_Pushing_the_Frontiers_2015_CVPR_paper.pdf>
6. <https://arxiv.org/pdf/1708.08197.pdf>
7. <https://susanqq.github.io/UTKFace/>