ФЕДЕРАЛЬНО ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО»  
(УНИВЕРСИТЕТ ИТМО)**

Дисциплина

«Продвинутое машинное обучение»

**ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ № 2**

**Выполнение экспериментов с моделью YOLOv5**

**Выполнил:**

Студент гр. M4221 Широков М. А.

**Проверил:**

Рюмина Елена Витальевна

Санкт-Петербург

2023 г.

**1 Цели и задачи**

Цель работы: проведение экспериментов с гиперпараметрами модели YOLOv5.

Задача работы состоит в применении нескольких способов обучения модели распознавания эмоций по лицам.

Использованные данные дикторов для:

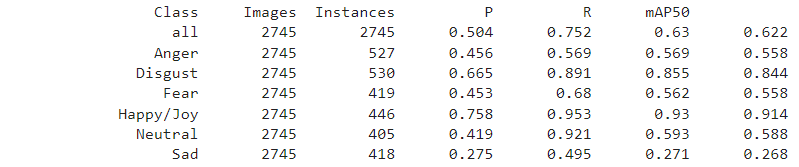
* Обучения – 1001-1077 и 1084-1091
* Тестирования – 1078-1083

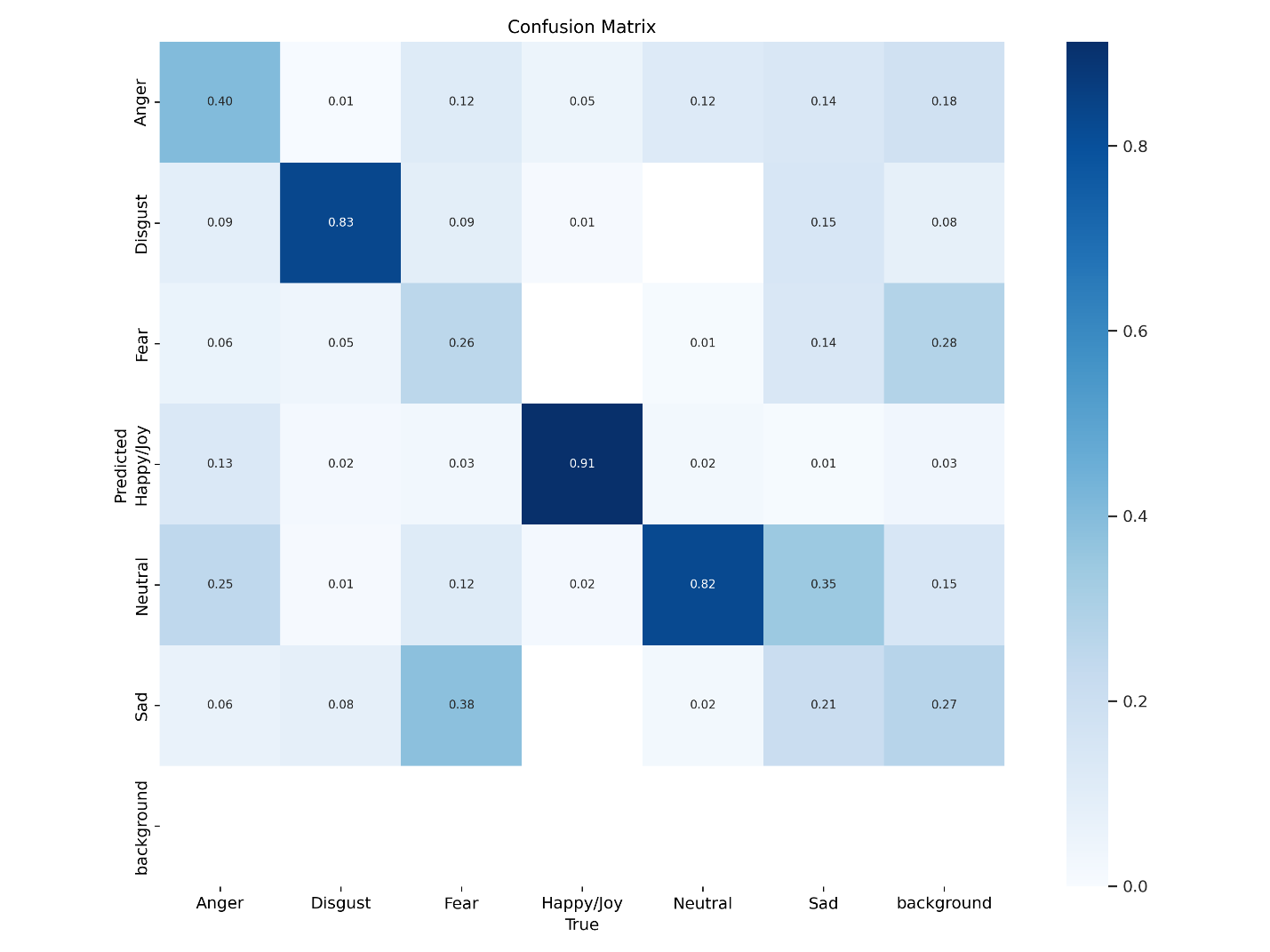
**2 Базовая модель**

Параметры:

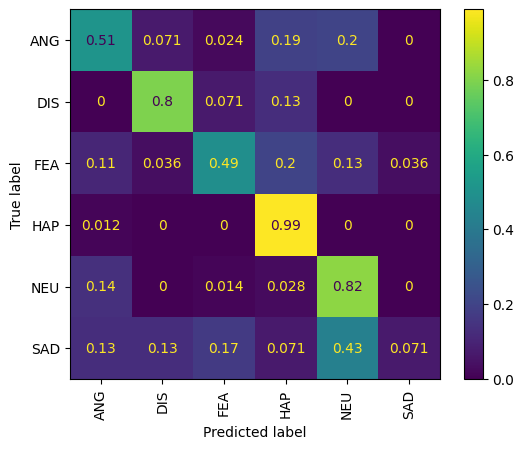
* lr0: 0.001
* lrf: 0.01
* momentum: 0.937
* weight\_decay: 0.0005
* warmup\_epochs: 3.0
* warmup\_momentum: 0.8
* warmup\_bias\_lr: 0.1
* box: 0.05
* cls: 0.5
* cls\_pw: 1.0
* obj: 1.0
* obj\_pw: 1.0
* iou\_t: 0.20
* anchor\_t: 4.0
* fl\_gamma: 0.0
* hsv\_h: 0.015
* hsv\_s: 0.7
* hsv\_v: 0.4
* degrees: 0.0
* translate: 0.1
* scale: 0.5
* shear: 0.0
* perspective: 0.0
* flipud: 0.0
* fliplr: 0.5
* mosaic: 0.5
* mixup: 0.0
* copy\_paste: 0.0
* batch: 16
* epochs: 10
* weights: yolov5l (large)
* img: 512
* optimizer: Adam

Результаты этого эксперимента показаны ниже.





UAR = 0.571. Матрица неточностей по целому видео ниже на картинке.



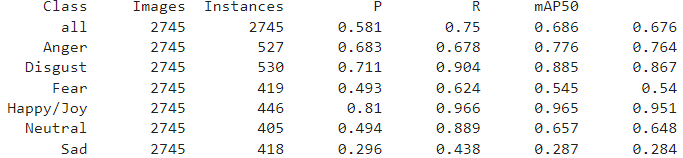
UAR=0.61

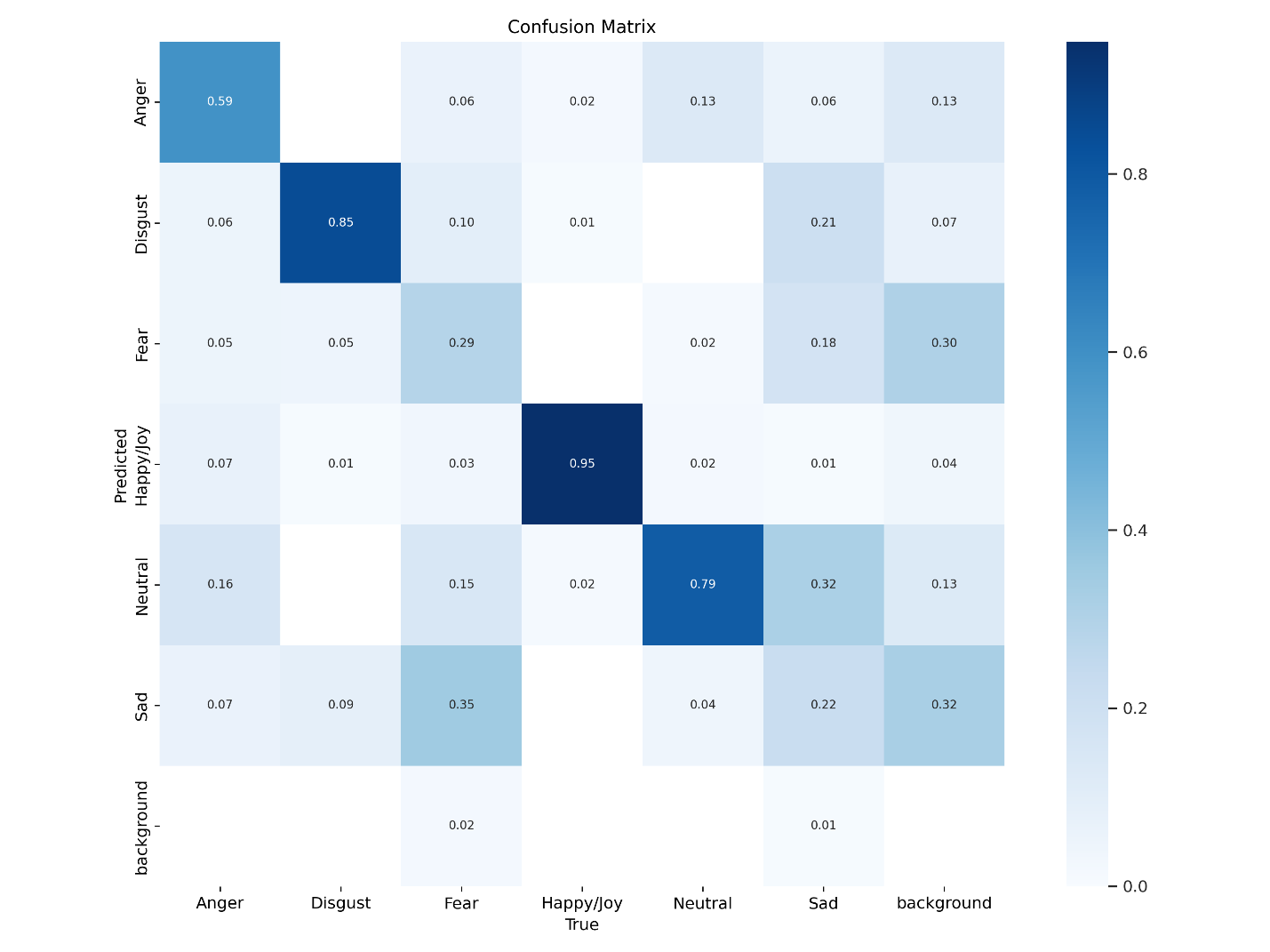
**3 Модель с SGD**

Параметры:

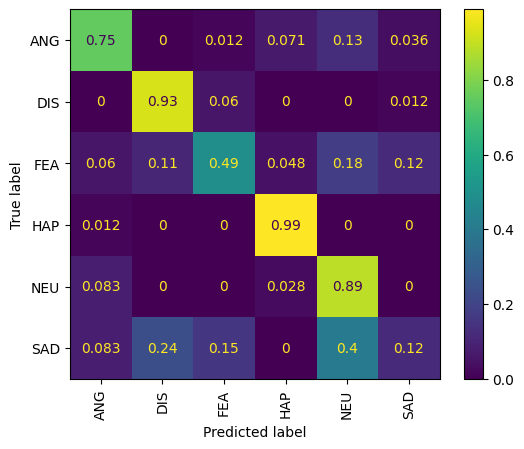
* lr0: 0.001
* lrf: 0.01
* momentum: 0.937
* weight\_decay: 0.0005
* warmup\_epochs: 3.0
* warmup\_momentum: 0.8
* warmup\_bias\_lr: 0.1
* box: 0.05
* cls: 0.5
* cls\_pw: 1.0
* obj: 1.0
* obj\_pw: 1.0
* iou\_t: 0.20
* anchor\_t: 4.0
* fl\_gamma: 0.0
* hsv\_h: 0.015
* hsv\_s: 0.7
* hsv\_v: 0.4
* degrees: 0.0
* translate: 0.1
* scale: 0.5
* shear: 0.0
* perspective: 0.0
* flipud: 0.0
* fliplr: 0.5
* mosaic: 0.5
* mixup: 0.0
* copy\_paste: 0.0
* batch: 16
* epochs: 10
* weights: yolov5l (large)
* img: 512
* optimizer: SGD

Результаты этого эксперимента показаны ниже.





UAR = 0.615. Матрица неточностей по целому видео ниже на картинке.



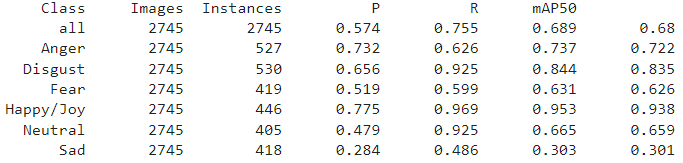
UAR=0.695

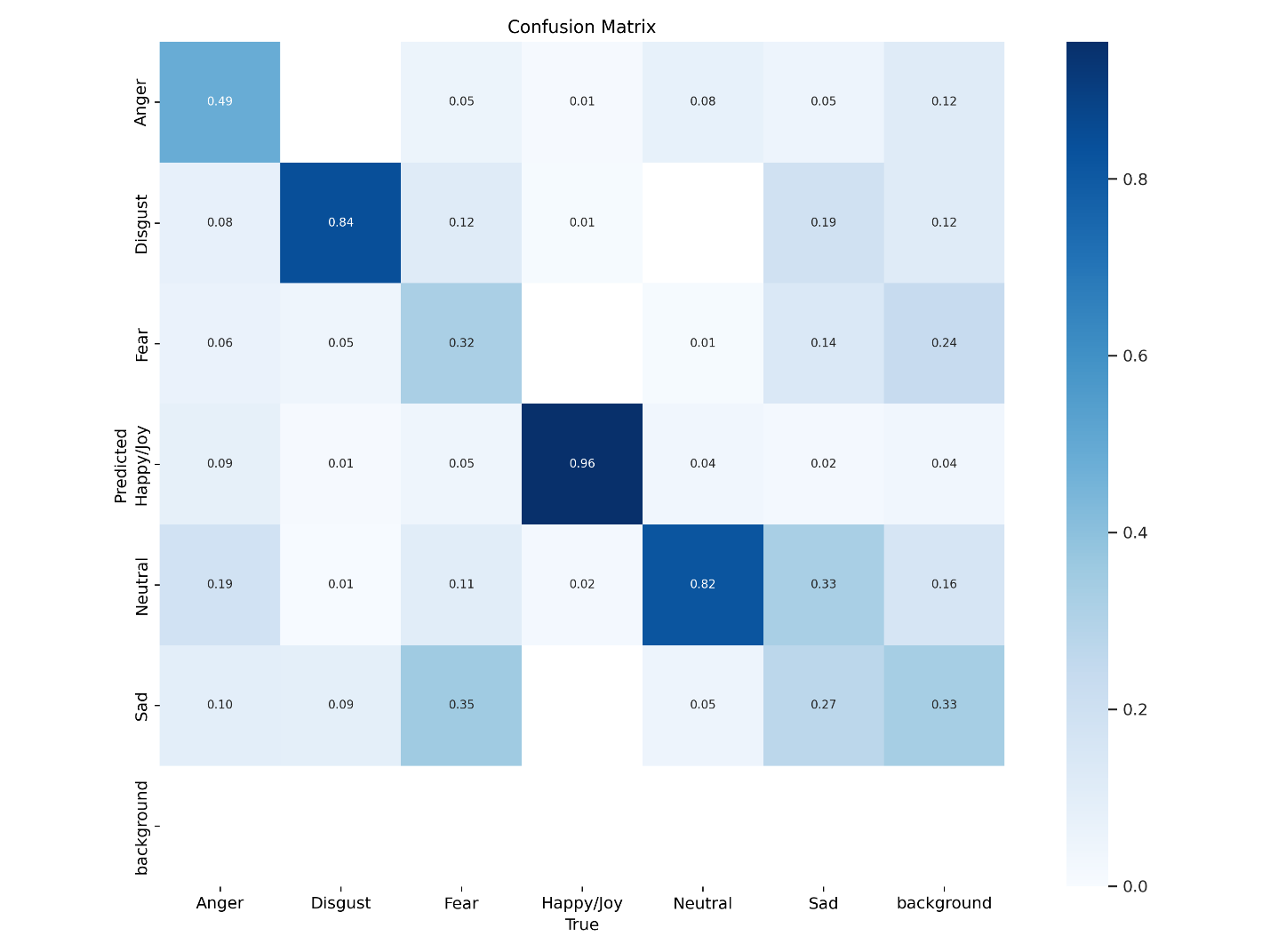
**4 Модель с AdamW**

Параметры:

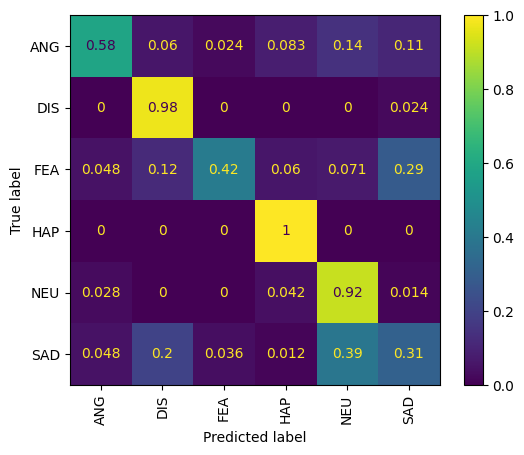
* lr0: 0.001
* lrf: 0.01
* momentum: 0.937
* weight\_decay: 0.0005
* warmup\_epochs: 3.0
* warmup\_momentum: 0.8
* warmup\_bias\_lr: 0.1
* box: 0.05
* cls: 0.5
* cls\_pw: 1.0
* obj: 1.0
* obj\_pw: 1.0
* iou\_t: 0.20
* anchor\_t: 4.0
* fl\_gamma: 0.0
* hsv\_h: 0.015
* hsv\_s: 0.7
* hsv\_v: 0.4
* degrees: 0.0
* translate: 0.1
* scale: 0.5
* shear: 0.0
* perspective: 0.0
* flipud: 0.0
* fliplr: 0.5
* mosaic: 0.5
* mixup: 0.0
* copy\_paste: 0.0
* batch: 16
* epochs: 10
* weights: yolov5l (large)
* img: 512
* optimizer: AdamW

Результаты этого эксперимента показаны ниже.





UAR = 0.616. Матрица неточностей по целому видео ниже на картинке.



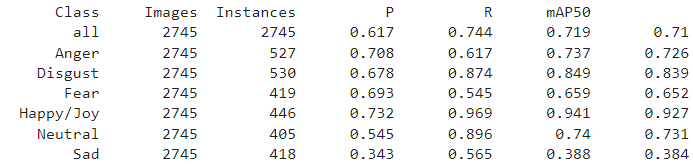
UAR=0.701

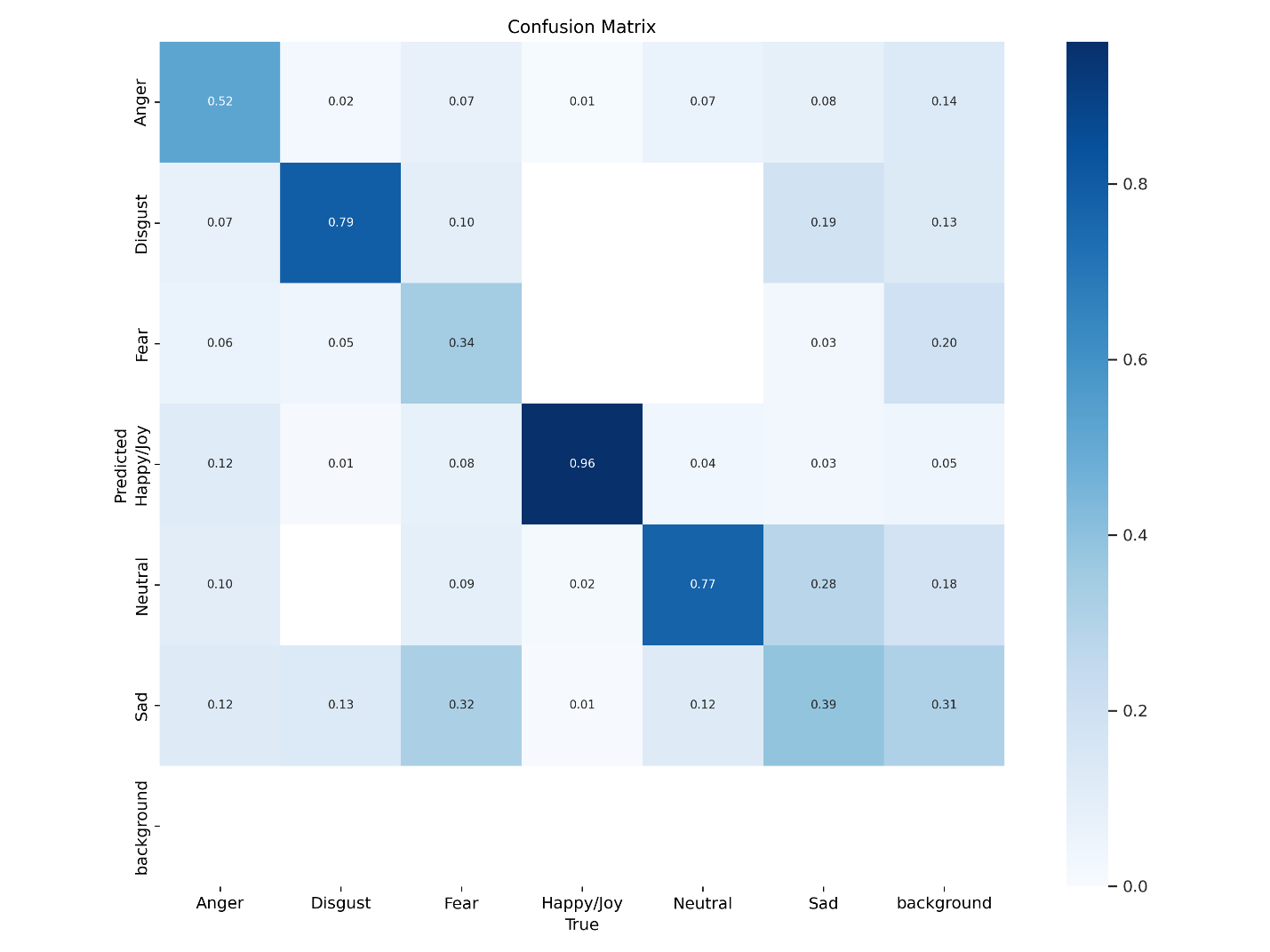
**5 Модель с 20 эпохами**

Параметры:

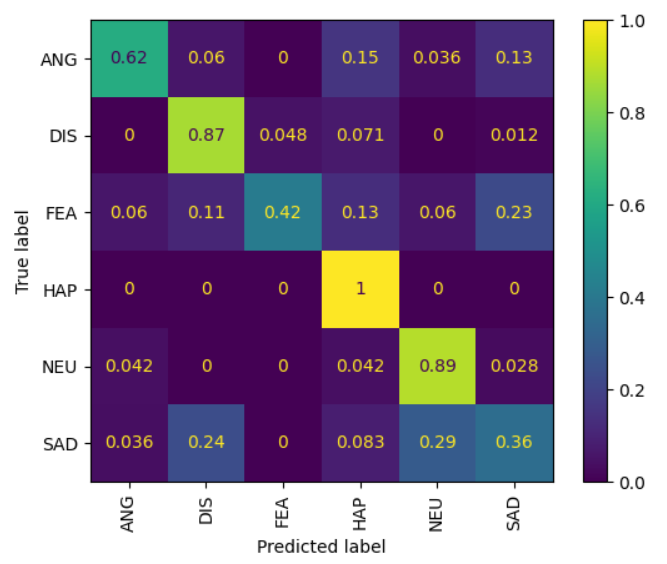
* lr0: 0.001
* lrf: 0.01
* momentum: 0.937
* weight\_decay: 0.0005
* warmup\_epochs: 3.0
* warmup\_momentum: 0.8
* warmup\_bias\_lr: 0.1
* box: 0.05
* cls: 0.5
* cls\_pw: 1.0
* obj: 1.0
* obj\_pw: 1.0
* iou\_t: 0.20
* anchor\_t: 4.0
* fl\_gamma: 0.0
* hsv\_h: 0.015
* hsv\_s: 0.7
* hsv\_v: 0.4
* degrees: 0.0
* translate: 0.1
* scale: 0.5
* shear: 0.0
* perspective: 0.0
* flipud: 0.0
* fliplr: 0.5
* mosaic: 0.5
* mixup: 0.0
* copy\_paste: 0.0
* batch: 16
* epochs: 20
* weights: yolov5l (large)
* img: 512
* optimizer: Adam

Результаты этого эксперимента показаны ниже.





UAR = 0.628. Матрица неточностей по целому видео ниже на картинке.



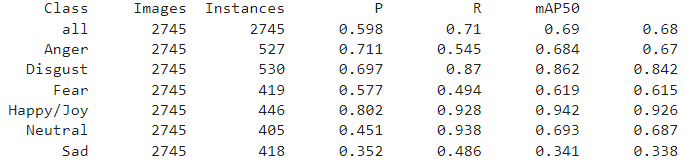
UAR=0.693

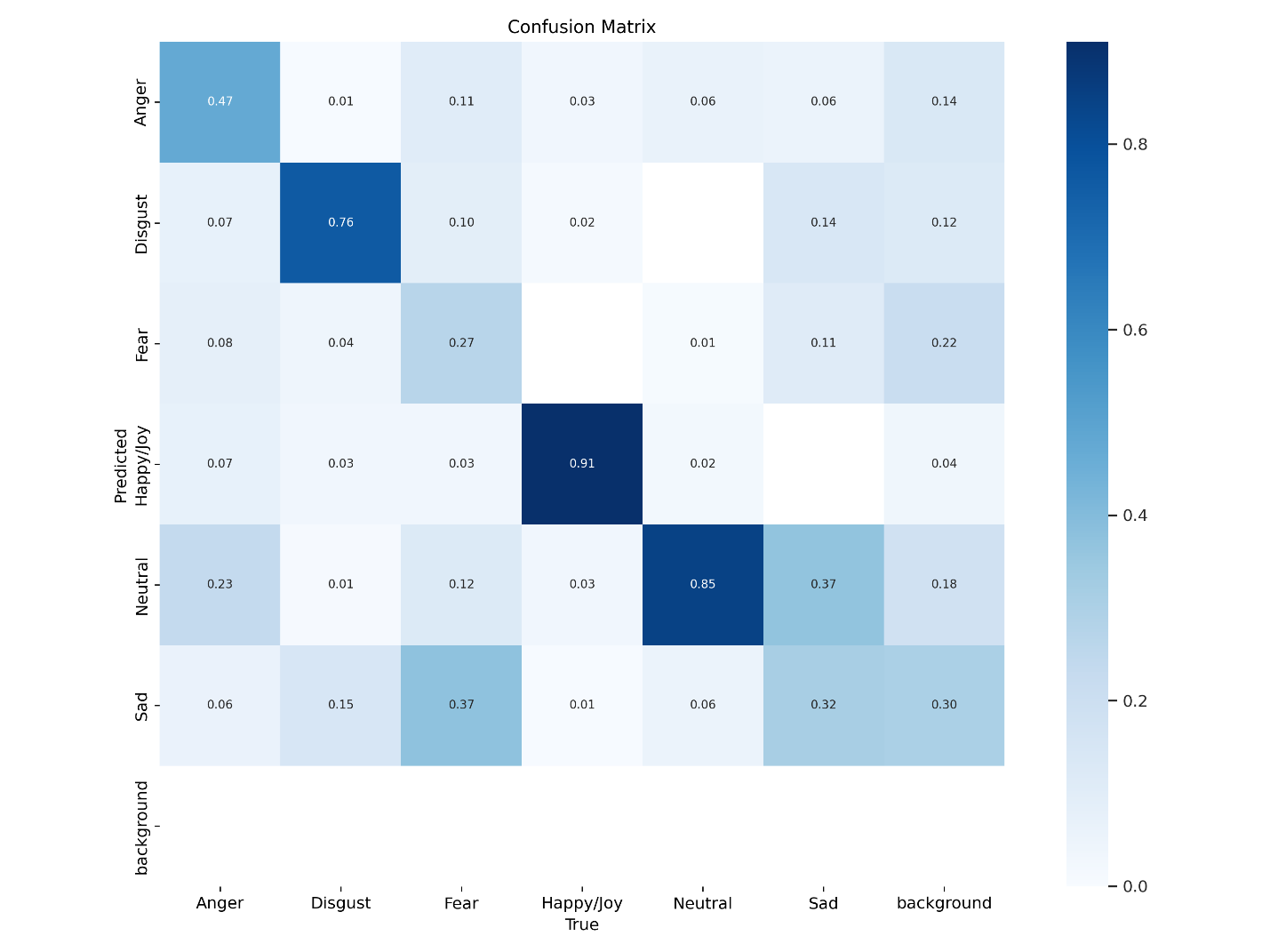
**6 Модель с mosaic 1.0**

Параметры:

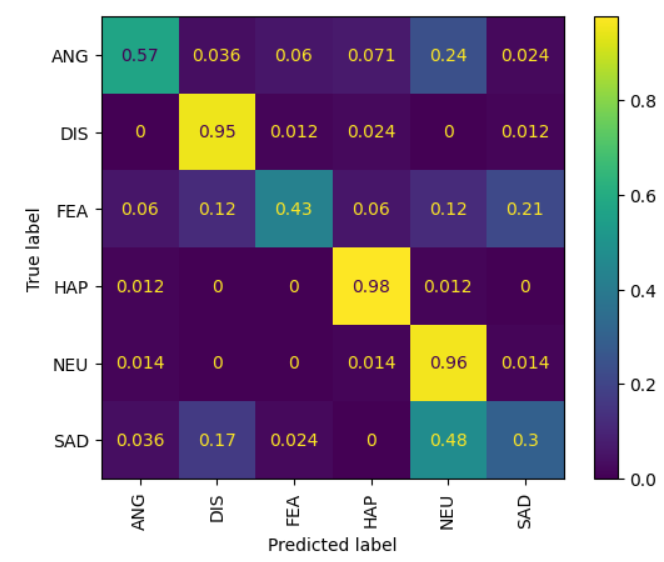
* lr0: 0.001
* lrf: 0.01
* momentum: 0.937
* weight\_decay: 0.0005
* warmup\_epochs: 3.0
* warmup\_momentum: 0.8
* warmup\_bias\_lr: 0.1
* box: 0.05
* cls: 0.5
* cls\_pw: 1.0
* obj: 1.0
* obj\_pw: 1.0
* iou\_t: 0.20
* anchor\_t: 4.0
* fl\_gamma: 0.0
* hsv\_h: 0.015
* hsv\_s: 0.7
* hsv\_v: 0.4
* degrees: 0.0
* translate: 0.1
* scale: 0.5
* shear: 0.0
* perspective: 0.0
* flipud: 0.0
* fliplr: 0.5
* mosaic: 1.0
* mixup: 0.0
* copy\_paste: 0.0
* batch: 16
* epochs: 10
* weights: yolov5l (large)
* img: 512
* optimizer: Adam

Результаты этого эксперимента показаны ниже.





UAR = 0.596. Матрица неточностей по целому видео ниже на картинке.



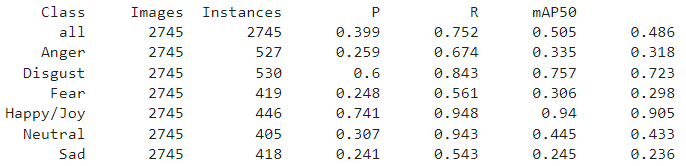
UAR=0.698

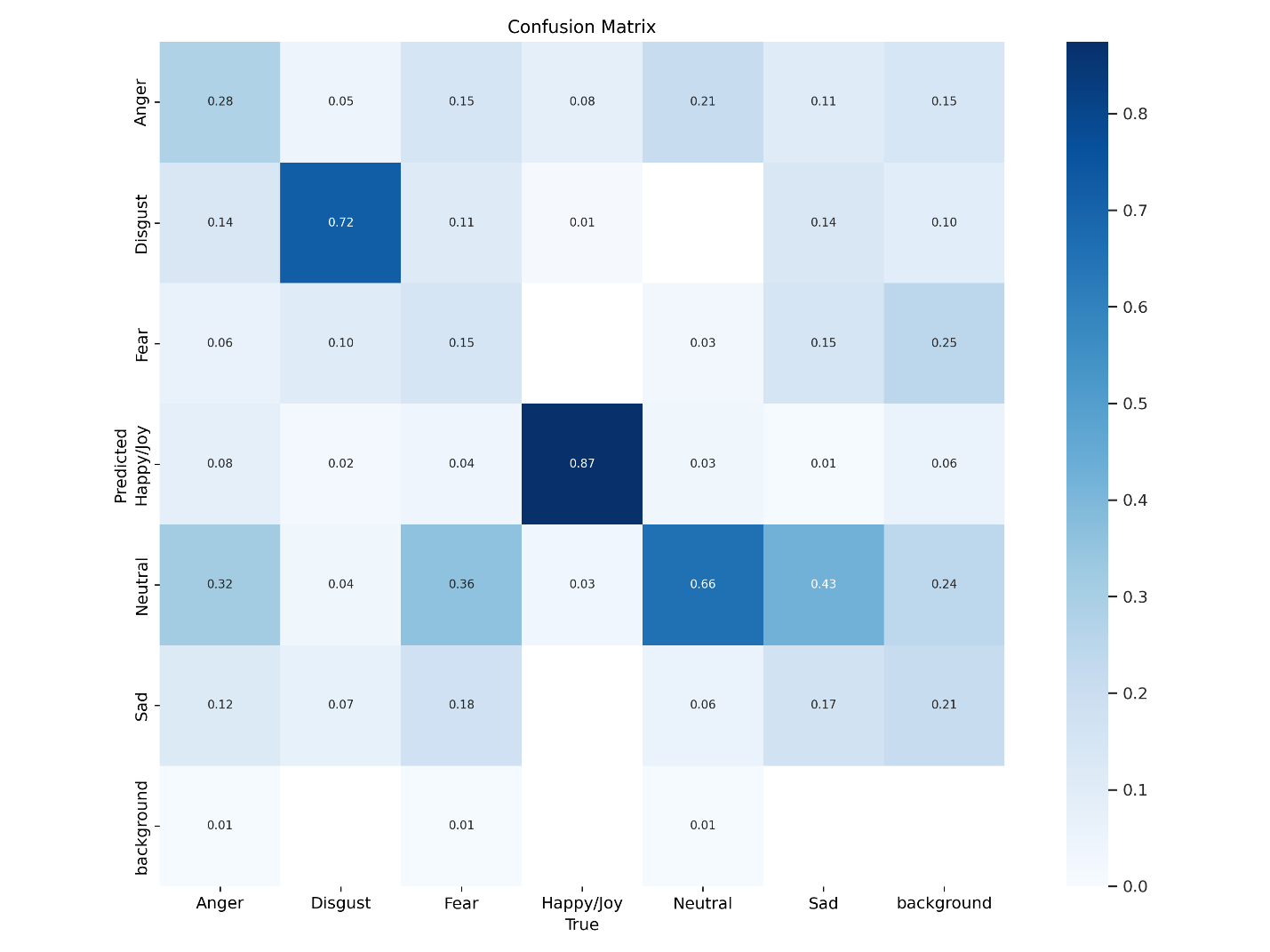
**7 Модель с lr0 0.01**

Параметры:

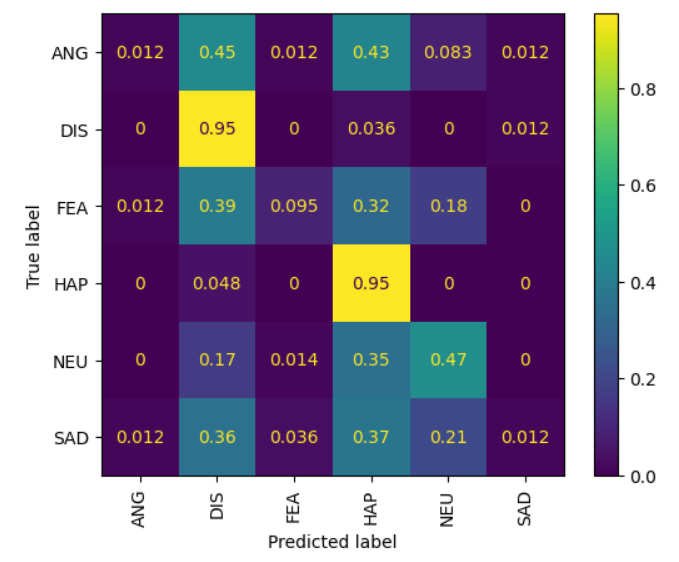
* lr0: 0.01
* lrf: 0.01
* momentum: 0.937
* weight\_decay: 0.0005
* warmup\_epochs: 3.0
* warmup\_momentum: 0.8
* warmup\_bias\_lr: 0.1
* box: 0.05
* cls: 0.5
* cls\_pw: 1.0
* obj: 1.0
* obj\_pw: 1.0
* iou\_t: 0.20
* anchor\_t: 4.0
* fl\_gamma: 0.0
* hsv\_h: 0.015
* hsv\_s: 0.7
* hsv\_v: 0.4
* degrees: 0.0
* translate: 0.1
* scale: 0.5
* shear: 0.0
* perspective: 0.0
* flipud: 0.0
* fliplr: 0.5
* mosaic: 0.5
* mixup: 0.0
* copy\_paste: 0.0
* batch: 16
* epochs: 10
* weights: yolov5l (large)
* img: 512
* optimizer: Adam

Результаты этого эксперимента показаны ниже.





UAR = 0.475. Матрица неточностей по целому видео ниже на картинке.



UAR=0.415

**Выводы**

В ходе лабораторной работы была достигнута поставленная цель. В процессе выполнения заданий были выполнены 6 экспериментов с моделью YOLOv5 large. Были проверены гипотезы зависимости результативности обучения от некоторых гиперпараметров изменения изображения, различных оптимизаторов и их параметров, а также количества эпох.

В результате было выявлено с какой точностью можно определять некоторые эмоции. Лучше всего во всех экспериментах определялись моделью машинного обучения радость, нейтральное состояние и отвращение. А вот остальные три эмоции, включая грусть, злость и страх были подвержены ошибкам. Можно увидеть, что при добавлении аугментации метрики увеличиваются, однако в какой-то момент это перестает происходить, что может быть связано с возможностями самой модели, а также с текущими значениями гиперпараметров. Более сложные методы оптимизации дают лучшие метрики, а также повышается качество при использовании аугментации. Необходимо также отслеживать поведение сети на валидации во время обучения для выявления подходящего количества эпох. В данном случае увеличение количества эпох привело улучшению результатов. Наилучший результат по метрике UAR был равен 0,71 или 71% по 6 классам, что является достаточным для применения на непрофессиональном уровне.