

УДК 004.932.2

**СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ УМЕНЬШЕНИЯ ДИСБАЛАНСА НАБОРА ДАННЫХ ДЛЯ ПОВЫШЕНИЯ ТОЧНОСТИ МУЛЬТИКЛАССОВОЙ КЛАССИФИКАЦИИ ВЫРАЖЕНИЙ ЛИЦ**

**Рюмина Е.В.**

(Санкт-Петербургский институт информатики и автоматизации Российской академии наук, Университет ИТМО)

**Научный руководитель – д.т.н., доцент, Карпов А.А.**

(Санкт-Петербургский институт информатики и автоматизации Российской академии наук)

В работе произведен сравнительный анализ методов уменьшения дисбаланса набора данных для повышения точности мультиклассовой классификации выражений лиц. Результаты исследования подкреплены экспериментальным путем.

Распознавание выражений лица является важной областью компьютерного зрения и искусственного интеллекта. В большинстве случаев наборы данных выражений лиц имеют несбалансированные данные. Не одинаковое представление классов делает многие алгоритмы машинного обучения менее эффективными. При имеющемся дисбалансе данных точность, как метрика оценки производительности системы, является не вполне корректной. Матрица-путаницы является хорошим медом для подведения итогов работы мультиклассовой классификации. В настоящий момент существуют три способа, которые помогают справиться с проблемой дисбаланса классов, а именно синтез, передискретизация и установка большего веса классам, имеющим меньше обучающих примеров.

Определение оптимального метода для уменьшения дисбаланса набора данных для повышения производительности классификации выражений лиц.

Экспериментальные исследования производились на базе данных CREMA-D, которая состоит из 7441 видео выражений лиц 91 актера и 6 разыгрываемых эмоций (злость, удивление, радость, печаль, страх, нейтральность). Перед уменьшением дисбаланса необходимо определить производительность мультиклассовой классификации выражений лиц на несбалансированных данных. Для этого необходимо осуществить ряд действий. С помощью метода гистограммы направленных градиентов (англ. Histogram of Oriented Gradients, HOG) обнаруживаются лица. Далее через обученную модель активной формы (англ. Active Shape Model, ASM) осуществляется обнаружение ключевых точек, которые очерчивают такие объекты лица как брови, глаза, нос, рот и овал лица. В качестве информативных признаков выступают: рассчитанное расстояние между ключевыми точками и их центр масс, расстояние между ключевыми точками и центром масс, а также угловое смещения. Затем полносвязная нейронная сеть (англ. Fully Connected Network, FCN) классифицирует выражения найденных лиц. Построение матрицы-путаницы позволяет оценить какое влияние на производительность оказывает дисбаланс входных данных выражений лиц. Заключительным этапом направлен на сравнение методов уменьшения дисбаланса классов и определение влияния того или иного метода на производительность мультиклассовой классификации выражений лиц, а также определение наиболее оптимального метода уменьшения дисбаланса для решения поставленной задачи.

Предложенный в работе сравнительный анализ позволяет определить оптимальный метод уменьшения дисбаланса. Данное экспериментальное исследование будет полезно для ряда других исследователей, столкнувшихся с проблемой дисбаланса классов.