

РЕФЕРАТ

Отчет 60 с., 32 рис., 3 лист., 10 табл., 21 источ., 6 прил.

КЛАССИФИКАТОР, МОДЕЛЬ, АЛГОРИТМ, ОБЪЕКТ, КЛАССЫ, МЕТКА, ТОЧНОСТЬ, ИЗОБРАЖЕНИЕ

Объект исследования и разработки – алгоритмы многоуровневой классификации объектов изображений.

Цель работы – разработка программного средства на основе новых алгоритмов машинного обучения для многоуровневой классификации объектов космической съемки, обеспечивающих повышение точности и скорости классификации.

В ходе выполнения этапа 2020 года разработаны программные модели на языках C++ и *Python* с использованием библиотеки *OpenCV* предобработки спутниковых снимков общедоступных репозиториях.

Разработана программная модель многоклассовой классификации объектов изображений на основе выходных кодов коррекции ошибок. Выполнен сравнительный анализ метрик модели. Экспериментально установлены оптимальные скорости выходных кодов коррекции ошибок для достижения максимальной точности классификации.

Разработана программная модель многоклассовой классификации по нескольким меткам (многозначной) на основе выходных кодов коррекции ошибок. Установлено, что в многозначной классификации, выходные коды коррекции ошибок, а также скорость кода не влияет на точность, полноту и F-меру.

Разработан метод многоуровневой классификации *REG* по трем признакам на основе композиции классификаторов *Nearest Neighbors (KNN)*, *Linear SVM (SVM)*, *Random Forest (RF)*. Разработана программная модель классификатора, позволяющая выполнять классификацию на основе классов и подклассов и анализ производительности.

ВВЕДЕНИЕ

В результате выполнения предыдущего годового этапа был выполнен сравнительный анализ алгоритмов многоклассовой (multiclass) и многозначной (multipliable) классификации объектов. Рассмотрены методы бинарной классификации, выполнена постановка задачи многоклассовой классификации на основе композиции и выходных кодов коррекции ошибок. Описаны методы многозначной (многофакторной) классификации. Выполнен обзор подходов для оценки качества классификации; выбраны метрики оценки для анализа методов классификации.

Разработана модель многоуровневой классификации объектов изображений на основе выходных самокорректирующихся кодов (*ECOC* - Error-Correcting Output Code); в качестве составных элементов композиционной модели использовались многоклассовые классификаторы. Описан алгоритм кодирования классов и декодирования при выполнении процесса многоуровневой классификации на базе ансамбля N компонентных классификаторов.

Разработано программное средство для оценки качества работы отдельно реализованных классификаторов с использованием библиотеки *Sklearn*. Рассмотрены *UCI* репозитории машинного обучения, библиотеки машинного обучения и нейронных сетей. Разработано первая версия программной модели на основе *ECOC* стратегии кодирования. Выполнен анализ трех наборов данных с разным числом классов. Установлено, что модифицированные методы на основе *ECOC* улучшили показатель точность в 83 % случаев.

Проанализированы базы снимков дистанционного зондирования Земли для получения датасетов при решении задач классификации.