

Машинное обучение в задаче инвентаризации лесных насаждений.

М.Р. Вагизов ФГБОУ ВО «Санкт-Петербургский государственный лесотехнический университет им. С. М. Кирова»

Научный руководитель: д.г.н. проф. Алексеев А.С.

Согласно последним научным данным, площадь лесов занимаемой в Российской Федерации составляет более 840 млн. гектар, Обладая большим потенциалом лесных ресурсов, лесное хозяйство в нашей стране имеет устаревшую информацию о количественном и качественном учёте лесов. Обладая полноценной, актуальной информацией о состоянии и составе лесов, становится возможным качественное управление лесным хозяйством и более точная оценка о состоянии леса как природного ресурса. Большой объём накопленной информации таксационных описаний является хорошей информативной базой для использования технологий машинного обучения в задаче инвентаризации лесных насаждений. С развитием цифровых технологий и методов искусственного интеллекта появляется новое научное направление – цифровая таксация лесов.

Цифровая таксация лесов – является междисциплинарной областью, сочетающая в себе методы, способы и программно-технические инструменты для разработки высокоточных способов анализа и оценки лесов при помощи информационных технологий. Перспективными цифровыми технологиями для решения задачи инвентаризации лесов является класс методов искусственного интеллекта – машинное обучение. В качестве основных источников обработки геопространственной информации для задачи обучения системы являются снимки дистанционного зондирования Земли. По аналогии с методами обработки так называемых больших данных (Big data) требуется разработка технологии анализа при помощи машинного обучения посредством обработки и анализа изображений территории лесного фонда страны.

Сложность процедуры анализа изображений лесных снимков полученными средствами дистанционного зондирования Земли в отличие от обработки числовых данных заключается в следующих обрабатываемых характеристиках:

- большой объём анализа данных
- сложность обработки многомерных недетерминированных объектов
- необходимость создания большой выборки эталонов для процесса обучения
- проработка архитектуры системы и цифровых алгоритмов системы распознавания

Процесс распознавания необходимо разделить на технологические этапы. К первоначальным, сформировавшимся этапам можно отнести формализацию признаков лесных насаждений и формирование шаблонов и правил распознавания. Необходимо отметить, что процедура распознавания должна проводиться с необходимой точностью конечных результатов, при этом требуется разработка двухуровневой системы обучения – стандартная система обработки информации на снимке и глубокая. В задачу стандартной системы машинного обучения входят задачи типизации и кластеризации признаков насаждений на изображениях. В задачи глубоко обучения входит анализ взаимосвязанных

совокупностей таксационных данных и связь этих данных с дешифровочными признаками насаждений на изображениях.

На сегодняшний день реализуется проект по применению технологий и методов машинного обучения в задаче обработки геопространственной информации о лесе на кафедре лесной таксации лесоустройства и геоинформационных систем СПбГЛТУ. Предварительные результаты исследований позволяют сделать вывод, о возможности полной реализации комплексной технологии инвентаризации лесов и расчёта их характеристик на основе методов машинного обучения. Программно-технические решения требуют всесторонней апробации и тестирования на качественных материалах дистанционного зондирования Земли. Решение данной задачи позволит повысить информационное обеспечение данными лесного фонда.