

УДК 004.048

**РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ВАЛИДНОЙ РАЗМЕТКИ
МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ ФОРМАТА DICOM**

Попов А.А. (Университет ИТМО), **Радченко И.А.** (Университет ИТМО), **Яркин А.С.**
(Университет ИТМО)

Научный руководитель – кандидат технических наук, доцент Радченко И.А.
(Университет ИТМО)

Цель работы - разработать программное обеспечение для разметки медицинских изображений формата DICOM. К разрабатываемому ПО сформированы следующие требования: должна реализована возможность работы с форматом DICOM, разметка изображений должна осуществляться с помощью единого унифицированного подхода, понятного для врачей, в результате которого должны генерироваться данные, хранимые в виде метаданных формата DICOM и применимые для решения широкого спектра задач анализа медицинских изображений.

Введение. В медицинских учреждениях по всему миру генерируются петабайты данных, значительной частью которых являются медицинские изображения. К сожалению, обычно эти изображения беспорядочно разбросаны, и в таком виде не поддаются какому-либо анализу.

В последнее время для обработки такого объема данных специалисты стараются внедрить методы машинного обучения. Однако для построения точных моделей требуются качественно подготовленные данные. И в отличие от других направлений интеллектуального анализа изображений, при построении моделей для обработки медицинских изображений в подавляющем большинстве случаев качественные данные могут подготовить только врачи. Но у врачей и без того очень непростая работа и врач не обязан быть экспертом в области машинного обучения и компьютерного зрения. Следовательно, для решения образовавшейся проблемы необходимо разработать инструмент, позволяющий качественно подготовить данные человеку далекому от современных способов автоматизированного анализа изображений, и при этом подготовить их так, чтобы эти особенно ценные данные можно было использовать в максимально широком спектре задач анализа медицинских изображений.

Медицинские изображения хранятся в формате DICOM, который является медицинским отраслевым стандартом создания, хранения, передачи и визуализации цифровых медицинских изображений и документов обследованных пациентов. Найти качественный инструмент для разметки таких изображений крайне трудно. Либо нужно конвертировать изображения в более консервативный формат (JPEG, PNG), при этом теряя важную часть метаданных, либо искать платный инструмент, который будет достаточно дорого стоить и будет представлять собой веб-приложение (например, TrainigData.io, LabalBox и т.п.). Использование такого инструмента небезопасно с учетом специфики формата DICOM и положения о единой государственной информационной системе в сфере здравоохранения. А самое главное, будет достаточно непросто в обращении для врачей, что очень сильно отразится на качестве подготовляемых данных и вообще на желании врачей заниматься разметкой.

Основная часть. На основании вышесказанного к разрабатываемому программному обеспечению сформированы следующие требования: в разрабатываемом ПО должна быть осуществлена возможность работы именно с форматом DICOM, разметка изображений должна осуществляться единым унифицированным подходом, в результате которого должны генерироваться данные, хранимые в виде метаданных формата DICOM и применимые для решения широкого спектра задач анализа медицинских изображений.

Унифицированный подход к разметке заключается в следующем: врачу необходимо выбрать на изображении область с объектом интереса, выделить данную область ограничивающей рамкой и указать имя класса для выделенного объекта. После этого на

выделенный рамкой объект автоматическим образом будет наложена маска, врачу же останется только немного уточнить ее, для повышения качества наложенной маски. Уточнение маски производится максимально простым образом, который заключается в изменении двух параметров, и, следовательно, может быть легко реализовано в графическом пользовательском интерфейсе. Такой подход позволяет получить и хранить в метаданных DICOM-файлов информацию, которая может быть использована для построения качественных моделей классификации, сегментации и обнаружения объектов на медицинских изображениях.

Выводы. Таким образом, программное обеспечение для разметки медицинских изображений формата DICOM, в котором будет реализован, описанный выше унифицированный подход к разметке, может стать удобным инструментом для врачей, что будет способствовать значительному повышению общего количества качественно подготовленных данных для решения задач интеллектуального анализа медицинских изображений.