

УДК 004.85

## РАСПОЗНАВАНИЕ КЛЮЧЕВЫХ ОБЪЕКТОВ ПО ВИДЕОЗАПИСЯМ ХИРУРГИЧЕСКИХ ОПЕРАЦИЙ

Крючков М.И. Университет ИТМО, Ефимова В.А. Университет ИТМО  
Научный руководитель – Фильченков А.А., кандидат физико-математических наук,  
доцент Факультет информационных технологий и программирования,  
Университет ИТМО

В данном докладе представлена работа по распознаванию и локализации ключевых объектов на видеозаписях хирургических операций с целью создания обучающего видеоматериала для врачей. Задачами работы являются локализация лица пациента и дальнейшее его размытие, а также определение моментов времени начала и окончания операции.

**Введение.** Широко известна задача локализации и трекинга объектов на видеозаписях. Особое внимание уделяется локализации лиц, и существуют технологии (Faster RCNN и другие), успешно решающие данную задачу. В контексте данной работы задача становится несколько сложнее, поскольку во время операции врачи-хирурги вынуждены использовать маски и колпаки, которые оставляют открытым разрез глаз и пытаются максимально закрыть оставшуюся поверхность лица, лишая таким образом обычные модели локализации лиц возможности распознать ключевые черты, такие как нос, рот и волосы. При решении этой задачи в нескольких работах предлагалось подстраивать предобученные модели локализации лиц под задачу, что подразумевает дообучение модели на размеченной выборке, полученной искусственно или с помощью ручной разметки. Отметим, что непосредственная задача данной работы является еще более специфичной, поскольку требуется локализовать в первую очередь лицо пациента, которое может аналогично отличаться от лиц, локализуемых обычными моделями, в силу положения пациента на больничной койке, присутствия кислородной маски и т. д. В связи с этим после нахождения всех лиц требуется определить, какое из них принадлежит пациенту, что приводит к задаче классификации

**Основная часть.** Предлагаемое решение состоит в настройке предобученных общих моделей на лица врачей и пациентов в операционных комнатах по выборке картинок, полученной в результате нарезки самостоятельно размеченных видеозаписей, для локализации всех лиц на видео. Данная технология описана в нескольких работах. Далее предлагается использовать сверточные сети для классификации локализованных лиц, обученные на выборке, полученной аналогичным образом.

**Выводы.** Применение технологий компьютерного зрения может оказать значительную пользу в задаче обработки видеозаписей хирургических операций для использования их в качестве учебного материала для врачей. Данная работа призвана решить проблемы автоматической анонимизации, а также автоматической обработки видеозаписей: обрезки информативной части видео, на которой непосредственно происходит хирургическая операция.

Крючков М.И

Ефимова В.А.

Фильченков А.А.