

## Рецензия на доклад Владислава Шахуро «Порождающие модели распознавания изображений»

Доклад посвящен порождающему подходу в задачах распознавания изображений. По большей части рассказ состоял из краткого обзора существующих подходов в этой области, что и понятно — автор находится в самом начале исследований. Доклад был хорошо подготовлен (как презентация, так и устная часть), а также понятно рассказан. Актуальность рассматриваемой проблемы не вызывает сомнений: большинство задач компьютерного зрения сейчас решаются хорошо лишь дискриминативными подходами, которые требуют больших размеченных обучающих выборок, что не всегда возможно.

Во вступлении были рассмотрены две работы из области компьютерного зрения: предсказание позы по картам глубины Kinect и распознавание человека по фотографии. Для задачи распознавания позы был сгенерирован миллион изображений; для распознавания лиц использовалось 4.4 миллиона фотографий. Это лишний раз подтверждает актуальность задачи.

В основной части были представлены пример работы и общая схема порождающего подхода в задачах распознавания изображений. Также были подробно рассмотрены вероятностные модели описания сцен. Стоит отметить подробный рассказ о дифференцируемом движении OpenDR. Насколько я понял, эта особенность позволяет в уже готовой трехмерной модели изменить положение объекта, то есть получить новую сцену с соответствующим преобразованием разметки. Тут было бы интересно услышать, на сколько хорошо модели, обученные на подобных синтетических примерах, затем работают на реальных данных.

Далее были подробно описаны методы генерирования гипотез. Так как изначально сцена описана вероятностным языком, для сэмплирования гипотез предлагается использовать различные подходы, основанные на создании цепей Маркова, например, алгоритм Метрополиса-Гастингса. Также в случаях, когда возможно вычислить производную по параметрам сцены, предлагается использовать метод hamiltonian Monte Carlo, позволяющий уменьшить корреляцию между удачными сэмплами для увеличения скорости сходимости метода.

В завершающей части был показан пример работы порождающего подхода. На этом слайде хотелось бы увидеть весь процесс сходимости алгоритма из начального приближения к готовой модели и оценки скорости сходимости. Еще отмечу, что слайд с резюме приводит некоторые проблемы представленного подхода, однако решения этих проблем отсутствуют. Я понимаю, что работа в этом направлении как раз является частью будущей диссертации автора, однако хотелось бы увидеть какие-нибудь идеи решения сформулированных проблем.

Основной целью доклада было знакомство слушателей с областью применения порождающих моделей в задачах распознавания изображений, с чем автор отлично справился. К небольшим недостаткам я бы отнес небольшое количество примеров (компьютерное зрение является той редкой областью, где очень интересно и полезно изучать визуализации работы алгоритмов), а также отсутствие собственных мыслей по поводу решения проблем рассматриваемой области. К счастью, первый недочет легко решается благодаря приведенным ссылкам на оригинальные работы.