

УДК 004.853

## ИСПОЛЬЗОВАНИЕ АВТОКОДИРОВЩИКОВ В ЗАДАЧЕ ВИЗУАЛЬНОЙ УНИВЕРСАЛЬНОЙ ДОМЕННОЙ АДАПТАЦИИ

Бобров О.Б. (Университет ИТМО)

Научный руководитель – к.ф.-м.н., доцент Фильченков А.А,  
(Университет ИТМО)

В работе исследуется применение автокодировщиков для решения задачи универсальной доменной адаптации в компьютерном зрении.

**Введение.** Задача универсальной доменной адаптации — это частный случай задачи неконтролируемой доменной адаптации без ограничений на соотношение между множествами классов исходного и целевого доменов: постановка задачи подразумевает, что помимо общих классов, присутствующих в обоих доменах, у каждого из доменов могут быть еще свои частные классы, отсутствующие в другом домене.

Существует несколько решений, предлагаемых для решения задачи универсальной доменной адаптации в компьютерном зрении. Часть из них основывается на сочетании состязательной доменной адаптации и неопределенности предсказания для детекции частных классов целевого домена. Также были предложены решения основанные на использовании поклассовых бинарных классификаторов и на сопоставлении кластеров латентных представлений.

Все существующие решения используют глубокие методы обучения, требующие больших вычислительных ресурсов. В данной работе мы предлагаем использовать неглубокие автокодировщики для решения задачи визуальной универсальной доменной адаптации с использованием меньшего количества ресурсов.

**Основная часть.** В данной работе предлагается решение на основе неглубоких нейронных сетей. Предлагаемая архитектура решения содержит энкодер, декодер и четыре классификатора. Энкодер и декодер формируют автокодировщик и используются для детекции частных классов целевого домена. Латентное представление энкодера делится на две части: семантическую и доменную. Семантическое латентное представление отвечает за кодирование информации о содержании картинки, в то время как доменное латентное представление кодирует информацию об особенностях изображения, специфических для того или иного домена. Разделение латентного представления на две части, отвечающие за универсальную и специфическую для доменов информацию, обеспечивается обучением классификатора классов и классификатора доменов на каждом из них. Однако градиенты от классификатора доменов на семантическом латентном пространстве и классификатора классов на доменном передаются в энкодер с умножением на отрицательный коэффициент. Для детекции частных классов целевого домена предлагается использовать семантическое латентное представление, так как именно оно несет информацию о классе объекта. Были опробованы различные способы для детекции новых классов на основе семантических латентных векторов, такие как метод ближайших соседей, различные методы кластеризации, детекция с помощью ошибки восстановления автокодировщика.

**Выводы.** Использование предлагаемой архитектуры позволило достичь точности 55% на стандартном наборе данных Office-Home с разбиением на 10 общих классов, 5 частных классов для исходного домена и 50 частных классов для целевого домена. При этом данная точность была достигнута с использованием всего лишь 600Мб врам.

Бобров О.Б. (автор)

Подпись

Фильченков А.А. (научный руководитель)

Подпись