

Разработка и исследование информационной системы распознавания трёхмерного объекта по его изображению

Д.И. Берлин¹, Е.В. Гошин¹

¹Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34а, Самара, Россия, 443086

Аннотация

В работе представлена информационная система детектирования трехмерного объекта на изображении при помощи искусственных нейронных сетей. Поставленная задача сведена к задаче регрессии, решение которой предоставляется информационной системой, с возможностью дальнейшей оценки и корректировки.

Ключевые слова

Искусственные нейронные сети, трехмерный объект, компьютерная графика, CNN

1. Введение

Задачи распознавания и детектирования объектов решает большое количество исследователей различными способами, это одна из наиболее востребованных задач нашего времени. Как результат, на данный момент существует множество методов распознавания и детектирования объектов и их классификации.

В общих чертах все они, так или иначе, решают задачу кибернетики, получая изображение или видео, как входные данные, обрабатывают их и выдают результат в виде текста или выделенной области на исходном изображении [1, 2]. Зачастую, они не представляют каких-либо возможностей для обратной связи, а количество информации об исходном объекте минимально. Полученный текст, является всего лишь текстом и работа с ним – это отдельный класс задач. Обладая большей информацией об искомом объекте, а именно трехмерной моделью, становится возможным решать задачу кибернетики на более высоком уровне.

Основное внимание в данной работе уделяется построению и анализу информационной системы позволяющей обнаружить заранее заданный трехмерный объект на двухмерном изображении.

2. Постановка задачи

Исходный трехмерный объект храниться в памяти в виде нормализованной полигональной модели. Для того чтобы определить его положение на изображении необходимо определить матрицу преобразований [3] трехмерного объекта относительно некоторого нулевого положения. В частности данная матрица определяет такие характеристики объекта в пространстве, как изменение его положения, углы поворота вдоль осей, а так же масштабирование. Таким образом, задача детектирования трехмерной модели сводиться к задаче определения произвольных численных преобразований, то есть задачи регрессии.

3. Архитектура системы

Информационная система представлена в виде трех связанных модулей с разделением ответственности.

1. Первый, отвечает за подготовку набора данных по предоставленному трехмерному объекту, для дальнейшего обучения, а так же за первичную настройку модели. Пример начальных данных приведен на рисунке 1.

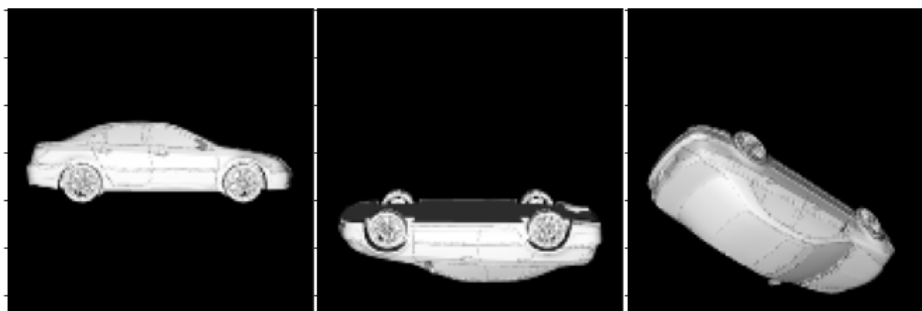


Рисунок 1: Входные изображения

2. Второй, проводит обучение модели, выдает результаты обучения, а так же оценивает полученную модель.
3. Третий, основной модуль, использующийся для решения поставленной задачи, то есть детектирования трехмерного объекта.

4. Проведение вычислительного эксперимента

Для оценки модели был взят тестовый объект, примеры которого изображены на рисунке 1. Обучающая выборка была сгенерирована исходя из равномерного вращения объекта с градусным шагом 5° . Тестовая выборка представляет собой объект, повернутый на случайную величину, таким образом, в среднем 80% экземпляров тестовой выборки являются уникальными. Для обучения модели использовалась видеокарта NVIDIA GeForce GTX 1050 TI. Время обучения модели составило ~ 13 мин, после обучения модели, среднее абсолютное отклонение (MAE) на тренировочной и тестовой выборке составили $\sim 8^\circ$, $\sim 10^\circ$ соответственно, пример отклонения объекта на 10° по каждой из двух координат изображены на рисунке 2.

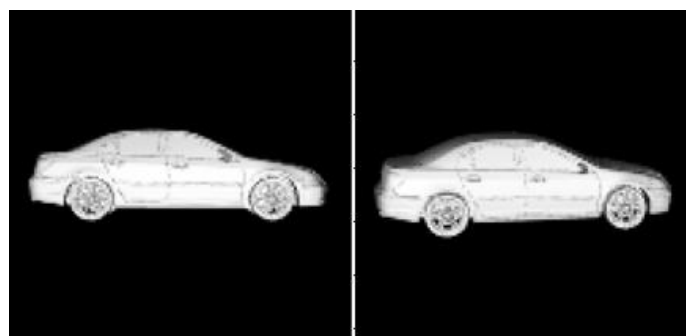


Рисунок 2: Поворот объекта на 10 градусов

5. Заключение

В ходе проделанной работы была построена система для обучения и дальнейшего использования нейронной сети, позволяющей детектировать трехмерный объект на изображении. Произведен качественный анализ построенной системы.

6. Литература

- [1] Liu, L. Learning a Rotation Invariant Detector with Rotatable Bounding Box / L. Liu, Z. Pan, B. Lei // Institute of Electronics, Chinese Academy of Sciences. – Cornell University, 2017. – URL: <https://arxiv.org/abs/1711.09405> (21.12.2020).
- [2] Zhao, X. 3D Object Detection Using Scale Invariant and Feature Reweighting Networks / X. Zhao, Z. Liu, R. Hu, K. Huang. – Cornell University, 2019. – URL: <https://arxiv.org/abs/1901.02237> (21.12.2020).

- [3] Энджел, Э. Интерактивная компьютерная графика. Вводный курс на базе OpenGL. –М.: «Вильямс», 2001. – С. 592.
- [4] Bataineh M. Neural network for regression problems with reduced training sets / M. Bataineh, T. Marler // Neural networks: the official journal of the International Neural Network Society. – 2017. – URL: https://www.researchgate.net/publication/319023995_Neural_network_for_regression_problems_with_reduced_training_sets (21.12.2020).
- [5] OpenCV modules [Электронный ресурс]. – Режим доступа: <https://docs.opencv.org/4.3.0/> (21.12.2020).