

ПРИМЕНЕНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ЗАДАЧ КЛАССИФИКАЦИИ ДОРОЖНЫХ ЗНАКОВ

В.Н. Сичкар, С.А. Колюбин

Научный руководитель – д.т.н., профессор С.А. Колюбин

Университет ИТМО

Системы распознавания и классификации дорожных знаков предоставляют дополнительный уровень безопасности как для водителей, так и для других участников дорожного движения. Такая система может использоваться для помощи водителям, предупреждая о наличии дорожных знаков, чтобы снизить риски в ситуациях отвлечения внимания, усталости, меняющейся освещенности, сложных погодных условий. Несмотря на то, что уже существуют и описываются различные методы распознавания дорожных знаков, разработка надежного алгоритма все еще остается актуальной задачей. Видимость дорожного знака в реальной среде усложняет процесс его классификации по нескольким причинам. Поскольку изображения получены с подвижных датчиков, установленных на транспортном средстве, и которые подвержены вибрации и тряске, неизбежно размытие сцен. Кроме того, различная освещенность также добавляет сложность в классификации, а именно яркое солнце, или блики от фар в ночное время. Задача усложняется наличием объектов, похожих по цвету и форме, таких как здания и другие транспортные средства. Поэтому система должна быть способна иметь дело с видимостью дорожных знаков в широком диапазоне погодных условий, в условиях различной освещенности – туман, дождь, облака и солнце.

В данной работе проводятся исследования надежности работы Сверточной Нейронной сети для классификации дорожных знаков. Цель работы – поиск эффективного метода предобработки набора данных для обучения Сверточной Нейронной сети, который обеспечивает точность валидации более 95%.

Для достижения поставленной цели был выбран набор данных для обучения, а именно German Traffic Sign Benchmarks, содержащий 43 класса знаков с различным числом примеров в каждом из них. Перед тем, как были применены различные методы предобработки набора данных, количество примеров изображений в каждом классе было выравнено путем добавления изображений со случайным образом измененной яркостью и поворотом изображения в двухмерной плоскости относительно центра влево или вправо в пределах 10 градусов. Таким образом, количество примеров в целом во всем наборе данных увеличилось с 40 000 до 86 000 изображений. В качестве архитектуры Сверточной Нейронной сети была выбрана для использования следующая модель:

- один Сверточный слой, состоящий из 32 фильтров размером 7 на 7 пикселей, с шагом 1 и фреймом для обработки границ равным 1;
- один слой уменьшения размерности с фильтром 2 на 2, с порогом обработки путем выбора максимального значения и шагом 2;
- один скрытый Полносвязный слой с количеством нейронов равным 500;
- выходной слой с количеством нейронов равным 43, в соответствие с количеством классов для обучения.

Для решения задачи предварительной обработки обучающего набора данных и определения наиболее эффективного метода, были выбраны следующие подходы: нормализация путем деления на 255.0 (сокращенно /255.0 Normalization), нормализация путем вычитания среднего изображения (сокращенно Mean Image), нормализация путем поиска и деления на стандартное отклонение (сокращенно STD Image), улучшение качества изображения путем локального гистограммного выравнивания (Local Histogram Equalization).

Таким образом, были сформированы и подготовлены следующие наборы данных для обучения Сверточной Нейронной сети:

1. RGB, /255 Normalization, Mean Image
2. RGB, /255 Normalization, Mean Image, STD Image
3. GrayScale, Local Histogram Equalization, /255 Normalization, Mean Image
4. GrayScale, Local Histogram Equalization, /255 Normalization, Mean and STD Image

Обучение Сверточной Нейронной сети проводилось одновременно (путем применения параллельных вычислений) с набором из 50 изображений, выбранных случайным образом, на 10 эпох, и, соответственно на 18 000 итераций в целом.

Основными результатами являются показатели точности при проверке уже обученной Сверточной Нейронной сети разными наборами данных, а именно разными методами предварительной обработки наборов данных. Проверка осуществлялась с помощью валидационного набора данных, который был выделен из обучающего набора данных и не использовался в процессе обучения. Результаты представлены ниже:

1. Точность достигла 88.1%
2. Точность достигла 88.9%
3. Точность достигла 95.3%
4. Точность достигла 92.4%

Таким образом, наилучшим методом для предварительной обработки обучающего набора данных оказался метод, где используется следующая последовательность операций: конвертация в GrayScale, применение Local Histogram Equalization для улучшения качества изображения, нормализация путем деления на 255 и нормализация путем вычитания среднего изображения.