

ПРИМЕНЕНИЕ РАЗЛИЧНЫХ ТИПОВ ИЗОБРАЖЕНИЙ ДЛЯ ОБУЧЕНИЯ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ

Д.А. Булыгин, Т.Е. Мамонова
Томский политехнический университет
E-mail: 128dmitriy128@gmail.com

Введение

Выбор изображений для обучения свёрточной нейронной сети является очень важной задачей. На основе выбранных изображений будет определяться скорость обучения, а также возможность распознавания образов на изображении.

Целью данной работы является выбор типа изображений для обучения свёрточной нейронной сети. В работе будут проанализированы следующие типы изображений: бинарное, полутоновое и цветное. В работе будут проанализированы все сильные и слабые стороны каждого типа изображений. Рассматриваемые цветные изображения имеют формат RGB.

Постановка задачи

Первостепенной задачей является получение изображений, пригодных для обработки, на которых будет возможно выделение признаков для классификации изображений. В основе алгоритма распознавания лежит свёрточная нейронная сеть, обученная по принципу “обучение с учителем”, поэтому выбор наиболее подходящего типа изображений очень важен для обучения и дальнейшей работы СНС.

Одним из основных параметров цифрового изображения является его цветность, согласно которой изображения можно классифицировать:

- Бинарные (только два возможных состояний пикселя);
- Полутоновые (изображение в тонах (оттенках) определенного цвета);
- Цветные (многоцветные).

Бинарные изображения

Бинарное изображение (двухуровневое, двоичное) — разновидность цифровых растровых изображений, когда каждый пиксел может представлять только один из двух цветов [1].

Значения каждого пикселя условно кодируются, как «0» и «1». Значение «0» условно называют задним планом или фоном, а «1» — передним планом.

Бинарное изображение можно получить при помощи отсечки по яркости. То есть, алгоритмически задаётся значение яркости пикселя, выше которых значению пикселя присваивается единица (пиксель становится белого цвета), а если яркость пикселя ниже этого порогового значения, то присваивается ноль (пиксель становится чёрного цвета).

Обычно пороговая граница устанавливается в зависимости от задачи. Например, если стоит

задача выделить объекты красного цвета, то на бинарном изображении все объекты красного цвета будут закрашены в белый цвет. Всё остальное будет закрашено в чёрный цвет.

Преимуществами этих изображений является низкий объём информации, который необходим для их хранения, простота обработки компьютером, следовательно, быстродействие алгоритма по обработке и классификации изображений будет очень высокой.

Из недостатков этих изображений можно выделить необходимость установки порогового значения и малое количество признаков, по которым свёрточная нейронная сеть сможет классифицировать изображение.

При работе с бинарными изображениями появляется проблема отсечки по яркости. Так как при изменении освещённости меняется цвет объектов, часть информации теряется. Например, руку видно не полностью. Также если фон изображения будет примерно совпадать с цветом объекта, который необходимо выделить, то в итоге получится почти полностью белое изображение.

Пример бинарного изображения приведён на рисунке 1.



Рис. 1. Пример бинарного изображения

Полутоновые изображения

Полутоновое изображение — это изображение, имеющее множество значений тона, и их непрерывное, плавное изменение.

Множество возможных полутонов называют уровнями серого (англ. gray scale), независимо от того, полутона какого цвета или его оттенка передаются. Таким образом, уровни серого не отличаются по спектральному составу (оттенку цвета), но отличаются по яркости. Количество возможных полутонов в данном случае есть глубина цвета, которую часто передают не в количестве самих полутонов, а в количестве бит на пиксел. Какое из значений в допустимом диапазоне будет считаться самым ярким, а какое самым тёмным не имеет значения, т. к. число, являющееся значением каждого пикселя — всего лишь

условный код яркости. Достаточно указать направление отсчёта [2].

Например, могут существовать полутоновые растры, где на каждый пиксел отведено 8 бит, изображение имеет 256 полутонов, а пикселы со значением 0 или 255 являются черными, и наоборот, пикселы со значением 255 или 0 — белыми, остальные полутона серого будут равномерно распределены между данными значениями цветового индекса.

Преимуществом такого типа изображений является большое количество информации, получаемое с изображения. На полутоновых изображениях можно рассмотреть не только примерную форму объекта, но и его конкретные очертания и прочие более мелкие детали.

Но данный тип изображений не лишён недостатков. Несмотря на увеличившийся объём информации, по сравнению с бинарными изображениями, этой информации не всегда хватает для успешной классификации изображений свёрточной нейронной сетью. Например, эти изображения подвержены влиянию света, в следствие чего, объекты на изображении начинают сливаться. Данное явление приводит к нежелательному результату: нейронная сеть неверно распознаёт изображение и неверно его классифицирует.

Пример полутонового изображения приведён на рисунке 2.



Рис. 2. Пример полутонового изображения

Цветные изображения формата RGB

RGB (аббревиатура английских слов red, green, blue — красный, зелёный, синий) или КЗС — аддитивная цветовая модель, как правило, описывающая способ кодирования цвета для цветовоспроизведения с помощью трёх цветов, которые принято называть основными. Выбор основных цветов обусловлен особенностями физиологии восприятия цвета сетчаткой человеческого глаза.

RGB-модель является аддитивной, где цвета получаются путём добавления (англ. addition) к чёрному цвету. При отсутствии краски нет никакого цвета — чёрный, максимальное смешение даёт белый [3].

Преимуществом цветных изображений является возможность выделять свёрточной нейронной сетью признаки не только по форме, но и по цвету, что существенно увеличивает вероятность успешной классификации изображений.

Из недостатков можно отметить то, что эти изображения содержат большой объём информации, из-за чего обработка таких изображений более затруднительна, в сравнении с полутоновыми и бинарными изображениями.

Пример цветного изображения приведён на рисунке 3.



Рис. 3. Пример цветного изображения

Выбор типа изображений для обучения свёрточной нейронной сети

Для выбора наиболее подходящего типа изображений была исследована точность нейронной сети при следующих параметрах обучения:

- Количество эпох обучения равно трём;
- Обучающая выборка состоит из 5000 изображений, разделённых на два класса (по 1000 изображений каждого типа и 2000 изображений с отсутствием объекта, который необходимо обнаружить);
- Функция активации в каждом слое ReLu;
- Количество фильтров в каждом слое равно 10, каждый из которых имеет размерность 3x3.

В результате, точность сети, обученной на бинарных изображениях, равна 97%, на монотонных — 93%, и 95% на цветных.

Заключение

Из полученных результатов, можно сделать вывод, что свёрточная нейронная сеть лучше обнаруживает и классифицирует изображения в следствие выделения признаков как по форме объекта, так и по цвету. Бинарные изображения имеют более высокую точность, но очень сильно подвержены влиянию света и нуждаются в постоянной настройке порогового значения.

Список использованных источников

1. Бинарное изображение [Электронный ресурс] — URL: <http://ru.knowledgr.com/00066610/%D0%91%D0%B8%D0%BD%D0%B0%D1%80%D0%BD%D0%BE%D0%B5%D0%98%D0%B7%D0%BE%D0%B1%D1%80%D0%B0%D0%B6%D0%B5%D0%BD%D0%B8%D0%B5> (Дата обращения: 02.10.2019).
2. Полутоновые изображения [Электронный ресурс]. — URL: <https://mash-xxl.info/info/193831/> (Дата обращения: 30.09.2019).
3. Цвет и его модели [Электронный ресурс] — URL: <https://compuart.ru/article/23772> (Дата обращения: 27.09.2019).