

РАСПОЗНАВАНИЕ СТИЛЯ ПРОИЗВЕДЕНИЙ ЖИВОПИСИ ПО ИХ ИЗОБРАЖЕНИЯМ С ПОМОЩЬЮ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ

Ковалев Валентин Юрьевич

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: kovalev.vyu@google.com

Научный руководитель — Шишкин Алексей Геннадиевич

В изобразительном искусстве существует огромное разнообразие стилей, начиная с ренессанса и заканчивая стрит-артом. Во многих случаях они не имеют четко определенных границ, что в ряде случаев затрудняет их идентификацию даже экспертами. Поэтому задача автоматического определения стиля картин является весьма актуальной в настоящее время. В целом ее можно рассматривать как составную часть другой задачи, заключающейся в определении авторства картин по их цифровым изображениям.

Целью данной работы является построение модели и разработка на ее основе программной системы для определения стиля произведения живописи по его цифровому изображению, а также изучение различных факторов, влияющих на эффективность решения поставленной задачи.

В качестве входного набора данных был использован большой набор картин [1] с частичной разметкой. Большая часть данных собрана с сайта wikiart.org. Всего в наборе представлено 103250 картин, которые относятся к 136 стилям. Все данные были предварительно обработаны: картины, которые не были полностью размечены, исключались из дальнейшего анализа (их количество – 27392). Для оставшихся изображений были изменены размеры (меньший размер равен 224 пикселям). При изучении распределения картин по стилям выяснено, что их большая часть относится к 5 стилям, которые и были выбраны в качестве целевых классов. Окончательное число изображений, пригодных для анализа, составило 43242, каждое из которых относилось к одному из 5 стилей: импрессионизм, реализм, романтизм, экспрессионизм, постимпрессионизм. Вследствие ограниченности имеющихся вычислительных мощностей было оставлено только 25% имеющихся данных – 10810 изображений с сохранением распределения картин по стилям.

Для создания классификатора использовалась заранее обученная для классификации на 1000 классов на наборе данных ImageNet

[2] архитектура сверточной сети ResNet [3], последний слой которой был изменен на полносвязный слой с числом выходов, равному числу стилей – в данном случае – 5. Веса данной нейросети были изменены при обучении в соответствии с имеющимися данными. После завершения процесса обучения, были проведены эксперименты по ускорению и уменьшения размера графа вычислений, используя алгоритмы обрезания весов (pruning) нейронной сети.

Помимо этого, были проведены эксперименты с использованием дополнительных данных о картине, из которых видно, что их использование позволяет повысить эффективность классификации. На основе нейросетевой модели была разработана программная система, позволяющая для входного изображения произведения живописи быстро и с хорошей точностью определять его стиль. Также, в качестве готового решения представлен Telegram бот, который способен провести классификацию входного изображения по пяти стилям используя предложенный алгоритм.

В результате данной работы был реализован алгоритм, который решает поставленную задачу с точностью 52,03% (табл. 1) и обладает рядом существенных отличий от уже имеющихся решений. Так, для определения стиля произведения живописи были использованы методы глубокого обучения. Кроме того, большой объем входных данных был собран из открытых источников, в то время как в аналогичных исследованиях обычно использовались проприетарные базы данных, с относительно небольшим числом изображений.

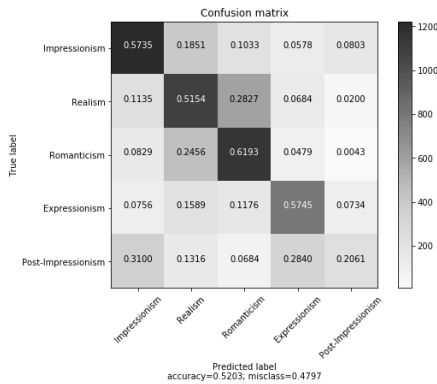


Таблица 1. Матрица ошибок реализованного алгоритма.

Литература

1. Painter by numbers kaggle comptetion
www.kaggle.com/c/painter-by-numbers/
2. ImageNet database
<http://www.image-net.org>
3. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Deep Residual Learning for Image Recognition // arXiv:1512.03385 [cs.CV], 2015