

УДК 681.518.3

Н.В. Смирнов, А.С. Трифонов

Петрозаводский государственный университет, Петрозаводск

ВЛИЯНИЕ ПРИСУТСТВИЯ НЕСКОЛЬКИХ КЛАССОВ ОБЪЕКТОВ НА РЕЗУЛЬТАТЫ КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ СВЕРТОЧНЫМИ НЕЙРОННЫМИ СЕТЯМИ

Человеческий фактор обуславливает риск возникновения ошибки распознавания изображения. Автоматизация процесса классификации фотографий объектов, приходящих на перерабатывающий завод, уменьшает риск возникновения ошибки классификации и соответствующие финансовые затраты. Создан собственный набор изображений, объекты на которых аналогичны оригинальным. На изображениях с объектами целевого класса также может присутствовать некоторое количество объектов нецелевого класса. В работе приведены результаты классификации фотографий, содержащих два класса объектов одновременно.

Ключевые слова: классификация изображений; сверточные нейронные сети; глубокое обучение.

N.V. Smirnov, A.S. Trifonov

Petrozavodsk State University, Petrozavodsk

INFLUENCE OF SEVERAL CLASSES OBJECTS PRESENCE ON THE RESULTS OF IMAGE CLASSIFICATION BY CONVOLUTIONAL NEURAL NETWORKS

The human factor causes the risk of image recognition error. Automating the process of classifying photographs of materials coming to a processing plant reduces the risk of errors and the corresponding financial costs. The authors created their own dataset of images, the objects on which are similar to the original. Images with objects of the target class may also contain a number of objects of the non-target class. The paper presents the results of classification of photographs containing two classes of objects simultaneously.

Keywords: image classification; convolutional neural networks; deep learning.

Введение. Задачи классификации изображений и распознавания объектов часто появляются в различных сферах человеческой деятельности. Рассмотрим перерабатывающий завод, на который в железнодорожных вагонах поступают заранее отсортированные по категориям объекты (металлолом). Человек может допускать ошибки при распознавании образов, поэтому возможным решением задачи классификации фотографий вагонов является применением методов машинного и глубокого обучения. При решении задачи классификации часто возникает проблема отсутствия необходимого количества изображений для обучения нейронных сетей. Также довольно часто в вагон с объектами одного типа попадает некоторое количество объектов другого типа, что затрудняет задачу классификации изображения.

Чтобы решения проблемы недостаточного количества исходных изображений в тренировочной выборке была предпринята попытка создать дополнительные изображения. Изображения, полученные с использованием различных электронных средств [1], не содержат искажений изображения, которые получаются при фотографировании, поэтому в подавляющем большинстве не обеспечивают приемлемую точность классификации на реальных фотографиях. Авторами был создан набор фотографий, имитирующих объекты, которые в большой степени совпадают с объектами в настоящем наборе фотографий по форме, цвету, размерам.

Набор изображений. Созданные фотографии (рис. 1) можно рассматривать как текстурные изображения. Фотографии были нарезаны на необходимый для применения рассмотренных сверточных нейронных сетей размер, и для повышения степени схожести с настоящими изображениями, было применено размытие путем свертки с матрицей из единиц размером 5x5. Полученный набор изображений [2], содержащий 6 классов был разбит на: тренировочную, валидационную, и тестовую, соответствующие размеры которых составляют 840, 105, 105 изображений каждого класса объектов.



Рисунок 1 – Пример исходной фотографии, класс обрезки

Для проведения классификации фотографий вагонов по преобладающему (целевому) классу объектов был сформирован набор смешанных изображений. Этот набор создавался путем наложения фрагментов фотографий нецелевого класса объектов на фотографии с объектами целевого класса, так что объекты целевого и нецелевого классов находятся в соотношении: 90/10, 80/20, 70/30, 60/40 процентов соответственно.



Рисунок 2 – Типы объектов

Результаты классификации

Общепринятыми метриками в задаче классификации изображений являются:

— точность (*precision*) – отношение количества объектов целевого класса, которые классифицированы как объекты этого класса, к общему количеству объектов, которые классифицированы как объекты этого класса.

— *recall* (*полнота*) – отношение количества объектов, которые классифицированы как объекты целевого класса, к общему количеству объектов этого класса,

— среднее гармоническое *f1*:

$$f1 = \frac{2 * (precision * recall)}{precision + recall}$$

В качестве метрики выбора наилучшего классификатора выбрано среднее значение *f1*:

$$avg_f1 = \sum_i^n f1_i,$$

$$avg_f1 \rightarrow \max,$$

Где *n* – количество классов в тестовом наборе, *f1_i* – *f1* *i*-го класса, *i* = 1, 2, ..., *n*.

Сверточные нейронные сети Xception, ResNet50, InceptionV3 [3] на тестовой выборке сформированного набора данных показали значения точности, полноты и метрики *f1* равные 0,99 [2]. Эти нейронные сети были выбраны для классификации смешанного набора изображений. Результаты классификации представлены в таблице 1 и для сети Xception на тепловых картах (рис. 3-6). У каждой нейронной сети при увеличении процента присутствия объектов нецелевого класса на изображениях с объектами целевого класса происходит уменьшение % верно распознанных изображений.

Таблица 1 – Результаты классификации изображений

Название нейронной сети	Процент верно классифицированных изображений целевого класса, при заданных пропорциях площади фотографии, занимаемой объектами целевого класса/нецелевого класса				
	100%/0%	90%/10%	80%/20%	70%/30%	60%/40%
ResNet50	99,00%	99,37%	98,41%	93,17%	87,94%
InceptionV3	99,00%	99,68%	97,78%	92,54%	72,54%
Xception	99,00%	99,04%	98,25%	96,50%	88,57%

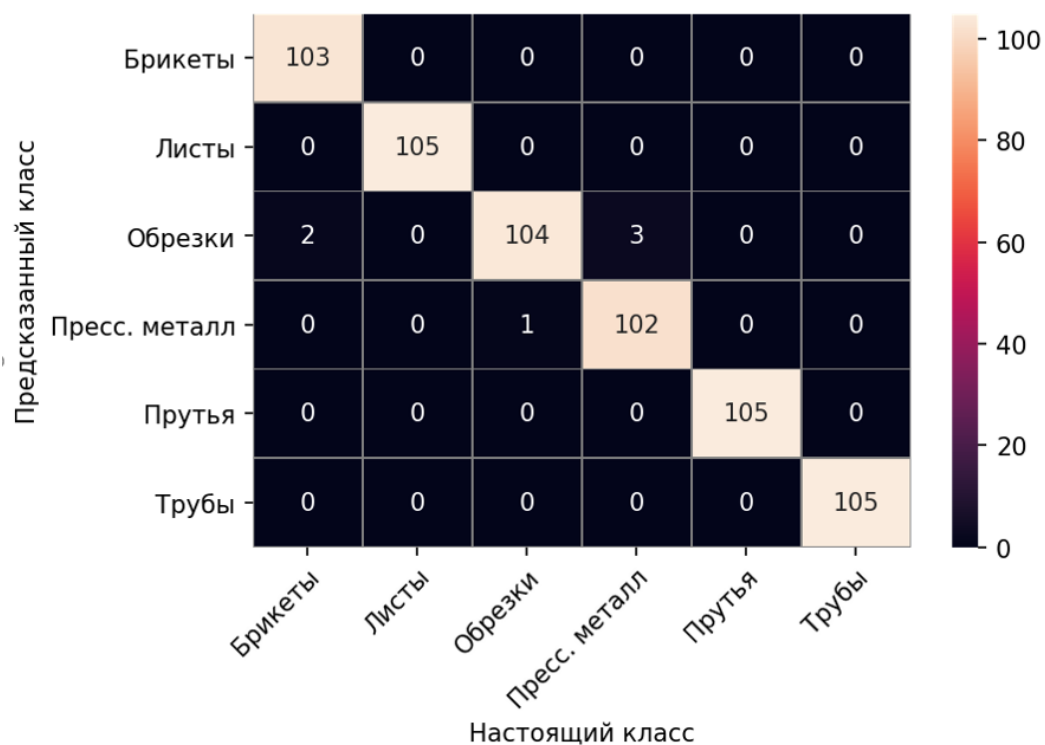


Рисунок 3 – Тепловая карта смешивания классов 90%/10%

Заключение. Нейронные сети показали довольно высокую точность классификации на созданном тестовом наборе изображений, содержащих два класса объектов: целевого и нецелевого. Наибольшее количество верно классифицированных изображений были получены при применении нейронных сетей ResNet50, Xception.

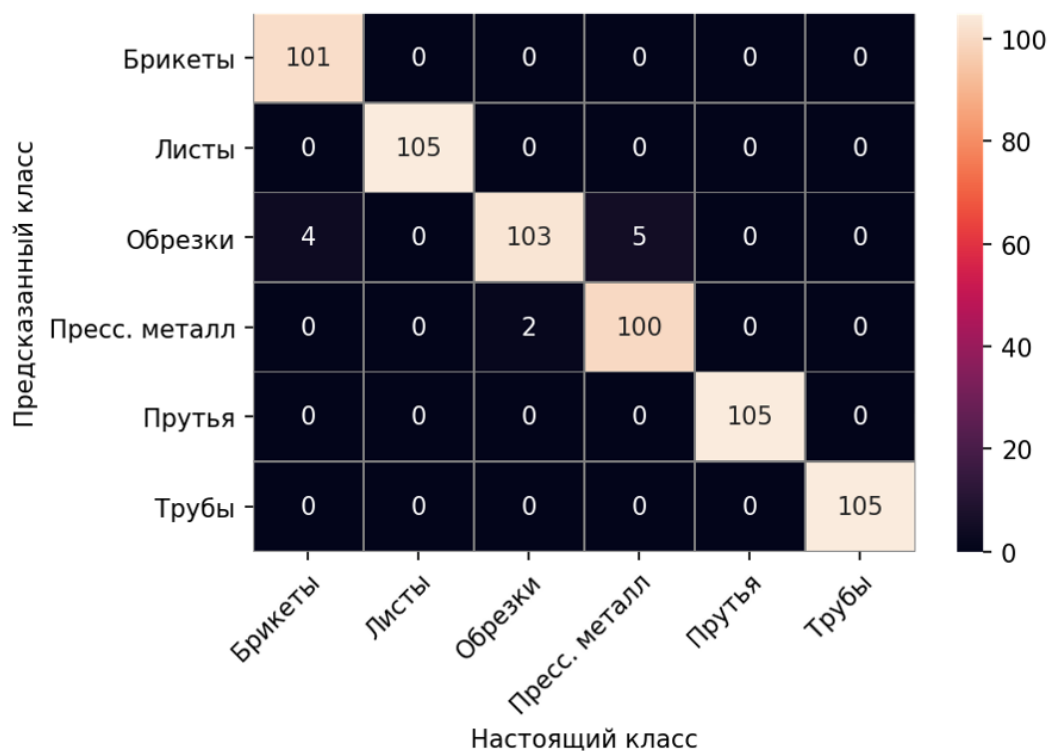


Рисунок 4 – Тепловая карта смешивания классов 80%/20%

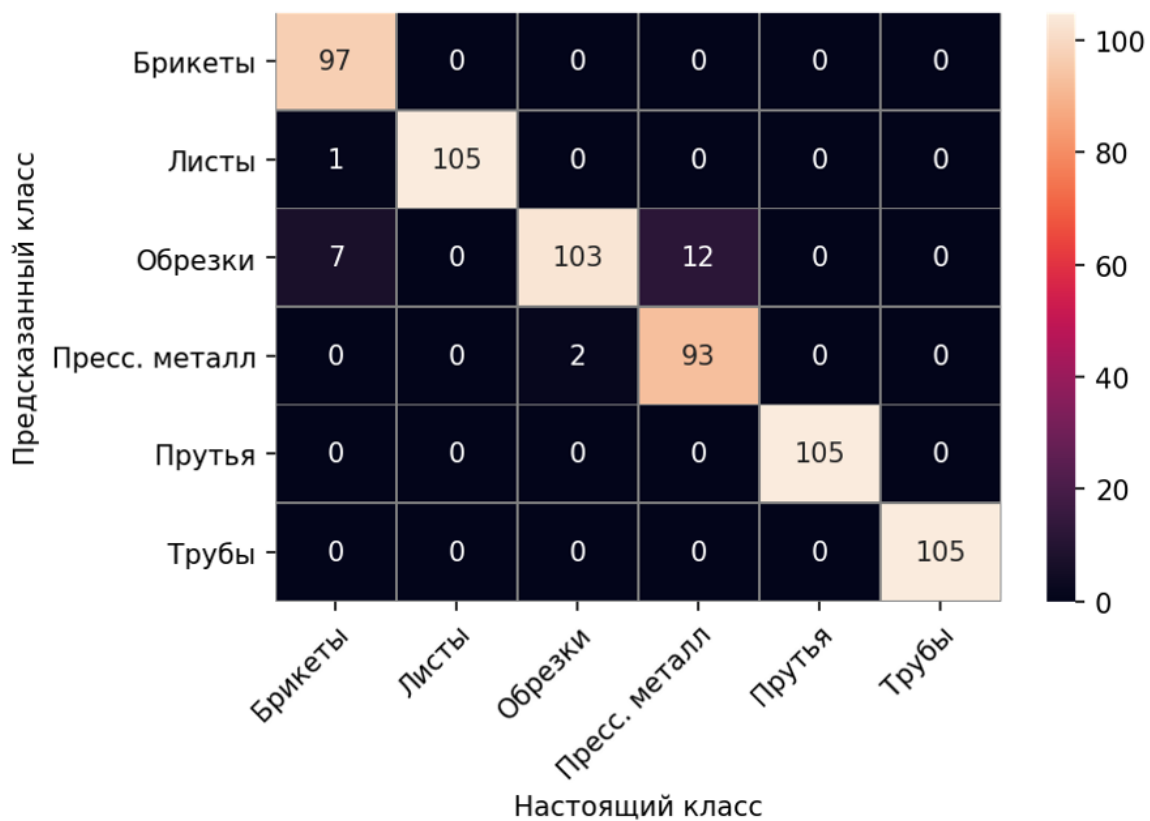


Рисунок 5 – Тепловая карта смешивания классов 70%/30%

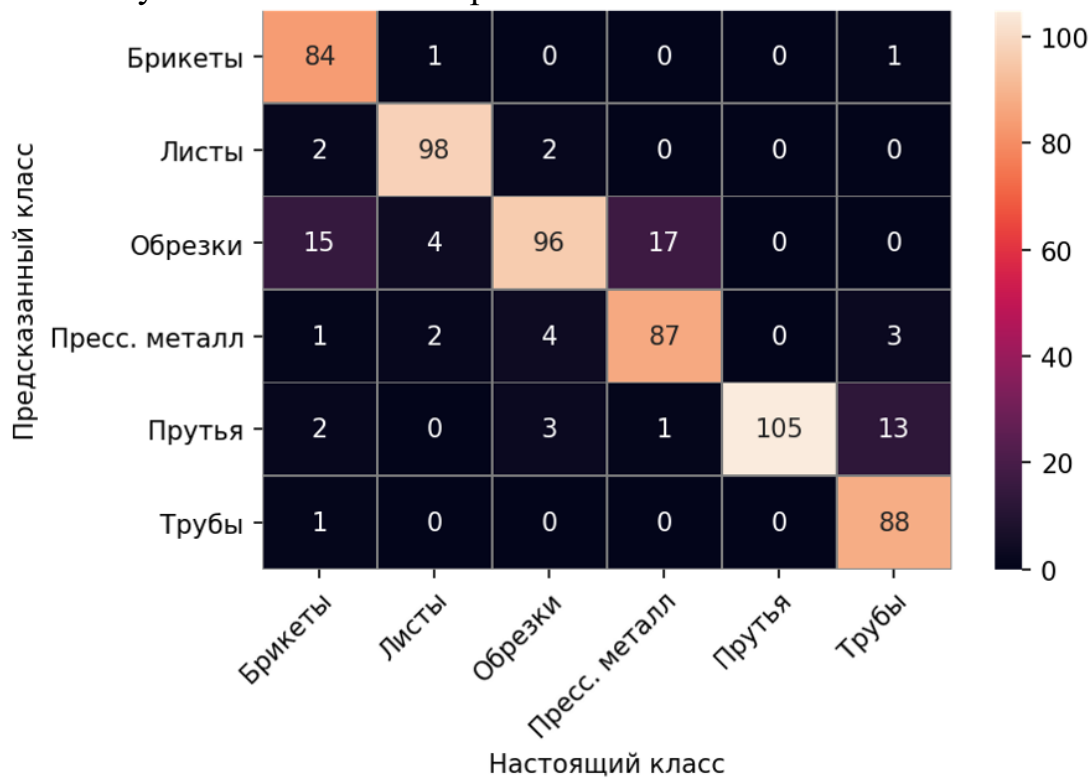


Рисунок 6 – Тепловая карта смешивания классов 60%/40%

Список литературы

1. Смирнов Н. В., Рыбин Е. И. Исследование влияния смешивания различных классов объектов на результаты классификации изображений // Шаг в будущее: искусственный интеллект и цифровая экономика. Революция в

управлении: новая цифровая экономика или новый мир машин: материалы II Международного научного форума. – М.: Издательский дом ФГБОУ ВО ГУУ, 2018. – С. 111-116.

2. Смирнов Н. В., Трифонов А. С. Применение сверточных нейронных сетей в задаче классификации текстурных изображений. [Электронный ресурс] // Цифровые технологии в образовании, науке, обществе: материалы XIV всероссийской научно-практической конференции. – Петрозаводск, 2020. – С. 146-149. URL: <https://it2020.petsu.ru/doc/it2020.pdf>. – (06.12.2020).

3. Keras Applications [Электронный ресурс]. URL: <https://keras.io/api/applications/>

Сведения об авторах

Смирнов Николай Васильевич – кандидат технических наук, доцент кафедры «Теория вероятностей и анализа данных» Петрозаводского государственного университета, г. Петрозаводск, e-mail: smirnovn@petsu.ru

Трифонов Алексей Сергеевич – студент Петрозаводского государственного университета, гр. 22307, г. Петрозаводск, e-mail: lexa.2tri@yandex.ru

About the authors

Smirnov Nikolai Vasilievich – Candidate of Technical Sciences, Associate Professor of the Department of Probability and Data Analysis, Petrozavodsk State University, Petrozavodsk, e-mail: smirnovn@petsu.ru

Trifonov Alexey Sergeevich – student of Petrozavodsk State University, gr. 22307, Petrozavodsk, e-mail: lexa.2tri@yandex.ru