

# Классификация текстурных изображений с помощью методов глубокого обучения

М.А. Головастикова<sup>1</sup>, А.В. Гайдель<sup>1,2</sup>

<sup>1</sup>Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

<sup>2</sup>Институт систем обработки изображений РАН - филиал ФНИЦ «Кристаллография и фотоника» РАН, Молодогвардейская 151, Самара, Россия, 443001

**Аннотация.** В работе исследуется зависимость достоверности классификации текстурных изображений от архитектуры нейронной сети и исходных данных. Предложенная архитектура сети, представляющая собой каскад сверточных и полносвязных слоев, принимает на вход текстурные изображения. С ее помощью была достигнута достоверность классификации в 92%. Также в работе рассматривается аналогичный вариант классификации изображений, основанный на использовании полносвязной нейронной сети. Используя сеть, принимающую на вход вектор признаков изображения, удалось добиться существенной экономии времени при незначительной потере достоверности классификации.

## 1. Введение

Текстурный анализ широко применяется в обработке изображений, например, в промышленности для контроля поверхностей, в биомедицине для обнаружения патологий тканей [1], для классификации и сегментации наземных фотографий спутников, а также для поиска изображений по указанному содержанию.

Текстурному изображению, несмотря на то что люди его легко могут распознать, довольно сложно дать определение. В литературе было дано много описаний для текстуры, но до сих пор нет общепринятого определения [2]. Были проведены многочисленные исследования по классификации текстурных изображений, но эта задача все еще остается трудной. Фундаментальные проблемы заключаются в том, чтобы извлечь избыточные признаки для описания текстуры и использовать правильный тип метрики при сравнении векторов признаков.

Классический подход классификации текстурных изображений состоит в переходе от самого изображения текстуры к описывающему его вектору признаков при помощи некоторого набора фильтров, после чего следует провести сглаживание вектора. Было проведено огромное количество исследований по поиску наиболее оптимального набора признаков для составления векторов для классификации текстурных изображений [3 - 5].

Двумя более ранними широко использовавшимися подходами к классификации текстур являются матрица совместного вхождения [6] и Марковские случайные поля [7]. Однако такие подходы к классификации не всегда показывают удовлетворительные результаты для всех наборов текстур [8]. Позже для классификации и сегментации текстур были разработаны улучшенные методы – фильтры Габора и вейвлет-фильтры [9 - 11]. Одним из недостатков

данных подходов является то, что сгенерированные фильтры обычно распознают широкий спектр текстурных изображений, что требует дополнительной оптимизации или осуществления ручного отбора фильтров для каждого набора текстур. Более того, некоторые методы оптимизации подходят только для бинарной классификации или могут быть вычислительно сложными [12].

В статье [13] было предложено объединить выделение признаков текстурных изображений и классификацию в единую структуру, внедрив принципы многоканальной схемы фильтрации в архитектуру нейронной сети. Первая разработанная авторами сеть – трехслойная, включая входной уровень, с прямой связью, то есть каждый входной узел полностью связан с небольшой областью размера  $M \times M$  во входном изображении. Вторая сеть использовала механизм распределения весов скрытых нейронов для связи с предыдущим слоем. Последняя сеть имеет около 5000 весов и была обучена с использованием онлайн алгоритма обратного распространения [14].

В данной статье мы создаем сверточную нейронную сеть для классификации текстурных изображений. Сеть основана на тех же принципах построения, что и VGGNet [15], но с более простой структурой, что приводит к сокращению времени обучения и классификации.

Для сравнения времени и результатов классификации в следующем разделе мы протестируем работу нейронной сети, на вход которой подаются не сами изображения, а векторы признаков, составленные на основе этих же текстур, которые были получены в нашей предыдущей работе [16].

## 2. Архитектура сверточной нейронной сети

Сверточная нейронная сеть, которую мы использовали для классификации текстурных изображений, представляет собой каскад сверточных и полносвязных слоев. На вход нейронная сеть принимает квадратные изображения текстур, которые предварительно были уменьшены до размера  $144 \times 144$  пикселя. На выходе мы получаем вектор размерности 28 – по числу классов изображений в наборе данных. Архитектура нейронной сети представлена на рисунке 1.

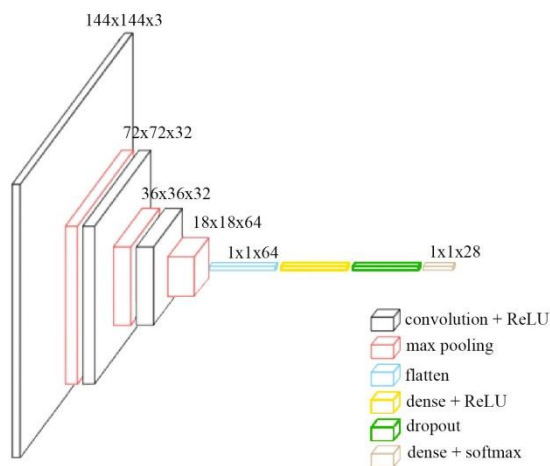


Рисунок 1. Архитектура нейронной сети.

Сверточные слои (Conv2D) нацелены на эффективное распознавание образов. Они предназначены для выделения характерных признаков изображений. Сверточные слои чередуются подвыборочными слоями (MaxPooling2D) для уменьшения размерности предыдущего слоя. Так как на этапе предыдущей свертки уже были выявлены некоторые признаки, то для дальнейшей обработки настолько подробное изображение не нужно, поэтому оно уплотняется до менее подробного. Более точные характеристики нейросети представлены в таблице 1.

Первый сверточный слой нейронной сети имеет размерность входного изображения, но на выходе имеет размерность  $144 \times 144 \times 32$ , где последнее число отражает количество фильтров. В данной архитектуре мы используем полулинейную функцию активации ReLU. Затем слой

MaxPooling2D уменьшает размер – ширину и высоту до значений  $72 \times 72 \times 32$ . Следующие два каскада повторяют этот принцип расположения слоев, последовательно уменьшая размерность до  $18 \times 18 \times 64$ .

**Таблица 1.** Детали структуры нейронной сети.

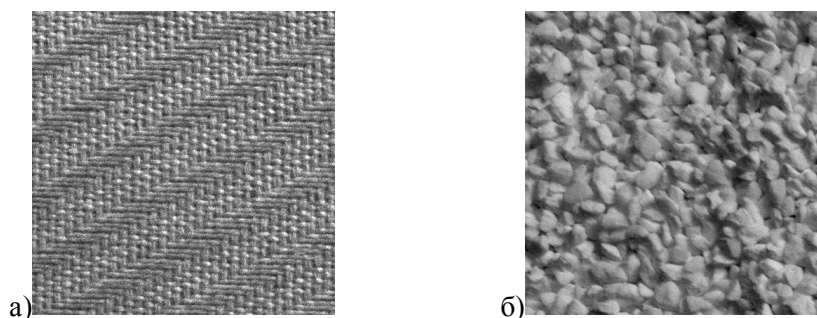
№	Входная размерность			Выходная размерность			Слой	Ядро	
1	144	144	3	144	144	32	Conv2D	3	3
2	144	144	32	72	72	32	MaxPooling2D	2	2
3	72	72	32	72	72	32	Conv2D	3	3
4	72	72	32	36	36	32	MaxPooling2D	2	2
5	36	36	32	36	36	64	Conv2D	3	3
6	36	36	64	18	18	64	MaxPooling2D	2	2
7	18	18	64	1	1	64	Flatten		
8	1	1	64	1	1	64	Dense	1	1
9	1	1	64	1	1	64	Dropout		
10	1	1	64	1	1	28	Dense	1	1

Для того чтобы перейти к полносвязной части нейронной сети, нужно перейти от двумерной размерности данных к одномерной. Для этого используется слой Flatten, после прохождения которого размерность становится  $1 \times 1 \times 64$ .

После сверточной части идет полносвязанная часть нейронной сети, которая отвечает за классификацию текстур. Она состоит из двух полносвязанных слоев. Первый включает в себя 64 нейрона, используется также полулинейная функция активации. Затем следует слой Dropout для уменьшения переобучения. Для выходного полносвязного слоя нейронной сети используется функция активации SoftMax – мягкого максимума. На выходе нейронной сети мы получаем вероятность принадлежности к определенному классу для каждой метки.

### 3. Обучение и оценка нейронной сети

В качестве базы текстурных изображений был использован Kylberg Texture Dataset v. 1.0 [17]. Он содержит 28 классов текстур по 160 фотографий размером  $576 \times 576$  пикселей с математическим ожиданием равным 127 и дисперсией 40, примеры которых приведены на рисунке 2.



**Рисунок 2.** Примеры текстурных изображений: (а) – одеяло, (б) – гравия.

Для правильной классификации изображений необходимо провести некоторые подготовительные шаги. Интенсивность пикселей из диапазона  $[0, 255]$  была переведена в непрерывный вещественный диапазон  $[0, 1]$ . Метки, представляющие собой целочисленные значения от 1 до 28, характеризующие к каким классам принадлежат изображения, должны быть приведены к виду вектора размерности 28, где значение единицы указывает

принадлежность к классу, а 0 дополняет вектор до нужной размерности. Выполняется это с помощью подхода One-Hot Encoding или прямого кодирования.

Обучение сети происходит за 20 эпох, что занимает 3425,62 секунды. С помощью использования сети такой архитектуры удалось добиться достоверности классификации в 92%. Вероятности принадлежности к определенному классу для каждой метки представлены на рисунке 3.

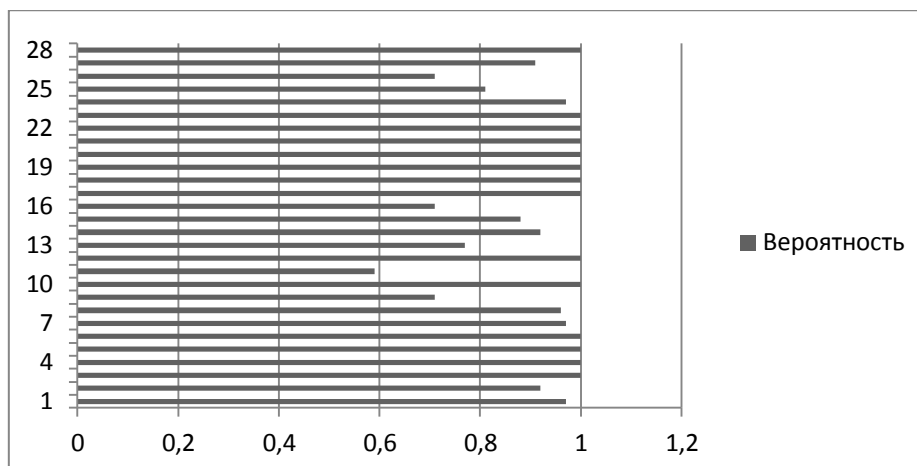


Рисунок 3. Вероятности классификации.

#### 4. Сравнение результатов классификации

В статье [14] нами была проведена работа по составлению вектора признаков из корреляционных признаков текстурных изображений и коэффициентов из уравнений регрессии.

Для сравнения результатов времени и достоверности классификации с предыдущими, было решено сделать тестовую полносвязную нейронную сеть, на вход которой подаются полученные ранее вектора признаков, а не сами изображения текстур, как в предыдущей нейронной сети.

Количество слоев варьировалось от трех до пяти. Также подбирались количества нейронов в слоях от 30 до 360 с постоянным шагом. Для каждой архитектуры оценивались результат работы сети и время, затраченное на классификацию, в зависимости от числа слоев. Полученные результаты представлены в таблице 3.

Таблица 2. Результаты классификации тестовой нейронной сети.

Количество нейронов в слое				Достоверность классификации, %	Время классификации, с
30	360	360	28	84	70,67
30	270	360	270	28	82,99
120		90	28	79	38,25
30	30	30	28	79	35,07
90	18	180	28	79	50,6
30	180	360	180	28	69,25

Вероятности принадлежности к определенному классу для каждой метки для архитектуры сети 30-360-360-28, показавшей наилучший результат классификации, представлены на рисунке 4. Оказалось, что использование сети намного более простой архитектуры – из 4 полносвязных слоев с количеством нейронов 30-360-360-28 обеспечивает адекватный результат классификации текстур с результатом в 84% верно распознанных изображений.

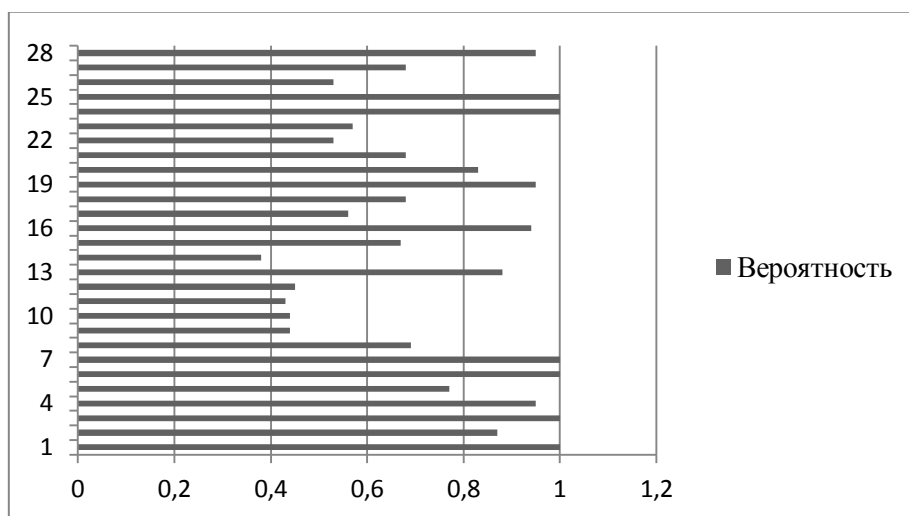


Рисунок 4. Вероятности классификации тестовой нейронной сети.

## 5. Заключение

В результате при классификации текстурных изображений с помощью сверточной нейронной сети, описанной выше архитектуры, максимальная достигнутая достоверность классификации составила 92%. Время классификации – 3425,62 секунды.

При классификации с помощью полносвязной нейронной сети, на вход которой подавались наборы векторов признаков текстурных изображений, была достигнута достоверность классификации в 84%. Сеть состояла из 4 слоев по 30-360-360-28 нейронов в слое соответственно. Время классификации составило 70,67 секунд.

Использование нейронной сети указанной простой архитектуры, с подаваемыми на вход векторами признаков, составленными по текстурным изображениям, позволяет экономить время в более чем 48 раз.

## 6. Благодарности

Работа выполнена при поддержке грантов РФФИ № 19-29-01235 мк и № 19-29-01135 мк, экспериментальные исследования – в рамках госзадания ИСОИ РАН – филиала ФНИЦ «Кристаллография и Фотоника» РАН (соглашение № 007-ГЗ/Ч3363/26).

## 7. Литература

- [1] Гайдель, А.В. Возможности текстурного анализа компьютерных томограмм в диагностике хронической обструктивной болезни / А.В. Гайдель, П.М. Зельтер, А.В. Капишников, А.Г. Храмов // Компьютерная оптика. – 2014. – Т. 38, № 4. – С. 843-850.
- [2] Tuceryan, M. Texture analysis. The Handbook of Pattern Recognition and Computer Vision / A.K. Jain, C. Chen, L.F. Pau, P. Wang – Singapore: World Scientific, 1998. – 207-248.
- [3] Rahmaninia, M. OSFSMI: Online Stream Feature Selection Method based on Mutual Information / M. Rahmaninia, P. Moradi // Applied Soft Computing Journal. – 2018. – Vol. 68. – P. 733-746. DOI: 10.1016/j.asoc.2017.08.034.
- [4] Ma, B. A tribe competition-based genetic algorithm for feature selection in pattern classification / B. Ма, Y. Xia // Applied Soft Computing Journal. – 2017. – Vol. 58. – P. 328-338.
- [5] Кутикова, В.В. Исследование методов отбора информативных признаков для задачи распознавания текстурных изображений с помощью масок Лавса / В.В. Кутикова, А.В. Гайдель // Компьютерная оптика. – 2015. – Т. 39, № 5. – С. 744-750. DOI: 10.18287/0134-2452-2015-39-5-744-750.
- [6] Haralick, R.M. Statistical and Structural Approaches to Texture // Proceedings IEEE. – 1979. – Vol. 67. – P. 786-804.
- [7] Cross, G.R. Markov Random Field Texture Models / G.R. Cross, A.K. Jain // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1983. – Vol. 5(1). – P. 25-39.

- [8] Liu, X. Texture Classification Using Spectral Histograms / X. Liu, D. Wang // IEEE Transactions on Image Processing. – 2003. – Vol. 12(6). – P. 661-670.
- [9] Jain, A. K. Unsupervised Texture Segmentation Using Gabor Filters / F. Farrokhnia // Pattern Recognition. – 1991. – Vol. 24(12). – P. 1167-1186.
- [10] Randen, T. Multichannel Filtering For Image Texture Segmentation / J.H. Husøy // Optical Engineering. – 1994. – Vol. 33. – P. 2617-2625.
- [11] Ng, B.W. Supervised Texture Segmentation Using DWT and a Modified K-NN Classifier / A. Bouzerdoum // Proceedings of the 15th International Conference on Pattern Recognition. – 2000. – Vol. 2. – P. 545-548.
- [12] Randen, T. Filtering for Texture Classification: A Comparative Study / J.H. Husøy // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1999. – Vol. 21(4).
- [13] Jain, A.K. Learning Texture Discrimination Masks / K. Karu // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1996. – Vol. 18(2). – P. 195-205.
- [14] Tivive, F.H.C. Texture Classification Using Convolutional Neural Networks / A. Bouzerdoum // TENCON IEEE Region 10 Conference. – 2006. – P. 1-4.
- [15] Muhammad, U. Pre-trained VGGNet Architecture for Remote-Sensing Image Scene Classification / W. Wang, Sh.P.C. Chattha, S. Ali // 24th International Conference on Pattern Recognition (ICPR) Beijing, China, 2018.
- [16] Туркова, М.А. Корреляционные признаки для классификации текстурных изображений / М.А. Туркова, А.В. Гайдель // Сборник трудов IV международной конференции и молодежной школы «Информационные технологии и нанотехнологии» – Самара: Новая техника, 2018. – С. 595-599.
- [17] Kylberg, G. The Kylberg Texture Dataset v. 1.0 // Centre for Image Analysis, Swedish University of Agricultural Sciences and Uppsala University, External report (Blue series). – 35. – URL: <http://www.cb.uu.se/~gustaf/texture/>.

## Texture images classification using deep learning techniques

M.A. Golovastikova<sup>1</sup>, A.V. Gaidel<sup>1,2</sup>

<sup>1</sup>Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

<sup>2</sup>Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardejskaya street 151, Samara, Russia, 443001

**Abstract.** In this work we investigate the dependence of the reliability of the classification of texture images on the architecture of the neural network and the source data. The proposed network architecture, which is a cascade of convolutional and fully connected layers, receives texture images as input. The classification accuracy of 92% was achieved. Also, a similar version of image classification based on the use of a fully connected neural network is considered. Using a network that receives an input vector of image features, it was possible to achieve significant time savings with a slight loss of classification reliability.