

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ, РЕЧИ, ТЕКСТА И РАСПОЗНАВАНИЕ ОБРАЗОВ

SIGNAL, IMAGE, SPEECH, TEXT PROCESSING AND PATTERN RECOGNITION



УДК 004.89
<https://doi.org/10.37661/1816-0301-2023-20-3-37-49>

Оригинальная статья
Original Paper

Распознавание состава ткани предметов одежды на изображении в сфере электронной коммерции с использованием нейронных сетей

В. В. Сорокина

*Белорусский государственный университет,
пр. Независимости, 4, Минск, 220050, Беларусь
E-mail: viktoria.sorokina.96@gmail.com*

Аннотация

Цели. Поставлена цель разработать новый подход к распознаванию состава ткани предметов одежды на изображении в сфере электронной коммерции путем использования генеративно-сопоставительной сети для создания синтетических изображений предметов одежды с известным составом ткани, используемых затем для обучения сверточной нейронной сети при классификации состава ткани реальных изображений одежды. Вместо классического изображения предмета одежды генерируется копия, у которой материал увеличен до волокон и структуры ткани.

Методы. Основными методами распознавания состава ткани предметов одежды на изображении в сфере электронной коммерции являются создание и аннотация наборов данных для обучения нейронных сетей, синтез изображений ткани предметов одежды, выбор архитектуры и ее модификация, валидация и проведение тестов, а также интерпретация результатов.

Результаты. Результаты экспериментов, проведенных с помощью предложенного подхода, показывают его эффективность при точном определении состава ткани предметов одежды в сфере электронной коммерции, что позволяет использовать данный метод для улучшения поиска и просмотра на веб-сайтах.

Заключение. При помощи генеративно-сопоставительной сети был синтезирован набор данных товаров электронной коммерции, произведена его аннотация, построены нейронные сети для распознавания состава ткани предметов одежды, проведено сравнение результатов. Результаты исследования показали, что новый подход для распознавания ткани предметов одежды обладает высокой точностью в сравнении с уже известными методами. Дополнительное использование модели внимания также дает хорошие результаты, что отражается в улучшении метрик.

Ключевые слова: классификация состава ткани, генеративно-сопоставительная сеть, сверточная нейронная сеть, электронная коммерция, синтез изображений, модель внимания

Для цитирования. Сорокина, В. В. Распознавание состава ткани предметов одежды на изображении в сфере электронной коммерции с использованием нейронных сетей / В. В. Сорокина // Информатика. – 2023. – Т. 20, № 3. – С. 37–49. <https://doi.org/10.37661/1816-0301-2023-20-3-37-49>

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Поступила в редакцию | Received 27.02.2023

Подписана в печать | Accepted 06.04.2023

Опубликована | Published 29.09.2023

Recognition of fabric composition of clothing in an image in e-commerce using neural networks

Viktoriya V. Sorokina

Belarusian State University,
av. Nezavisimosti, 4, Minsk, 220050, Belarus
E-mail: viktoriya.sorokina.96@gmail.com

Abstract

Objectives. Development of new approach for recognizing the fabric composition of clothing in e-commerce images by using generative adversarial network (GAN) to generate synthetic images of clothing with known fabric composition, to be used to train the CNN to classify the fabric composition of real clothing images. Instead of a classic clothing image, a copy is generated with the material zoomed to fibers and fabric structure.

Methods. The main methods to recognize the fabric composition of the clothing image in the e-commerce are the creation and annotation of a dataset for the neural network training, synthesis of the fabric of clothing, the choice of architecture and its modification, validation and testing, and interpretation of the results.

Results. Experimental results with the constructed method show that it is effective for accurately recognizing the fabric composition of e-commerce clothing to be used to improve search and browsing on websites.

Conclusion. In the course of the experiment, using a generative adversarial network, a data set of e-commerce products was synthesized and annotated, neural networks were built to recognize the composition of the fabric of clothing items. The results of the study showed that the new approach for recognizing the fabric of clothing provides higher accuracy in comparison with already known methods, in addition, the use of the attention model also gives good results to improve the metrics.

Keywords: classification of fabric composition, generative adversarial network, convolutional neural network, e-commerce, image synthesis, attention model

For citation. Sorokina V. V. *Recognition of fabric composition of clothing in an image in e-commerce using neural networks*. Informatika [Informatics], 2023, vol. 20, no. 3, pp. 37–49 (In Russ.). <https://doi.org/10.37661/1816-0301-2023-20-3-37-49>

Conflict of interest. The author declares of no conflict of interest.

Введение. В последние годы можно наблюдать значительный рост популярности интернет-магазинов, что влечет за собой и быстрое развитие индустрии одежды в сфере электронной коммерции. Между тем одной из основных проблем при покупках одежды в онлайн-магазинах является отсутствие возможности потрогать одежду, что затрудняет для покупателей определение состава ее ткани. Состав ткани – важный фактор для многих потребителей при совершении покупок, так как он влияет на долговечность, комфорт и уход за одеждой.

Очень часто онлайн-площадки, где множество продавцов представляют свой товар, не имеют унифицированного стандарта описания товара, поэтому состав ткани предмета одежды может отсутствовать. Процесс подготовки полноценного описания является трудозатратной задачей, включающей в себя наличие физического предмета одежды и соответствующего специалиста для определения свойств ткани и ее классификации. Настоящая работа направлена на автоматизацию этого процесса без привлечения человеческих ресурсов для детального представления товара.

Классификация изображений – одна из важных задач компьютерного зрения, которая включает в себя присвоение метки или класса изображению на основе его визуального содержимого. Целью данной задачи является точное определение объектов или сцен, присутствующих на изображении. Процесс классификации изображений обычно включает следующие этапы:

- сбор и аннотирование набора данных изображений, где каждое изображение помечено соответствующим классом или меткой;
- обучение модели, например сверточной нейронной сети (Convolutional Neural Network, CNN), на аннотированном наборе данных;
- использование обученной модели для классификации новых изображений путем присвоения им метки на основе прогнозов модели.

Решение задачи классификации тканей предполагает присвоение метки класса образцу ткани, но из-за специфики изображений выделить такой образец ткани, чтобы классификация была точной, почти невозможно. При этом определение состава ткани по входным изображениям тканевых поверхностей является более простой задачей. Так, например, трикотажные и вязаные ткани легко отличить друг от друга. Петлевая структура трикотажных тканей и переплетение основных и уточных нитей на тканых материалах позволяют легко идентифицировать две структуры.

Цель исследования – повысить точность классификации тканей предметов одежды на изображении, что является важной задачей в индустрии моды и текстиля. Традиционно классификация тканей проводилась путем изучения характеристик изображения предметов одежды, т. е. самой структуры ткани. Однако этот метод затруднителен, если изображения низкого качества или сняты с большого расстояния. В таких случаях существующие подходы могут быть недостаточно точными для классификации тканей только на основе характеристик изображения.

Предлагаемое в статье решение состоит в том, чтобы при классификации тканей сместить акцент с характеристик изображения на структуру ткани, вместо изучения особенностей изображения предмета одежды сосредоточиться на выявлении внутренней структуры самой ткани. С помощью данного подхода можно достичь более высокого уровня точности в классификации тканей даже в тех случаях, когда элементы изображения трудно различимы.

Подход включает использование такого передового метода анализа изображений, как CNN, для выявления и анализа структурных особенностей тканей на изображениях. Этот метод может помочь системе распознать уникальные узоры и текстуры, характерные для различных типов тканей, и использовать данную информацию для их точной классификации. Важно, имея изображение предмета одежды на снимке низкого качества или снятого с большого расстояния, распознать состав ткани путем генерации образца ткани, отображающего его структуру.

В работе предлагается новый подход для определения состава ткани предметов одежды в сфере электронной коммерции, основанный на комбинации генеративно-состязательной сети (Generative Adversarial Network, GAN) и CNN, основная идея которого заключается в замене исходного материала предмета одежды его детализированной версией для более точной классификации и упрощения модели.

Основной вклад автора в решение поставленной задачи заключается в следующем:

- разработка нового подхода к определению состава ткани при помощи условных генеративно-состязательных сетей;
- представление самой задачи классификации как простой структуры, достаточной для получения результатов высокой точности (модификация нейросети VGG-19 с помощью модели внимания [1]);
- ускорение обучения GAN за счет использования архитектуры U-Net [2] в генераторе.

1. Анализ существующих подходов. Классификация тканей важна в таких отраслях, как текстильная промышленность, индустрия моды и электронная коммерция, где необходимо классифицировать различные типы тканей на основе их визуальных характеристик. Она может включать определение типа ткани (например, хлопок, полиэстер, шелк и т. д.), рисунка переплетения (саржа, атлас и т. д.), текстуры (гладкая, грубая и т. д.) или других параметров по ви-

зуальным свойствам. Задача классификации может быть решена с помощью методов на основе анализа изображений, где в качестве входных данных используются изображения тканей, а классификатор обучается распознавать различные типы тканей по их визуальным характеристикам. Кроме того, задача классификации может быть с учителем [3, 4], когда модель обучается на размеченном наборе данных изображений ткани, или без учителя [5], когда обучение осуществляется на неразмеченном наборе данных и необходимо изучать особенности тканей.

В публикации [3] для определения тканей авторы использовали как текстурные признаки, так и цветовые признаки HSV (Hue, Saturation, Value). В работе [4] предложен метод, сочетающий цветовые моменты и так называемые признаки GIST для классификации 700 изображений тканей. В статье [5] представлен алгоритм сопоставления изображений ткани путем создания набора визуальных слоев с использованием алгоритма k -средних. Общая проблема указанных методов заключается в том, что они разрабатывались и тестировались на наборе данных небольшого размера и это повлияло на их точность. В работе [6] авторам удалось добиться высокой точности за счет использования обучающего набора данных, включающего около 10 000 образцов ткани, и CNN с функцией потерь *Center Loss*.

В перечисленных выше работах применяются схожие подходы для решения задач классификации и идентификации тканей. Однако все они базируются на таком определении тканей, когда в качестве входного изображения используются образцы ткани, т. е. изображение, отображающее детализированные структурные детали.

В настоящей статье для генерации более детализированного изображения материала ткани используется сеть $pix2pix$, которая относится к классу условных генеративно-состязательных сетей cGAN, а для классификации полученных образцов ткани – CNN VGG-19, модифицированная моделью внимания. Основная идея работы заключается в том, чтобы при классификации тканей перейти от определения материала предмета одежды (например, хлопка) на изображениях, где структура ткани не видна, к определению типа ткани из ее структуры, а затем решить задачу классификации при помощи CNN. Поставлена цель построить алгоритм, способный по классическому изображению предмета одежды в каталоге товаров электронной коммерции определить его состав.

2. Предлагаемый метод. Алгоритм определения состава ткани предметов одежды изображен на рис. 1, схема последовательности блоков алгоритма – на рис. 2.



Рис. 1. Алгоритм определения состава ткани

Fig. 1. Algorithm of fabric composition recognition



Рис. 2. Схема последовательности блоков алгоритма

Fig. 2. Scheme of algorithm's blocks sequence

В работе используются три нейронные сети. Хотя можно объединить модели pix2pix и VGG-19 в единую архитектуру, есть несколько причин, по которым в этом исследовании они применяются по отдельности.

Во-первых, сеть pix2pix предназначена для преобразования изображения в изображение, в частности для создания высококачественных изображений из входных данных. Объединение задачи, выполняемой сетью pix2pix, с задачей классификации ткани может быть неоптимальным, поскольку при этом не используется вся мощь архитектуры модели.

Модель VGG-19 представляет собой хорошо зарекомендовавшую себя архитектуру для решения задач классификации изображений и успешно используется во многих приложениях. В рамках данного исследования модификация ее с помощью модели внимания показала, что модель VGG-19 может лучше выделять важные функции и отфильтровывать шум. Это приводит к более точной классификации.

Во-вторых, использование отдельных моделей обеспечивает бóльшую гибкость и модульность алгоритма. В дальнейшем в случае построения более эффективной модели для создания изображений или решения задачи классификации ее можно легко заменить, не затрагивая другую часть алгоритма.

Наконец, разделение моделей также упрощает интерпретацию и отладку алгоритма. Если есть проблемы с генерацией или классификацией изображений, проще изолировать источник проблемы и устранить его, не внося изменений в другую часть алгоритма.

Рассмотрим каждый этап алгоритма:

Сегментация. Для повышения точности представленного в статье метода можно предварительно сегментировать изображение, чтобы отделить передний план (т. е. сам объект) от фона. Для этого использовалась созданная и описанная автором в работе [7] модель на основе CNN YOLACT и стандартизации весов.

Генерация карты текстуры. Задача генерации карты текстуры заключается в создании более детализированного изображения материала ткани на основе классического изображения товара. Входное и выходное изображения алгоритма генерации карты текстуры показаны на рис. 3.



Рис. 3. Примеры изображений модуля генерации карты текстуры: *a)* входное; *b)* выходное
Fig. 3. Examples of image of texture map generation module: a) input; b) output

Для решения задачи использовалась генеративно-состязательная сеть, для которой был необходим соответствующий набор данных для обучения. Построение набора данных для обучения и самой модели описано в разд. 3.

Выделение образца ткани. Задача выделения образца ткани представляет собой генерацию изображения размером 50×50 пикселей, которое отображает текстуру ткани. Для решения этой задачи следует определить доминантный цвет материала, на основе которого выделяется соответствующий фрагмент из исходного изображения предмета одежды в сфере электронной коммерции. Входное и выходное изображения алгоритма выделения образца ткани показаны на рис. 4.



Рис. 4. Примеры изображений модуля выделения образца ткани: *a)* входное; *b)* выходное
Fig. 4. Examples of image of tissue sampling module: a) input; b) output

Классификация образца ткани. Задача классификации образца ткани направлена на распознавание одного из 10 типов тканей предметов одежды (хлопок, шерсть, эластан и др.) в сфере электронной коммерции. На вход алгоритма поступает изображение образца ткани, на выход –

соответствующая метка материала. Для решения задачи был подготовлен набор данных для обучения, а также построена CNN. Вход и выход алгоритма представлены на рис. 5.



Рис. 5. Пример входа и выхода алгоритма классификации образца
 Fig. 5. Example of input and output of sample classification

3. Обучающее множество: синтез и основные характеристики. Для построения методов генерации карты текстуры и классификации образца ткани необходимы наборы данных для обучения. Набор данных для задачи генерации карты текстуры представляет собой оригинальное изображение в паре с изображением с более детализированным материалом предмета одежды. Набор данных для классификации образца – изображение образца ткани и соответствующая ему метка класса.

3.1. Синтез изображений для обучающего множества. Поскольку задачу определения состава ткани с помощью комбинации GAN и CNN до этого не решали, необходимо было создать набор данных для обучения, который включал бы исходное изображение предмета одежды и изображение с детализированным материалом ткани. Процесс ручного создания такого набора является трудозатратным, поэтому было принято решение синтезировать его автоматически. Для этого использовалась база изображений образцов тканей с открытым исходным кодом, которая создавалась специально для решения задачи классификации образцов в исследовании [8], а также набор данных предметов одежды, разбитый на четыре класса, вместе с соответствующими бинарными масками, где 255 – объект, 0 – фон. Представленные классы:

- 1) майки (рубашки);
- 2) платья;
- 3) брюки;
- 4) юбки.

Примеры образца ткани и предмета одежды показаны соответственно на рис. 6 и 7.

Образцы тканей также были разделены на классы в соответствии с типом одежды и материалом (например, юбка-синтетика-85-хлопок-15).

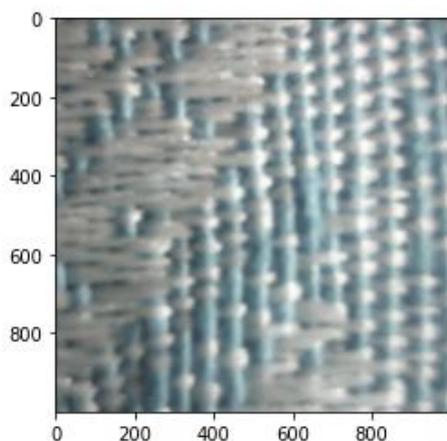


Рис. 6. Образец ткани
 Fig. 6. Sample of fabric

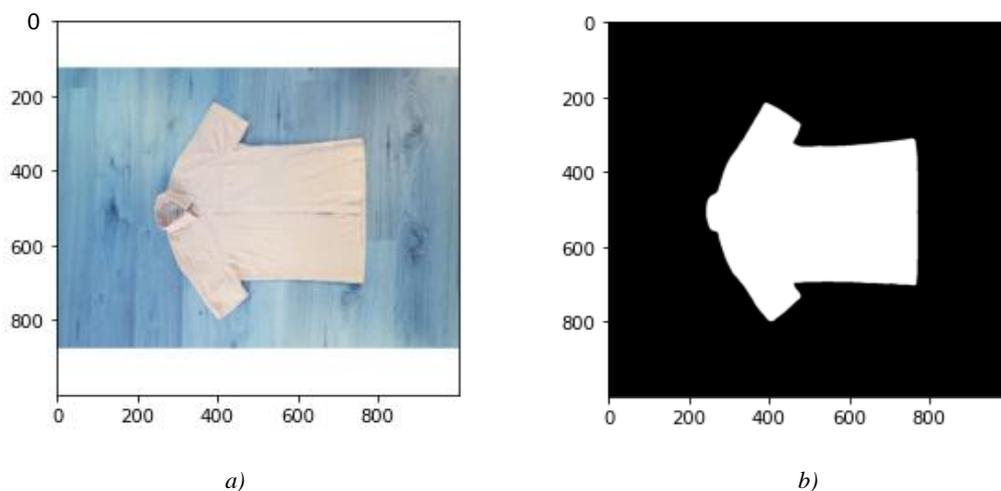


Рис. 7. Предмет одежды (a) и его бинарная маска (b)

Fig. 7. Clothing item (a) and its binary mask (b)

Чтобы автоматически синтезировать набор данных, для каждого предмета одежды выбирался соответствующий классу данного предмета материал, определялся доминантный цвет предмета одежды и материал окрашивался в этот цвет.

3.2. Алгоритм определения доминантного цвета.

Вход: сегментированное RGB-изображение.

Используется метод k -means – алгоритм обучения без учителя для кластеризации данных в K кластеров. Идея метода заключается в минимизации суммарного квадратичного отклонения точек кластеров от центров этих кластеров.

Алгоритм k -means для определения доминирующего цвета включает в себя следующие шаги:

1. Преобразование изображения в двумерный массив, где каждая строка представляет пиксел изображения.
2. Инициализация K центроидов кластеров случайным образом.
3. Итерация по всем пикселям изображения и назначение каждому пикселу его ближайшего центроида на основе евклидова расстояния между пикселом и центроидом.
4. Пересчет позиции K центроидов через среднее значение всех пикселей, назначенных каждому из этих кластеров.
5. Повтор шагов 3 и 4 до тех пор, пока центроиды не перестанут меняться или не будет достигнуто максимальное количество итераций.

6. Сопоставление каждого пиксела изображения с его ближайшим центроидом.

Конечным результатом является определение K доминирующих цветов в изображении, представленных K центроидами кластеров.

Выход: доминантный цвет товара в формате RGB.

При помощи синтеза изображений был создан набор данных, включающий 3938 RGB-изображений размером 800×800 пикселей, из которых в обучающей выборке было 2355 изображений, в валидационной – 1299 и в тестовой – 284.

4. Генерация карты текстуры. Для решения задачи необходимо было выбрать нейросеть, способную генерировать высококачественные снимки, неотличимые от реальных. Была выбрана архитектура pix2pix [9].

Сеть pix2pix, также известная как Image-to-Image Translation with Conditional Adversarial Networks, представляет собой тип условной GAN, который используется для задач преобразования изображения в изображение. Она состоит из двух основных компонентов: генератора и дискриминатора. Генератор принимает входное изображение и генерирует выходное, в то время как дискриминатор принимает как входное изображение, так и сгенерированное выходное изображение и решает, является выходное изображение реальным или поддельным. Гене-

ратор и дискриминатор обучаются путем состязания: генератор пытается генерировать изображения, которые могут обмануть дискриминатор, а дискриминатор пытается правильно идентифицировать настоящие и поддельные изображения. Математически это можно выразить следующим образом: пусть дано входное изображение x , выходное y и вектор шума z . Тогда отображение $G: \{x, z\} \rightarrow y$ задает генератор G , который будет оцениваться дискриминатором D .

Основными преимуществами $pix2pix$ являются:

1. Высококачественные изображения: сеть $pix2pix$ известна созданием высококачественных изображений, неотличимых от реальных.
2. Универсальность: сети $pix2pix$ можно использовать для широкого спектра задач преобразования изображения в изображение, таких как перенос стиля, преобразование объектов и дорисовка изображения.
3. Условная генерация: сеть $pix2pix$ можно использовать для условной генерации изображений, когда выходное изображение создается на основе определенного условия или ввода.
4. Обработка структурированных данных: сеть $pix2pix$ может обрабатывать структурированные данные, такие как карты и рисунки.
5. Обработка отсутствующих данных: сеть $pix2pix$ может обрабатывать отсутствующие данные и дорисовывать изображение.

В настоящей работе для улучшения работы генератора вместо классического подхода encoder-decoder, в котором входные данные проходят через серию слоев, постепенно понижающих дискретизацию, до слоя узкого места (bottleneck), после чего процесс меняется на обратный (такая сеть требует, чтобы весь информационный поток проходил через все слои, включая bottleneck), использовалась архитектура U-Net (рис. 8). В частности, добавлялись пропущенные соединения между каждым слоем i и слоем $n - i$, где n – общее количество слоев. Каждое соединение с пропуском просто объединяет все каналы на уровне i с каналами на уровне $n - i$.

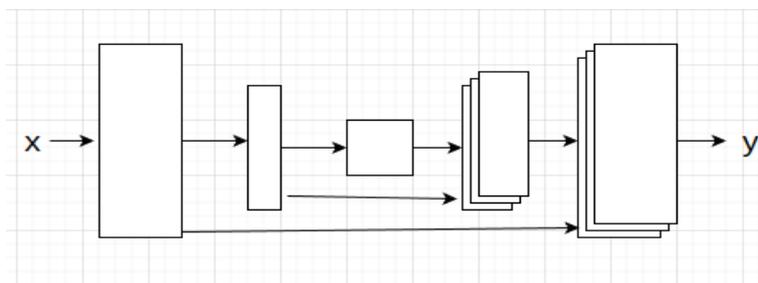


Рис. 8. Архитектура генератора
 Fig. 8. Generator architecture

Обучение сети $pix2pix$ с сетью U-Net в качестве генератора проводилось с использованием фреймворка PyTorch с видеокартой GPU NVIDIA T4, размер пакета ($batch_size$) – 2, количество эпох – 50, архитектура генератора – U-Net 256.

5. Выделение и классификация образца ткани. Следующий шаг после генерации карты текстур – это выделение образца ткани и его дальнейшая классификация. Выделение образца ткани осуществлялось при помощи алгоритма определения доминантного цвета с переходом в цветовое пространство LAB путем задания размера блоков, содержащих доминантный цвет. Выход алгоритма – изображение 50×50 пикселей, содержащее структуру ткани.

Для классификации образца ткани была построена модель на основе VGG-19 [10] и модели внимания [11] по аналогии с работой [12].

Общая структура сети VGG-19 с моделью внимания показана на рис. 9. Для предотвращения переобучения и улучшения способности модели к обобщению была использована предобученная сеть VGG-19 с удаленными полносвязными слоями. Блоки внимания добавлены соответственно после третьего, четвертого и пятого слоев Pooling сети VGG-19. Наконец, выходные данные трех блоков внимания объединяются вместе, чтобы сформировать окончательный вектор признаков.

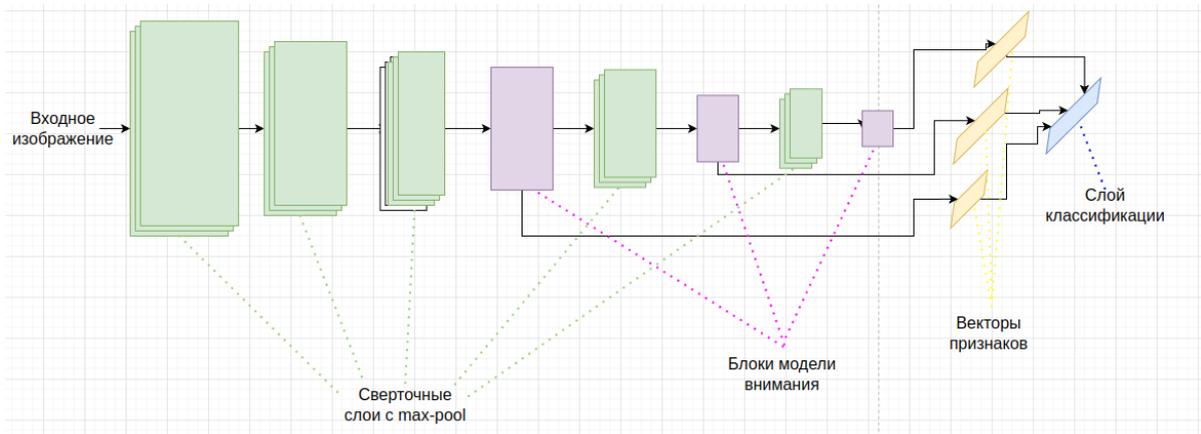


Рис. 9. Архитектура модели классификации

Fig. 9. Classification model architecture

Входом блока внимания является промежуточный признак $F(C \times H \times W)$ сети VGG-19, где F представляет собой промежуточный признак (feature map) размерности $C \times H \times W$, полученный из слоя сети VGG-19 (C обозначает количество каналов (или признаков), H – высоту, а W – ширину этого признака.). На первом шаге пространственный модуль в блоке внимания генерирует пространственную карту внимания $M_S(1 \times H \times W)$. Далее M_S умножается на промежуточный признак F , чтобы получить пространственно уточненный признак F' . Затем в блоке внимания генерируется карта внимания канала $M_C(C \times 1 \times 1)$ путем умножения M_C на F' , чтобы получить окончательный уточненный признак F'' . Таким образом, процесс расчета в блоке внимания можно представить следующим образом:

$$F' = M_S(F) \cdot F, F'' = M_C(F) \cdot F',$$

где знак \cdot обозначает поэлементное умножение.

Модель обучалась в течение 100 эпох с использованием фреймворка PyTorch с видеокартой GPU NVIDIA T4.

6. Результаты работы алгоритмов. Для классификации обучались три сети: классическая VGG-19, VGG-19 с моделью внимания и VGG-19 с моделью внимания и сетью pix2pix. Обучение было направлено на распознавание 10 типов тканей предметов одежды в сфере электронной коммерции. Для обучения применялась видеокарта GPU NVIDIA T4. В качестве метрик оценки моделей использовались Precision, Recall и F1-score.

Примеры генерации карт текстуры и дальнейшей их классификации приведены на рис. 10. Точность классификации представлена в таблице.

Средняя точность классификации обученных сетей
Average classification result of the trained neural networks

Сеть Network	Метрика Metrics		
	Precision	Recall	F1-score
VGG-19	0,68	0,75	0,71
VGG-19 с моделью внимания	0,73	0,84	0,78
VGG-19 с моделью внимания и сетью Pix2pix	0,97	0,95	0,95

Результатом применения модели внимания стало увеличение скорости обучения нейронной сети. Для получения равнозначных со стандартной сетью VGG-19 результатов понадобилось в 2,5 раза меньше итераций при обучении.

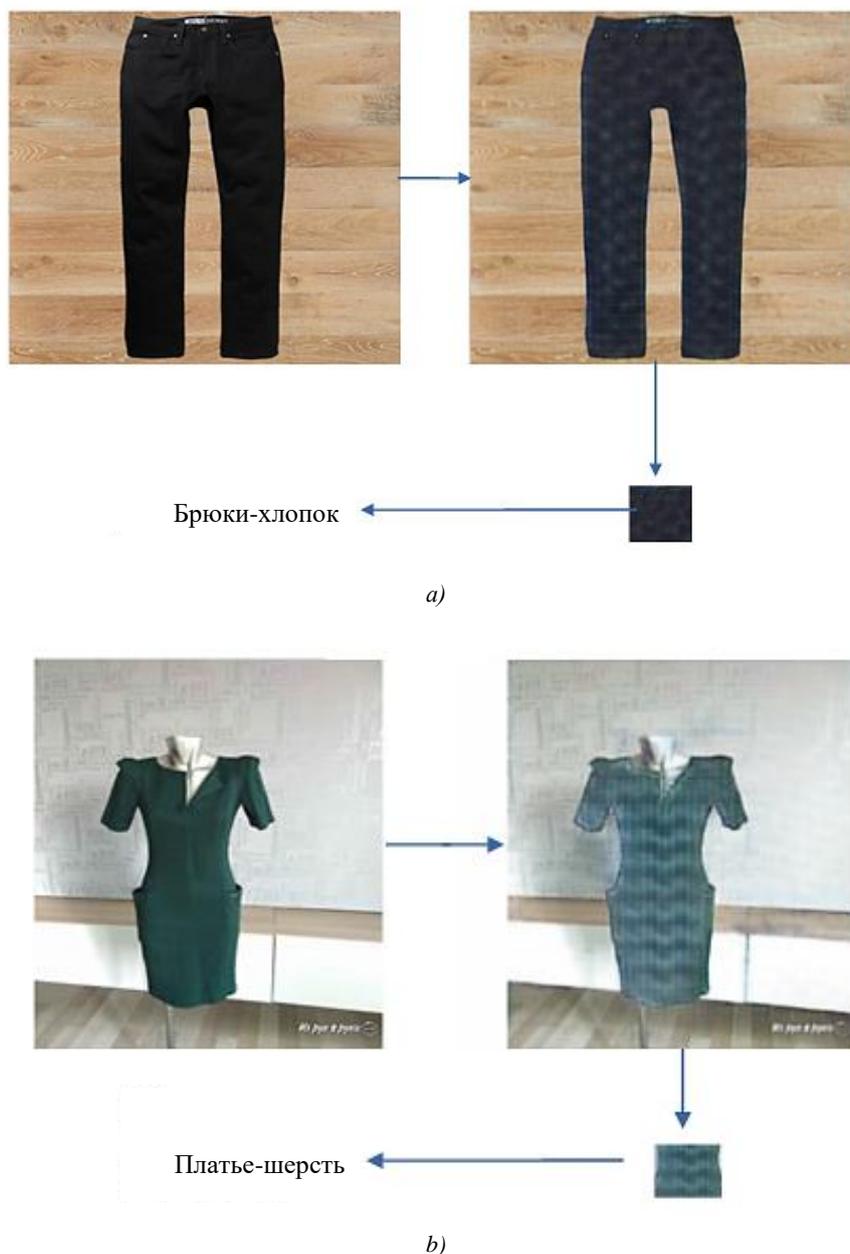


Рис. 10. Результаты работы алгоритма на примере предмета одежды «брюки» (a) и «платье» (b)
Fig. 10. The results of the algorithm on the example of the item of clothing "trousers" (a) and "dress" (b)

Результаты экспериментов показали, что построенная модель обладает хорошими свойствами: стабильностью, более быстрой сходимостью и высокой точностью. Тем не менее можно выделить следующие направления для дальнейших исследований:

1. Генерация и классификация комбинированных материалов.
2. Решение проблемы объема данных. (Метод может не идентифицировать ткани, которые не включены в синтетические изображения, но присутствуют в изображениях реального мира.)
3. Создание единой модели, осуществляющей как генерацию, так и последующую классификацию образца ткани.

Возможным путем решения проблемы комбинированных материалов является их предварительная сегментация, проблема объема данных решается за счет расширения обучающей выборки.

На текущий момент создание единой модели для генерации и классификации образца ткани ограничено двумя факторами:

– недостаточностью вычислительной мощности для обучения, так как описанная архитектура классификатора с моделью внимания при встраивании в генеративно-состязательную сеть на этапе обучения требует увеличения видеопамяти RAM GPU в несколько раз;

– невозможностью достичь того же уровня точности при использовании более простой модели классификации.

Заключение. В ходе исследования был подготовлен набор данных путем автоматического синтеза, произведена его аннотация, построены три сети VGG-19: классическая, с использованием модели внимания и сети pix2pix. Данные модели были обучены на построенном наборе данных, проведены валидация и тестирование. Установлено, что переход от классического изображения предмета одежды к структуре ткани значительно повышает точность модели. Модель внимания позволяет сети лучше сосредоточиться на признаках объекта, что влияет на качество обученной сети и скорость сходимости.

Было выявлено, что использование архитектуры U-Net в качестве основы для генератора сети pix2pix позволяет ускорить модель и помогает ей обойти слой узкого места (bottleneck).

Предложенный метод может быть применен и к другим архитектурам нейронных сетей, а также использован для решения задачи регрессии, когда необходимо установить процентный состав материала.

Список использованных источников

1. Сорокина, В. В. Выделение отдельных участков тела человека на изображении с использованием нейронных сетей и модели внимания / В. В. Сорокина, С. В. Абламейко // Журнал Бел. гос. ун-та. Математика. Информатика. – 2022. – № 2. – С. 94–106.

2. Ronneberger, O. U-Net: convolutional networks for biomedical image segmentation / O. Ronnenberger, P. Fischer, T. Brox // 18th Intern. Conf. on Medical Image Computing and Computer-Assisted Intervention, MICCAI 2015, Munich, Germany, 5–9 Oct. 2015. – Munich, 2015. – Part III. – P. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

3. Suciati, N. Fractal-based texture and HSV color features for fabric image retrieval / N. Suciati, D. Herumurti, A. Y. Wijaya // 5th IEEE Intern. Conf. on Control System, Computing and Engineering, ICCSCE 2015, Batu Ferringhi, Penang, Malaysia, 27–29 Nov. 2015. – Batu Ferringhi, 2015. – P. 178–182.

4. A new method of printed fabric image retrieval based on color moments and gist feature description / J. Jing [et al.] // Textile Research J. – 2016. – Vol. 86, iss. 11. – P. 1137–1150. <https://doi.org/10.1177/0040517515606378>

5. SKL algorithm based fabric image matching and retrieval / Y. Cao [et al.] // Ninth Intern. Conf. on Digital Image Processing (ICDIP 2017), Hong Kong, China, 19–22 May 2017. – Hong Kong, 2017. – Vol. 10420. – P. 104201F.

6. Wang, X. Fabric identification using convolutional neural network / X. Wang, G. Wu, Y. Zhong // Artificial Intelligence on Fashion and Textiles : Proceedings of the Artificial Intelligence on Fashion and Textiles (AIFT) Conf. 2018, Hong Kong, 3–6 July 2018. – Hong Kong, 2018. – P. 93–100. https://doi.org/10.1007/978-3-319-99695-0_12

7. Sorokina, V. Neural network training acceleration by weight standardization in segmentation of electronic commerce images / V. Sorokina, S. Ablameyko // Studies in Computational Intelligence. – 2020. – Vol. 976. – P. 237–244.

8. Fine-grained material classification using micro-geometry and reflectance / C. Kampouris [et al.] // 14th European Conf. on Computer Vision, Amsterdam, the Netherlands, 11–14 Oct. 2016. – Amsterdam, 2016. – Part V. – P. 778–792.

9. Image-to-image translation with conditional adversarial networks / P. Isola [et al.] // 2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017. – Honolulu, 2017. – P. 1125–1134.

10. Simonyan, K. Very deep convolutional networks for large-scale image recognition / K. Simonyan, A. Zisserman // Intern. Conf. on Learning Representations, San Diego, 7–9 May 2015. – San Diego, 2015. – P. 1137–1149.

11. Bahdanau, D. Neural machine translation by jointly learning to align and translate / D. Bahdanau, K. Cho, Y. Bengio // 3rd Intern. Conf. on Learning Representations, ICLR 2015, San Diego, CA, USA, 7–9 May 2015. – San Diego, 2015.

12. Sorokina, V. Extraction of human body parts from the image using convolutional neural network and attention mode / V. Sorokina, S. Ablameyko // Proc. of 15th Intern. Conf. "Pattern Recognition and Information Processing". – Minsk : UIIP NASB, 2021. – P. 84–88.

References

1. Sorokina V., Ablameyko S. *Selection of individual parts of the human body in the image using neural networks and the attention model*. Zhurnal Belorusskogo gosudarstvennogo universiteta. Matematika. Informatika [Journal of the Belarusian State University. Mathematics. Computer Science], 2022, no. 2, pp. 94–106.

2. Ronneberger O., Fischer P., Brox T. U-Net: convolutional networks for biomedical image segmentation. *18th International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI 2015, Munich, Germany, 5–9 October 2015*. Munich, 2015, part III, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

3. Suciati N., Herumurti D., Wijaya A. Y. Fractal-based texture and HSV color features for fabric image retrieval. *5th IEEE International Conference on Control System, Computing and Engineering, ICCSCE 2015, Batu Ferringhi, Penang, Malaysia, 27–29 November 2015*. Batu Ferringhi, 2015, pp. 178–182.

4. Jing J., Li Q., Li P., Zhang L. A new method of printed fabric image retrieval based on color moments and gist feature description. *Textile Research Journal*, 2016, vol. 86, iss. 11, pp. 1137–1150. <https://doi.org/10.1177/0040517515606378>

5. Cao Y., Zhang X., Ma G., Sun R., Dong D. SKL algorithm based fabric image matching and retrieval. *Ninth International Conference on Digital Image Processing (ICDIP 2017), Hong Kong, China, 19–22 May 2017*. Hong Kong, 2017, vol. 10420, p. 104201F.

6. Wang X., Wu G., Zhong Y. Fabric identification using convolutional neural network. *Artificial Intelligence on Fashion and Textiles : Proceedings of the Artificial Intelligence on Fashion and Textiles (AIFT) Conference 2018, Hong Kong, 3–6 July 2018*. Hong Kong, 2018, pp. 93–100. https://doi.org/10.1007/978-3-319-99695-0_12

7. Sorokina V., Ablameyko S. Neural network training acceleration by weight standardization in segmentation of electronic commerce images. *Studies in Computational Intelligence*, 2020, vol. 976, pp. 237–244.

8. Kampouris C., Zafeiriou S., Ghosh A., Malassiotis S. Fine-grained material classification using micro-geometry and reflectance. *14th European Conference on Computer Vision, Amsterdam, the Netherlands, 11–14 October 2016*. Amsterdam, 2016, part V, pp. 778–792.

9. Isola P., Zhu J.-Y., Zhou T., Efros A. A. Image-to-image translation with conditional adversarial networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017*. Honolulu, 2017, pp. 1125–1134.

10. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations, San Diego, 7–9 May 2015*. San Diego, 2015, pp. 1137–1149.

11. Bahdanau D., Cho K., Bengio Y. Neural machine translation by jointly learning to align and translate. *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, 7–9 May 2015*. San Diego, 2015.

12. Sorokina V., Ablameyko S. Extraction of human body parts from the image using convolutional neural network and attention mode. *Proceedings of 15th International Conference "Pattern Recognition and Information Processing"*. Minsk, UIIP NASB, 2021, pp. 84–88.

Информация об авторе

Сорокина Виктория Вадимовна, аспирант кафедры веб-технологий и компьютерного моделирования механико-математического факультета, Белорусский государственный университет.
<https://orcid.org/0000-0002-2128-1943>

Information about the author

Viktoria V. Sorokina, Postgraduate Student of web-Technologies and Computer Modeling Department of Mechanics and Mathematics Faculty, Belarusian State University.
<https://orcid.org/0000-0002-2128-1943>