

СОЗДАНИЕ СИСТЕМЫ РАСПОЗНАВАНИЯ ПОЛА ЧЕЛОВЕКА ПО ЕГО ПОЧЕРКУ

Мясоутов Рамиль Хамзиевич

Аспирант, Волгоградский Государственный

Университет

ramilmyasoutov@yandex.ru

CREATION OF A SYSTEM FOR RECOGNIZING A PERSON'S GENDER BY HIS HANDWRITING

R. Myasoutov

Summary. This paper proposes a solution to the problem of identifying a person's gender by his handwriting. To solve this problem, the author suggests using a system complex consisting of several parts using neural networks and probabilistic — static decision-making methods. The work demonstrates a prototype of this system and also describes the overall interaction of submodules. Various methods for solving the classification problem using neural networks are considered and a comparative characteristic of the operation of each of the methods is given.

Keywords: convolutional neural networks, decision-making methods, handwriting analysis, computer vision.

Аннотация. В данной работе предлагается решение проблемы идентификации пола человека по его почерку. Для решения данной проблемы автор предлагает использовать системный комплекс, состоящий из нескольких частей, использующих нейронные сети и вероятностно — статические методы принятия решений. В работе продемонстрирован прототип данной системы, а также описано общее взаимодействие подмодулей. Рассматриваются разные методы решения задачи классификации с помощью нейронных сетей, а также приведена сравнительная характеристика работы каждого из методов.

Ключевые слова: сверточные нейронные сети, методы принятия решений, анализ рукописного текста, компьютерное зрение.

Введение

В процессе расследования преступлений нередки случаи получения потерпевшими анонимных писем с угрозами, оставленными злоумышленниками. Также нередки ситуации, когда на местах преступления находятся некие заметки, расписки или записи подозреваемых. В подобных случаях следователям приходится прибегать к почерковедческим исследованиям, так как это может помочь получить дополнительную информацию о подозреваемом. К сожалению, данный процесс может занимать до 15 дней, в зависимости от сложности текста.

Сегодня методы компьютерного зрения позволяют решать разнообразные задачи, которые позволяют ускорить рутинную работу. Цель данной работы заключается в создании системы, которая могла бы с минимальным участием человека решать задачу распознавания пола человека на основе рукописного текста.

I Структура системы

На рисунке 1. приведена диаграмма компонентов проектируемой системы. Кратко опишем работу каждого блока. Модуль ввода данных обеспечивает загрузку исследуемого снимка рукописного текста. Далее, полученное изображение передается в модуль

сегментации. Данный модуль состоит из двух основных подмодулей. С помощью алгоритма сегментирования текста на слова[1] модуль сегментации автоматически разбивает текст на слова, предоставляя возможность пользователю скорректировать результат, в случае необходимости. После этого пользователь вручную разбивает слова на буквы, идентифицируя эти буквы вручную. Выполнив все шаги на этапе сегментирования, полученные данные передаются в модуль анализа букв. На этом этапе с помощью нейронных сетей[2] формируется набор гипотез относительно того, кем были написаны буквы в этом тексте. Собрав все необходимые операции, результат работы передается в модуль принятия решений. На этом этапе система с помощью вероятностно-статических методов принимается решение относительно изучаемого образца. В данной статье будет подробно описана работа модулей анализа букв и принятия решений.

II Сбор данных

Для решения поставленной задачи, на первом шаге было необходимо собрать достаточное количество материала для обучения нейронной сети. Согласно методике определения пола исполнителей кратких рукописных текстов[3], большая часть признаков, которые определяют различие между мужскими и женскими текстами, содержатся в местах соединяющие буквы, либо в распо-

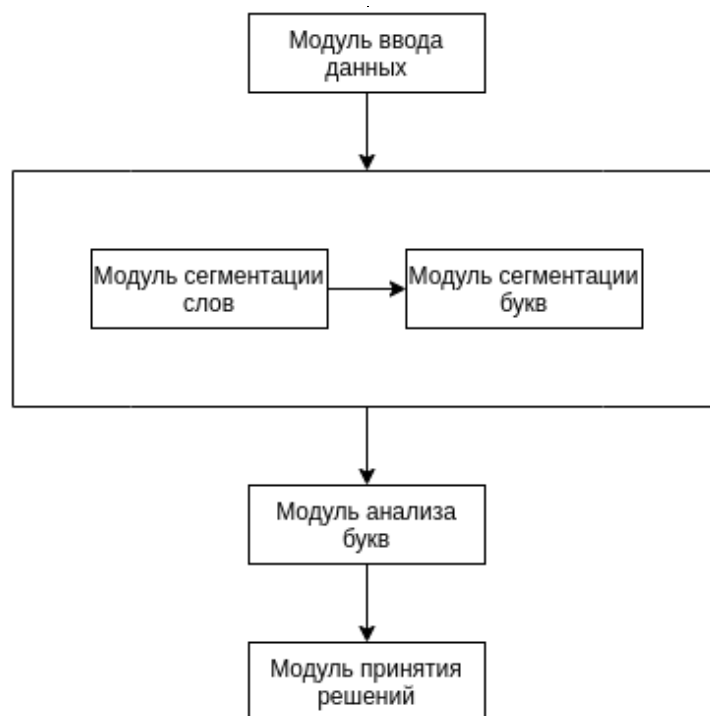


Рис. 1

58	Пол (М/Ж):	М	Возраст:	24
<p>Разрезая носом воды, ходят в море пароходы. Дуют ветры яростные, гонят лодки парусные.</p>				
<p><i>Разрезая носом воды, ходят в море пароходы Дуют ветры яростные, гонят лодки парусные</i></p>				

Рис. 2

ложении самих букв, в слове или тексте. Исходя из этого, для подбора лучшей выборки для обучения необходимо было обучать нейронную сеть на буквах, которые были выделены из текста, а не на буквах, которые были написаны отдельно. Также, согласно сформированному списку признаков[3], был определен набор букв, которые имеют наибольшее количество признаков. Это 22 буквы: а, б, в, г, д, е, ж, з, и, к, л, м, н, о, п, р, с, т, у, ч, ш, я. На основе заданных требований были сформированы анкеты, в которых опрашиваемым предлагалось переписать вручную указанный текст, а также указать свой пол и возраст. Пример анкеты (Рис. 2):

В результате было собрано 420 анкет, на основе которых далее проводились исследования, среди кото-

рых 215 были заполнены мужчинами, а 205 женщинами в возрасте от 19 до 50 лет.

Далее с помощью алгоритма сегментирования[1] полученные тексты были разбиты на слова, а после, вручную, на буквы. В результате были получены изображения букв, содержащие информацию о поле и возрасте человека, написавшего эту букву.

III Обучение нейронной сети

На основе полученных данных была сформирована выборка для обучения. План обучения заключался в том, чтобы для каждой буквы собрать экземпляры написанные мужчинами и женщинам, а после передать этот спи-



Рис. 3

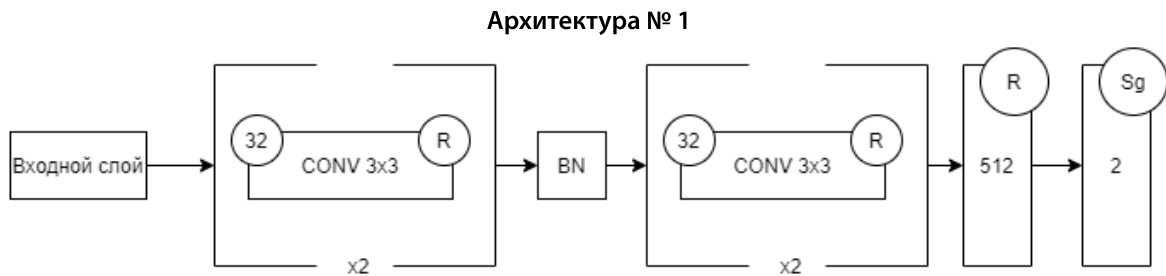


Рис. 4

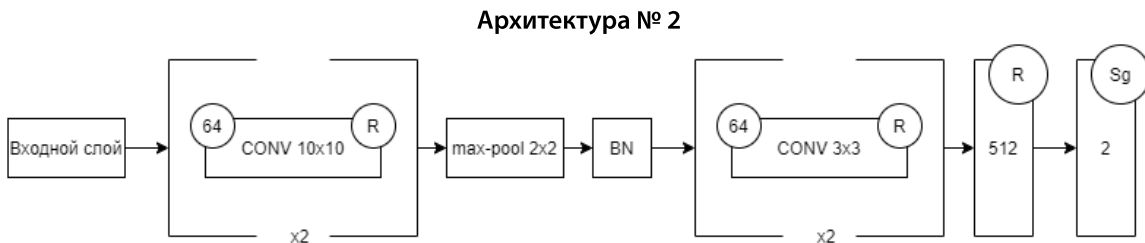


Рис. 5

сок на вход нейронной сети. Было рассмотрено 5 разных архитектур: 3 широко используемых для классификации изображений — LeNet[4], AlexNet[5], VGG16[5], и 2 собственных. Рассмотрим архитектуры собственных нейронных сетей, согласно легенде (Рис. 3)

Для каждой буквы была обучена нейронная сеть, поэтому общая точность выбранной архитектуры вычислялась как среднее арифметическое всех результатов для выбранной архитектуры.

Для каждой из сети были использованы разные алгоритмы оптимизации с различным количеством эпох для обучения. Для вычисления точности использовалась метрика “ассигасу” [6], которая вычисляла отношение количество верных предсказаний по отношению ко всем предсказаниям. Для функции потерь[7] была выбрана средняя квадратичная ошибка:

$$\frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2.$$

Таблица 1

Архитектура	Точность на обучающей выборке	Точность на тестовой выборке
AlexNet	72%	51%
LeNet	63%	51%
VGG16	54%	49%
Архитектура № 1	72%	53%
Архитектура № 2	72%	49%

В итоге были выбраны следующие 5 лучших результатов:

Анализ результатов показывает, что не всегда усложнение архитектуры приводит к улучшению результатов, довольно простая Архитектура № 1 дала лучший результат. В итоге было принято решение использовать для каждой буквы ту архитектуру, которая давала наиболее точный результат. Такой подход позволил повысить точность до 77% и 57% на обучающей и тестовой выборках соответственно.

IV Алгоритм принятия решения

После получения результатов для каждой буквы, необходимо принять решение относительно всех результатов, поэтому было решено воспользоваться вероятностно-статическими методами принятия решения. Был рассмотрен метод Байеса[8].

Результаты работы нейронной сети на предыдущем этапе можно представить в виде двух векторов: $\vec{m} = (m_1, \dots, m_n)$ и $\vec{f} = (f_1, \dots, f_n)$, где векторы \vec{m} и \vec{f} содержат вероятности того, что символ S_i был написан мужчиной или женщиной. Для каждой конкретной буквы может содержаться несколько результатов, так как буквы могут повторяться. На первом шаге сгруппируем все буквы по классам. Классом будем называть буквы в алфавите. Иными словами, если в тексте встречаются три буквы "а", две буквы "б" и несколько букв "с", то будет полагать, что в тексте существуют только три класса K_i . Для каждого класса вычислим вероятность того, что все его элементы были написаны мужской или женской рукой. Положим, что диагноз D_1 говорит о том, что текст написан мужчиной, а D_2 женщиной. Исходя из этих утверждений согласно обобщенной формуле Байеса:

$$P(D_i / K^*) = \frac{P(D_i) P(K^* / D_i)}{P(D_s) P(K^* / D_s)}$$

вычислим вероятности диагнозов D_i для каждого класса K_i :

$$P(D_1 / K^*) = \frac{P(D_1) P(K^* / D_1)}{P(D_2) P(K^* / D_2) + P(D_1) P(K^* / D_1)}$$

$$P(D_2 / K^*) = \frac{P(D_2) P(K^* / D_2)}{P(D_2) P(K^* / D_2) + P(D_1) P(K^* / D_1)}$$

Получим вектора вероятностей для каждого класса K_i : $\vec{M} = (P(K_1), \dots, P(K_n))$ и $\vec{F} = (P(K_1), \dots, P(K_n))$. Далее для каждого класса вычислим его вес. Вес будем вычислять по формуле:

$$W_{k_i} = \frac{l_{k_i}}{L},$$

где l_{k_i} — количество повторений буквы в классе K_i , а L — общее число букв в тексте. После этого, вычислим вероятность написания всего текста мужчиной или женщиной. Воспользуемся критерием Байеса. Будем полагать, что диагноз D_1 говорит о том, что текст написан мужчиной, а D_2 женщиной, тогда:

$$D_1 = \sum_{i=0}^{len(K_i)} W_{K_i} M_i$$

$$D_2 = \sum_{i=0}^{len(K_i)} W_{K_i} F_i$$

тогда корректным диагнозом будет $\max(D_1, D_2)$.

Таким образом, в данной работе была рассмотрена проблема распознавания пола человека на основе рукописного текста. Была описана работа основных компонентов

системы, а также их взаимодействие на каждом из этапов. Конечно, точность идентификации на данный момент нельзя назвать высокой, но работа по улучшению продолжается.

ЛИТЕРАТУРА

1. Мясоутов Р. Х. Анализ рукописного текста в задаче идентификации пола на основе рукописного текста. // Сборник трудов ежегодной научно-практической конференции «Интеллектуальные информационные системы» ИИС-2019, Воронеж, 10–13 декабря 2019 г. ВГУ, 2019. Ч. 1. С. 43–46.
2. Хайкин С. Нейронные сети. Полный курс — М.: Вильямс, 2006. — 995с.
3. Методика определения пола исполнителя кратких рукописных текстов. Министерство внутренних дел СССР. Учебное пособие. 1990 г.
4. Sheng Fang. Review of LeNet-5: How to design the architecture of CNN // Towards Data Science. 2020. URL: <https://towardsdatascience.com/review-of-lenet-5-how-to-design-the-architecture-of-cnn-8ee92ff760ac> (дата обращения 20.04.2020)
5. Raimi Karim. Illustrated: 10 CNN Architectures // Towards Data Science. 2019. URL: <https://towardsdatascience.com/illustrated-10-cnn-architectures-95d78ace614d> (дата обращения 20.04.2020)
6. Accuracy metrics // Keras. URL: https://keras.io/api/metrics/accuracy_metrics/ (дата обращения 21.04.2020)
7. Losses // Keras. URL: https://keras.io/api/losses/regression_losses/#meansquarederror-class (дата обращения 22.04.2020)
8. А.П. Науменко, И. С. Кудрявцева, А. И. Одинец, Вероятностно-статистические методы принятия решений: теория, примеры, задачи. — Омск: ОмГТУ, 2018—8с.

© Мясоутов Рамиль Хамзяевич (ramilmyasoutov@yandex.ru).

Журнал «Современная наука: актуальные проблемы теории и практики»



Волгоградский государственный университет