

Технологическое обеспечение почерковедческих исследований с внедрением нейронных сетей на основе рукописного материала

 **К.А. Чернышев**

ФГБОУ ВО «Московский государственный юридический университет имени О.Е. Кутафина (МГЮА)», Москва 125993, Россия

Аннотация. Статья затрагивает вопросы автоматизации и компьютеризации судебной почерковедческой экспертизы и почерковедческих исследований с внедрением в производство новых информационных технологий – искусственных нейронных сетей. Основной целью работы является разработка методов исследования почерка, помогающих следователям и судьям распознавать неразборчивый почерк из материалов дел. Обосновывается идея создания автоматизированной информационно-поисковой системы объектов почерковедческой экспертизы, содержащей образцы почерка лиц, сдававших единый государственный экзамен. Обсуждаются преимущества и недостатки таких образцов почерка с точки зрения их полноты, достоверности, достаточности и сопоставимости. Для создания соответствующего репозитория предлагается принять необходимые управленческие и законодательные решения. Кроме обеспечения почерковедческих экспертиз, образцами почерка для сравнительного исследования репозиторий образцов почерка будет выполнять и другую важную задачу – хранить большой объем данных, позволяющих обучать нейронные сети, которые впоследствии поспособствуют решению различных идентификационных и диагностических задач судебного почерковедения. Все обучающие данные подлежат разметке и систематизации, потому в статье также предлагаются основания для их разметки. Представлен алгоритм двухэтапного распознавания неразборчивого почерка нейронными сетями: на основе графических характеристик рукописи, а затем морфологических и синтаксических признаков письменной речи с помощью технологий компьютерного зрения и языковой модели русского языка соответственно. Автор описывает собственные системы диагностических и идентификационных признаков и правила их расчета, применение которых будет способствовать решению поставленных перед экспертами задач.

Ключевые слова: *нейронные сети, искусственный интеллект, судебная экспертология, автоматизация, большие данные, образцы для сравнительного исследования, судебное почерковедение, экспертное исследование, методы и методики, распознавание почерка*

Для цитирования: Чернышев К.А. Технологическое обеспечение почерковедческих исследований с внедрением нейронных сетей на основе рукописного материала // Теория и практика судебной экспертизы. 2025. Т. 20. № 3. С. 72–84. <https://doi.org/10.30764/1819-2785-2025-3-72-84>

Technological Support of Forensic Examination of Handwritten Material with Introduction of Neural Networks

 **Kirill A. Chernyshev**

Kutafin Moscow State Law University (MSAL), Moscow 125993, Russia

Abstract. The article deals with the issues of automation and computerization of forensic handwriting examination and handwriting studies with regard to the introduction of new information technologies – the artificial neural networks – into the expert production. The main purpose of the scientific work is to develop handwriting examination methods that help investigators and judges recognize illegible handwriting in case materials. It substantiates the idea of setting up an automated information retrieval system (IRS) for handwriting examination objects containing handwriting reference samples of those persons who passed the Unified

State Exam. The advantages and disadvantages of such handwriting samples are discussed in terms of their completeness, reliability, sufficiency and comparability. Necessary management and legislative decisions are proposed to be taken in order to set up the relevant repository. In addition to providing handwriting reference samples for comparative analysis in the framework of handwriting examination, the handwriting samples repository solves another important task – it contains big data allowing for neural networks training which, in its turn, is capable of solving various identification and diagnostic tasks of forensic handwriting examination. The training data are subject to markup and systematization, and the article suggests the basis for their markup. An algorithm for two-stage recognition of illegible handwriting by neural networks is proposed: the one based on the graphical characteristics of the manuscript and the other one related to morphological and syntactic features of the written speech, which is based on computer vision technologies and the language model of the Russian language. The author presents his own systems of diagnostic and identification features and rules for their calculation, the application of which contributes to solving the tasks assigned to experts.

Keywords: *neural networks, artificial intelligence, forensic examination, automation, big data, samples for comparative analysis, forensic handwriting examination, expert study, methods and techniques, handwriting recognition*

For citation: Chernyshev K.A. Technological Support of Forensic Examination of Handwritten Material with Introduction of Neural Networks. *Theory and Practice of Forensic Science*. 2025. Vol. 20. No. 3. P. 72–84. (In Russ.). <https://doi.org/10.30764/1819-2785-2025-3-72-84>

Введение

Вступление в эру развития информационных технологий отразилось на всех сферах общественных отношений. Так, например, на смену бумажному документообороту пришел электронный, а рукописный почерк все чаще стал отображаться в документах в цифровом виде. По мнению Е.Р. Россинской и Н.Ф. Бодрова, «уже сейчас можно с уверенностью утверждать, что процессы цифровизации приводят к трансформации объектов судебно-экспертного исследования, которые, в сравнении с рукописями на бумаге, обладают значительно большим потенциалом для идентификации» [1, с. 124]. Названные обстоятельства ставят новые задачи перед судебной экспертизой и влияют на ее практические аспекты, вынуждая адаптироваться к новым условиям, совершенствовать существующие и разрабатывать новые методы и методики исследования почерка.

С конца 1990-ых годов создание баз данных и автоматизированных информационно-поисковых систем (АИПС) по конкретным объектам экспертизы, в числе которых объекты криминалистических и справочно-вспомогательных учетов [2, с. 22], является одним из приоритетных направлений развития судебной экспертизы. Уже существуют базы данных объектов

дактилоскопической [3, с. 15], взрывотехнической [4, с. 69], баллистической [5, с. 237] и иных экспертиз. В настоящей статье предлагается создать АИПС по объектам судебной почерковедческой экспертизы, представляющую собой базу данных использованных при проведении единого государственного экзамена (ЕГЭ) бланков с образцами рукописного почерка экзаменуемых.

Основная часть

Бланки ЕГЭ (рис. 1) являются машиночитаемыми формами, подлежащими автоматизированной обработке аппаратно-программным комплексом. В процессе такой обработки внесенная в поля бланков информация преобразуется в текст посредством специальных аппаратно-программных средств.

Основным препятствием при создании хранилища данных объектов почерковедческой экспертизы является отсутствие технической и законодательной базы, обеспечивающей сохранность бланков ЕГЭ. Согласно пункту 17 Порядка разработки, использования и хранения контрольных измерительных материалов (КИМ) при проведении государственной итоговой аттестации (ГИА), разновидностью которой является ЕГЭ, по образовательным про-



Апробация

Дополнительный бланк ответов № 2

Код региона 77	Код предмета 12	Название предмета Общ	Резерв 6
--------------------------	---------------------------	---------------------------------	--------------------

Дополнительный бланк ответов № 2 **2741195592557** Лист **3**



Перепишите значения полей: Код региона Код предмета Название предмета из БЛАНКА РЕГИСТРАЦИИ

Отвечая на задания с РАЗВЕРНУТЫМ ОТВЕТОМ пишите аккуратно и разборчиво, соблюдая разметку страницы.

Не забудьте указать номер задания, на которое Вы отвечаете, например: 31

Условия задания переписывать не нужно

ВНИМАНИЕ! Данный бланк использовать только после заполнения основного бланка ответов № 2

124 (продолжение)

3 - При устройстве на работу человек уравнивает новые правила и установки трудового коллектива, что сказывается на его самосознании.

125 1) политическая система общества - совокупность политических институтов, а также формы их взаимодействия, обеспечивающие управление обществом и его стабильное функционирование.

2) Существует 4 подсистемы политической системы общества: институциональная, коммуникативная, функциональная, идеологическая.

3) Одной из функций идеологической подсистемы является распространение идеологии, то есть политических установок и ценностей и слоганов общества на их основе.

127. 1 - судопроизводство: гражданское

2 - стороны: истец - Ташка. Ответчик - Бывший муж Александры

3 - другие категории дел: деление имущества при разводе, установление наследства.

Оборотная сторона бланка НЕ ЗАПОЛНЯЕТСЯ. Попросите дополнительный бланк ответов № 2

Рис. 1. Общий вид бланка второй части ЕГЭ по обществознанию за 2018 год
Fig. 1. General view of the form of the second part of the Unified State Exam in social studies as of 2018

граммам среднего общего образования, утвержденному приказом Федеральной службы по надзору в сфере образования и науки от 11.08.2022 № 871¹, после

¹ Приказ Рособрнадзора от 11.08.2022 № 871 (ред. от 19.01.2024) «Об утверждении Порядка разработки, использования и хранения контрольных измерительных материалов при проведении государственной итоговой аттестации по образовательным программам среднего общего образования и Порядка разработки, использования и хранения контрольных измерительных материалов при проведении государственной итоговой аттестации по образовательным программам среднего общего образования» // КонсультантПлюс.

проведения ГИА неиспользованные и использованные КИМ и критерии оценивания направляются в места, определенные органами исполнительной власти (ОИВ), учредителями, загранучреждениями для обеспечения их хранения. Там они остаются до 1 марта года, следующего за годом проведения экзамена. По истечении указанного срока перечисленные материалы уничтожаются лицами, определенными ОИВ, учредителями, загранучреждениями с оформлением соответствующего акта.

Таким образом, учитывая, что проведение ЕГЭ в сложившейся практике приходится на лето каждого года, срок хранения бланков ЕГЭ и потенциальных образцов почерка составляет около полугода. В этой связи считаем необходимым принять на государственном уровне решение о создании базы данных почерковых объектов и определить государственный орган либо подведомственную такому органу организацию, отвечающую за создание, ведение, обслуживание и функционирование *репозитория образцов почерка* (РОП).

Для наполнения РОП данными и обеспечения их хранения предлагаем внести поправки в вышеназванное законодательство, возложив на Рособрнадзор и иных субъектов обработки КИМ обязанность обеспечить передачу КИМ и бланков ЕГЭ по всем учебным предметам (за исключением предметов: «иностраный язык», ввиду несоответствия требованиям, предъявляемым к сравнительным образцам почерка, и «математика» (базовый уровень); а также «информатика», ввиду отсутствия соответствующих бланков в принципе) в государственные органы и подведомственные им организации, ответственные за ведение РОП.

После получения бланков ЕГЭ выбранные государственные органы и подведомственные им организации будут осуществлять предварительную проверку образцов почерка перед включением бланков в репозиторий. Исключению подлежат только те бланки ЕГЭ, в которых отсутствуют образцы почерка либо их объем является недостаточным для использования в качестве сравнительных образцов почерка, а также бланки, чья достоверность объективно вызывает сомнение. К таким, например, относятся бланки ЕГЭ по разным учебным предметам, формально принадлежащие одному лицу, но заполненные почерками разных лиц.

На следующем этапе все образцы почерка подлежат группировке и разметке. Данные систематизируются в группы по принадлежности конкретному лицу, являющемуся исполнителем почерковых записей. Так, при обнаружении бланков ЕГЭ по русскому языку, истории и обществознанию, выполненных А.А. Ивановым, все образцы его почерка помещаются в одну «папку», а все имеющиеся бланки ЕГЭ Б.Б. Петрова – в другую и т.д. Внутри каждой «именной» папки бланки классифицируются по учебным

предметам и датам отбора образцов почерка, располагаясь при этом в хронологическом порядке их отбора (в последовательности сдачи экзаменов).

Классификация по учебным предметам имеет значение при определении темы письменной речи, если образцы понадобятся для производства автороведческих или лингвистических экспертиз. При маркировании данных обязательно указываются ФИО лица, которому принадлежат образцы почерка, его пол, возраст, ведущая пишущая рука (правша или левша – при наличии такой информации), общее количество почеркового материала (в количестве листов бланков ЕГЭ), дата получения образцов почерка (дата проведения экзамена) для бланков ЕГЭ каждого экзамена, дата внесения образцов почерка в РОП, сведения об использовании образцов почерка в качестве обучающих и/или тестовых данных для нейросетей (при использовании). Столь подробная детализация данных в перспективе может быть полезна при обучении алгоритмов искусственного интеллекта распознаванию мужского и женского почерка, возрастных характеристик пишущего лица и т.д.

На текущий момент преобладающим способом развития нейросетей является обучение, требующее размеченных (структурированных) данных. В зарубежной литературе данные делят на три категории в зависимости от типа их организации. К структурированным относятся те данные, которые систематизированы в хранилище по конкретным форматам, позволяющим беспрепятственно осуществлять их поиск и выборку. Неструктурированными считаются неупорядоченные и не соответствующие предопределенным моделям данные. Они часто недоступны для поиска, и их трудно анализировать. Под полуструктурированными понимается тип данных, обладающий определенными организационными характеристиками, облегчающими их обработку. К ним относят теги или другие маркеры, зачастую включенные в сами данные для их организации [6, с. 2].

В представленном на рисунке 1 бланке видно, что отображенный почерк обладает такими признаками, как индивидуальность, динамическая устойчивость, избирательная изменчивость и вариационность. Это дает основание считать материал пригодным для судебно-экспертного исследова-

ния, а объекты – имеющими ценность для решения идентификационных и диагностических задач судебной почерковедческой экспертизы.

Говоря о преимуществах применения бланков ЕГЭ, отметим, что упомянутый экзамен, как правило, сдается обучающимися по достижении 17–18 лет, когда формирование письменно-двигательного навыка близко к своему завершению, а почерк обучающихся можно охарактеризовать как средне- или высоковыработанный.

Названные рукописи содержат все виды объектов почерковедческой экспертизы: тексты больших объемов, подписи, цифровые записи. Имеется как печатное, так и письменное выполнение различных букв и цифр. Такие образцы почерка подойдут для производства любых почерковедческих экспертиз. Отсутствует необходимость проверять достоверность полученных образцов: ЕГЭ имеет строгий порядок проведения, включающий проверку удостоверяющего личность документа, и видеофиксацию.

Преимуществом использования бланков ЕГЭ в качестве сравнительных образцов является также достаточное количество материала, поскольку экспертная практика независимо от объектов исследования рекомендует, чтобы объем образцов составлял не менее пяти страниц стандартного формата [7, с. 89]. Действующим законодательством предусмотрена обязательная сдача двух экзаменов: «русский язык» и «математика» (базовый уровень). Для поступления в вуз требуется, как правило, сдать еще два экзамена по выбору. За один экзамен учащийся в среднем заполняет 4–6 бланков, соответственно, за 4 экзамена используются от 16 до 24 бланков ЕГЭ, содержащих образцы почерка. Таким образом, даже в случае заполнения базы только бланками, полученными по

итогах сдачи обязательного для всех обучающихся ЕГЭ по русскому языку, эксперт будет обеспечен достаточным количеством сравнительного материала.

В судебной экспертологии все образцы для сравнительного исследования классифицируются на свободные, условно-свободные и экспериментальные. К свободным относятся такие образцы, которые возникли до возбуждения уголовного дела и вне связи с судебной экспертизой [8, с. 31]. Образцы бланков ЕГЭ в полной мере подпадают под определение свободных, что сводит к минимуму вероятность умышленного искажения почерка, имитации почерка иных лиц и сокрытия собственных общих и частных признаков. Более того, сдающие ЕГЭ лица стараются проявлять особую старательность и прилежность при заполнении бланков, поскольку осведомлены, что первая часть экзамена проверяется компьютером в автоматическом режиме, а вторая – экспертами-проверяющими. Компьютер и человек могут не распознать неразборчиво внесенные в бланк рукописные записи, что может привести к понижению количества выставляемых за экзамен баллов.

Еще одним достоинством является то обстоятельство, что эксперт-почерковед освобождается от решения задачи по установлению вида пишущего устройства, поскольку заполнение бланков исключительно капиллярной или гелевой ручкой с чернилами черного цвета (рис. 2) – одно из требований экзамена. Зная об этом и учитывая особенности названных пишущих устройств, эксперт заранее может скорректировать свое исследование и максимально объективно оценить признаки почерка.

Однако стоит сказать и о недостатках получения образцов подобным способом. В первых, экзаменуемые в той или иной мере

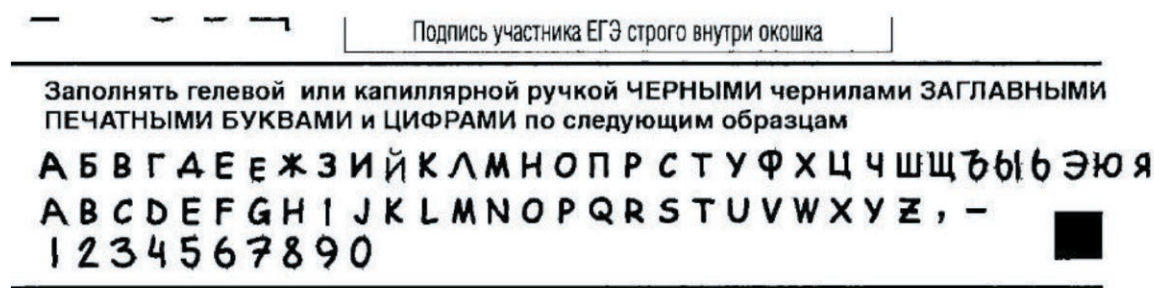


Рис. 2. Требования к заполнению бланков при решении первой части ЕГЭ
Fig. 2. Requirements for filling out the Unified State Exam forms when solving its first part

находятся в стрессовой ситуации, что является фактором, иногда приводящим к искажению рукописи. Во-вторых, заполнение бланков сопряжено с соблюдением строгих правил. При решении первой части ЕГЭ обучающийся должен вносить записи строго в границах очерченных клеток бланка, что снижает ценность таких признаков, как размер и ширина букв. Более того, сами символы должны соответствовать тем образцам, которые указаны на бланке (рис. 2). Таким образом, если у сдающего ЕГЭ характерной чертой выполнения цифры «четыре» является слитное написание двух ее верхних черт («4»), то данный признак будет утрачен.

Как уже было сказано ранее, образцы почерка при сдаче ЕГЭ отбираются у лиц 17–18-летнего возраста. Критерий сопоставимости образцов по времени их выполнения не позволит применять бланки ЕГЭ, заполненные учеником старших классов, если спорный документ выполнен взрослым либо пенсионером.

Наличие базы данных образцов почерка значительно упрощает процесс назначения судебной экспертизы. Но несмотря на существование нормы, предусматривающей право следователя изымать образцы для сравнительного исследования (ст. 202 УПК РФ), на практике получить образцы почерка при противодействии проверяемого лица практически невозможно. Имея вышеназванный репозиторий, при необходимости проведения почерковедческой экспертизы следователю или суду достаточно будет направить запрос о предоставлении образцов почерка конкретного лица, что позволит существенно ускорить процесс производства экспертизы и повысить ее доказательственную значимость.

Нельзя не согласиться с мнением А.Ф. Купина и А.С. Коваленко в отношении использования нейросетей в почерковедческих исследованиях. По их словам, «представляется актуальным использование технологий искусственного интеллекта для распознавания содержания рукописных текстов, фрагменты букв и цифр которых частично не дописаны либо искажены в силу естественных или искусственных причин» [9, с. 30]. Однако если упомянутые авторы делают акцент на проведении сравнительных исследований посредством применения технологий ИИ, то мы сконцентрируемся на решении диагностической задачи по установлению содержания неразборчивого рукописного текста.

Под нейросетями в судебной экспертологии понимаются основанные на технологии ИИ математические модели, обученные принимать автономные решения на базе соответствующих данных и применяемые при назначении, производстве и оценке судебных экспертиз [10, с. 245–246].

Рассмотрим существующие подходы к производству судебных почерковедческих экспертиз при помощи нейросетей.

Представители пензенской школы работали среду моделирования «БиоНейроАвтограф», способную решить вопрос о подлинности или подложности подписи. В основе работы нейросети лежит 480 параметров, которые преобразуются ею в двоичный код, где каждому параметру присваивается значение «0» или «1». При выполнении подписи отличным от предполагаемого лицом подразумевается, что параметры примут другие значения, а нейросеть констатирует другого исполнителя рукописи [11, с. 251–253]. Но стоит отметить, что такой способ исследования непригоден для решения задач по исследованию длинных текстовых записей.

Как было сказано выше, бланки ЕГЭ способны не только обеспечить базу данных почерковых объектов конкретных лиц, которые могут быть использованы в качестве сравнительных образцов при производстве почерковедческих экспертиз, но и предоставить огромный объем обучающих данных для искусственных нейронных сетей.

В литературе часто обсуждается задача распознавания нейросетью десяти арабских рукописных цифр. В теории нейронных сетей это является задачей классификации, а в теории судебной экспертологии – диагностической задачей по установлению содержания рукописной записи, выполненной неразборчивым почерком. С этой целью была создана соответствующая база данных MNIST, (датасет [data set]) образцов рукописного написания цифр – 60 000 изображений размером 28 × 28 пикселей (рис. 3). MNIST выступает в качестве обучающих данных – на нем нейронная сеть учится распознавать рукописные цифры, и при надлежащем обучении точность распознавания составляет более 95% [12, с. 21].

Обращаясь к использованным бланкам ЕГЭ, мы представляем их как датасет заглавных и строчных прописных букв русского языка, арабских цифр и подписей. Учитывая потенциальный объем рукописного



Рис. 3. Фрагмент образцов базы MNIST
Fig. 3. A fragment of MNIST database samples

материала, необходимого и достаточного для обучения нейросетей, мы полагаем, что в первую очередь должны быть оцифрованы нормы прописи русского алфавита с пометкой эталонных образцов для того, чтобы нейросеть понимала корректное исполнение букв и, соответственно, могла правильно распознавать вариативности их выполнения. В качестве примера нами подготовлен мини-макет базы данных выполнения строчной рукописной буквы «д» (табл. 1). Эмпирический материал составили 7 исписанных бланков второй части ЕГЭ по обществознанию за 2018 год. По аналогии могут быть созданы датасеты для других строчных и заглавных букв русского языка, арабских цифр.

Представленные изображения бланков ЕГЭ являются цифровыми изображениями в формате TIFF (tagged image file format). Это графический контейнер, хранящий в себе растровые изображения. Формат отличается высоким качеством картинки с повышенной глубиной цвета и несколькими вариантами компрессии данных: с потерями в качестве и без потерь. Формат TIFF «без использования сжатия» применяют для растровых изображений – он считается лучшим форматом для их сохранения с максимальной детализацией. Как видно из рисунка 4, все рукописные штрихи при переводе их в электронную форму отображаются в виде маленьких черных квадратов или же пикселей. Данное обстоятельство открывает но-

вые возможности для обучения нейросетей распознаванию почерка и методам его исследования.

Еще самые первые сверточные нейронные сети, предназначенные для работы с изображениями, в том числе распознавания рукописных символов, были основаны именно на этой технологии: распознавание символов по конфигурации пикселей. На рисунке 5 изображена модель такой искусственной нейронной сети.

Для обучения нейросети распознаванию каждого символа, а в нашем случае – каждой буквы, Ян Лекун, французский и американский ученый, использовал матрицы [13, с. 88–89]. В каждой ячейке матрицы находилось значение +1 или -1. Черные клетки, используемые при выполнении символа, помечались значением +1, незаполненные клетки – значением -1. В рамках матрицы для каждой цифры и буквы набор этих плюсов и минусов был разным, что позволило искусственному интеллекту отличать одну букву от другой. Этот метод применялся при обучении перцептрона – простейшей из всех существующих нейронных сетей. Современное развитие технологий и методов обучения не оставляет сомнений в том, что упомянутая ранее задача без особых трудностей может быть решена с применением глубоких нейронных сетей и современных методов их обучения.

В аспекте решения судебно-экспертной задачи распознавания почерка было бы

Таблица 1. Датасет оцифрованной строчной буквы «д» (извлечение)
Table 1. Dataset of the digitized lowercase letter "d" (extraction)

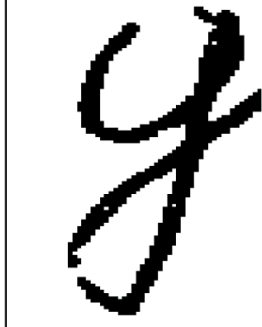
		
Образец №1	Образец №2	Образец №3
		
Образец №4	Образец №5	Образец №9
		
Образец №11	Образец №15	Образец №21



Рис. 4. Увеличенный фрагмент бланка ЕГЭ
Fig. 4. Enlarged fragment of the Unified State Exam form

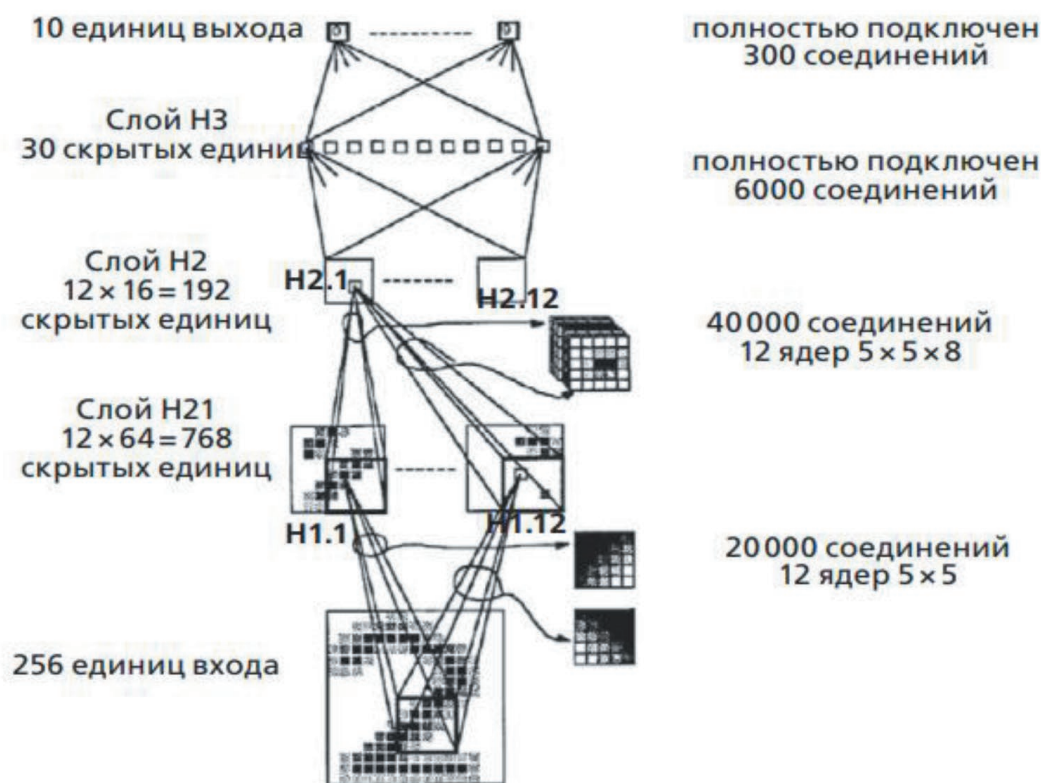


Рис. 5. Первая сверточная сеть для распознавания рукописных символов, созданная Яном Лекуном в 1988 году²

Fig. 5. The first convolutional network for handwritten character recognition created by Yann LeCun in 1988

оптимально использовать не только анализ взаиморасположения знаков и элементов букв, отразив тем самым двигательные навыки пишущего лица, но и (при затрудненном распознавании) морфологические и синтаксические признаки письменной речи. Для этого в обучаемой нейросети должна быть воплощена языковая модель русского языка.

Примерная схема работы подобной нейронной сети представлена на рисунке 6.

Взаимосвязь между расположением конкретных букв и морфологическими, синтаксическими признаками можно проследить на основе следующих закономерностей русского языка:

- если в предложении выявлено наречие, то оно должно относиться к глаголу в этом предложении;
- все наречия в русском языке оканчиваются на букву «о»;
- если неразборчивое слово является глаголом, то с большой вероятностью в кон-

це слова расположены согласные буквы «л (а)», «т (ь)», «т», «м»; перед каждой из указанных согласных букв обязательно будет стоять гласная буква;

- если в предложении выявлен глагол в форме инфинитива, оканчивающийся на «ть», то в предложении должен быть еще один глагол;

– если в предложении есть имя прилагательное или причастие, то оно должно относиться к имени существительному;

- имя существительное и зависимые от него имена прилагательные и причастия имеют одинаковые род, число, падеж;

– если неразборчивое слово является именем прилагательным, то с большой вероятностью на конце слова расположены буквы «ая», «яя» – для прилагательного женского рода, «ый», «ий» – для прилагательного мужского рода, «ое» – для прилагательного среднего рода, «ые», «ие» – для прилагательного во множественном числе;

- если в предложении имеются союзы «и», «или», то, вероятно, они соединяют однородные члены предложения, и нераспознанные слова принадлежат к одной и той же части речи.

² Изображение взято из источника: Лекун Я. Как учится машина: революция в области нейронных сетей и глубокого обучения / Пер. с фр. Е. Арсеновой. М.: Альпина ПРО, 2021. С. 48.

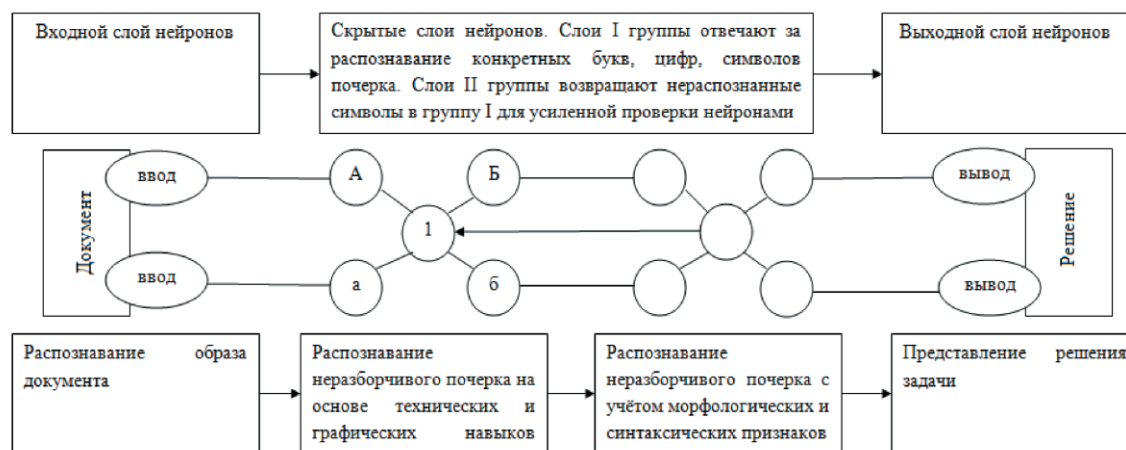


Рис. 6. Строение и схема работы нейронной сети с двухэтапным распознаванием почерка
Fig. 6. The structure and operation scheme of the neural network with two-stage handwriting recognition

Приведем пример работы представленной нейронной сети. Допустим, алгоритму предоставили краткую запись, сделанную неразборчивым почерком, содержащую следующий текст: *работать тяжело и трудно*. После проверки нейронами первой группы, где каждый нейрон отвечает за распознавание конкретной буквы или цифры, нейросеть распознала запись следующим образом: «*работать XXXXXX и трудно*». Затем данная рукопись передается нейронам второй группы, ими совершаются следующие операции-рассуждения: 1) после нераспознанного слова следует союз «и», который соединяет однородные члены предложения; 2) нераспознанное слово соединено союзом «и» со словом «*трудно*», которое является наречием; 3) следовательно, нераспознанное слово с большой вероятностью тоже является наречием; 4) все наречия в русском языке заканчиваются на букву «о», перед которой стоит согласная буква; 5) нераспознанное слово подлежит повторной усиленной проверке нейронами первой группы, отвечающими за распознавание буквы «о» и согласных букв на конце слова; 6) семантически нераспознанное слово связано с глаголом «*работать*», значит, при повторном распознавании нейронами первой группы это слово подлежит сравнению с наречиями, которые, согласно Национальному корпусу русского языка, наиболее часто употребляются с глаголом «*работать*».

Иными словами, распознавание по морфологическим и синтаксическим признакам письменной речи по существу представляет собой определение содержания рукописного текста исходя из контекста написанного, заранее установленного экс-

пертом. Таким образом, распознавание неразборчивого почерка осуществляется в два этапа. На первом этапе исследуются письменно-двигательные навыки, что выражается в изучении цифровых образов букв, расположения пикселей относительно друг друга. При невозможности решения задачи распознавания некоторых символов нейросеть переходит к анализу морфологических и синтаксических признаков. Для такого всестороннего исследования почерка необходимы нейронные сети с технологией глубокого обучения. Нейроны первой группы функционируют на основе компьютерного зрения [14, с. 71] и представляют собой сверточную нейронную сеть. Нейроны второй группы взаимодействуют на основе языковой модели русского языка [15, с. 136–138].

Несмотря на то, что в настоящей публикации сделан акцент на решении диагностической задачи почерковедческих исследований, полагаем, что нейросети могут быть применены схожим образом и для решения идентификационной задачи по установлению исполнителя рукописи.

При этом оцифрованное изображение в пиксельном формате позволяет исследовать новые математические характеристики почерка, анализ которых невозможен при применении исключительно традиционных методов судебной почерковедческой экспертизы.

В качестве новых признаков для исследования мы предлагаем учитывать:

1) общее количество пикселей при выполнении надстрочных элементов букв (сумма пикселей всех надстрочных элементов всех букв);

2) общее количество пикселей при выполнении подстрочных элементов букв (сумма пикселей всех подстрочных элементов всех букв);

3) общее количество недостающих пикселей в конкретных буквах в сравнении с нормами прописи (эталонные образцы) этих букв (вычисляется как сумма разностей между количеством пикселей в эталонном выполнении буквы и каждом конкретном выполнении этой же буквы);

4) общее количество избыточных пикселей в конкретных буквах в сравнении с нормами прописи (эталонные образцы) этих букв (вычисляется как сумма разностей между количеством пикселей в эталонном выполнении буквы и каждом конкретном выполнении этой же буквы);

5) среднее количество пикселей при выполнении конкретной буквы (вычисляется общее количество пикселей при выполнении, к примеру, всех строчных букв «к», а затем делится на общее количество строчных букв «к» в рукописи)³;

6) общее количество пикселей при выполнении знаков препинания (сумма пикселей всех знаков препинания).

Не исключается возможность подсчета иных количественных характеристик почерка. Указанные признаки высчитываются нейросетью в спорной рукописи и сравнительных образцах в автоматическом режиме. Затем эксперт находит индекс, который представляет собой разность между признаками⁴ в спорной рукописи и признаками в сравнительных образцах. Математические знаки при получении результатов игнорируются. Если полученный индекс больше значения X (оно условно равно 50 пикселям), то признак считается различающимся, а в случае получения индекса меньше X , имеет место совпадение признака. Отметим, что значение $X=50$ взято условно для демонстрации алгоритма; его конкретное значение должно устанавливаться эмпирическим путем в процессе апробации методики.

Предложенные признаки почерка в той или иной мере соответствуют традиционным частным признакам. Рассчитывая количество пикселей, их расположение в той

или иной букве, ее элементах, мы фактически исследуем конструктивное строение, протяженность движений, количество движений и их относительное размещение. Таким образом, новые методы исследования гармонично дополняют уже существующие, придают экспертному исследованию комплексный и всесторонний характер.

Представленный способ исследования, разумеется, еще подлежит испытанию на практике, но уже сейчас трудно не согласиться с тем, что такие компьютерно-математические методы исследования почерка с применением искусственного интеллекта являются хорошим вспомогательным инструментом для эксперта-почерковеда, повышая доказательственное значение и научную обоснованность его выводов.

Выводы

Подводя итог, отметим, что наши во многом инновационные предложения, во-первых, отвечают за разные аспекты совершенствования технологии производства почерковедческих исследований и судебной экспертизы, во-вторых, взаимосвязаны и дополняют друг друга. Репозиторий образцов почерка, который предлагается создать путем накопления и систематизации использованных в процессе проведения ЕГЭ бланков, выполняет двойную функцию: является базой данных образцов почерка, отвечающих требованиям, предъявляемым к образцам для сравнительного исследования, а также представляет собой большой объем данных с возможностью их использования для обучения искусственных нейронных сетей решению задач почерковедческих исследований, для чего соответствующие данные систематизируются и размечаются. В дополнение ко второй задаче его функционирования предлагается двухэтапная процедура распознавания почерка, в первую очередь – на основании письменно-двигательных навыков и графических признаков, а при наличии трудностей в распознавании неразборчивого почерка – с учетом морфологических и синтаксических закономерностей русского языка. Кроме того, в настоящей статье предложены новые признаки почерка, основанные на пиксельной структуре символов, изучение которых облегчит проведение идентификационного исследования.

³ Данный признак предлагается устанавливать для наиболее часто встречаемых букв: «а», «о», «и», «м», «н», «к», «с», «т» и др.

⁴ Для каждого признака индекс вычисляется отдельно.

СПИСОК ЛИТЕРАТУРЫ

1. Россинская Е.Р., Бодров Н.Ф. Современное состояние и перспективы исследования образов цифровых следов в судебной почерковедческой экспертизе // Криминалистика: вчера, сегодня, завтра. 2022. Т. 21. № 1. С. 121–135.
<https://doi.org/10.55001/2587-9820.2022.44.98.011>
2. Россинская Е.Р. Система теории цифровизации судебно-экспертной деятельности // Теория и практика судебной экспертизы. 2024. Т. 19. № 3. С. 20–32.
<https://doi.org/10.30764/1819-2785-2024-3-20-32>
3. Захарова В.Д., Ермилова Е.А. Возможности использования нейросетей при проведении дактилоскопических исследований // Молодой ученый. 2023. № 50 (497). С. 14–16.
4. Вербицкая В.А. Цифровые технологии при проведении взрывотехнической экспертизы // Международный журнал гуманитарных и естественных наук. 2020. № 12-3 (51). С. 68–70.
5. Чубарь И.А. Потенциальные возможности использования автоматизированных баллистических идентификационных систем при раскрытии и расследовании преступлений // Вестник экономической безопасности. 2022. № 5. С. 236–239.
<https://doi.org/10.24412/2414-3995-2022-5-236-239>
6. Breiting F., Jotterand A. Sharing Datasets for Digital Forensic: A Novel Taxonomy and Legal Concerns // Forensic Science International: Digital Investigation. 2023. Vol. 45.
<https://doi.org/10.1016/j.fsidi.2023.301562>
7. Почерковедение и почерковедческая экспертиза: Курс лекций / Под ред. В.В. Серегина. 5-е изд. СПб: МВД России, НПСЭП, 2015. 229 с.
8. Россинская Е.Р., Галяшина Е.И. Настольная книга судьи: судебная экспертиза: теория и практика, типичные вопросы и нестандартные ситуации. Монография / Под ред. Е.Р. Россинской. 2-е изд., перераб. и доп. М.: Проспект, 2024. 448 с.
9. Купин А.Ф., Коваленко А.С. К вопросу о возможностях применения систем искусственного интеллекта при криминалистическом исследовании документов и их реквизитов // Теория и практика судебной экспертизы. 2023. Т. 18. № 4. С. 28–35.
<https://doi.org/10.30764/1819-2785-4-28-35>
10. Чернышев К.А. Нейросетевые технологии в аспекте судебно-экспертной деятельности // Криминалистика: вчера, сегодня, завтра. 2024. Т. 32. № 4. С. 239–252.
11. Иванов А.И., Газин А.И., Качайкин Е.И., Андреев Д.Ю. Автоматизация почерковедческой экспертизы, построенная на обучении больших искусственных нейронных сетей // Модели, системы, сети в экономике, технике, природе и обществе. 2016. № 1 (17). С. 249–257.
12. Самылкина Н.Н., Калинин И.А. Изучение работы нейронной сети на примере распознавания рукописных цифр с использованием

REFERENCES

1. Rossinskaya E.R., Bodrov N.F. The Current State and Prospects for the Study of Digital Trace Images in Forensic Handwriting Expertise. *Forensics: yesterday, today, tomorrow*. 2022. Vol. 21. No. 1. P. 121–135. (In Russ.).
<https://doi.org/10.55001/2587-9820.2022.44.98.011>
2. Rossinskaya E.R. The System of Forensic Activity Digitalization Theory. *Theory and Practice of Forensic Science*. 2024. Vol. 19. No. 3. P. 20–32. (In Russ.).
<https://doi.org/10.30764/1819-2785-2024-3-20-32>
3. Zakharova V.D., Ermilova E.A. Possibilities of Using Neural Networks in Conducting Fingerprinting Examination. *Young Scientist*. 2023. No. 50 (497). P. 14–16. (In Russ.).
4. Verbitskaya V.A. Digital Technologies in Explosion Expertise. *International Journal of Humanities and Natural Sciences*. 2020. No. 12-3 (51). P. 68–70. (In Russ.).
5. Chubar I.A. Potential Possibilities of Using Automated Ballistic Identification Systems in the Detection and Investigation of Crime. *Bulletin of Economic Security*. 2022. No. 5. P. 236–239. (In Russ.).
<https://doi.org/10.24412/2414-3995-2022-5-236-239>
6. Breiting F., Jotterand A. Sharing Datasets for Digital Forensic: A Novel Taxonomy and Legal Concerns. *Forensic Science International: Digital Investigation*. 2023. Vol. 45.
<https://doi.org/10.1016/j.fsidi.2023.301562>
7. *Handwriting Studies and Handwriting Examination: Course of Lectures* / V.V. Seregin (ed.). 5th ed. Saint Petersburg: MVD Rossii, NPSEP, 2015. 229 p. (In Russ.).
8. Rossinskaya E.R., Galyashina E.I. *Judge's Handbook: Forensic Examination: Theory and Practice, Typical Issues and Non-Standard Situations: Monograph* / E.R. Rossinskaya (ed.). 2nd ed. Moscow: Prospekt, 2024. 448 p. (In Russ.).
9. Kupin A.F., Kovalenko A.S. On the Question of Applicability of Artificial Intelligence Systems to Forensic Examination of Documents and Their Requisites. *Theory and Practice of Forensic Science*. 2023. Vol. 18. No. 4. P. 28–35. (In Russ.).
<https://doi.org/10.30764/1819-2785-4-28-35>
10. Chernyshev K.A. Neural Network Technologies in the Aspect of Forensic Activities. *Criminalistics: yesterday, today, tomorrow*. 2024. Vol. 32. No. 4. P. 239–252. (In Russ.).
11. Ivanov A.I., Gazin A.I., Kachaykin E.I., Andreev D.Yu. Automation of Graphologic Examination Based on Teaching Large Artificial Neural Networks. *Models, Systems, and Networks in Economics, Technology, Nature, and Society*. 2016. No. 1 (17). P. 249–257. (In Russ.).
12. Samylkina N.N., Kalinin I.A. Exploring the Operation of a Neural Network Using the Example of Handwritten Digits Recognition in an Advanced

- языка программирования Python 3.8 в углубленном курсе информатики // Информатика в школе. 2021. № 10 (173). С. 19–27.
<https://doi.org/10.32517/2221-1993-2021-20-10-19-27>
13. Le Cun Y. *Quand la Machine Apprend: la Révolution des Neurones Artificiels et de L'apprentissage Profound*. Paris: Odile Jacob, 2019. 394 p.
<https://doi.org/10.32517/2221-1993-2021-20-10-19-27>
14. Аскерова Л.Ф. Возможности применения программного обеспечения «MATLAB» для судебно-экспертного исследования видеозображений // Современные научные исследования и инновации. 2018. № 1 [Электронный ресурс].
15. Петрова И.М. Потенциал поисковой системы Google при проведении исследований в рамках когнитивной корпусной лингвистики // Теоретическая и прикладная лингвистика. 2019. Т. 5. № 3. С. 127–142.
https://doi.org/10.22250/2410-7190_2019_5_3_127_142
- Informatics Course Using the Python 3.8 Programming Language. *Informatics in school*. 2021. No. 10 (173). P. 19–27. (In Russ.).
<https://doi.org/10.32517/2221-1993-2021-20-10-19-27>
13. Le Cun Y. *Quand la Machine Apprend: la Révolution des Neurones Artificiels et de L'apprentissage Profound*. Paris: Odile Jacob, 2019. 394 p.
14. Askerova L.F. Possibilities of Using MATLAB Software for Forensic Examination of Video Images. *Modern Scientific Research and Innovations*. 2018. No. 1 (Electronic resource). (In Russ.).
15. Petrova I.M. The Google Search Engine Potential for Conducting Examination in the Framework of Cognitive Corpus Linguistics. *Theoretical and Applied Linguistics*. 2019. Vol. 5. No. 3. P. 127–142. (In Russ.).
https://doi.org/10.22250/2410-7190_2019_5_3_127_142

ИНФОРМАЦИЯ ОБ АВТОРЕ

Чернышев Кирилл Александрович – аспирант кафедры судебных экспертиз Московского государственного юридического университета имени О.Е. Кутафина (МГЮА);
e-mail: mr.kirillch2000@mail.ru

ABOUT THE AUTHOR

Chernyshev Kirill Aleksandrovich – Postgraduate student, Department of Forensic Science, Kutafin Moscow State Law University (MSAL);
e-mail: mr.kirillch2000@mail.ru

Статья поступила: 27.02.2025

После доработки: 14.04.2025

Принята к печати: 16.05.2025

Received: February 27, 2025

Revised: April 14, 2025

Accepted: May 16, 2025