

СРАВНИТЕЛЬНЫЙ АНАЛИЗ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ ДЛЯ РАСПОЗНАВАНИЯ ДИНАМИКИ РУКОПИСНЫХ ПОДПИСЕЙ

Э.С. Анисимова^{1,2}, И.В. Аникин¹

¹ Казанский национальный исследовательский технический университет
им. А.Н. Туполева-КАИ
Российская Федерация, 420111, г. Казань, ул. К. Маркса, д. 10

² Елабужский институт (филиал) федерального государственного автономного
образовательного учреждения высшего образования
"Казанский (Приволжский) федеральный университет"
Российская Федерация, 423604, г. Елабуга, ул. Казанская, д. 89

Аннотация. В работе исследуется применение нейронных сетей для распознавания рукописных подписей с учётом динамики их ввода. Рассматриваются три типа архитектур: полносвязные сети (MLP), сети долгой краткосрочной памяти (LSTM) и свёрточные нейронные сети (CNN). Проведён сравнительный анализ эффективности данных архитектур на наборе данных МСЧТ Signature 100, содержащем подписи 100 авторов и их подделки. В ходе исследования варьировались ключевые гиперпараметры, такие как оптимизатор (RMSProp, Adam, SGD), размер пакета и количество слоёв, для каждой из архитектур. Результаты показывают, что свёрточные нейронные сети (CNN) достигают наилучшей точности, превосходя MLP и LSTM. Предложенный метод имеет практическую ценность для автоматизации верификации подписей и может быть применён в работе систем безопасности и электронного документооборота.

Ключевые слова: распознавание рукописных подписей, нейронные сети, LSTM, CNN, полносвязные сети, динамика подписи, верификация, биометрическая аутентификация, предобработка данных, гиперпараметры.

Введение

В современном цифровом мире аутентификация личности играет ключевую роль в обеспечении безопасности общества. Среди различных биометрических методов верификация по рукописной подписи сохраняет свою актуальность, особенно в юридически значимых документах и финансовых операциях [1]. Однако традиционные методы распознавания подписей [2, 3], основанные на статическом анализе изображения, уязвимы к подделкам высокого качества. Поэтому всё большее внимание сегодня привлекает анализ динамики рукописной подписи, учитывающий такие параметры, как координаты, давление, азимут, наклон пера [4-6]. Этот подход позволяет получить более полную информацию о процессе написания подписи и существенно повысить надёжность систем аутентификации.

Для распознавания рукописных подписей по динамике их ввода традиционно применяются такие методы, как метод опорных векторов [7], скрытые марковские модели [8], вейвлет-преобразования [9], теория нечётких множеств [10-11]. Однако в последние годы нейронные сети продемонстрировали выдающиеся результаты в различных областях, включая компьютерное зрение и анализ временных рядов. Их способность к обучению на больших объёмах данных и извлечению сложных зависимостей делает нейронные сети перспективным инструментом для распознавания динамики рукописных подписей. Несмотря на имеющиеся исследования, потенциал нейросетевого подхода в этой области остаётся не до конца исследованным. В частности, актуальными задачами являются разработка и исследование новых архитектур нейронных сетей, адаптированных к специфике данных о динамике подписи, а также оптимизация параметров обучения для достижения максимальной

точности распознавания и устойчивости к подделкам. В данной статье рассматривается разработка и исследование нейросетевых моделей, включая полносвязную (MLP, Multilayer Perceptron), свёрточную (CNN, Convolutional Neural Network) и LSTM (Long Short-Term Memory) сети, для распознавания рукописных подписей с учётом динамики их ввода. Проводится сравнительный анализ эффективности предложенных моделей и исследуется их чувствительность к различным параметрам.

Материалы и методы

Для решения задачи распознавания рукописных подписей с учётом динамики их ввода был использован набор данных MCYT Signature 100 [12], содержащий рукописные подписи 100 авторов. Разделение набора данных на обучающую и валидационную выборки было проведено в стандартном соотношении 4:1 (4000 и 1000 подписей соответственно), что является общепринятой практикой в задачах машинного обучения и позволяет обеспечить достаточный объём данных для обучения модели и адекватную оценку её качества. Рукописные подписи были введены с использованием графического планшета Wacom INTUOS A6 USB. В процессе ввода рукописных подписей осуществлялось считывание следующих параметров: значений координат (X и Y), давления на перо (Z), азимута (Γ), угла наклона пера (H) в каждой точке подписи. Параметры рабочей области графического планшета 127x97, чувствительность к нажиму – 1024 уровня. Таким образом, диапазоны значений параметров: X : [0; 12700]; Y : [0; 9700]; Z : [0; 1024]; Γ : [0; 3600]; H : [300; 900] (Рисунок 1).

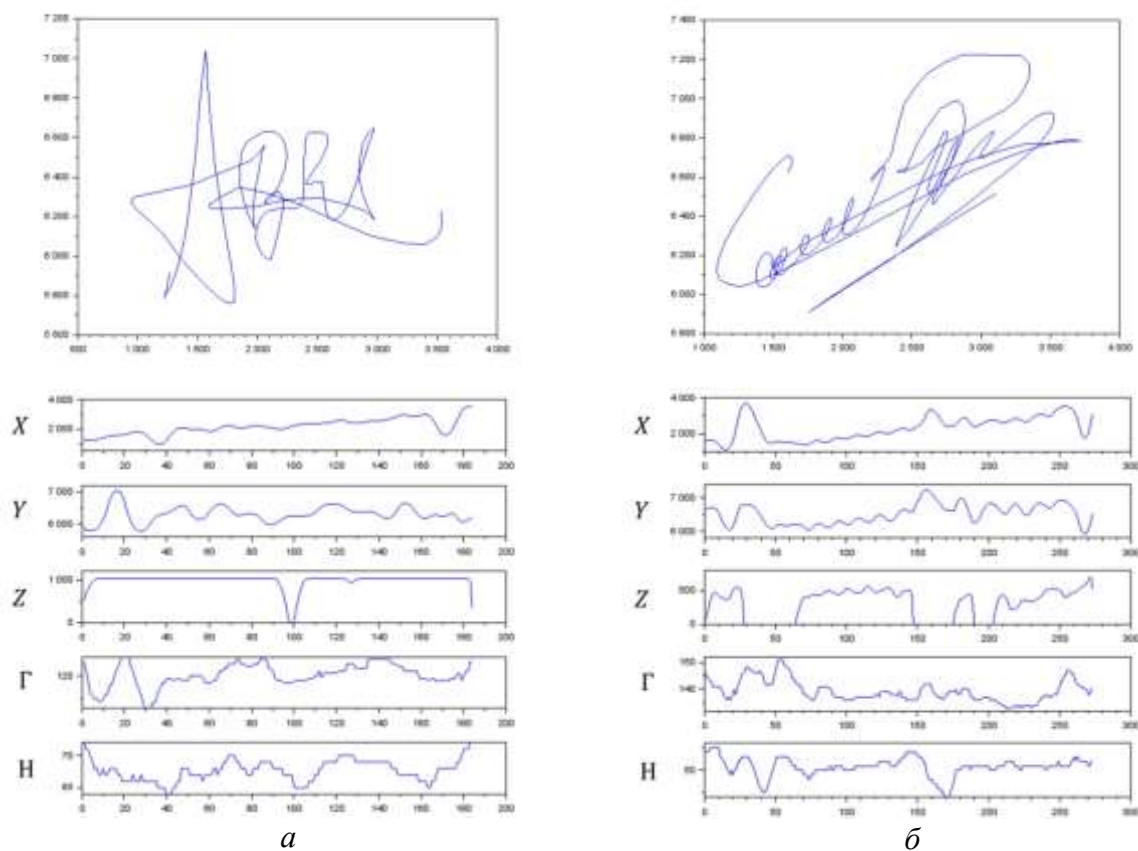


Рис.1. Визуализация рукописных подписей пользователей 0002 (а) и 0036 (б) с параметрами из набора данных MCYT Signature 100

В процессе подготовки данных временные ряды были нормализованы с учётом характеристик графического планшета. Поскольку рукописные подписи содержат разное количество точек, для корректного сравнения признаков и последующей обработки данных была проведена интерполяция, унифицирующая длину последовательностей. Это необходимо, так как алгоритмы машинного обучения, используемые для распознавания, часто требуют входных последовательностей одинаковой длины.

В исследовании рассматривались как подлинные подписи, так и их подделки. Набор поддельных подписей был сформирован путём сбора 25 образцов для каждого из участников. Подлинные подписи каждого автора формировали отдельный класс, а все поддельные подписи были объединены в один класс, что в итоге привело к 101 классу. Обучающая выборка содержала по 20 подлинных подписей каждого автора. Каждая подпись представлена многомерным временным рядом размерности 1000×5 , где 1000 – длина временного ряда (количество отсчётов после проведения интерполяции), а 5 – количество регистрируемых параметров: X, Y, Z, G, H .

Для решения задачи классификации были разработаны и исследованы три типа нейронных сетей: полносвязная, свёрточная и LSTM. Выбор этих архитектур обусловлен их эффективностью в задачах анализа временных рядов. Полносвязные сети хорошо зарекомендовали себя в задачах классификации, LSTM сети способны учитывать долговременные зависимости во временных последовательностях, а свёрточные сети эффективно извлекают локальные признаки.

Экспериментальные исследования

Представим описание структуры и результатов экспериментов с каждой моделью.

1. Полносвязная нейронная сеть (MLP)

Для исследования влияния различных гиперпараметров на точность классификации полносвязной нейронной сети (MLP) была проведена серия экспериментов. В ходе экспериментов варьировались следующие гиперпараметры: оптимизатор (RMSProp, Adam, SGD), размер пакета данных (от 50 до 300) и количество слоёв (2, 3). Результаты экспериментов представлены в таблице 1. В каждом эксперименте фиксировалась архитектура сети, количество нейронов в каждом слое и использовалась функция активации ReLU. Это позволило оценить влияние выбранных гиперпараметров на производительность модели в задаче распознавания динамики рукописных подписей.

Таблица 1. Точность MLP при различных оптимизаторах, размерах пакета и количестве слоёв

Количество слоёв	Количество нейронов	Оптимизатор	Размер пакета	Функция активации	Точность (%)
2	100	RMSProp	50	ReLU	90,6
2	100	RMSProp	100	ReLU	91,8
2	100	RMSProp	150	ReLU	93,3
2	100	RMSProp	200	ReLU	92,8
2	100	RMSProp	250	ReLU	91,9
2	100	RMSProp	300	ReLU	92,1
2	100	Adam	50	ReLU	92,3
2	100	Adam	100	ReLU	84,9

2	100	Adam	150	ReLU	91,7
2	100	Adam	200	ReLU	92,5
2	100	Adam	250	ReLU	92
2	100	Adam	300	ReLU	91,8
2	100	SGD	50	ReLU	88,8
2	100	SGD	100	ReLU	87,3
2	100	SGD	150	ReLU	86
2	100	SGD	200	ReLU	87,1
2	100	SGD	250	ReLU	86,6
2	100	SGD	300	ReLU	85,9
3	100	RMSProp	50	ReLU	89,6
3	100	RMSProp	100	ReLU	91,1
3	100	RMSProp	150	ReLU	87,9
3	100	RMSProp	200	ReLU	93,6
3	100	RMSProp	250	ReLU	92,1
3	100	RMSProp	300	ReLU	92,6
3	100	Adam	50	ReLU	60,7
3	100	Adam	100	ReLU	81,4
3	100	Adam	150	ReLU	93,1
3	100	Adam	200	ReLU	92,7
3	100	Adam	250	ReLU	92,9
3	100	Adam	300	ReLU	92,5
3	100	SGD	50	ReLU	89,7
3	100	SGD	100	ReLU	87,7
3	100	SGD	150	ReLU	86,5
3	100	SGD	200	ReLU	84,7
3	100	SGD	250	ReLU	84,7
3	100	SGD	300	ReLU	85,8

2. LSTM нейронная сеть

Для исследования влияния оптимизатора, размера пакета и количества слоёв на точность классификации LSTM нейронной сети была проведена серия экспериментов. Количество нейронов в каждом слое было зафиксировано на 100, а коэффициент dropout – на 0.2. Варьировались следующие гиперпараметры: оптимизатор (RMSProp, Adam), размер пакета (50, 100, 150, 200, 250, 300) и количество слоёв (2, 3). Результаты экспериментов, демонстрирующие влияние этих параметров на точность, представлены в Таблице 2.

Таблица 2. Точность LSTM-сети при различных оптимизаторах, размерах пакета и количестве слоёв

Количество слоёв	Количество нейронов	Оптимизатор	Размер пакета	Коэффициент dropout	Точность (%)
2	100	RMSProp	50	0,2	88,7
2	100	RMSProp	100	0,2	88,8
2	100	RMSProp	150	0,2	83,5
2	100	RMSProp	200	0,2	88,1
2	100	RMSProp	250	0,2	87,2
2	100	RMSProp	300	0,2	87,2
2	100	Adam	50	0,2	89,2
2	100	Adam	100	0,2	88,8
2	100	Adam	150	0,2	88,1
2	100	Adam	200	0,2	87,8
2	100	Adam	250	0,2	85,8
2	100	Adam	300	0,2	86,5
3	100	RMSProp	50	0,2	87,8
3	100	RMSProp	100	0,2	86,9
3	100	RMSProp	150	0,2	87,3
3	100	RMSProp	200	0,2	85,8
3	100	RMSProp	250	0,2	85,6
3	100	RMSProp	300	0,2	85,1
3	100	Adam	50	0,2	84,6
3	100	Adam	100	0,2	85,8
3	100	Adam	150	0,2	87,8
3	100	Adam	200	0,2	87,2
3	100	Adam	250	0,2	86,7
3	100	Adam	300	0,2	86,2

3. Свёрточная нейронная сеть (CNN)

Для анализа чувствительности CNN к количеству слоёв, оптимизатору и размеру пакета был проведен ряд экспериментов. В Таблице 3 представлены результаты, демонстрирующие влияние следующих гиперпараметров: количество слоёв (2, 3), оптимизатор (RMSProp, SGD, Adam) и размер пакета (от 50 до 300). Функция активации (ReLU), количество фильтров (20), размер ядра свёртки (4x4), шаг свёртки (4) и размер окна пулинга (2x2) были зафиксированы.

Таблица 3. Точность CNN при различных оптимизаторах, размерах пакета и количестве слоёв

Количество слоёв	Оптимизатор	Размер пакета	Функция активации	Количество фильтров	Размер ядра свёртки	Шаг свёртки	Размер окна пулинга	Точность (%)
2	RMSProp	50	ReLU	20	4x4	4	2x2	93
2	RMSProp	100	ReLU	20	4x4	4	2x2	93,5
2	RMSProp	150	ReLU	20	4x4	4	2x2	93,5
2	RMSProp	200	ReLU	20	4x4	4	2x2	93,2
2	RMSProp	250	ReLU	20	4x4	4	2x2	94
2	RMSProp	300	ReLU	20	4x4	4	2x2	93
2	Adam	50	ReLU	20	4x4	4	2x2	91,9
2	Adam	100	ReLU	20	4x4	4	2x2	93,1
2	Adam	150	ReLU	20	4x4	4	2x2	93,2
2	Adam	200	ReLU	20	4x4	4	2x2	91
2	Adam	250	ReLU	20	4x4	4	2x2	93,5
2	Adam	300	ReLU	20	4x4	4	2x2	92,7
2	SGD	50	ReLU	20	4x4	4	2x2	92,7
2	SGD	100	ReLU	20	4x4	4	2x2	92
2	SGD	150	ReLU	20	4x4	4	2x2	92,8
2	SGD	200	ReLU	20	4x4	4	2x2	92,7
2	SGD	250	ReLU	20	4x4	4	2x2	91,7
2	SGD	300	ReLU	20	4x4	4	2x2	91,1
3	RMSProp	50	ReLU	20	4x4	4	2x2	94,3
3	RMSProp	100	ReLU	20	4x4	4	2x2	93,9
3	RMSProp	150	ReLU	20	4x4	4	2x2	93,5
3	RMSProp	200	ReLU	20	4x4	4	2x2	93,3
3	RMSProp	250	ReLU	20	4x4	4	2x2	93,5
3	RMSProp	300	ReLU	20	4x4	4	2x2	92,8
3	Adam	50	ReLU	20	4x4	4	2x2	91,9
3	Adam	100	ReLU	20	4x4	4	2x2	90,4
3	Adam	150	ReLU	20	4x4	4	2x2	91,9
3	Adam	200	ReLU	20	4x4	4	2x2	92,5
3	Adam	250	ReLU	20	4x4	4	2x2	92,5
3	Adam	300	ReLU	20	4x4	4	2x2	93,2
3	SGD	50	ReLU	20	4x4	4	2x2	93,5

3	SGD	100	ReLU	20	4x4	4	2x2	93,4
3	SGD	150	ReLU	20	4x4	4	2x2	93,1
3	SGD	200	ReLU	20	4x4	4	2x2	92,5
3	SGD	250	ReLU	20	4x4	4	2x2	92,4
3	SGD	300	ReLU	20	4x4	4	2x2	91,5

Обсуждение полученных результатов

В данной работе исследовалась задача распознавания рукописных подписей с использованием набора данных MCYT Signature 100, содержащего подписи 100 авторов, а также их подделки. Для решения задачи распознавания были разработаны и исследованы три типа нейронных сетей: полносвязная (MLP), долгой краткосрочной памяти (LSTM) и свёрточная (CNN). Выбор этих архитектур обусловлен их эффективностью в задачах анализа временных рядов, к которым относится и распознавание динамики рукописных подписей.

1. Полносвязная нейронная сеть (MLP). Анализ экспериментов с полносвязной нейронной сетью (MLP) выявил зависимость точности классификации от выбора оптимизатора и размера пакета данных. RMSProp показал наилучшую производительность, достигнув точности 93,6% с размером пакета 200 и трёхслойной архитектурой. Adam и SGD продемонстрировали менее стабильные результаты, кроме того, Adam оказался чувствителен к изменению архитектуры сети: его эффективность заметно снижалась при переходе от двух- к трёхслойной модели. В целом, результаты подчеркивают важность выбора подходящего оптимизатора и размера пакета для достижения оптимальной производительности MLP в задаче распознавания рукописных подписей.

2. LSTM нейронная сеть. Наилучшая точность (89,2%) была достигнута с использованием двухслойной архитектуры, оптимизатора Adam и размера пакета 50. При использовании RMSProp наибольшая точность составила 88,8% с размером пакета 100 (в двухслойной модели). Увеличение количества слоев до трёх в целом привело к снижению точности. Изменение размера пакета показывает незначительные колебания точности, не выявляя чёткой зависимости.

3. Свёрточная нейронная сеть (CNN). Наилучшая точность (94,3%) была достигнута при использовании трёхслойной CNN, оптимизатора RMSProp и размера пакета 50. Для двухслойной CNN максимальная точность составила 94% с оптимизатором RMSProp и размером пакета 250. Оптимизаторы Adam и SGD продемонстрировали более низкую точность по сравнению с RMSProp. Влияние размера пакета на точность было относительно небольшим, однако наблюдались некоторые колебания в зависимости от используемого оптимизатора.

Практическая ценность: Предложенный метод распознавания рукописных подписей с использованием нейронных сетей имеет высокую практическую ценность, поскольку позволяет автоматизировать процесс верификации подписей, что может быть применено в различных областях, таких как банковское дело, электронный документооборот, системы безопасности и другие. Автоматизация процесса верификации подписей повышает эффективность работы, снижает вероятность ошибок и мошенничества, а также уменьшает затраты на ручной труд.

Условия применимости: Для эффективного применения предложенного метода необходимо соблюдение следующих условий:

- Наличие достаточного объёма данных для обучения.
- Предварительная обработка данных.

- Выбор подходящей архитектуры нейронной сети и настройка гиперпараметров.
- Аппаратные ресурсы: Обучение и применение глубоких нейронных сетей может требовать значительных вычислительных ресурсов.

Заключение

Проведенное исследование подтвердило эффективность нейронных сетей для распознавания динамики рукописных подписей. Свёрточная нейронная сеть (CNN) продемонстрировала наилучший результат (94,3%), превосходя полносвязную (MLP) и LSTM нейронную сеть. Для CNN ключевым фактором достижения высокой точности стал выбор оптимизатора и количества слоёв. Влияние оптимизатора и размера пакета также оказалось важным для MLP и LSTM. Дальнейшие исследования могут быть направлены на разработку более сложных архитектур, комбинированных моделей, а также на исследование новых методов предобработки данных и техник расширения набора данных для повышения обобщающей способности моделей и их устойчивости к вариациям в стиле подписи. Это позволит создать более надёжные и точные системы распознавания динамики рукописных подписей.

Список литературы

1. Самотуга А.Е. Распознавание субъектов и их психофизиологических состояний на основе параметров подписи для защиты документооборота / А.Е. Самотуга // Системная инженерия и информационные технологии. – 2023. – Т.2. – С.56-65.
2. Diaz M. Investigating the common authorship of signatures by off-line automatic signature verification without the use of reference signatures / M. Diaz, M.A. Ferrer, S. Ramalingam, R. Guest // IEEE Transactions on Information Forensics and Security. - 2020. Т. 1. – С. 487-499.
3. Mo L.-F., Zhu Y.-L., Mamat H., Ubul K. Off-line handwritten signature recognition based on discrete curvelet transform / L.-F. Mo, Y.-L. Zhu, H. Mamat, K. Ubul // Lecture Notes in Computer Science. – 2019. – P. 424-434.
4. Riesen K. Online signature verification based on string edit distance / K. Riesen, R. Schmidt // International Journal on Document Analysis and Recognition. - 2019. – V.1. – P. 41-54.
5. Sae-Bae N. A Simple and Effective Method for Online Signature Verification / Sae-Bae N., Memon Sae-Bae N. // BIOSIG 2013 – Proceedings of the 12th International Conference of the Biometrics Special Interest Group. - 2013. – P. 147-158.
6. Tolosana R. ICDAR 2021 competition on on-line signature verification / R. Tolosana, R. Vera-Rodriguez, C. Gonzalez-Garcia, Fierrez J. et al. // Lecture Notes in Computer Science. – 2021. – P. 723-737.
7. Elzeki O. Signature recognition based on support vector machine and deep convolutional neural networks for multi-region of interest / O. Elzeki, M. Shams, N. Hikal et al. // Journal of Theoretical and Applied Information Technology. – 2020. - V.98. – P. 3887-3897.
8. Maiorana E., Martinez-Diaz M., Campisi P., Ortega-Garcia J., Neri A. Template Protection for HMM-based On-Line Signature Authentication / E. Maiorana, M. Martinez-Diaz, P. Campisi et al. // CVPR Workshop. – 2008. - P.1-6.
9. Kumari M. Signature Recognition using 2D Discrete Wavelet Transforms International / M. Kumari, A. Kumar, M. Saxena // Journal of Innovative Technology and Exploring Engineering (IJITEE). – 2020. – V. 9(7). – P. 528-532.
10. Anisimova E.S. Fuzzy Sets Theory Approach for Recognition Handwritten Signatures / E.S. Anisimova, I.V. Anikin // Lecture Notes in Electrical Engineering. – 2021. –P. 969-982.
11. Anikin I. Framework for Biometric User Authentication Based on a Dynamic Handwritten Signature / I. Anikin, E. Anisimova // Studies in Systems, Decision and Control. – 2022. – 417. – P. 219-231.

12. Ortega-Garcia J. MCYT Baseline Corpus: a Bimodal Biometric Database / Ortega- J. Garcia, J. Fierrez-Aguilar, D. Simon et al. // IEEE Proceedings - Vision Image and Signal Processing. – 2003. – V.150. – P. 395-401.

COMPARATIVE ANALYSIS OF NEURAL NETWORK MODELS FOR HANDWRITTEN SIGNATURE DYNAMICS RECOGNITION

E.S. Anisimova^{1,2}, I.V. Anikin¹

¹ Kazan National Research Technical University named after A. N. Tupolev-KAI
10, st. Karl Marx, Kazan, 420111, Russian Federation

² Elabuga Institute of Kazan Federal University
89, st. Kazanskaya, Elabuga, 423604, Russian Federation

Annotation. In this work, we investigated the application of neural networks for handwritten signature recognition, taking into account the dynamics of their input. We considered three types of architectures: Multilayer Perceptrons (MLPs), Long Short-Term Memory networks (LSTMs), and Convolutional Neural Networks (CNNs). We conducted a comparative analysis of the effectiveness of these architectures on the MCYT Signature 100 dataset, which contains signatures from 100 authors and their forgeries. Key hyperparameters, such as the optimizer (RMSProp, Adam, SGD), batch size, and number of layers, were varied for each architecture during the study. The results show that Convolutional Neural Networks (CNNs) achieve the best accuracy, outperforming MLPs and LSTMs. The proposed method has practical value for automating signature verification and can be applied in security systems and electronic document management.

Keywords: handwritten signature recognition, neural networks, LSTM, CNN, fully connected networks, signature dynamics, verification, biometric authentication, data preprocessing, hyperparameters.

Статья представлена в редакцию 9 декабря 2024 г.