Комплексная биометрическая аутентификация пользователей информационной системы с применением нейронных сетей

А.С. Исмагилова, Н.Д. Лушников

Уфимский университет науки и технологий, Уфа

Аннотация: Результатом исследования является метод комплексной биометрической аутентификации. Метод реализован в виде программного комплекса, состоящего из подсистемы биометрической аутентификации по изображению лица и подсистемы биометрической аутентификации по голосу. Обучающая выборка, состоящая из сохраненных файлов биометрических образов (изображения лица и аудиозаписи), позволяет уменьшить показатели ошибок первого и второго рода при распознавании пользователей. Предложенный метод биометрической аутентификации предназначен для повышения эффективности процессов распознавания пользователей.
Ключевые слова: аутентификация, биометрия, архитектура нейронной сети, обучающая выборка.

Введение

Процессы идентификации и аутентификации относятся к числу наиболее эффективных инструментов для противодействия несанкционированному доступу. На данный момент к числу актуальных проблем в области защиты информации следует отнести хранение, обработку и передачу биометрических персональных данных. Применение архитектур нейронных сетей позволило достичь высоких показателей эффективности биометрических систем с разными наборами биометрических характеристик. Для достижения такого уровня эффективности были сформированы необходимые обучающие выборки, а также использованы ранее зарекомендовавшие себя лучшие российские и международные программные решения. Так, Бейкер Д., Рабинер Л.Р. и Цзюан Б.Х. в своих трудах проводили исследования по распознаванию речи с применением статистических данных и метода скрытых марковских моделей [1]. Васильевым В.И. были предложены новые методы и алгоритмы машинного обучения в процессах идентификации биометрических систем на базе статических признаков [2]. Шелупановым А.А. и Сабановым А.Г. приведен анализ цифровой идентификации и аутентификации субъектов доступа

применительно к задаче управления доступом к информационным ресурсам, а также предложены критерии доверия к результатам идентификации и аутентификации [3]. В работах Ложникова П.С. представлены методы машинного обучения и новые подходы к применению искусственного интеллекта в процессах биометрической аутентификации [4]. Тодиско М. в своих исследованиях нашел новый извлеченный признак – Q-константный кепстральный коэффициент [5].

Целью данного исследования является повышение точности аутентификации пользователей информационной системы по извлеченным биометрическим признакам. Центральным результатом работы является метод распознавания пользователей с помощью комплексной биометрической аутентификации. На основе предложенного метода разработано программное обеспечение, состоящее из модулей биометрической аутентификации на основе архитектур нейронных сетей на разных наборах биометрических данных. Данный программный комплекс дополнен модулем противодействия методу синтеза изображений лица и аудиозаписей (дипфейк). Все программные модули и скомпилированные модели запущены в качестве исполняемого скрипта на языке программирования Python 3.8 и Python 3.10 с установленными библиотеками (Tensorflow, Keras, Spafe, Librosa, Matplotlib, Pytorch, PyQT5 и TKinter).

Система биометрической аутентификации пользователей

При запуске программного обеспечения необходимо пройти первый этап биометрической аутентификации – распознавание личности по изображению лица в режиме онлайн. При проведении данного этапа распознавания личности необходимо извлечь основные биометрические признаки, к которым относятся локальные бинарные шаблоны и гистограммы направленных градиентов [6].

Значения локальных бинарных шаблонов возможно вычислить следующим образом:

$$LBP_{Y,R} = \sum_{p=0}^{Y-1} s(g_p - g_c)2^p$$

где $s(x)$ – пороговая функция, g_p – значение интенсивности –ого пикселя, g_c – значение интенсивности центрального пикселя, p – номер пикселя, Y – окрестность с пикселями, R – радиус, $p = 0, \dots, Y-1$.

В каждой точке изображения приближенное значение величины градиента можно вычислить путем использования полученных приближенных значений производных:

$$G = \sqrt{G_x^2 + G_y^2}$$

где G – величина градиента, G_x и G_y – координаты изображения.

Для учета изменений освещения и контрастности силы градиента должны быть локально нормализованы, что требует группировки ячеек вместе в более крупные, пространственно связанные блоки. Дескриптор гистограммы направленных градиентов представляет собой конкатенированный вектор компонентов нормализованных гистограмм ячеек из всех областей блока.

В случае успешного прохождения первого этапа необходимо перейти к распознаванию личности по голосу. Для достижения поставленной цели исследования значения биометрических признаков по голосу можно вычислить следующим образом:

$$H(z) = \frac{G}{1 - \sum_{k=1}^{p} \alpha_k z^{-k}}$$

где $H(z)$ – передаточная функция линейной системы, z – значение передаточной функции, G – коэффициент усиления возбуждения, α_k –

коэффициент линейного предсказания, k – индекс частоты, p – порядок предыдущих значений линейного предсказания.

Данное программное обеспечение разработано по принципу Zero Trust («Нулевое доверие»), концепция которого заключается в отсутствии доверенных или проверенных пользователей. Полагаясь на соответствующую политику минимального доступа к ресурсам системы, авторы разработали программный модуль шифрования биометрических персональных данных [6]. Поэтому в данном релизе отсутствует возможность подмены файлов биометрических персональных данных (фотографии лица, аудиозаписи).

Описание структуры нейронных сетей биометрической аутентификации пользователей

Для повышения точности распознавания личности и для повышения качества обработки биометрических персональных данных были применены и разработаны архитектуры нейронных сетей для обучения на датасетах. При создании нейронных сетей на начальном этапе были сформированы датасеты изображений и аудиозаписей на основе категориальной кросс-энтропии. В рамках данного исследования используется категориальная кросс-энтропия, так как общее количество классов равно трем (два класса авторизованных пользователей, класс неавторизованных пользователей). Для формирования датасета были созданы папки с тренировочной (train), валидационной (val) и тестовой (test) обучающими выборками. Каждая из этих папок, в свою очередь, состоит из папок с наименованием классов (User1, User2, Unknown). Подсчет общего количества нейронов в скрытом слое обусловлен ранее полученными результатами многослойных нейронных сетей, обучаемых с помощью алгоритма обратного распространения. Таким образом, общее количество коэффициентов, представляющих собой веса синаптических связей нейронных сетей, соответствует минимальному числу нейронов в

скрытом слое в соответствии с нижней границей сложности нейронных сетей [7, 8].

Объем обучающей выборки для каждой архитектуры нейронной сети по акустическим признакам составляет 450 аудиозаписей двух пользователей. Длительность каждой аудиозаписи составляет 8, 15 и 25 секунд соответственно. Аудиозапись длительностью 8 секунд представлена в виде файла с записанным голосом, в котором производится счет от одного до пяти. В аудиозаписи длительностью 15 секунд производится счет от одного до десяти, а в аудиозаписи длительностью 25 секунд – от одного до двадцати.

Далее в скрипте компилируются модели с функцией оптимизации Adam и количеством указанных эпох обучения. В рамках поставленной задачи исследования выбранное количество эпох является приемлемым для получения необходимого уровня точности биометрической аутентификации и качества проводимого обучения нейронных сетей. Далее необходимо сгенерировать папки обучающих выборок с помощью функции datagen. После этого происходит процесс генерации скомпилированной модели на основе сгенерированных папок обучающих выборок. В завершении обучения архитектур нейронных сетей выводится результат точности работы (accuracy) и функции потерь (loss). Также в рамках проведения апробации исследования была скомпилирована и протестирована модель архитектуры нейронной сети Wav2vec с использованием таких датасетов аудиозаписей, как TIMIT и VoxCeleb [9,10]. Архитектуры нейронных сетей с выводом авторизованных пользователей (АП) и неавторизованных пользователей (НАП) представлены на рис. 1.

Показатели обучения нейронных сетей

Для оценки эффективности извлечения дескрипторов и характеристик изображений лица пользователей на рис. 2 приведены показатели ошибок

первого и второго рода на разных наборах сформированной обучающей выборки, состоящей из 100, 200 и 600 изображений.

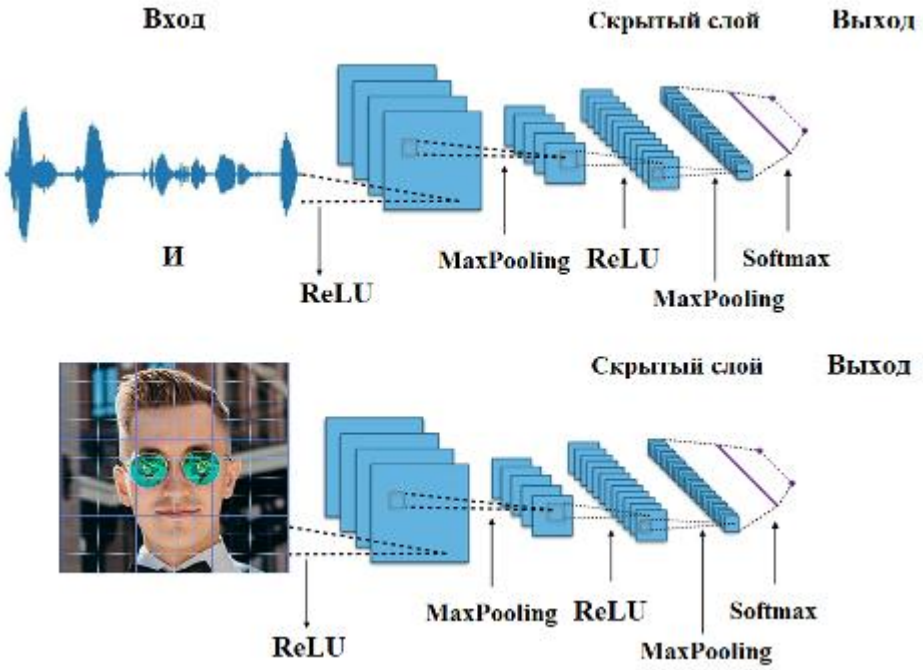

Рис. 1. – Нейронные сети комплексной биометрической аутентификации пользователей

Ошибка первого рода вычисляется по следующей формуле:

$$FRR = \frac{a}{N},$$

где FRR – значение ошибки первого рода, a – показатели некорректного распознавания авторизованных пользователей системы, N – общее количество проведенных экспериментов.

Ошибка второго рода вычисляется по формуле:

$$FAR = \frac{b}{N},$$

где FAR – значение ошибки второго рода, b – показатели некорректного распознавания неавторизованных пользователей системы, N – общее количество проведенных экспериментов.

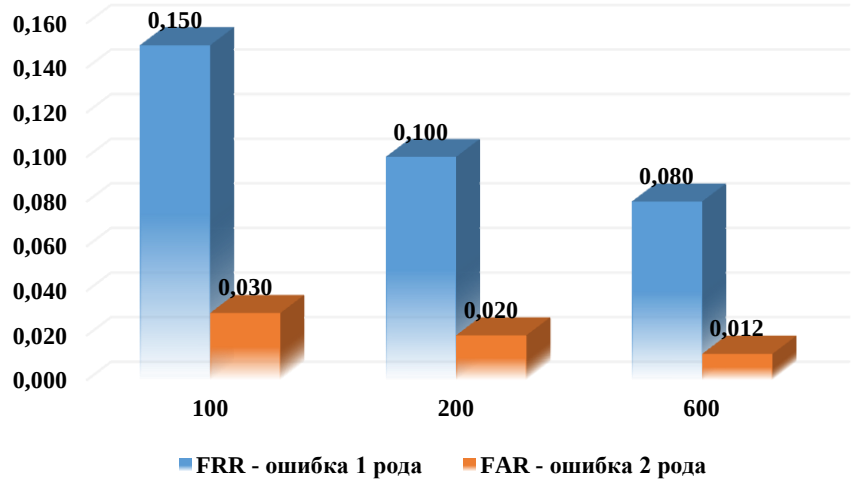

Рис. 2. – Ошибки первого рода и ошибки второго рода при распознавании пользователей по изображению лица на разных наборах обучающей выборки

При оценке точности и качества биометрической аутентификации пользователей по извлеченным акустическим признакам применяются показатели ошибок первого и второго рода. Как известно, также рекомендуется учесть показатели ошибок разделения дикторов, средней чистоты кластеров (содержание в аудиозаписи только речи диктора) и средней чистоты дикторов (содержание в аудиозаписи речи пользователей по отдельности), которые указаны в таблице № 1. Данные в таблице были подсчитаны в соответствии с формулой, по которой вычисляется ошибка разделения дикторов:

$$E_{spkr} = \frac{\sum_{seg}(T(seg)\times \min(N_{ref}(seg), N_{sys}(seg)) - N_{correct}(seg))}{\sum_{seg} T(seg)\times N_{ref}(seg)},$$

где E_{spkr} – ошибка разделения дикторов, $T(seg)$ – длительность речевого сегмента seg, $N_{ref}(seg)$ – количество дикторов (эталонная разметка), $N_{sys}(seg)$ – количество дикторов (оцениваемая система), $N_{correct}(seg)$ – количество верно отнесенных дикторов.

Также в таблице приведены значения средней чистоты кластеров:

$$ACP_c = \frac{\sum_{s=1}^{S} n_{sc}^2}{(N_c^{cluster})^2},$$

$$ACP = \frac{1}{N}\sum_{c=1}^{M} ACP_c \times N_c^{cluster},$$

И значения средней чистоты дикторов:

$$ASP_s = \frac{\sum_{c=1}^{M} n_{sc}^2}{(N_c^{cluster})^2},$$

$$ASP = \frac{1}{N}\sum_{s=1}^{S} ASP_s \times N_s^{speaker},$$

где ACP – средняя чистота кластеров, ASP – средняя чистота дикторов, S – количество дикторов (эталонная разметка), M – полученное количество кластеров, n_{sc} – количество данных в кластере c, которые принадлежат диктору s, $N_c^{cluster} = \sum_{s=1}^{S} n_{sc}$ – количество данных в кластере c, $N_s^{speaker} = \sum_{c=1}^{M} n_{sc}$ – количество данных, принадлежащих диктору s, $N = \sum_{s=1}^{S}\sum_{c=1}^{M} n_{sc}$ – количество всех данных.

В таблице в качестве итоговой совокупной оценки системы разделения пользователей используется среднее геометрическое значений ASP и ACP, обозначается, как K:

$$K = \sqrt{ACP \times ASP}$$

Таблица № 1

Показатели ошибок разделения дикторов и оценка системы разделения дикторов

Набор обучающей выборки	MFCC		LPC		PLP		CQCC		SCF	
	E_{spkr} (%)	K	E_{spkr} (%)	K	E_{spkr} (%)	K	E_{spkr} (%)	K	E_{spkr} (%)	K
DataSet 100	8,26	0,862	7,60	0,873	9,42	0,845	7,52	0,793	8,34	0,853
DataSet 300	8,01	0,82	7,99	0,832	8,87	0,839	7,49	0,769	8,1	0,821
DataSet 450	7,67	0,79	7,8	0,801	8,34	0,821	7,21	0,742	7,99	0,8

Результаты

Таким образом, в данном исследовании сформирован репрезентативный набор обучающей выборки. Рассмотрена задача

комплексирования подсистем биометрической аутентификации пользователей. Проведено тестирование системы биометрической аутентификации пользователей на основе извлеченных биометрических признаков. В ходе анализа системы было выявлено, что достигнут баланс между вероятностями ошибки первого рода и ошибки второго рода. Точность работы (accuracy) системы биометрической аутентификации составляет 91,82%; показатели потерь (loss) равны 8,18% на протяжении 100 эпох.

Литература

1. Lawrence R. Rabiner, Ronald W. Schafer. Theory and Applications of Digital Speech Processing. Prentice Hall, 2010. 1056 p.

2. Васильев В.И., Жумажанова С.С., Ложников П.С. и др. Оценка идентификационных возможностей биометрических признаков от стандартного периферийного оборудования // Вопросы защиты информации. 2016. №1. С. 12-20.

3. Сабанов А.Г., Шелупанов А.А. Идентификация и аутентификация в цифровом мире. М.: Научно-техническое издательство «Горячая линия-Телеком», 2022. 356 с.

4. Ложников П.С., Сулавко А.Е., Еременко А.В. и др. Экспериментальная оценка надежности верификации подписи сетями квадратичных форм, нечеткими экстракторами и персептронами // Информационно-управляющие системы. 2016. №5. С. 73-85.

5. Todisco M., Delgado H., Evans N. A new feature for automatic speaker verification anti-spoofing: Constant Q cepstral coefficients // Odyssey 2016: The Speaker and Language Recognition Workshop. Bilbao: ISCA SIG, 2016. pp. 283–290.

6. Валеев С.С., Кондратьева Н.В., Гузаиров М.Б. и др. Иерархическая динамическая система управления информационной безопасностью информационной системы предприятия // Инженерный вестник Дона, 2023, №11 URL: ivdon.ru/ru/magazine/archive/n11y2023/8802.

7. Гузаиров М.Б., Исмагилова А.С., Лушников Н.Д. Аутентификация пользователей информационной системы по изображению лица // Моделирование, оптимизация и информационные технологии. 2023. №4. URL: moitvivt.ru/ru/journal/pdf?id=1465.

8. Хайкин С. Нейронные сети: полный курс. М.: И.Д. Вильямс, 2006. 1104 с.

9. Марьев А.А. Метод интерпретации результатов измерений параметров речевого сигнала в задачах диагностики психоэмоционального состояния человека по его речи // Инженерный вестник Дона, 2011, №4. URL: ivdon.ru/ru/magazine/archive/n4y2011/538.

10. Орлова Ю.А., Дмитриев А.С., Колчева Д.В. Адаптация модели распознавания речи Google Cloud Speech для упрощения редактирования исходного кода программ для ЭВМ с мобильных устройств // Инженерный вестник Дона, 2021, №2 URL: ivdon.ru/ru/magazine/archive/n2y2021/6822.

References

1. Lawrence R. Rabiner, Ronald W. Schafer. Theory and Applications of Digital Speech Processing. Prentice Hall, 2010. 1056 p.

2. Vasil'ev V.I., Zhumazhanova S.S., Lozhnikov P.S., Sulavko A.E. Voprosy zashhity informacii. 2016. №1. pp. 12-20.

3. Sabanov A.G., Shelupanov A.A. Identifikacija i autentifikacija v cifrovom mire [Identification and authentication in the digital world]. M.: Nauchno-tekhnicheskoye izdatel'stvo «Goryachaya liniya-Telekom», 2022. 356 p.

4. Lozhnikov P.S., Sulavko A.E., Yeremenko A.V. i dr. Informacionno–upravljajushhie sistemy. 2016. №5. pp. 73-85.

5. Todisco M., Delgado H., Evans N. Odyssey 2016: The Speaker and Language Recognition Workshop. Bilbao: ISCA SIG, 2016. pp. 283–290.

6. Valeyev S. S., Kondrat′yeva N. V., Guzairov M. B. i dr. Inzhenernyj vestnik Dona, 2023, №11 URL: ivdon.ru/ru/magazine/archive/n11y2023/8802.

7. Guzairov M.B., Ismagilova A.S., Lushnikov N.D. Modelirovaniye, optimizatsiya i informatsionnyye tekhnologii. 2023. №4 URL: moitvivt.ru/ru/journal/pdf?id=1465

8. Khaykin S. Neyronnyye seti: polnyy kurs [Neural Networks: a comprehensive foundation]. M.: I.D. Vil′yams, 2006. 1104 p.

9. Mar'ev A.A. Inzhenernyj vestnik Dona, 2011, №4 URL: ivdon.ru/ru/magazine/archive/n4y2011/538.

10. Orlova Ju.A., Dmitriev A.S., Kolcheva D.V. Inzhenernyj vestnik Dona, 2021, №2. URL: ivdon.ru/ru/magazine/archive/n2y2021/6822.

Дата поступления: 7.12.2023

Дата публикации: 18.01.2024