

Торайғыров университетінің
ҒЫЛЫМИ ЖУРНАЛ

НАУЧНЫЙ ЖУРНАЛ
Торайғыров университета

ТОРАЙҒЫРОВ УНИВЕРСИТЕТІНІҢ ХАБАРШЫСЫ

Физика, математика және компьютерлік
ғылымдар сериясы
1997 жылдан бастап шығады



ВЕСТНИК ТОРАЙҒЫРОВ УНИВЕРСИТЕТА

Серия: Физика, математика
и компьютерные науки
Издается с 1997 года

ISSN 2959-068X

№ 2 (2024)
Павлодар

**НАУЧНЫЙ ЖУРНАЛ
ТОРАЙГЫРОВ УНИВЕРСИТЕТА**

Серия: Физика, математика и компьютерные науки
выходит 4 раза в год

СВИДЕТЕЛЬСТВО

о постановке на переучет периодического печатного издания,
информационного агентства и сетевого издания
№ KZ91VPY00046988

выдано

Министерством информации и общественного развития
Республики Казахстан

Тематическая направленность

публикация материалов в области физики, математики,
механики и информатики

Подписной индекс – 76208

<https://doi.org/10.48081/PALS7230>

Бас редакторы – главный редактор

Тлеукинов С. К., *д.ф-м.н., профессор*

Заместитель главного редактора Испулов Н. А., *к.ф-м.н., профессор*

Ответственный секретарь Жумабеков А. Ж., *PhD доктор*

Редакция алкасы – Редакционная коллегия

Esref Adali,	<i>PhD доктор, профессор (Турция);</i>
Abdul Qadir Rahimoon,	<i>PhD доктор, профессор (Пакистан);</i>
Донбаев К. М.,	<i>д.ф-м.н., профессор;</i>
Демкин В. П.,	<i>д.ф-м.н., профессор (Российская Федерация);</i>
Жумадиллаева А. К.,	<i>к.т.н., профессор;</i>
Ибраев Н. Х.,	<i>д.ф-м.н., профессор;</i>
Косов В. Н.,	<i>д.ф-м.н., профессор;</i>
Сенцова С. М.,	<i>д.пед.н., профессор;</i>
Шоканов А. К.,	<i>д.ф-м.н., профессор</i>
Омарова А. Р.,	<i>технический редактор</i>

За достоверность материалов и рекламы ответственность несут авторы и рекламодатели

Редакция оставляет за собой право на отклонение материалов

При использовании материалов журнала ссылка на «Вестник Торайгыров
университета» обязательна

© Торайгыров университет

***М. К. Бейсенов, Н. К. Токжигитова**

Торайгыров университет, Республика Казахстан, г. Павлодар

*e-mail: meirbek-dev@pm.me

АРХИТЕКТУРЫ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РАСПОЗНАВАНИЯ ЛИЦ

Распознавание лиц является важной технологией, которая имеет множество применений, включая мониторинг посещаемости. В последние годы нейронные сети стали доминирующим подходом к распознаванию лиц, предлагая высокую точность и гибкость. Однако существует множество различных архитектур нейронных сетей для распознавания лиц.

В этой статье исследуются компромиссы между тремя популярными архитектурами нейронных сетей для распознавания лиц в контексте систем мониторинга посещаемости. Исследование углубляется в сильные и слабые стороны сверточных нейронных сетей, FaceNet и сиамских сетей, анализируя их эффективность в решении таких факторов, как вычислительные ресурсы и доступность данных. Анализируя их возможности и ограничения, статья призвана осветить оптимальный архитектурный выбор для конкретных приложений, будь то приоритет высокой точности, эффективности использования ресурсов или устойчивости к изменениям в позе и освещении.

Работа изучает сложную среду архитектур распознавания лиц, предлагая тонкое понимание их пригодности для мониторинга посещаемости на основе различных ограничений и желаемых результатов и направлена на то, чтобы предоставить технологическим архитекторам информацию, необходимую для принятия обоснованных решений при выборе архитектуры, которая лучше всего соответствует их конкретным потребностям.

Ключевые слова: распознавание лиц, мониторинг посещаемости, архитектуры нейронных сетей, сверточные нейронные сети, сиамские сети, искусственный интеллект.

Введение

Распознавание лиц – это технология, созданная Вудро Уилсоном (Вудди) Бледсоу в 1966 году, которая позволяет сопоставлять человеческие лица с помощью цифровых изображений или видеозаписей через базу данных лиц. Распознавание лиц стало идеей, позволяющей компьютерам быстро и точно находить и распознавать человеческие лица [1]. Для повышения производительности и точности распознавания лиц было разработано множество алгоритмов, со своими слабыми и сильными сторонами, и соответствующими предпочтительными сценариями использования [2]. Выбор архитектуры нейронной сети имеет первостепенное значение при разработке системы распознавания лиц, поскольку он глубоко влияет на производительность и общую эффективность системы.

Актуальность темы обусловлена недостатком исчерпывающих ответов на вопросы, связанные с выбором оптимальной архитектуры нейронных сетей в контексте распознавания лиц. Несмотря на продвинутые технологии в области компьютерного зрения, существующие системы часто сталкиваются с низкой скоростью обработки, неудовлетворительной точностью и чувствительностью к изменениям условий освещения.

Практическая значимость статьи заключается в том, что она может быть использована для разработки более эффективных систем распознавания лиц. Результаты исследования могут быть использованы в таких областях, как биометрическая аутентификация, контроль доступа, видеонаблюдение и т.д.

Материалы и методы

Теоретический анализ и обзор литературы по распознаванию лиц с использованием нейронных сетей; анализ преимуществ и недостатков различных архитектур нейронных сетей для распознавания лиц; сбор и подготовка набора данных; сбор изображений лиц из различных источников; обучение нейронных сетей; оценка точности нейронных сетей.

Результаты и обсуждение

Свёрточные нейронные сети (CNN) представляют собой разумный выбор при разработке систем распознавания лиц, основанный на их присущей способности извлекать и различать сложные закономерности в изображениях. Эта сверточная архитектура, вдохновленная организацией зрительной коры в биологических системах, продемонстрировала высокую эффективность в задачах обработки изображений, благодаря способности справляться с пространственными иерархиями и абстракцией функций, что делает свёрточные нейронные сети подходящими для распознавания лиц – области, требующей распознавания сложных черт лица и их пространственных отношений [3]. Принцип работы CNN предполагает применение сверточных

слоев для захвата локальных шаблонов и иерархических представлений, что позволяет сети автономно изучать отличительные признаки.

Преимущества использования CNN в распознавании лиц выходят за рамки извлечения признаков. Трансферное обучение, метод, при котором предварительно обученные модели на обширных наборах данных адаптируются для конкретных задач, может быть использован для смягчения проблемы нехватки данных, часто возникающей в приложениях распознавания лиц. Этот подход позволяет модели наследовать знания из более широких наборов данных изображений, расширяя ее возможности обобщения. Кроме того, включение слоев объединения помогает в пространственной понижающей дискретизации, уменьшая вычислительную нагрузку при сохранении основных функций. Методы пакетной нормализации и исключения способствуют регуляризации, предотвращая переобучение и повышая надежность модели. Эти соображения в совокупности подчеркивают актуальность CNN в контексте распознавания лиц.

Тем не менее, к присущим сверточным нейронным сетям ограничениям и потенциальным недостаткам относятся:

1 **Вычислительная интенсивность.** Архитектурная сложность CNN, требует значительных вычислительных ресурсов во время обучения и вывода, а следовательно, ограничивающим фактором в средах с ограниченными ресурсами, потенциально препятствуя развертыванию в реальном времени.

2 **Зависимость от данных.** CNN процветают благодаря большим объемам размеченных данных для эффективного обучения. В сценариях распознавания лиц получение разнообразного и репрезентативного набора данных может оказаться сложной задачей. Ограниченные данные могут привести к переобучению, когда модель не может обобщить незнакомые лица и привести низкой точности.

3 **Устойчивость к состязательным атакам.** CNN продемонстрировали уязвимость к состязательным атакам, при этом небольшие искажения входных изображений могут привести к неправильной классификации. В контексте распознавания лиц такая уязвимость создает проблемы безопасности, поскольку злоумышленники могут использовать эти уязвимости для обмана системы [4].

4 **Специфика домена.** CNN проявляют ограничения при столкновении с изменением освещения, позы или окклюзии.

Архитектура FaceNet, разработанная для распознавания лиц с использованием сетей тройных потерь, представляет собой новаторский подход в области систем распознавания лиц. FaceNet, разработанный исследователями Google, использует глубокие нейронные сети и функцию триплетной потери для отображения черт лица в многомерном пространстве,

облегчая точную проверку и идентификацию лица [5]. Использование триплетной потери гарантирует, что сеть изучает вложения так, что расстояние между вложениями одного и того же человека сводится к минимуму, а расстояние между вложениями разных людей максимизируется [6]. FaceNet обычно использует архитектуру глубоких нейронных сетей, часто основанную на исходной архитектуре, для захвата сложных черт лица в различных масштабах. Архитектура состоит из нескольких сверточных слоев, слоев пула и полностью связанных слоев. Примечательно, что сеть выводит представления (векторы признаков), которые представляют уникальные характеристики каждого лица. Во время обучения FaceNet предоставляется тройка изображений лиц, состоящая из основного изображения, положительного изображения и отрицательного изображения. Модель оптимизирована для минимизации потерь триплетов в наборе данных. Этот процесс заставляет сеть изучать отличительные особенности для точного распознавания лиц.

Плюсы архитектуры FaceNet:

1 Устойчивость к изменениям: формулировка тройной потери позволяет сети быть устойчивой к изменениям в освещении, позе и выражениях, поскольку она фокусируется на функциях обучения, которые инвариантны в пределах одной и той же личности и различаются у разных личностей [7].

2 Обучение работе с небольшими наборами данных: сети триплетных потерь могут быть эффективны даже с относительно небольшими наборами размеченных данных. Использование триплетов помогает эффективно использовать информацию в обучающих выборках.

Минусы архитектуры FaceNet:

1 Чувствительность гиперпараметра: эффективность Triplet Loss Networks зависит от тщательной настройки гиперпараметров, включая запас в функции потерь [8].

2 Дисбаланс данных: дисбаланс в распределении положительных и отрицательных пар в наборе обучающих данных может повлиять на производительность сети.

3 Необходимость в качественных аннотациях: обучение сети тройных потерь требует точных и надежных аннотаций для положительных и отрицательных пар. Получение высококачественных размеченных данных может быть сложной и ресурсоемкой задачей.

Сиамская архитектура нейронных сетей, известна своей эффективностью в задачах, связанных с оценкой сходства или несходства, является привлекательным решением сложных проблем, связанных с распознаванием лиц, благодаря своей тандемной структуре, в которой две идентичные подсети имеют общий вес и архитектуру [9]. Эти подсети одновременно

обрабатывают разные входные данные, что позволяет извлекать отличительные признаки. В контексте распознавания лиц эта архитектура превосходно распознает тонкие нюансы черт лица, необходимые для точной идентификации.

Суть сиамской сети заключается в ее способности обучаться надежному пространству встраивания, в котором похожие лица расположены проксимально, а несходные лица четко разделены. Это достигается за счет формулировки функции потерь, которая способствует минимизации расстояний между вложениями идентичных граней при одновременной максимизации расстояний для несовпадающих пар. Следовательно, сеть развивает врожденную способность различать черты лица с высокой степенью точности.

Способность сиамской архитектуры эффективно работать с ограниченными размеченными данными [10], учитывая проблемы, связанные с получением обширных наборов данных для приложений распознавания лиц. Более того, сиамская сеть превосходна в сценариях, где количество положительных образцов значительно меньше, чем отрицательных образцов.

Процесс обучения включает в себя представление в сети пар изображений лиц вместе с соответствующими метками, обозначающими, совпадают ли лица или нет. Посредством итеративной оптимизации сеть уточняет свои параметры, чтобы создать различительное пространство для встраивания, способствующее точному распознаванию лиц. После обучения система может эффективно сравнивать и идентифицировать лица на основе изученных вложений.

Архитектура сиамских нейронных сетей представляет заметные преимущества в контексте систем распознавания лиц, но имеет недостатки охватывающие различные аспекты сиамской архитектуры:

1. Сложность проектирования и реализации: внедрение и тонкая настройка сиамской сети требует детального понимания архитектуры нейронных сетей и методологий обучения, что приводит к высокой сложности разработки системы.

2. Ограниченное обобщение для вариаций поз: сиамские сети могут столкнуться с проблемами, когда сталкиваются со значительными вариациями поз в изображениях лиц. Хотя сети можно научить обрабатывать незначительные изменения позы, экстремальные изменения могут повлиять на их способность точно идентифицировать совпадающие лица.

Выводы

В статье рассмотрены преимущества и недостатки трех архитектур нейронных сетей, являющихся популярными решениями для разработки систем мониторинга посещаемости с использованием распознавания лиц.

На основании анализа преимуществ и недостатков различных архитектур нейронных сетей для распознавания лиц можно сделать следующие выводы:

CNN представляют собой эффективный подход к распознаванию лиц, но им требуется значительная вычислительная мощность и большие наборы данных для достижения высокой точности.

Архитектура FaceNet демонстрирует высокую точность распознавания лиц в различных условиях, но она требует тщательной настройки гиперпараметров и качественных аннотаций.

Сиамские нейронные сети являются эффективным решением для распознавания лиц с ограниченными наборами данных, но они могут столкнуться с проблемами при распознавании лиц с большими вариациями поз.

REFERENCES

1 **Khawla Alhanaee, Mitha Alhammadi, Nahla Almenhali, Maad Shatnawi.** Face Recognition Smart Attendance System using Deep Transfer Learning // *Procedia Computer Science*. – Volume 192. – 2021. – P. 4093–4102, ISSN 1877-0509. – <https://doi.org/10.1016/j.procs.2021.09.184>.

2 **Andre Budiman, Fabian, Ricky Aryatama Yaputera, Said Achmad, Aditya Kurniawan.** Student attendance with face recognition (LBPH or CNN): Systematic literature review // *Procedia Computer Science*. – Volume 216. – 2023. – P. 31–38. – ISSN 1877-0509. – <https://doi.org/10.1016/j.procs.2022.12.108>.

3 **Surya Teja Chavali, Charan Tej Kandavalli, T. M. Sugash, R Subramani.** Smart Facial Emotion Recognition with Gender and Age Factor Estimation // *Procedia Computer Science*. – Volume 218. – 2023. – P. 113–123. – ISSN 1877-0509. – <https://doi.org/10.1016/j.procs.2022.12.407>.

4 **Deepti Saraswat, Pronaya Bhattacharya, Tirth Shah, Rushi Satani, Sudeep Tanwar.** Anti-spoofing-enabled Contactless Attendance Monitoring System in the COVID-19 Pandemic // *Procedia Computer Science*. – Volume 218. – 2023. – P. 1506–1515. – ISSN 1877-0509. – <https://doi.org/10.1016/j.procs.2023.01.129>.

5 **Anwar Shah, Bahar Ali, Masood Habib, Jaroslav Frnda, Inam Ullah, Muhammad Shahid Anwar.** An ensemble face recognition mechanism based on three-way decisions // *Journal of King Saud University – Computer and Information Sciences*. – Volume 35. – Issue 4. – 2023. – P. 196–208. – ISSN 1319-1578. – <https://doi.org/10.1016/j.jksuci.2023.03.016>.

6 **Joseph A. Mensah, Justice K. Appati, Elijah K. A. Boateng, Eric Ocran, Louis Asiedu.** FaceNet recognition algorithm subject to multiple constraints:

Assessment of the performance // Scientific African. – Volume 23. – 2024. – e02007. – ISSN 2468-2276. – <https://doi.org/10.1016/j.sciaf.2023.e02007>.

7 **Lingling Yang, Xingshi Xu, Jizheng Zhao, Huaibo Song.** Fusion of RetinaFace and improved FaceNet for individual cow identification in natural scenes // Information Processing in Agriculture. – 2023. – ISSN 2214-3173. – <https://doi.org/10.1016/j.inpa.2023.09.001>.

8 **Joseph A. Mensah, Justice K. Appati, Elijah K.A Boateng, Eric Ocran, Louis Asiedu.** FaceNet recognition algorithm subject to multiple constraints: Assessment of the performance // Scientific African. – Volume 23. – 2024. – e02007. – ISSN 2468-2276. – <https://doi.org/10.1016/j.sciaf.2023.e02007>.

9 **I. Melekhov, J. Kannala and E. Rahtu,** «Siamese network features for image matching», 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico. – 2016. – P. 378–383, <https://doi.org/10.1109/ICPR.2016.7899663>.

10 **John Atanbori, Samuel Rose.** MergedNET: A simple approach for one-shot learning in siamese networks based on similarity layers // Neurocomputing, Volume 509. – 2022. – P. 1–10. – ISSN 0925-2312. – <https://doi.org/10.1016/j.neucom.2022.08.070>.

Поступило в редакцию 01.02.24.

Поступило с исправлениями 01.02.24.

Принято в печать 20.02.24.

***М. К. Бейсенов, Н. Қ. Токжигитова**

Торайгыров университет, Қазақстан Республикасы, Павлодар қ.

01.02.24 ж. баспаға түсті.

01.02.24 ж. түзетулерімен түсті.

20.02.24 ж. басып шығаруға қабылданды.

БЕТТІ ТАНУҒА АРНАЛҒАН НЕЙРОНДЫҚ ЖЕЛІ АРХИТЕКТУРАСЫ

Бетті тану-бұл көптеген қолданыстары бар маңызды технология, соның ішінде сабаққа қатысуды бақылау. Соңғы жылдары нейрондық желілер жоғары дәлдік пен икемділікті ұсына отырып, бетті танудың басым тәсіліне айналды. Дегенмен, бетті тануға арналған көптеген әртүрлі нейрондық желі архитектуралары бар. Бұл мақалада трафикті бақылау жүйелері контекстінде бетті тануға арналған үш танымал нейрондық желі архитектурасы арасындағыссалар қарастырылады. Зерттеу конвульсиялық нейрондық желілердің, FaceNet және сиамдық

жселілердің күшті және әлсіз жақтарын зерттеп, олардың есептеу ресурстары мен деректердің қолжетімділігі сияқты факторларды шешудегі тиімділігін талдайды. Олардың мүмкіндіктері мен шектеулерін талдай отырып, мақала жоғары дәлдікке, ресурстарды пайдалану тиімділігіне немесе поза мен жарықтың өзгеруіне төзімділікке басымдық беру болсын, Нақты қолданбалар үшін оңтайлы архитектуралық таңдауды көрсетуге арналған. Жұмыс бет-әлпетті тану архитектурасының күрделі ортасын зерттейді, олардың әртүрлі шектеулер мен қажетті нәтижелерге негізделген сабаққа қатысуды бақылауға жарамдылығын Мұқият түсінуді ұсынады және технологиялық сәулетшілерге олардың нақты қажеттіліктеріне сәйкес келетін архитектураны таңдауда негізделген шешімдер қабылдау үшін қажетті ақпаратты беруге бағытталған.

Кілтімі сөздер: бетті тану, трафикті бақылау, нейрондық желі архитектурасы, конволюциялық нейрондық желілер, сиамдық желілер, жасанды интеллект.

***M. K. Beisenov, N. K. Tokzhigitova**

Toraighyrov University, Republic of Kazakhstan, Pavlodar.

Received 01.02.24.

Received in revised form 01.02.24.

Accepted for publication 20.02.24.

NEURAL NETWORK ARCHITECTURES FOR FACE RECOGNITION

Facial recognition is an important technology that has many applications, including attendance monitoring. In recent years, neural networks have become the dominant approach to facial recognition, offering high accuracy and flexibility. However, there are many different neural network architectures for face recognition.

This paper examines the trade-offs between three popular neural network architectures for facial recognition in the context of attendance monitoring systems. The study delves into the strengths and weaknesses of convolutional neural networks, FaceNet and Siamese networks, analyzing their effectiveness in addressing factors such as computational resources, data availability. By analyzing their capabilities and limitations, the article aims to highlight optimal architectural choices for specific

applications, whether prioritizing high precision, resource efficiency, or resilience to changes in pose and lighting.

The work examines the complex landscape of facial recognition architectures, offering a nuanced understanding of their suitability for attendance monitoring based on various constraints and desired outcomes, and aims to provide technology architects with the information needed to make informed decisions when selecting the architecture that best suits their specific needs.

Keywords: face recognition, attendance monitoring, neural network architectures, convolutional neural networks, Siamese networks, artificial intelligence.

Теруге 03.06.2024 ж. жіберілді. Басуға 28.06.2024 ж. қол қойылды.

Электрондық баспа

7,50 Mb RAM

Шартты баспа табағы 10,01. Таралымы 300 дана. Бағасы келісім бойынша.

Компьютерде беттеген: Е. Е. Калихан

Корректор: А. Р. Омарова

Тапсырыс № 4247

Сдано в набор 03.06.2024 г. Подписано в печать 28.06.2024 г.

Электронное издание

7,50 Mb RAM

Усл.печ.л. 10,01. Тираж 300 экз. Цена договорная.

Компьютерная верстка Е. Е. Калихан

Корректор: А. Р. Омарова

Заказ № 4247

«Toraighyrov University» баспасынан басылып шығарылған

«Торайгыров университеті» КЕ АҚ

140008, Павлодар қ., Ломов к., 64, 137 каб.

«Toraighyrov University» баспасы

«Торайгыров университеті» КЕ АҚ

140008, Павлодар қ., Ломов к., 64, 137 каб.

+7(718)267-36-69

e-mail: kereku@tou.edu.kz

www.vestnik.tou.edu.kz

<https://vestnik-pm.tou.edu.kz/>