

УДК 004.946.2 + 004.932.72 + 004.272

Сравнительное исследование инструментов детекции поз на мобильных платформах

Ярко Егор Игоревич^(*)

yarkoeg@gmail.com

SPIN-код: 7619-5077

ИТИС КФУ, Казань, Россия

Аннотация. Определение позы человека (Human Pose Estimation, HPE) является одной из ключевых задач в области компьютерного зрения, находящей широкое применение в видеонаблюдении, медицине и анализе спортивной активности. С начала 2000-х годов интерес к этой теме значительно возрос, что привело к появлению большого количества специализированных библиотек. Особенно заметный прогресс наблюдается с 2017 года — с развитием скелетно-ориентированных алгоритмов, реализованных в удобных программных решениях для исследовательских и прикладных целей. Проведен сравнительный анализ четырех современных библиотек для HPE, способных эффективно функционировать на мобильных устройствах: Lightweight OpenPose, PoseNet, MoveNet и BlazePose. Рассмотрены их преимущества и недостатки с точки зрения точности, производительности и удобства интеграции в реальные приложения.

Ключевые слова: поза человека, Human Pose Estimation, HPE, детектирование позы, компьютерное зрение, мобильные устройства, дополненная реальность, Lightweight OpenPose, PoseNet, MoveNet, BlazePose, скелетная модель

Comparative study of pose detection tools on mobile platforms

Yarko Yegor Igorevich^(*)

yarkoeg@gmail.com

SPIN-code: 7619-5077

Kazan Federal University, Kazan, Russia

Abstract. Human Pose Estimation (HPE) stands as one of the key challenges in the field of computer vision, finding broad application in surveillance, medicine, and sports activity analysis. Since the early 2000s, interest in this area has grown significantly, leading to the emergence of numerous specialized libraries. Particularly notable progress has been observed since 2017, marked by the development of skeleton-based algorithms implemented in user-friendly software solutions for both research and practical purposes. This study presents a comparative analysis of four modern HPE libraries optimized for efficient performance on mobile devices: Lightweight OpenPose, PoseNet, MoveNet, and BlazePose. Their strengths and weaknesses are examined in terms of accuracy, performance, and ease of integration into real-world applications.

Keywords: human posture, Human Pose Estimation (HPE), pose detection, computer vision, mobile devices, augmented reality, Lightweight OpenPose, PoseNet, MoveNet, BlazePose, skeletal model

Введение. Оценка позы человека (Human Pose Estimation, HPE) — одна из ключевых задач компьютерного зрения, направленная на определение координат суставов и других ключевых точек человеческого тела. Она востребована в разнообразных прикладных областях: от реабилитации и мониторинга физической активности до видеонаблюдения и дополненной реальности. В условиях ограниченных ресурсов мобильных устройств задача становится особенно сложной — необходимо обеспечить баланс между точностью, скоростью и вычислительной эффективностью. В последние годы появилось множество HPE-библиотек, ориентированных на мобильные и встраиваемые платформы. Цель данной работы — сравнительно проанализировать четыре популярные библиотеки: PoseNet, MoveNet, Lightweight OpenPose и BlazePose, выявив их сильные и слабые стороны в мобильном контексте.

Материалы и методы. Анализ ограничен 2D-оценкой позы одного человека на изображениях или видео без использования глубинных сенсоров. Рассматривались следующие параметры: нейросетевая архитектура, число распознаваемых ключевых точек, стратегия локализации (top-down или bottom-up), вычислительная эффективность и применимость на мобильных устройствах. В качестве тестовых наборов данных использовались COCO и Penn Action — одни из наиболее распространенных в HPE-исследованиях. Качество распознавания оценивалось с помощью метрики PDJ (Percentage of Detected Joints), вычисляемой на основе Евклидова расстояния между аннотированными и предсказанными точками [1].

Результаты. PoseNet представляет собой одну из первых попыток адаптировать HPE под браузерные и мобильные среды. Он использует MobileNet или ResNet и позволяет отслеживать до 17 ключевых точек. Прост в использовании, но по точности уступает более современным решениям [2]. MoveNet — разработка Google, представленная в двух вариантах: Lightning (быстрая, менее точная) и Thunder (более точная, но ресурсоемкая). Он обеспечивает высокую скорость работы и точность, хорошо подходя для мобильных приложений [3]. Lightweight OpenPose адаптирует классический OpenPose под слабые устройства, используя MobileNet вместо VGG-19. Она сохраняет bottom-up подход, но уступает в детализации [4]. BlazePose — наиболее функциональное решение, выделяющее до 33 ключевых точек, включая лицо, кисти и стопы. Модель использует двухэтапную архитектуру и тесно интегрирована с фреймворком MediaPipe [5]. Анализ библиотек для оценки позы человека представлен в таблице.

Обсуждение полученных результатов. PoseNet выгоден в образовательных и браузерных проектах благодаря легкости интеграции и документации, но не подходит для задач, требующих высокой точности [2]. MoveNet предлагает хорошее соотношение точность/производительность и прост в развертывании благодаря тесной связи с TensorFlow и поддержке TF Lite [3]. Однако он использует ограниченный набор ключевых точек, чего может быть недостаточно, например, для медицинских упражнений или спортивного анализа. Lightweight OpenPose ориентирован на компромисс между точностью и

быстродействием, но требует дополнительных усилий при интеграции в мобильные приложения [4]. BlazePose уверенно лидирует по количеству отслеживаемых точек и общему качеству предсказаний. Высокая производительность достигается благодаря оптимизированной архитектуре и интеграции с MediaPipe, что делает его особенно актуальным для задач телереабилитации, спортивной аналитики и фитнеса [5].

Характеристики библиотек оценки позы человека

Название библиотеки	Год выпуска	Макс. число ключевых точек	Расположение ключевых точек в частях тела	Метод определения	Нейросетевая архитектура
PoseNet 2.0	2022	17	Глаза, уши, плечи, локти, запястья, бедра, колени, лодыжки	Bottom-up	ResNet
MoveNet	2021	17	Глаза, уши, плечи, локти, запястья, бедра, колени, лодыжки	Top-down	MobileNetV2
Lightweight OpenPose	2018	18	Включает шею, нос и дополнительные суставы рук	Bottom-up	MobileNet
BlazePose	2020	33	Полное тело с пальцами рук и ног	Top-down	BlazePose CNN

Заключение. Проведенный анализ показал, что каждая библиотека НРЕ имеет свои преимущества: PoseNet — прост и доступен, MoveNet — стабилен и быстр, Lightweight OpenPose — универсален, а BlazePose — наиболее точен и информативен. В контексте мобильных устройств наилучший результат показывает BlazePose, объединяя детальную локализацию, высокую точность и адаптацию под слабое оборудование. В дальнейшем планируется проведение экспериментального тестирования на мобильных устройствах с использованием PDJ как основной метрики качества.

Список источников

- [1] Chung J.-L., Ong L.-Y., Leow M.-C. Comparative Analysis of Skeleton-Based Human Pose Estimation. Future Internet, 2022, vol. 14, no. 12, pp. 1–19. <https://doi.org/10.3390/fi14120380>
- [2] Papandreou G., Zhu T., Kanazawa N. et al. Towards accurate multi-person pose estimation in the wild. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4903–4911. <https://doi.org/10.1109/CVPR.2017.520>

- [3] MoveNet: Ultra fast and accurate pose detection model. TensorFlow Blog, 2021. URL: <https://blog.tensorflow.org/2021/05/movenet-and-posture-analysis.html> (дата обращения 03.04.2025).
- [4] Cao Z., Simon T., Wei S.-E., Sheikh Y. Realtime multi-person 2D pose estimation using part affinity fields. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 7291–7299. <https://doi.org/10.1109/CVPR.2017.143>
- [5] Bazarevsky V., Zhang F. BlazePose: On-device real-time body pose tracking. URL: <https://ai.googleblog.com/2020/08/on-device-real-time-body-pose-tracking.html> (accessed May 15, 2025).