

**УДК 159.944:159.947:796.01:004.93'1**

**КУМСКОВ Михаил Иванович**

доктор физико-математических наук, профессор, кафедра «Вычислительная математика»,  
Московский государственный университет им. М.В. Ломоносова, г. Москва, Россия  
e-mail: kumskov@math.msu.ru

**МАМУРОВ Баходир Бахшуллоевич**

доктор педагогических наук (DSc), профессор, Ректор Бухарского государственного  
педагогического института, г. Бухара, Узбекистан  
e-mail: b.b.mamurov@mail.ru

На основе рецензии: Маманазаров А.Б., доктор педагогических наук (DSc), доцент,  
заведующий кафедрой «Общегуманитарные и социально-экономические дисциплины»,  
Филиал МГУ им. М.В. Ломоносова в г. Ташкенте

**МЕТОДОЛОГИЯ ИСПОЛЬЗОВАНИЯ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ОЦЕНКИ  
КАЧЕСТВА ОБУЧЕНИЯ В СПОРТЕ**

**Аннотация**

В статье предложена методология анализа спортивных движений с использованием нейронных сетей, объединяющая этапы выделения позы, построения цифрового скелета движения и сравнения с эталонными траекториями. Подход основан на использовании современных архитектур нейросетей (HRNet, OpenPose, ViTPose, PoseFormer, KASportsFormer) и профильных спортивных датасетов для обеспечения высокой точности распознавания поз спортсмена и интерпретации динамических фаз выполнения спортивных упражнений. Разработана методология, ориентированная на формирование и применение требований к информационной системе с перспективой ее использования в спортивном образовании и тренерской практике.

**Ключевые слова:** нейронные сети, распознавание позы, спортивная аналитика, оценка качества, датасеты, методология, компьютерное зрение, биомеханика.

**SPORTDA TA'LIM SIFATINI BAHOLASH UCHUN NEYRON TARMOQLARDAN FOYDALANISH USLUBIYATI**

**Annotatsiya**

Maqolada neyron tarmoqlardan foydalanish asosida sport harakatlarini tahlil qilish uslubiyati taklif etilgan bo'lib, u pozani ajratib olish, harakatning raqamli skeletini qurish va etalon trayektoriyalar bilan taqqoslash bosqichlarini birlashtiradi. Yondashuv zamonaviy neyron tarmoq arxitekturalari (HRNet, OpenPose, ViTPose, PoseFormer, KASportsFormer) hamda ixtisoslashtirilgan sport datasetlaridan foydalanishga asoslanib, sportchining pozasini aniqlash va sport mashqlarini bajarishning dinamik fazalarini talqin qilishda yuqori aniqlikni ta'minlaydi. Axborot tizimiga qo'yiladigan talablarni shakllantirish va qo'llashga yo'naltirilgan uslubiyat ishlab chiqilgan bo'lib, undan sport ta'limi va murabbiylik amaliyotida foydalanish istiqboli mavjud.

**Kalit so'zlar:** neyron tarmoqlar, pozani aniqlash, sport analitikasi, sifatni baholash, datasetlar, metodologiya, kompyuter ko'rishi, biomexanika.

**METHODOLOGY FOR THE USE OF NEURAL NETWORKS TO ASSESS THE QUALITY OF TRAINING IN SPORTS**

**Abstract**

The article proposes a methodology for analyzing sports movements using neural networks, integrating the stages of pose estimation, construction of a digital movement skeleton, and comparison with reference trajectories. The approach is based on the use of modern neural network architectures (HRNet, OpenPose, ViTPose, PoseFormer, KASportsFormer) and specialized sports datasets to ensure high accuracy in recognizing an athlete's pose and interpreting the dynamic phases of performing sports exercises. A methodology has been developed focused on the formulation and application of requirements for an information system with the prospect of its use in sports education and coaching practice.

**Key words:** neural networks, pose estimation, sports analytics, quality assessment, datasets, methodology, computer vision, biomechanics.

В последние годы исследование спортивных движений вошло в новую фазу: стало ясно, что методы компьютерного зрения могут не просто фиксировать позу спортсмена, но и анализировать его технику. Этот переход от «картинки» к «анализу» оказался возможным благодаря нескольким крупным линиям развития нейросетевых моделей. Анализ показывает, что именно архитектуры OpenPose [4], HRNet [7], ViTPose [13] и трансформерные решения для 3D-моделирования [12] задали основу современного анализа движений.

Научная основа данной работы была представлена авторами в ряде научных публикаций, которые определили дальнейший вектор отечественных исследований в области автоматизированной спортивной аналитики. Первый доклад описывал практическую схему выделения ключевых точек спортсмена из реального изображения, фактически формируя этап «построения цифрового скелета». Второй доклад предложил переход от визуального распознавания позы спортсмена к оценке качества выполнения спортивного упражнения. В этих работах была зафиксирована двухэтапная концепция анализа данных: сначала построить точную модель позы, затем использовать её для анализа и сравнения траекторий [1,2].

Современные исследования продолжают развивать эту линию. Модели, учитывающие анатомические зависимости (например, KASportsFormer [16]), физические ограничения (MultiPhys [11]) или биомеханические свойства движения (BioPose [8]), позволяют подходить к спорту не как к последовательности кадров, а как к системе взаимосвязанных фаз. Это особенно важно в обучении: именно фазы, а не статичные положения, показывают, где возникает ошибка. В спортивной практике это означает переход от субъективной визуальной оценки к объективной реконструкции движения – с возможностью фиксировать отклонения от эталонной траектории. При построении нейронных систем широко используются обширные датасеты. COCO Keypoints [6], MPII, Human3.6M, SportsPose [5], AthletePose3D [14], а также новые многокамерные коллекции типа Hi4D и TeamTrack [15], которые дают возможность анализировать движение в разных контекстах: от бытовой позы до сложных спортивных ситуаций. Рассмотрены метрики и форматы представления данных, которые определяют уровень совместимости проводимых исследований. SMPL-модели [3], выраженные представления тела и нейроархитектуры MixSTE [7] создают основу, на которой можно строить сопоставимые, проверяемые и воспроизводимые выводы. С учётом этих инструментов методология исследования качества спортивных движений превращается в логичную систему.

Таким образом, рассматриваются несколько ключевых точек: технологическая эволюция моделей позы, биомеханический контекст анализа техники и необходимость формирования целостной методологии. Именно эта логика становится базой для дальнейших разделов: сравнительного анализа архитектур [8, 10, 14], обзора датасетов [4, 5, 6, 8, 14, 15], выбора форматов и метрик [3, 7], постро-

ения методологической схемы и определения перспектив практической реализации данной работы [9, 11, 13].

В работе предложена методика использования современных нейронных сетей для анализа и оценки качества выполнения упражнений в спорте. Решены следующие задачи: проведен сравнительный анализ нейросетевых архитектур анализа человеческой позы, сделан анализ известных датасетов для обучения нейросетей, описана методика решения поставленной задачи и сформулированы перспективы практической реализации.

Развитие методов распознавания позы за последние десять лет прошло путь от первых CNN-регрессий к сложным трансформерным моделям, которые работают не только с изображением, но и с временной структурой движения. В спортивной аналитике это особенно важно: техника упражнения всегда зависит от последовательности фаз, а не от отдельного кадра. Поэтому при выборе архитектур исследование опирается на работы, которые отражают разные поколения подходов – от DeepPose [10] до современных ViTPose [13] и KASportsFormer [16].

Одним из первых успешных методов стала модель DeepPose, предложенная Toshev и Szegedy [10]. Она использовала CNN-регрессию и фактически задавала направление всей области. Следующий значимый шаг – архитектура OpenPose [4], где авторы ввели Part Affinity Fields, позволившие одновременно выделять позы нескольких человек. Это стало критично для многолюдных спортивных сцен.

Параллельно с этим появились маркерлесс-системы, такие как DeepLabCut, продемонстрировавшие, что точное выделение ключевых точек возможно без использования специализированной спортивной разметки. Этот подход стал важным аргументом в пользу применения нейросетевых моделей в реальных тренировочных условиях, где невозможно использовать маркеры или лабораторные системы.

Существенный рост точности произошёл с появлением HRNet [9]. Модель использует параллельные потоки разного разрешения и сохраняет детальность признаков на всех этапах. В спортивных видео, где важна точность локализации суставов, HRNet обеспечивает более стабильный результат. ViTPose [13] и ViTPose++ [12] продолжают эту линию, предлагая Vision Transformer как основу. Эти модели лучше обобщают данные и работают устойчиво даже на сложных ракурсах.

Для задач 3D-восстановления и анализа последовательностей большое значение имеют трансформерные решения. PoseFormer [12] использует спatio-темпоральный трансформер, который связывает кадры в единую динамическую структуру. Для спортивных движений это особенно ценно: ошибка в конкретной фазе может быть незаметна на статике, но критична в динамике.

Отдельного внимания заслуживают архитектуры, объединяющие биомеханику и искусственный интеллект. KASportsFormer [16] вводит анатомические ограничения, что особенно важно при анали-

зе коротких спортивных сцен. Модели MultiPhys [11] и BioPose [8] учитывают физические зависимости и строят точные 3D-реконструкции, согласованные с реальной кинематикой тела.

Сравнительный анализ показывает, что для построения информационных систем анализа спортивных упражнений лучше всего подходят: HRNet / ViTPose – для высокоточной 2D-локализации каркаса тела; PoseFormer – для анализа последовательностей и 3D-реконструкции; MultiPhys, BioPose и KASportsFormer – для задач, где важны биомеханические ограничения. Эти архитектуры формируют технологическую основу архитектуры ИС и позволяют перейти от простой детекции позы к полноценной оценке качества выполнения упражнения, где важны точность суставных координат, согласованность фаз и реалистичность движения.

Развитие архитектур прошло несколько этапов: от ранних CNN-моделей [10] к устойчивым bottom-up подходам [4], затем к высокоразрешённым нейросетям и современным трансформерным решениям [9,12, 16]. Для задач спортивной аналитики особенно важны модели, которые не просто фиксируют позу, но способны учитывать временную структуру движения и биомеханические ограничения. Именно поэтому нейросети HRNet и ViTPose являются оптимальными для точной 2D-локализации, PoseFormer обеспечивает более корректную работу с динамическими последовательностями [12], а архитектуры MultiPhys и BioPose [8, 11] формируют основу для анализа движения с учётом физических свойств тела.

В совокупности рассмотренные архитектуры задают технологическую основу – они позволяют переходить от простой детекции позы к полноценной оценке качества выполнения упражнения, где важны точность координат элементов тела спортсмена, согласованность фаз движения и биомеханическая реалистичность. Такой набор инструментов обеспечивает возможность построения устойчивой системы анализа техники спортсмена и служит фундаментом для последующих этапов формирования требований к ИС.

Развитие методологии анализа спортивных движений напрямую зависит от качества и разнообразия наборов данных для обучения нейросетей – датасетов. В отличие от задач компьютерного зрения спортивная аналитика требует таких данных, которые максимально приближены к реальным двигательным шаблонам выполнения спортивных упражнений. Именно поэтому можно рассматривать как классические базы данных, задающие общие стандарты ключевых точек, так и специализированные подборки, ориентированные на движение спортсменов.

К числу наиболее значимых датасетов относят COCO Keypoints [6] и MPII Human Pose [4]. Они стали основой для обучения практически всех современных моделей. COCO, благодаря большому числу изображений и 17 ключевым точкам, обеспечивает хорошее базовое качество локализации. MPII, в свою очередь, содержит более естественные бытовые и двигательные сцены, что помога-

ет моделям лучше обобщать данные. Несмотря на свою «общую» природу, оба набора до сих пор остаются фундаментальными, потому что формируют единый стандарт аннотаций.

Для задач спортивной аналитики важное значение имеют специализированные датасеты. Один из наиболее востребованных – SportsPose [5], где представлены динамические 3D-сцены. Он создавался как инструмент для анализа реальных спортивных движений, что делает его ценным при обучении моделей, ориентированных на фазы упражнения или высокие амплитуды. AthletePose3D [14] идёт ещё дальше: он содержит 1,3 миллиона кадров с 3D-аннотациями и кинематическими характеристиками. Такой уровень детализации позволяет моделям не только выделять позу, но и учитывать биомеханику движения, что важно для корректной оценки спортивной техники.

В контексте 3D-реконструкции особое место занимает датасет Human3.6M. Он представляет собой одну из самых структурированных и точных баз данных, где каждое движение зафиксировано множеством камер и снабжено точными трёхмерными координатами суставов. Несмотря на лабораторный характер набора, он остаётся эталоном: все крупные архитектуры 3D-позы оценивают свои результаты именно на Human3.6M.

Для анализа взаимодействий в командных видах спорта важны датасеты TeamTrack и Hi4D [15], включающие многокамерные сцены и сложную динамику нескольких спортсменов. Они полезны в ситуациях, когда требуется не только оценка индивидуальной техники, но и понимание группового движения.

Таким образом, классические наборы (COCO, MPII) обеспечивают стабильный фундамент, спортивные датасеты (SportsPose, AthletePose3D) дают возможность работать с реальными двигательными паттернами, а лабораторные базы (Human3.6M) позволяют достигать высокой точности в трёхмерных координатах. В совокупности эти ресурсы позволяют формировать надёжную и воспроизводимую систему анализа движений, на которой строятся все последующие этапы проекта по созданию ИС.

В совокупности рассмотренные датасеты создают базовый фундамент для построения устойчивой архитектуры ИС. Универсальные наборы (COCO, MPII) формируют стандарт ключевых точек и обеспечивают базовую сопоставимость моделей [4, 6]. Лабораторные ресурсы, такие как Human3.6M, задают точность для 3D-реконструкции и служат эталоном при сравнении архитектур. Специализированные спортивные базы (SportsPose [5], AthletePose3D [14]) позволяют моделям адаптироваться к реальным двигательным паттернам, включая фазы прыжков, изменения углов и вариативность техники. Многокамерные датасеты, такие как Hi4D и TeamTrack [15], обеспечивают возможность анализа взаимодействий спортсменов и усложнённых сцен. Это позволяет формулировать требования в ИС углублённой оценки качества движения в спортивных задачах. Правильная ком-

бинация этих ресурсов позволяет получать стабильные и воспроизводимые результаты, что критично для последующих этапов анализа.

При переходе от выделения ключевых точек к оценке качества движения важным становится вопрос формата представления позы спортсмена и выбора метрик, способных измерять различия между эталонными и фактическими траекториями движения тела. В спортивной аналитике формат данных определяет не только точность вычислений, но и интерпретируемость результата: тренер должен понимать, что именно делает спортсмен и почему модель движений фиксирует те или иные отклонения от «стандарта».

Одним из самых распространённых форматов представления позы человека является формат Skeleton, где каждая поза представлена набором суставов и связей между ними. Этот формат удобен тем, что он подходит для большинства алгоритмов сравнения поз. Однако, когда требуется более детальная реконструкция движения, используется модель SMPL [3], позволяющая описывать тело как трёхмерную поверхность с параметрами формы и позы. Её расширенная версия SMPL-X [7] учитывает дополнительно положение лица и кистей рук, что важно для видов спорта, где критична точность мелких движений.

Метрики оценки качества движения в спортивных задачах играют не меньшую роль, чем сами модели. Одной из таких ключевых метрик является MPJPE (Mean Per Joint Position Error), измеряющая среднюю ошибку в координатах суставов. Она подходит для ситуаций, когда важна точность пространственного положения тела. Но для спортивных упражнений, где критична динамика, MPJPE не всегда достаточна.

Также используется метрика DTW (Dynamic Time Warping), которая учитывает временные сдвиги. Метрика показывает, насколько совпадают фазы движения, даже если спортсмен выполняет

упражнение чуть быстрее или медленнее эталона. Такая особенность полезна, например, при анализе прыжков и акробатических упражнений, где важно не только итоговое положение тела спортсмена, но и траектория выполнения движений.

В задачах анализа угловых характеристик применяются метрики, основанные на сравнении направлений суставов и векторов сегментов тела. В этом контексте широко используется cosine similarity, которая позволяет учесть различия в ориентации тела даже при минимальных изменениях в позициях суставов. Иногда это критично, потому что отклонение на несколько градусов способно существенно изменить биомеханику движения.

Таким образом, выбор формата и метрик определяет, насколько глубоко и корректно модель сможет анализировать движение. Формат Skeleton обеспечивает лёгкость обработки, SMPL и SMPL-X [3, 7] – точную 3D-реконструкцию, MPJPE – пространственную точность, DTW – временную согласованность, а угловые метрики – биомеханическую интерпретацию. Совместное использование этих инструментов позволяет строить комплексную систему анализа техники, адаптированную к задачам спортивного обучения и исследования качества движений.

Формирование единой схемы анализа информации в будущей ИС становится важным шагом в переходе от разрозненных технических решений к единой системе анализа спортивных движений. Несмотря на разнообразие архитектур и наборов данных, именно структура процесса определяет, насколько технология сможет работать в реальном тренировочном цикле, а не только в лабораторных условиях [9, 11, 13]. Поэтому можно предположить последовательную цепочку этапов, где каждый шаг опирается на предыдущий и подготавливает данные для следующего (схема).



**Схема.** Показатели информационных потоков при анализе качества выполнения спортивных упражнений.

На первом этапе происходит сбор данных, который задаёт качество всей последующей обработки. Для спортивных задач важны не просто видеозаписи, а материалы с чётко фиксированным ракурсом, заданной частотой кадров и возможностью последующей синхронизации. Всё это позволяет нейросетям корректно выделять ключевые точки и избегать ошибок, связанных с размытием, перекрытием сегментов тела или быстрыми фазами движения. Выбор оборудования и формата записи определяет точность работы модели на следующем этапе.

Второй этап – выделение позы (Pose Estimation). Здесь используются архитектуры HRNet, OpenPose, ViTPose, PoseFormer и другие модели. Каждая из них по-своему преобразует видеокادر в структурированное описание – набор суставов и связей, который можно рассматривать как цифровой «скелет» движения. В спортивных задачах это описание особенно важно: отклонения углов, смещения траекторий или нарушенные последовательности фаз могут влиять на технику анализа качества выполнения упражнения. Точность выделения позы формирует основу для последующего анализа, а комбинация моделей, таких как HRNet и ViTPose, обеспечивает устойчивость к ракурсу и вариативности.

Третий этап – сравнение поз и траекторий, где происходит переход от визуальной формы движения к его аналитическому содержанию. Здесь применяются как классические метрики MPJPE и DTW, так и угловые показатели и модели SMPL/SMPL-X. На этом этапе можно выявить временные сдвиги, фазовые различия, нарушения траектории, что делает анализ не просто точечным, а динамическим. Для спортивных движений это важно: различия часто проявляются не в конкретной позе, а в переходе от одной фазы движения к другой.

Четвёртый этап – оценка качества выполнения упражнения, где аналитические данные переводятся в интерпретируемую тренером форму. Здесь объединяются вычислительная точность, биомеханические закономерности и результаты спортивного наблюдения. Современные модели, такие как AdaFuse и MixSTE, позволяют синтезировать данные с разных камер, сглаживать ошибки и формировать более устойчивое представление движения. На этом этапе цифровая аналитика превращается в инструмент тренера: модель выявляет ошибки, фиксирует прогресс, выделяет ключевые зоны, требующие корректировки, и делает процесс обучения более прозрачным. Важно подчеркнуть, что такая методологическая схема продолжает подход сопоставления эталонных и обучающихся траекторий как основы объективной оценки качества движения [2], а также концепцию накопления

данных и отслеживания динамики техники в процессе обучения [1]. Совмещение этих принципов с современными нейросетевыми моделями формирует единую систему, в которой вычислительная точность будет соединена с образовательной ценностью.

В итоге схема «сбор данных → выделение позы → сравнение траекторий → оценка качества» создаёт базовый фундамент для практического внедрения цифрового анализа движений. Она позволяет формировать объективные метрики, отслеживать прогресс спортсмена, выявлять ошибки и строить персонализированные траектории обучения, что делает процесс тренировочных занятий более эффективным и научно обоснованным.

Перспективы внедрения разработанной методологии связаны не столько с дальнейшим усложнением архитектур, сколько с возможностью построить на их основе реальные системы анализа техники, способные работать в учебных и тренировочных средах. Современные модели уже демонстрируют достаточную точность, чтобы фиксировать качество движения, и теперь ключевым становится вопрос интеграции – как объединить алгоритмы, датасеты и педагогические задачи в единый инструмент.

Наиболее очевидным направлением является возможное создание автоматизированных систем оценки спортивной техники, которые потенциально могут работать в режиме реального времени при наличии достаточных вычислительных ресурсов. Такие системы смогли бы фиксировать движения спортсмена, выделять фазовые ошибки, сравнивать траекторию с эталонной и формировать отчёты для тренеров. Подходы MultiPhys и BioPose открывают возможность учитывать физические закономерности движения, что приближает цифровую оценку к реальным биомеханическим параметрам.

Другим направлением становится интеграция методологии в образовательный процесс. Возможность фиксировать динамику прогресса через сравнение цифровых поз и накопление данных даёт тренеру инструмент, которого раньше не существовало. Система может показывать, какие элементы техники усвоены, где сохраняются ошибки, как изменяется траектория по мере учебно-тренировочных занятий. Такой метод переводит их в технически реализуемую форму [1, 2].

Не менее перспективным является использование гибридных моделей, таких как Physical Inertial Poser (PIP), которые объединяют видеоданные с инерциальными датчиками. Это позволяет работать даже в условиях недостаточного освещения, сложных ракурсов или ограниченного пространства. Подходы DeepMimic и другие методы, ос-

нованные на имитационном обучении, открывают возможность для создания симуляторов, в которых можно изучать технику, прогнозировать ошибки и предлагать варианты корректировки движения. Кроме того, методология может применяться в командных видах спорта, где важно не только индивидуальное движение, но и взаимодействие нескольких спортсменов. Датасеты Ni4D и TeamTrack задают основу для анализа групповой динамики, что позволяет расширять систему от индивидуальных техник к командным паттернам.

Наконец, перспективы связаны и с созданием единой базы данных движений спортсменов, в которой хранятся эталонные траектории, учебные примеры, комментарии тренеров и собранные метрики. Такая база станет инструментом не только для обучения моделей, но и для стандартизации техники и проведения сравнительных исследований между группами и поколениями спортсменов.

В совокупности эти направления демонстри-

руют, что разработанная методология может превратиться из теоретической схемы в полноценный инструмент, объединяющий нейросетевые модели, биомеханику и педагогическую практику. Это открывает возможности для повышения качества спортивного обучения, создания цифровых тренеров и формирования новых исследований в области анализа движений.

**Разработанная методология** может служить фундаментом для создания автоматизированных прототипов информационных систем оценки спортивной техники. Внедрение этих систем в тренировочный и образовательный процессы позволит тренерам и спортсменам получать объективную обратную связь, отслеживать динамику прогресса и строить персонализированные траектории обучения. **Таким образом**, методология объединяет вычислительную точность искусственного интеллекта с требованиями биомеханики и спортивной педагогики.

#### Литература

1. Кумсков М.И., Ильясова А.О., Проничкин Ю.В., Мамуров Б.Б. Комплексная система оценки качества обучения спортсменов на основе анализа нейронными сетями видеозаписей тренировок. // *Материалы конференции «Инновационные технологии в спорте и физическом воспитании подрастающего поколения»*. Бухара, 2019, С. 62–63.
2. Мамуров Б.Б., Кумсков М.И., Проничкин Ю.В., Ильясова А.О. Конструирование позы человека по фотографии на основе нейронных сетей глубокого изучения. // *Материалы конференции «Инновационные технологии в спорте и физическом воспитании подрастающего поколения»*. Бухара, 2019, С. 63–65.
3. Vogo F., Kanazawa A., Lassner C. et al. Keep It SMPL: Automatic Estimation of 3D Human Pose and Shape. // *In: Proc. European Conf. on Computer Vision (ECCV)*, 2016, Vol. 9909, P. 561–578.
4. Cao Z., Simon T., Wei S.E., Sheikh Y. Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. // *In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2017, P. 7291–7299.
5. Keilstrup C., Mikkelsen C., Hannemose M., Dahl A.B. SportsPose: a dynamic 3D sports pose dataset. // *In: Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2023, P. 5218–5227.
6. Lin T.Y., Maire M., Belongie S. et al. Microsoft COCO: Common Objects in Context. // *In: Computer Vision – ECCV*. Springer, 2014, P. 740–755.
7. Pavlakos G., Choutas V., Ghorbani N. et al. Expressive Body Capture: 3D Hands, Face, and Body from a Single Image. // *In: Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2019, P. 1328–1338.
8. Saleem M.U., Lee G.H., Cai Y., Koleini F. BioPose: Biomechanically-Accurate 3D Pose Estimation from Monocular Videos. // *In: Proc. IEEE/CVF Winter Conf. on Applications of Computer Vision (WACV)*, 2025, P. 6330–6339.
9. Sun K., Xiao B., Liu D., Wang J. Deep High-Resolution Representation Learning for Human Pose Estimation. // *IEEE Trans. Pattern Anal. Mach. Intell.*, 2021, Vol. 43, No. 10, P. 3349–3364.
10. Toshev A., Szegedy C. DeepPose: Human Pose Estimation via Deep Neural Networks. // *In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2014, P. 1653–1660.
11. Ugrinovic N., Pan B., Pavlakos G. et al. MultiPhys: Multi-person Physics-aware 3D Motion Estimation. // *In: Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2024, P. 2331–2340.
12. Xu Y., Zhang J., Qiming Z., Tao D. ViTPose++: Vision Transformer for Generic Body Pose Estimation. // *IEEE Trans. Pattern Anal. Mach. Intell.*, 2024, Vol. 46, No. 2, P. 1212–1230.
13. Xu Y., Zhang J., Zhang Q., Tao D. ViTPose: Simple Vision Transformer Baselines for Human Pose Estimation. // *Advances in Neural Information Processing Systems*, 2022, Vol. 35, P. 38571–38584.
14. Yeung C.K., Suzuki T., Tanaka R., Yin Z., Fujii K. AthletePose3D: Benchmark Dataset for 3D Human Pose Estimation. // *arXiv:2503.07499*, 2025 (Preprint).
15. Yin Y., Guo C., Kaufmann M., et al. Hi4D: 4D Instance Segmentation of Close Human Interaction. // *In: Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2023, P. 17016–17027.
16. Yin Z., Yeung C., Suzuki T., Tanaka R., Fujii K. KASportsFormer: Kinematic Anatomy Enhanced Transformer for 3D Human Pose Estimation on Short Sports Scene Video. // *arXiv:2507.20763*, 2025 (Preprint).