Development of Deep Learning Enabled Augmented Reality Framework for Monitoring the Physical Quality Training of Future Trainers-Teachers

Sarsenkul Tileubay[1]*, Meruert Yerekeshova[2]*, Altynzer Baiganova[3],
Dariqa Janyssova[4], Nurlan Omarov[5], Bakhytzhan Omarov[6], Zakhira Baiekeyeva[7]
Korkyt Ata Kyzylorda University, Kyzylorda, Kazakhstan[1, 4, 7]
K. Zhubanov Aktobe Regional University, Aktobe, Kazakhstan[2, 3]
Al-Farabi Kazakh National Unviersity, Almaty, Kazakhstan[5]
International University of Tourism and Hospitality, Turkistan, Kazakhstan[5, 6]

***Abstract*—The fusion of augmented reality (AR) and deep learning technologies has ushered in a transformative era in the realm of real-time physical activity monitoring. This research paper introduces a system that harnesses the capabilities of PoseNet-based skeletal keypoint extraction and deep neural networks to achieve unparalleled accuracy and real-time functionality in the identification and classification of a wide spectrum of physical activities. With an impressive accuracy rate of 98% within 100 training epochs, the system proves its mettle in precise activity recognition, making it invaluable in domains such as fitness training, physical education, sports coaching, and home-based fitness. The system's real-time feedback mechanism, bolstered by AR technology, not only enhances user engagement but also motivates users to optimize their exercise routines. This paper not only elucidates the system's architecture and functionality but also highlights its potential applications across diverse fields. Furthermore, it delineates the trajectory of future research avenues, including the development of advanced feedback mechanisms, exploration of multi-modal sensing techniques, personalization for users, assessment of long-term impacts, and endeavors to ensure accessibility, inclusivity, and data privacy. In essence, this research sets the stage for the evolution of real-time physical activity monitoring, offering a compelling framework to improve fitness, physical education, and athletic training while promoting healthier lifestyles and the overall well-being of individuals worldwide.**

Keywords*—PoseNET; MoveNET; deep learning; exercise; computer vision

I. Introduction

In the contemporary landscape of health and wellness, the significance of physical fitness and its correlation with a healthier lifestyle has gained substantial recognition. Engaging in regular physical activities is paramount in mitigating risks associated with chronic ailments like obesity, cardiovascular diseases, and diabetes, a narrative strongly supported by a plethora of scientific studies [1]. The benefits of such a regimen extend beyond mere physical well-being, encompassing enhancements in mental health, cognitive abilities, and even an elongated lifespan. Nonetheless, the crux of maintaining a steadfast exercise routine lies in the effective monitoring and progression tracking, a domain where traditional methodologies often fall short in terms of accessibility and efficiency.

Recent advancements in technology, particularly the integration of computer vision and deep learning, have ushered in a new era in exercise monitoring [2-4]. Leveraging these technological strides, this paper introduces a groundbreaking framework utilizing a PoseNet-enabled deep neural network, primarily aimed at real-time exercise monitoring of physical culture students [5]. PoseNet, a state-of-the-art model developed by Google, lies at the core of this system, enabling precise detection and tracking of human body movements during physical activities.

Traditional methods for monitoring exercise form and posture, often reliant on personal trainers or manual video analysis, are plagued by limitations such as high costs, time consumption, and restricted accessibility [6]. Our proposed framework seeks to dismantle these barriers, offering a cost-effective, real-time solution that does not necessitate additional human intervention [7]. The dual-component architecture of our system, comprising the PoseNet model and a sophisticated deep neural network, marks a significant leap forward in exercise monitoring technology. PoseNet's role is pivotal in identifying and tracking key body points, thereby facilitating the deep neural network in accurately discerning various exercises from the captured movements. This network, trained on an extensive exercise dataset, boasts a remarkable proficiency in recognizing a diverse range of physical activities.

The user-centric design of our system ensures its accessibility and ease of use for physical education students across all skill levels. Compatible with any standard camera-equipped device, such as smartphones, laptops, or tablets, the system allows users to either choose from a predefined exercise catalog or tailor their workout regimes [8]. Real-time feedback provided on form, posture, and motion range empowers users to make immediate adjustments, thus enhancing the effectiveness of their exercise routine.

A notable feature of this system is its adaptive learning capability. The deep neural network can be trained on new exercise datasets, thereby expanding the system's utility to various exercise forms. This adaptability not only customizes the system to cater to individual needs but also positions it as an ideal tool for personalized fitness training [9].

In summary, the development of this PoseNet-enabled deep neural network for real-time exercise monitoring symbolizes a paradigm shift in how physical education students engage with and monitor their fitness routines. The system's real-time feedback, accuracy, and adaptability significantly contribute to more effective and efficient achievement of fitness objectives. Its ease of use, affordability, and broad applicability render it a versatile tool, suitable for diverse settings including educational institutions, fitness centers, and home environments. This innovation not only aligns with the current digital transformation in fitness monitoring but also paves the way for future advancements in the domain of health and physical education.

II. Related Works

In the rapidly evolving landscape of fitness and health monitoring, the intersection of technology and physical well-being has garnered significant attention from researchers and practitioners alike. This section provides a comprehensive overview of the related works in this domain, tracing the evolution of fitness monitoring technologies from their nascent stages to the current state-of-the-art systems. By examining the progression from traditional methods to advanced technologies such as deep learning, computer vision, and augmented reality, we gain insights into the challenges, advancements, and future directions of fitness monitoring. This review not only contextualizes our research within the broader spectrum of technological innovations in fitness but also highlights the pivotal developments that have shaped current practices and are paving the way for future breakthroughs in this field.

A. *Evolution of Fitness Monitoring Approaches*

The journey of fitness monitoring has transitioned from traditional methods, like the use of personal trainers and self-reporting, to more sophisticated, technology-based approaches. Early studies in this field emphasized the personalized touch offered by human trainers, but noted limitations in terms of objectivity and continuity in monitoring physical activities [10]. These manual methods, while beneficial for personalized guidance, lacked the precision and consistency of data-driven approaches [11]. The advent of wearable technology marked a pivotal point in this evolution. Initial fitness trackers, focusing on basic metrics such as steps and heart rate, introduced a more quantifiable approach to fitness monitoring [12]. Subsequent enhancements, incorporating GPS and accelerometers, expanded these devices' capabilities, enabling a deeper analysis of physical exertion and movement [13]. However, these wearables faced challenges in capturing complex body movements with high accuracy, highlighting the need for more advanced monitoring solutions [14].

B. *Integration of Computer Vision in Exercise Monitoring*

Computer vision's integration into fitness monitoring has been a transformative development. Initial forays involved using cameras and basic algorithms for movement tracking, but were hindered by accuracy issues and the need for controlled environments [15]. The advent of deep learning propelled this field forward, significantly improving the accuracy of these systems in tracking complex human movements. Advanced algorithms, particularly those based on deep learning, enabled more precise tracking and analysis in dynamic settings [16]. These developments laid the groundwork for sophisticated applications like real-time exercise form monitoring and posture analysis [17]. Despite these advancements, challenges persisted, especially in terms of adapting these systems to varied and uncontrolled environments. This led to an increased focus on enhancing the robustness and versatility of computer vision applications in physical activity monitoring [18].

C. *Deep Learning and Physical Activity Recognition*

Deep learning models have become central to the advancement of physical activity recognition. These models, trained on extensive datasets encompassing a wide array of human movements, exhibit remarkable accuracy in classifying diverse physical activities. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have been particularly effective, adept at capturing the spatial and temporal dynamics of movement [19]. This has enabled more nuanced analysis and monitoring of exercises, far surpassing the capabilities of traditional fitness trackers. Research in this domain has explored various applications, from basic activity recognition to more complex analyses like form and technique assessment [20]. The ability of these models to learn and adapt to different movement patterns has opened up new possibilities for personalized exercise monitoring. However, the reliance on large, diverse datasets for training these models presents its own set of challenges, particularly in ensuring the representation of a broad range of movement types and exercise forms [21].

D. *Augmented Reality in Fitness Training*

Augmented Reality (AR) has emerged as a groundbreaking tool in enhancing fitness training experiences. Studies have shown that AR can create immersive and interactive environments, making workouts more engaging and effective [22]. The integration of AR with real-time data tracking has led to more interactive and personalized training experiences. This technology not only boosts engagement but also aids in proper technique adherence, reducing the risk of injury [23]. However, seamlessly integrating AR with accurate movement tracking technologies has been a challenge. The key has been to develop systems that are not only accurate but also intuitive and engaging for users. This has led to innovative approaches in AR application design, focusing on user-friendly interfaces and real-time feedback mechanisms. As AR technology continues to evolve, its potential in transforming fitness training methods and improving overall exercise effectiveness becomes increasingly evident [24].

E. *Pose Estimation Technologies in Exercise Monitoring*

Pose estimation technologies, especially those employing deep learning models like PoseNet, have significantly altered the landscape of exercise monitoring. PoseNet, for instance, excels at real-time tracking of human body movements, providing a detailed analysis of exercise form and posture [25]. This technology has been instrumental in enhancing the precision and effectiveness of fitness monitoring systems. Researchers have extensively explored its application across various fitness scenarios, demonstrating its potential in offering real-time feedback, which is crucial in preventing injuries and ensuring the effectiveness of exercise routines [26]. These pose estimation models stand out for their ability to discern subtle nuances in movement, a feat that was previously challenging with conventional monitoring systems. However, the application of such technologies is not without challenges. Ensuring accuracy in diverse and dynamic environments, along with maintaining user privacy, are areas that necessitate ongoing research and development [27]. The continuous improvement of these technologies is crucial for their wider adoption and effectiveness in real-world fitness monitoring scenarios.

F. Challenges and Limitations of Existing Systems

While significant advancements have been made in fitness monitoring technologies, several challenges and limitations persist. Accuracy in complex and uncontrolled environments remains a primary concern. Systems that perform well in laboratory settings often struggle in real-world scenarios, where variables such as lighting and background can affect performance [28]. Additionally, user privacy has emerged as a critical issue, especially with systems that rely on cameras and video analysis. There is a growing need to develop methods that respect user privacy while still providing accurate monitoring [29]. Another challenge is the extensive data required to train deep learning models effectively. These models often require large, diverse datasets to function optimally, which can be a hurdle in terms of data collection and processing [30]. Moreover, the accessibility and usability of these technologies for individuals with varying levels of fitness and technical proficiency remain areas for improvement. Ensuring that these systems are user-friendly and adaptable to different user needs is essential for their broader acceptance and effectiveness [31].

G. Personalization and Adaptability in Fitness Monitoring Systems

The trend towards more personalized and adaptable fitness monitoring systems is gaining momentum. Personalization in fitness technology is not just about tailoring to individual physical abilities, but also adapting to personal preferences and goals. Research has emphasized the importance of systems that can learn and adapt to individual user profiles and exercise routines [32]. Machine learning algorithms, particularly those capable of adaptive learning, are increasingly being integrated into fitness monitoring systems. These systems are designed to not only track and analyze physical activities but also learn from user behavior and preferences, thus enhancing the overall effectiveness of exercise routines [33]. The ability to customize these systems to individual needs not only improves user engagement but also ensures that the exercises are aligned with personal fitness goals. However, developing algorithms that can accurately adapt to a wide range of user profiles remains a challenge, requiring continuous research and development [34]. The ultimate goal is to create fitness monitoring systems that are not only technologically advanced but also deeply attuned to the unique needs and preferences of each user.

H. Future Directions and Emerging Technologies

The future of fitness monitoring is poised for further transformation with the emergence of new technologies and approaches. AI-powered virtual trainers and the integration of biometric sensors are among the most promising developments in this field [35]. These technologies have the potential to offer even more personalized and comprehensive monitoring of physical activities. AI-powered virtual trainers, for instance, can provide real-time feedback and coaching, tailored to individual performance and improvement areas. The integration of biometric sensors, on the other hand, can offer deeper insights into physiological responses during exercises, enabling a more holistic approach to fitness monitoring. Research in this area is focused not only on enhancing the technological capabilities of these systems but also on improving user engagement and overall health outcomes. The combination of AI, advanced sensor technology, and user-friendly interfaces is expected to lead to a new generation of fitness monitoring systems that are more accurate, engaging, and effective in promoting physical well-being [36]. As these technologies continue to evolve, they offer exciting possibilities for the future of personal fitness and wellness.

The future direction of fitness monitoring is geared towards even more personalized and adaptive systems. The integration of AI and biometric sensors is set to redefine the boundaries of what these systems can achieve. The goal is to develop fitness monitoring tools that are not only technologically advanced but also user-centric, catering to individual needs and preferences. As we look ahead, the potential for these technologies to transform personal fitness and health monitoring is immense, promising a future where fitness routines are more effective, engaging, and aligned with individual health goals.

III. DATA

The task of discerning physical activities encompasses a range of distinct yet interrelated subtasks. For clarity, the methodology of this research is depicted in Fig. 1, which illustrates the process as a systematic flowchart. The research design is segmented into three fundamental stages: identification of data requirements, collection of data, and its subsequent categorization.

Within the data requirements section, we define the specific attributes and characteristics of the patterns we aim to analyze. The data collection phase is critical, ensuring the procurement of appropriate video data. This phase involves the meticulous process of annotating videos according to predefined categories, converting them into .json format, and precisely extracting marked images and video sequences that depict various physical exercises, thereby creating a comprehensive dataset.

The final stage, categorization, involves a detailed breakdown of these videos into distinct classes. This stage is further subdivided into several key processes, including the preparation of data, extraction of pertinent features, training of the model, and its rigorous testing.

For the purpose of this study, we have meticulously compiled a dataset encompassing five distinct exercises: pull-ups, push-ups, squats, bicep workouts, and neck exercises. This dataset is derived from an extensive collection of video data, totaling 100 minutes for each exercise category. This rich dataset forms the backbone of our research, enabling detailed analysis and robust model training.

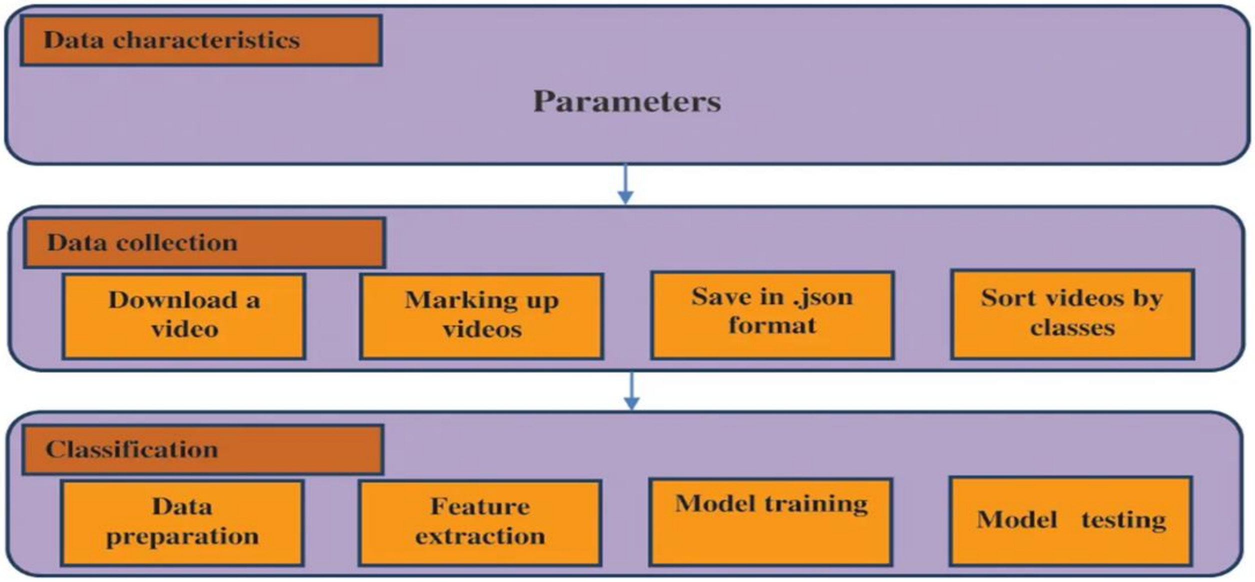

Fig. 1. Flowchart of the proposed framework.

IV. Materials and Methods

A. Proposed Approach

In this section, we shall elucidate the utilization of Deep Learning algorithms for the recognition of objects and postures, a fundamental aspect of the project's execution. The computational model undertakes a series of operations upon receiving data packets, which may include individual video sequences or audio segments. When configuring the computation process, the selection of the payload type for each port is a critical decision, as it determines the ingress and egress of data packets. Each computation module is equipped with ports that facilitate the ingress and egress of data packets. Throughout the execution of a graph, a sequence of actions is performed, encompassing the Open, Process, and Close methods in each computational module. The initialization of a calculator is achieved through the Open method, the continuous processing of new packets is managed by the Process method, and the finalization of the computational process is accomplished via the Close method. Fig. 2 provides an illustrative flowchart delineating the proposed pose detection system, a crucial component of the exercise monitoring framework.

The ensuing sections elucidate our proposed methodology, termed as skeleton-based classification of physical activities, a process comprehensively depicted in Fig. 3. This methodology dissects the overall challenge into three distinct yet interconnected subproblems, each playing a pivotal role in the classification process.

The initial phase involves the deployment of the PoseNet network on image sequences to ascertain body postures. This application of PoseNet to our input data is critical in predicting the stance of the body captured in each frame. Following this, the second phase focuses on the extraction of key points from each frame, represented as vectors. PoseNet is instrumental in this process, identifying a total of 17 critical points per frame. Consequently, this leads to the formation of vectors, each comprising 34 individual elements.

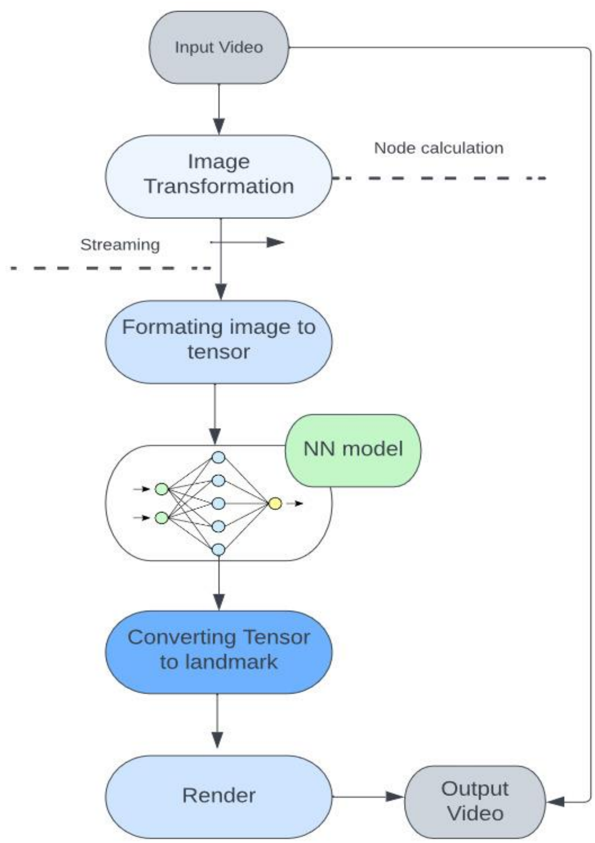

Fig. 2. Flowchart of the proposed pose detection system.

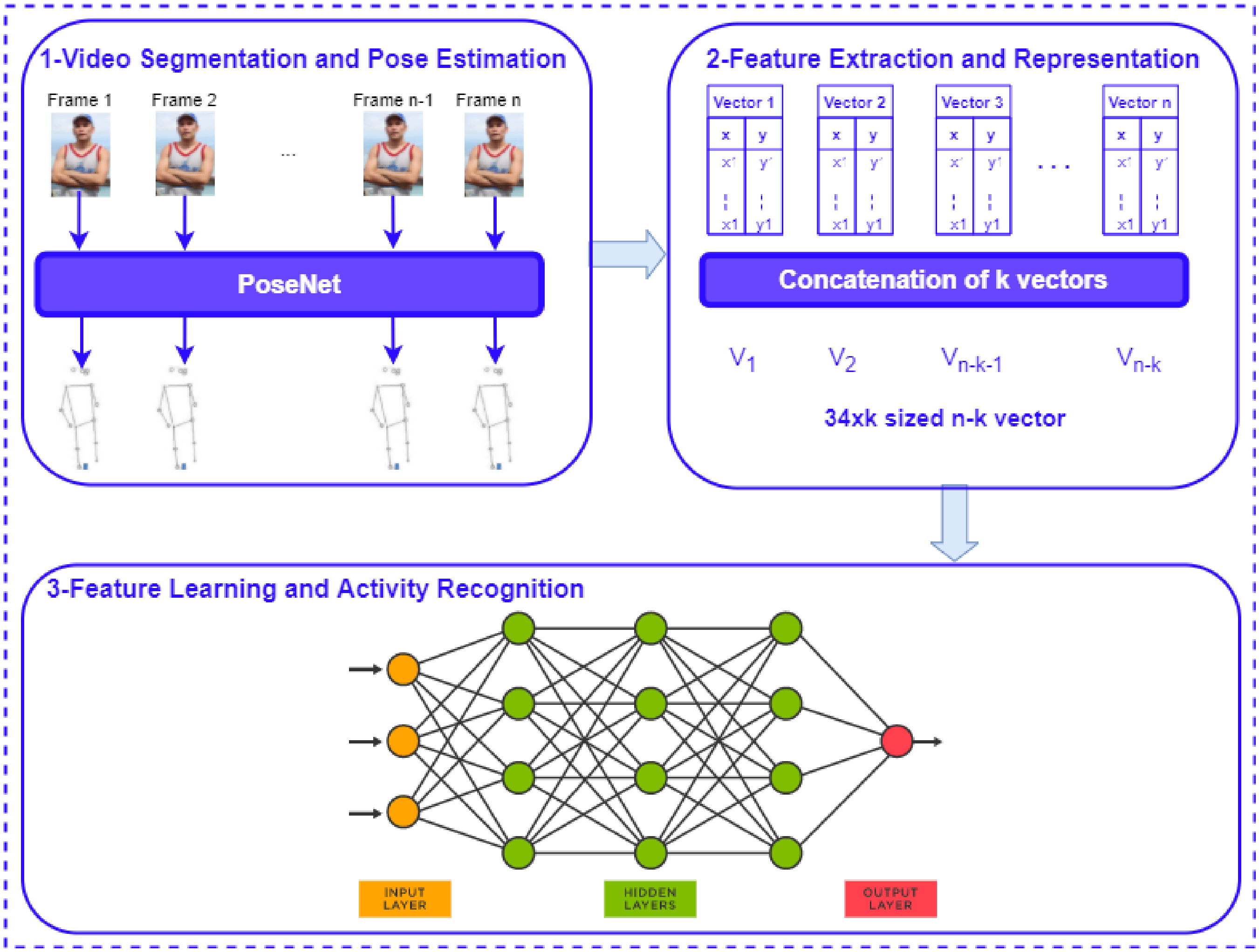

Fig. 3. The proposed framework architecture for action detection.

Subsequently, the methodology involves amalgamating these vectors (k vectors) into a singular comprehensive vector. This consolidated vector is then subjected to the next stage, involving feature extraction and the identification of physical activities. The final step in our methodology is the training of a Convolutional Neural Network (CNN) model, specifically tailored to address tasks associated with the classification of physical activities.

In the context of human body localization in RGB images, two primary approaches are recognized: top-down and bottom-up methods. Top-down approaches initiate with a human detector and proceed to analyze body joints within predetermined boundary boxes. Notable examples of top-down methods include PoseNet [36], HourglassNet [37], and Hornet [38]. Alternatively, bottom-up methods, such as Open space [39] and PifPaf [40], offer a different approach to body localization, each with their unique methodologies and applications.

In our chosen methodology, we have embraced a skeleton-based approach as the foundation for our training strategy. This approach has been strategically selected due to its inherent computational efficiency, which proves pivotal in the real-time assessment of human activities. Central to this approach is the utilization of a neural network architecture built upon PoseNet, a robust and well-established deep learning model. This PoseNet-based neural network serves as the linchpin of our system, facilitating the intricate and precise evaluation of a wide array of human activities.

The operationalization of this methodology entails the integration of a pre-trained PoseNet model into our framework. This pre-trained model stands as a testament to the efficiency and effectiveness of knowledge transfer from the input space to the specific target domain. By leveraging this pre-trained model, we streamline the learning process and empower our system to rapidly and accurately assess and classify human activities in real-time scenarios. This not only enhances the overall computational efficiency of our system but also ensures that it operates seamlessly and effectively, making it a valuable tool for applications such as fitness training, physical education, and sports coaching.

PoseNet's output is crucial in representing the human body, as it identifies 17 primary body points along with their respective positions and associated confidence levels. These key points encompass critical areas such as the face, eyes, ears, shoulders, elbows, wrists, hips, knees, and ankles [41]. Fig. 4 provides a visual representation of these 17 essential points as captured by PoseNet, illustrating the basis for training our neural network. The representation of these points in the coordinate space is achieved through the x and y coordinates, providing a spatial mapping essential for accurate activity analysis. This skeleton-based approach, underpinned by PoseNet's capabilities, forms the foundation of our training process, enabling a more nuanced understanding of human movement and posture in the context of physical activity classification.

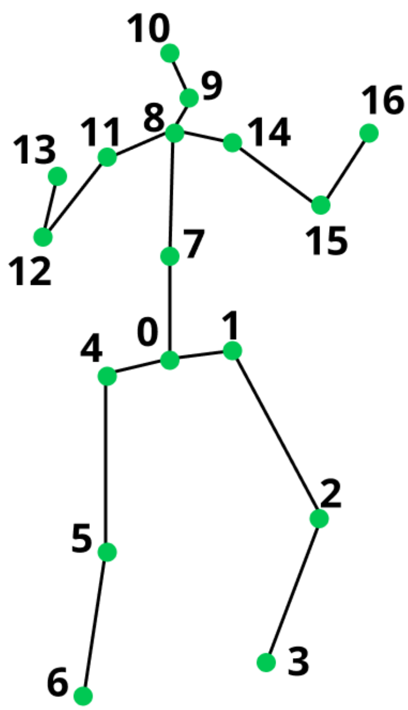

Fig. 4. PoseNET key points.

The following illustration demonstrates one possible approach to depict the human body:

$$r_b(x_i;\theta), \tag{1}$$

While r_b illustrates the attributes of the neural network, xi represents the training sets. In order to categorize the illustration of the human body, $r_b(x_i;\theta)$, , a layer of a completely linked neural network is introduced. The training of the additional neural network is facilitated by minimizing the class cross-entropy loss, a crucial step preceding the normalization of the network via the "Softmax" layer. Fig. 5 delineates the architecture of the PoseNet-based network employed in this process.

Initially, images depicting human activity are input into PoseNet, which is tasked with extracting key skeletal points. Subsequently, these extracted coordinates of the skeleton components are represented within the feature set. This representation plays a pivotal role in the next phase of the process. The extracted key points of the human skeleton, encapsulating vital information regarding human movement and posture, serve as the foundational data for training the neural network. This methodology ensures that the neural network is trained on accurate, spatially relevant data, enabling it to effectively identify and classify different human activities. The process, from the initial extraction of skeletal points to the final training of the neural network, is critical in achieving a robust and accurate system for activity classification.

The research initiates with the primary phase dedicated to data acquisition, feature extraction, class segmentation, and the subsequent construction of a dataset intended for utilization within the neural network. The subsequent phase of the study centers around the integration of the PoseNet model to effectively extract skeletal points, a pivotal component of the methodology aimed at classifying human activities with precision.

The culmination of this approach entails the development of a neural network tailored for the specific task of detecting physical activities. Subsequently, rigorous training and testing protocols are conducted to assess the viability and real-world applicability of the proposed approach. This comprehensive evaluation process is essential for gauging the effectiveness and suitability of the approach in practical, real-world scenarios.

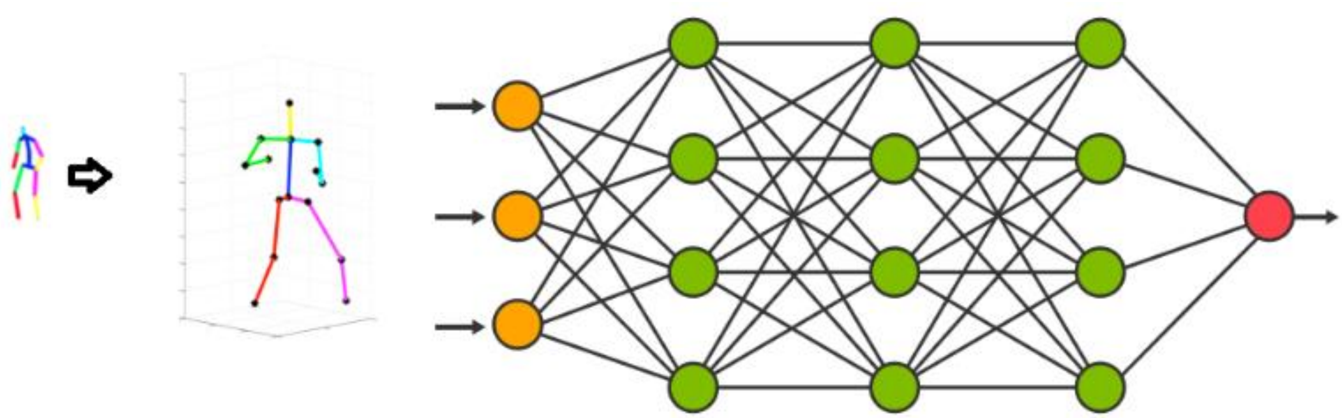

Fig. 5. Artificial neural network for physical activity classification.

B. Evaluation Parameters

To assess the performance of the proposed system, Accuracy is employed as one of the pivotal evaluation parameters. Accuracy measures the system's proficiency in correctly identifying and classifying physical activities, providing a quantitative representation of its correctness in predictions. This parameter quantifies the ratio of accurately identified activities to the total number of activities tested. In essence, Accuracy offers a fundamental gauge of how effectively the system aligns its predictions with the actual activities being performed. It serves as a vital metric in evaluating the overall performance of the proposed system, shedding light on its ability to make accurate classifications. However, it is important to note that while Accuracy provides valuable insights, a comprehensive evaluation may also consider additional metrics such as Precision, Recall, and the F1 Score to provide a more nuanced and complete assessment of the system's classification capabilities.

$$accuracy = \frac{TP + TN}{TP + FN + TN + FP}, \tag{2}$$

Precision is an evaluation metric that assesses the system's ability to minimize false positive errors when classifying physical activities. It quantifies the accuracy of positive predictions made by the system. In the context of activity classification, precision measures the proportion of correctly identified positive cases (true positives) among all the instances that the system predicted as positive (true positives plus false positives). Mathematically, precision is calculated as:

$$precision = \frac{TP}{TP + FP}, \tag{3}$$

Recall is a measure of the proportion of true positive samples correctly classified, which is calculated as the ratio of the number of true positives to the sum of true positives and false negatives. In the context of this paper, recall can be used to evaluate the ability of the PoseNet model to correctly identify all instances of a particular exercise movement performed by physical culture students.

$$recall = \frac{TP}{TP + FN}, \tag{4}$$

A high precision score indicates that the system is proficient at correctly identifying positive cases while minimizing incorrect positive identifications. In other words, it measures the system's ability to avoid labeling activities as positive when they are not.

$$F1 = \frac{2 \cdot precision \cdot recall}{precision + recall}, \tag{5}$$

The F1 Score is a metric that combines precision and recall into a single value, providing a balanced assessment of a system's classification performance. It is particularly useful when dealing with imbalanced datasets, where one class may significantly outnumber the other. The F1 Score is calculated as the harmonic mean of precision and recall.

V. RESULTS

In this section, we present the outcomes derived from our in-depth analysis of the primary challenges encountered during the processes of data acquisition, feature extraction, and the classification of physical activities. The subsequent paragraphs delineate the findings obtained in two distinct categories: the first section outlines the discoveries pertaining to the extraction of human skeleton points, while the subsequent section showcases the results obtained in the realm of physical activity detection. These findings represent the forefront of current research in this domain.

The assessment of these findings is conducted through the lens of comprehensive evaluation metrics, which include the utilization of confusion matrices, model accuracy, precision, recall, and the F1-score. These metrics serve as the cornerstone for a rigorous evaluation, enabling a thorough examination of the system's performance and its alignment with contemporary standards of excellence. The results garnered from these evaluations are indicative of the system's effectiveness in tackling the intricate challenges of data processing and physical activity classification, positioning it at the vanguard of cutting-edge research in the field.

A. Keypoints Extraction

This subsection presents the outcomes of the keypoint extraction process employing the PoseNet model. As depicted in Fig. 6, the proposed model's functionality in keypoint extraction is illustrated. Notably, the PoseNet model exhibits the capability to extract human body keypoints even in scenarios involving multiple individuals within the video frames. In such instances, each human presence in the video is assigned a distinct identification number. In the exemplified scenario, five individuals are denoted by IDs ranging from 1 to 5. This observation underscores the model's ability to effectively differentiate and identify various human entities within a given video context.

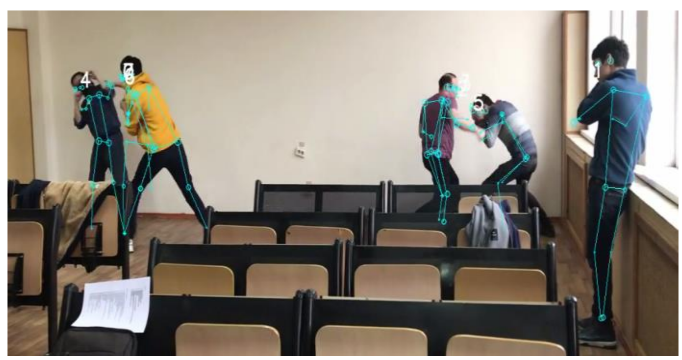

Fig. 6. Keypoints extraction from video.

B. Physical Activity Classification

In the methodology employed for this research, we have embraced a skeleton-based approach as the cornerstone of our training strategy. This strategic choice is rooted in the inherent computational efficiency it offers, a critical factor in enabling real-time assessment of human activities. At the heart of this approach lies the adoption of a neural network architecture built upon PoseNet, a powerful deep learning model renowned for its prowess in capturing human skeletal keypoints.

The practical implementation of this methodology involves the integration of a pre-trained PoseNet model into our system. This pre-trained model serves as a testament to the efficiency and effectiveness of knowledge transfer from the input space to our specific target domain. By leveraging this pre-trained model, we expedite the learning process and empower our system to swiftly and accurately evaluate and categorize a diverse range of human activities in real-time scenarios. This not only bolsters the overall computational efficiency of our system but also ensures its seamless and effective operation, making it an invaluable tool for applications spanning fitness training, physical education, and sports coaching.

In essence, our chosen methodology, with its skeleton-based approach and PoseNet-based neural network, forms the bedrock upon which our research findings and system capabilities rest. It is this methodology that empowers our system to offer precise and real-time assessments of human activities, contributing to advancements in fitness training, physical education, and various domains where the accurate monitoring of physical activities is of paramount importance.

C. Evaluation of the Proposed Model

In this section, we present the outcomes of the physical activity classification process. Fig. 7 and Fig. 8 offer graphical representations of model accuracy and model loss, respectively. Model loss, also referred to as training loss, serves as a metric quantifying the model's performance during the training phase on the training dataset. It is computed by comparing the model's predictions against the actual values within the training data. The primary objective during model training is to minimize this loss, signifying that the model is progressively improving its ability to make precise predictions based on the training data.

Conversely, validation loss assesses the model's performance on a distinct dataset known as the validation dataset, which was not utilized during the training phase. The purpose of validation is to ensure that the model does not exhibit overfitting, i.e., the tendency to memorize the training data rather than learning to generalize to new, unseen data.

Fig. 7 offers a visualization of both model accuracy and validation accuracy for the proposed model over 100 training epochs. The results indicate that within 100 epochs, the proposed model attains an impressive accuracy of 98%. Furthermore, the findings highlight that the model reaches a commendable accuracy of 90% after only 40 epochs of training. These outcomes underscore the model's effectiveness in accurately classifying physical activities, even with relatively modest training durations.

Fig. 8 provides a visual representation of both model loss and validation loss over the course of 100 learning epochs. The outcomes of this analysis reveal that within 100 epochs, the model loss diminishes to a value of 0.2. Additionally, it is noteworthy to emphasize that the proposed system operates in real-time, signifying its capacity to perform expeditiously and provide immediate feedback. This real-time functionality holds significance in the context of physical activity monitoring, as it enables users to engage seamlessly with the system while receiving timely assessments and guidance.

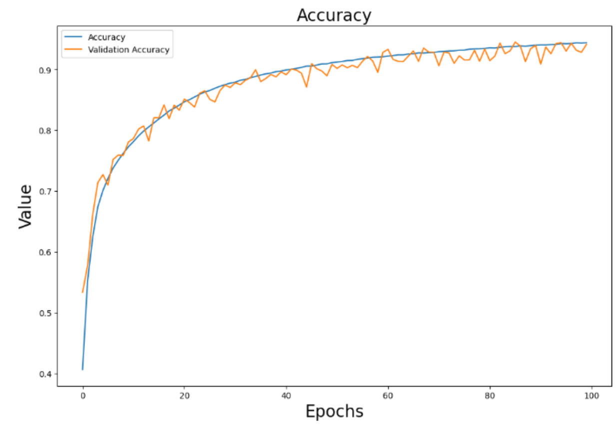

Fig. 7. Accuracy of the proposed model for 100 learning epochs.

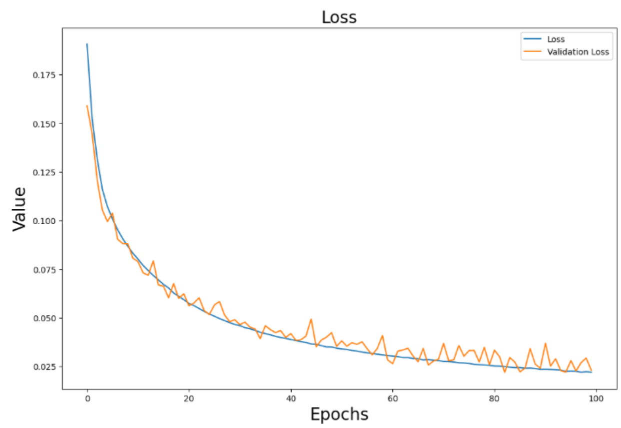

Fig. 8. Loss of the proposed model for 100 learning epochs.

VI. DISCUSSION AND FUTURE RESEARCH

The achieved accuracy of 98% within 100 training epochs underscores the robustness of the model in accurately classifying physical activities. This high level of accuracy is indicative of the system's proficiency in recognizing and distinguishing various exercise routines based on skeletal keypoint data. Such precision is a pivotal attribute, particularly in applications where the correctness of activity identification is critical, such as fitness training and rehabilitation programs.

Furthermore, the real-time operation of the proposed system is a noteworthy feature. The system's ability to provide immediate feedback to users during physical activities is an advantageous aspect that enhances user engagement and motivation [42]. This real-time feedback mechanism aligns with the principles of effective physical training, where timely corrections and adjustments to posture and form are essential for preventing injuries and optimizing the effectiveness of exercise routines. The integration of AR technology into the monitoring process elevates the user experience, making it both interactive and engaging [43].

The implications of the research findings extend to various domains where real-time physical activity monitoring can yield significant benefits. Below, we outline potential applications and the associated advantages.

A. Fitness Training and Rehabilitation

The proposed system holds immense promise in fitness training programs. It can serve as a virtual personal trainer, offering real-time guidance on exercise form, posture, and range of motion [44]. Individuals looking to improve their fitness levels can benefit from accurate feedback, reducing the risk of injuries and enhancing the effectiveness of workouts [44]. Additionally, the system can be adapted for use in rehabilitation programs, assisting patients in performing therapeutic exercises correctly and safely.

B. Physical Education in Schools

Incorporating the system into physical education classes in schools can revolutionize the way students learn and engage in physical activities. It can provide valuable feedback to both students and teachers, ensuring that exercise routines are performed with precision [45]. This can lead to increased interest and participation in physical education, ultimately promoting a healthier lifestyle among young individuals.

C. Sports Coaching

Coaches and athletes can leverage the system for sports training. It can assist in refining athletic techniques by offering real-time insights into movements and postures [46]. This can be particularly beneficial in sports where precise form is crucial, such as gymnastics, dance, and martial arts.

D. Home-Based Fitness

With the increasing popularity of home-based fitness routines, the proposed system can find application in guiding individuals through exercise regimens in the comfort of their homes [47]. It eliminates the need for expensive gym memberships and personal trainers while ensuring that users perform exercises correctly.

VII. Future Research Directions

A. Enhanced Feedback Mechanisms

Future research can focus on the development of more sophisticated feedback mechanisms. This may include integrating voice-based instructions and motivational cues to enhance the user experience further. Additionally, incorporating haptic feedback through wearables can provide tactile guidance during exercises.

B. Multi-Modal Sensing

Exploring multi-modal sensing techniques, such as combining visual data with data from wearable sensors, can improve the accuracy and comprehensiveness of physical activity monitoring [48]. This approach can enable the system to capture a broader range of information, including heart rate, muscle activity, and joint angles.

C. Personalization

Tailoring the system to individual users' needs and fitness levels is an area ripe for exploration. Machine learning algorithms can be employed to adapt the system's feedback and recommendations to each user's unique requirements, optimizing the training experience [49].

D. Long-Term Impact

Assessing the long-term impact of using the proposed system on individuals' fitness levels and overall health is a vital avenue for future research [50]. Longitudinal studies can track the progress and behavior changes of users over extended periods, providing insights into the sustained benefits of the technology.

E. Accessibility and Inclusivity

Research can focus on ensuring that the system is accessible and inclusive for individuals of diverse abilities and backgrounds. This involves addressing challenges related to accommodating various body types, physical conditions, and cultural preferences in exercise routines.

F. Privacy and Data Security

As with any technology that collects personal data, future research should emphasize robust privacy and data security measures [51]. Ensuring that user data is protected and used ethically is paramount.

In conclusion, the integration of augmented reality and deep learning for real-time physical activity monitoring holds immense potential for transforming fitness training, physical education, and sports coaching. The system's high accuracy and real-time feedback capabilities make it a valuable tool for improving exercise routines and promoting healthier lifestyles. Future research endeavors can further enhance the system's functionalities, personalize user experiences, and explore its long-term impacts on individuals' well-being.

VIII. Conclusion

In conclusion, the amalgamation of augmented reality (AR) and deep learning technologies has propelled the realm of real-time physical activity monitoring into an era of innovation and potential. The system presented in this research paper, leveraging the power of PoseNet-based skeletal keypoint extraction and deep neural networks, has demonstrated remarkable accuracy and real-time functionality. The implications of this work span across various domains, including fitness training, physical education, sports coaching, and home-based fitness.

The achieved accuracy rate of 98% within 100 training epochs underscores the system's prowess in precisely classifying a wide array of physical activities based on skeletal keypoint data. This level of accuracy is of paramount significance, particularly in applications where the correctness of activity identification is indispensable. Furthermore, the system's real-time operation stands as a testament to its utility, offering immediate feedback to users during their exercise routines. This feature fosters user engagement, motivation, and an interactive experience that is conducive to effective physical training.

As technology continues to advance and research in this field progresses, the prospects for further enhancements and applications are promising. Future research endeavors may delve into more sophisticated feedback mechanisms, multi-modal sensing techniques, personalized user experiences, and long-term impact assessments. Moreover, ensuring accessibility and inclusivity for individuals of diverse backgrounds and addressing privacy and data security concerns remain pivotal in the evolution of this technology.

The future of real-time physical activity monitoring holds immense potential, offering opportunities to revolutionize fitness training, physical education, and sports coaching. The work presented in this paper serves as a foundation upon which further innovations and advancements can be built, ultimately contributing to the promotion of healthier lifestyles and the well-being of individuals across the globe.

REFERENCES

[1] Aboamer, M. A., Sikkandar, M. Y., Gupta, S., Vives, L., Joshi, K., Omarov, B., & Singh, S. K. (2022). An investigation in analyzing the food quality well-being for lung cancer using blockchain through cnn. Journal of Food Quality, 2022.

[2] Wang, S., Zhang, J., Wang, P., Law, J., Calinescu, R., & Mihaylova, L. (2024). A deep learning-enhanced Digital Twin framework for improving safety and reliability in human–robot collaborative manufacturing. Robotics and computer-integrated manufacturing, 85, 102608.

[3] Narynov, S., Zhumanov, Z., Gumar, A., Khassanova, M., & Omarov, B. (2021, October). Chatbots and Conversational Agents in Mental Health: A Literature Review. In 2021 21st International Conference on Control, Automation and Systems (ICCAS) (pp. 353-358). IEEE.

[4] Sadhu, A., Peplinski, J. E., Mohammadkhorasani, A., & Moreu, F. (2023). A review of data management and visualization techniques for structural health monitoring using BIM and virtual or augmented reality. Journal of Structural Engineering, 149(1), 03122006.

[5] Rahman, A., Xi, M., Dabrowski, J. J., McCulloch, J., Arnold, S., Rana, M., ... & Adcock, M. (2021). An integrated framework of sensing, machine learning, and augmented reality for aquaculture prawn farm management. Aquacultural Engineering, 95, 102192.

[6] Lakshminarayanan, V., Ravikumar, A., Sriraman, H., Alla, S., & Chattu, V. K. (2023). Health care equity through intelligent edge computing and augmented reality/virtual reality: a systematic review. Journal of Multidisciplinary Healthcare, 2839-2859.

[7] Chen, J., Fu, Y., Lu, W., & Pan, Y. (2023). Augmented reality-enabled human-robot collaboration to balance construction waste sorting efficiency and occupational safety and health. Journal of Environmental Management, 348, 119341.

[8] Caiza, G., & Sanz, R. (2023). Digital Twin to Control and Monitor an Industrial Cyber-Physical Environment Supported by Augmented Reality. Applied Sciences, 13(13), 7503.

[9] A. Altayeva, B. Omarov, H.C. Jeong, Y.I. Cho. Multi-step face recognition for improving face detection and recognition rate. Far East Journal of Electronics and Communications 16(3), pp. 471-491.

[10] Omarov, B., Batyrbekov, A., Suliman, A., Omarov, B., Sabdenbekov, Y., & Aknazarov, S. (2020, November). Electronic stethoscope for detecting heart abnormalities in athletes. In 2020 21st International Arab Conference on Information Technology (ACIT) (pp. 1-5). IEEE.

[11] Aminizadeh, S., Heidari, A., Toumaj, S., Darbandi, M., Navimipour, N. J., Rezaei, M., ... & Unal, M. (2023). The applications of machine learning techniques in medical data processing based on distributed computing and the Internet of Things. Computer Methods and Programs in Biomedicine, 107745.

[12] Hong, F., Wang, L., & Li, C. Z. (2023). Adaptive mobile cloud computing on college physical training education based on virtual reality. Wireless Networks, 1-24.

[13] Di Capua, M., Ciaramella, A., & De Prisco, A. (2023). Machine learning and computer vision for the automation of processes in advanced logistics: The integrated logistic platform (ILP) 4.0. Procedia Computer Science, 217, 326-338.

[14] Gupta, Y. P., Mukul, & Gupta, N. (2023). Deep learning model based multimedia retrieval and its optimization in augmented reality applications. Multimedia Tools and Applications, 82(6), 8447-8466.

[15] Liu, C., Zhang, Z., Tang, D., Nie, Q., Zhang, L., & Song, J. (2023). A mixed perception-based human-robot collaborative maintenance approach driven by augmented reality and online deep reinforcement learning. Robotics and Computer-Integrated Manufacturing, 83, 102568.

[16] Finco, M. D., Dantas, V. R., & dos Santos, V. A. (2023). Exergames, Artificial Intelligence and Augmented Reality: Connections to Body and Sensorial Experiences. In Augmented Reality and Artificial Intelligence: The Fusion of Advanced Technologies (pp. 271-282). Cham: Springer Nature Switzerland.

[17] Kazanidis, I., Pellas, N., & Christopoulos, A. (2021). A learning analytics conceptual framework for augmented reality-supported educational case studies. Multimodal Technologies and Interaction, 5(3), 9.

[18] Sharma, M., & Sharma, S. (2023). A holistic approach to remote patient monitoring, fueled by ChatGPT and Metaverse technology: The future of nursing education. Nurse Education Today, 131, 105972.

[19] Hafidi, M. M., Djezzar, M., Hemam, M., Amara, F. Z., & Maimour, M. (2023). Semantic web and machine learning techniques addressing semantic interoperability in Industry 4.0. International Journal of Web Information Systems.

[20] Bu, X. (2023). Exploration of intelligent coaching systems: The application of Artificial intelligence in basketball training. Saudi Journal of Humanities and Social Sciences, 8(09), 290-295.

[21] Apostolopoulos, G., Andronas, D., Fourtakas, N., & Makris, S. (2022). Operator training framework for hybrid environments: an augmented reality module using machine learning object recognition. Procedia CIRP, 106, 102-107.

[22] Doskarayev, B., Omarov, N., Omarov, B., Ismagulova, Z., Kozhamkulova, Z., Nurlybaeva, E., & Kasimova, G. (2023). Development of Computer Vision-enabled Augmented Reality Games to Increase Motivation for Sports. International Journal of Advanced Computer Science and Applications, 14(4).

[23] Mohamed, K. S. (2023). Deep Learning-Powered Technologies: Autonomous Driving, Artificial Intelligence of Things (AIoT), Augmented Reality, 5G Communications and Beyond. Springer Nature.

[24] Latif, K., Sharafat, A., & Seo, J. (2023). Digital Twin-Driven Framework for TBM Performance Prediction, Visualization, and Monitoring through Machine Learning. Applied Sciences, 13(20), 11435.

[25] Kim, M., Choi, S. H., Park, K. B., & Lee, J. Y. (2021). A hybrid approach to industrial augmented reality using deep learning-based facility segmentation and depth prediction. Sensors, 21(1), 307.

[26] Mahariya, S. K., Kumar, A., Singh, R., Gehlot, A., Akram, S. V., Twala, B., ... & Priyadarshi, N. (2023). Smart campus 4.0: Digitalization of university campus with assimilation of industry 4.0 for innovation and sustainability. Journal of Advanced Research in Applied Sciences and Engineering Technology, 32(1), 120-138.

[27] Zhang, S., Suresh, L., Yang, J., Zhang, X., & Tan, S. C. (2022). Augmenting sensor performance with machine learning towards smart wearable sensing electronic systems. Advanced Intelligent Systems, 4(4), 2100194.

[28] Huang, D., & Hoon-Yang, J. (2023). Artificial intelligence combined with deep learning in film and television quality education for the youth. International Journal of Humanoid Robotics, 20(06), 2250019.

[29] Sampedro, G. A., Putra, M. A. P., & Abisado, M. (2023). 3D-AmplifAI: An Ensemble Machine Learning Approach to Digital Twin Fault Monitoring for Additive Manufacturing in Smart Factories. IEEE Access.

[30] Makhataeva, Z., & Varol, H. A. (2020). Augmented reality for robotics: A review. Robotics, 9(2), 21.

[31] Li, M., Feng, X., Han, Y., & Liu, X. (2023). Mobile augmented reality-based visualization framework for lifecycle O&M support of urban underground pipe networks. Tunnelling and Underground Space Technology, 136, 105069.

[32] Redžepagić, A., Löffler, C., Feigl, T., & Mutschler, C. (2020, November). A sense of quality for augmented reality assisted process guidance. In 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) (pp. 129-134). IEEE.

[33] Devagiri, J. S., Paheding, S., Niyaz, Q., Yang, X., & Smith, S. (2022). Augmented Reality and Artificial Intelligence in industry: Trends, tools, and future challenges. Expert Systems with Applications, 118002.

[34] Maharjan, D., Agüero, M., Mascarenas, D., Fierro, R., & Moreu, F. (2021). Enabling human–infrastructure interfaces for inspection using augmented reality. Structural Health Monitoring, 20(4), 1980-1996.

[35] Mohamed, K. S. (2023). Deep Learning for Spatial Computing: Augmented Reality and Metaverse “the Digital Universe”. In Deep Learning-Powered Technologies: Autonomous Driving, Artificial Intelligence of Things (AIoT), Augmented Reality, 5G Communications and Beyond (pp. 131-150). Cham: Springer Nature Switzerland.

[36] Mertes, J., Lindenschmitt, D., Amirrezai, M., Tashakor, N., Glatt, M., Schellenberger, C., ... & Schotten, H. D. (2022). Evaluation of 5G-capable framework for highly mobile, scalable human-machine interfaces in cyber-physical production systems. Journal of Manufacturing Systems, 64, 578-593.

[37] Park, K. B., Kim, M., Choi, S. H., & Lee, J. Y. (2020). Deep learning-based smart task assistance in wearable augmented reality. Robotics and Computer-Integrated Manufacturing, 63, 101887.

[38] Um, J., min Park, J., yeon Park, S., & Yilmaz, G. (2023). Low-cost mobile augmented reality service for building information modeling. Automation in Construction, 146, 104662.

[39] Goh, H. A., Ho, C. K., & Abas, F. S. (2023). Front-end deep learning web apps development and deployment: a review. Applied Intelligence, 53(12), 15923-15945.

[40] You, Z., & Feng, L. (2020). Integration of industry 4.0 related technologies in construction industry: a framework of cyber-physical system. Ieee Access, 8, 122908-122922.

[41] Nagy, M., Lăzăroiu, G., & Valaskova, K. (2023). Machine Intelligence and Autonomous Robotic Technologies in the Corporate Context of SMEs: Deep Learning and Virtual Simulation Algorithms, Cyber-Physical Production Networks, and Industry 4.0-Based Manufacturing Systems. Applied Sciences, 13(3), 1681.

[42] Rathore, M. M., Shah, S. A., Shukla, D., Bentafat, E., & Bakiras, S. (2021). The role of ai, machine learning, and big data in digital twinning: A systematic literature review, challenges, and opportunities. IEEE Access, 9, 32030-32052.

[43] Mourtzis, D., & Angelopoulos, J. (2020). An intelligent framework for modelling and simulation of artificial neural networks (ANNs) based on augmented reality. The International Journal of Advanced Manufacturing Technology, 111(5-6), 1603-1616.

[44] Omarov, N., Omarov, B., Baibaktina, A., Abilmazhinova, B., Abdimukhan, T., Doskarayev, B., & Adilzhan, A. (2023). Applying Artificial Intelligence and Computer Vision for Augmented Reality Game Development in Sports. International Journal of Advanced Computer Science and Applications, 14(8).

[45] Poonja, H. A., Shirazi, M. A., Khan, M. J., & Javed, K. (2023). Engagement detection and enhancement for STEM education through computer vision, augmented reality, and haptics. Image and Vision Computing, 104720.

[46] Yin, Y., Zheng, P., Li, C., & Wang, L. (2023). A state-of-the-art survey on Augmented Reality-assisted Digital Twin for futuristic human-centric industry transformation. Robotics and Computer-Integrated Manufacturing, 81, 102515.

[47] Ponnusamy, V., Natarajan, S., Ramasamy, N., Clement, J. C., Rajalingam, P., & Mitsunori, M. (2021). An IoT-Enabled Augmented Reality Framework for Plant Disease Detection. Rev. d'Intelligence Artif., 35(3), 185-192.

[48] Omarov, B., Nurmash, N., Doskarayev, B., Zhilisbaev, N., Dairabayev, M., Orazov, S., & Omarov, N. (2023). A Novel Deep Neural Network to Analyze and Monitoring the Physical Training Relation to Sports Activities. International Journal of Advanced Computer Science and Applications, 14(9).

[49] Mihai, S., Yaqoob, M., Hung, D. V., Davis, W., Towakel, P., Raza, M., ... & Nguyen, H. X. (2022). Digital twins: A survey on enabling technologies, challenges, trends and future prospects. IEEE Communications Surveys & Tutorials.

[50] Xu, M., Hoang, D. T., Kang, J., Niyato, D., Yan, Q., & Kim, D. I. (2022). Secure and reliable transfer learning framework for 6G-enabled Internet of Vehicles. IEEE Wireless Communications, 29(4), 132-139.

[51] Kumar, R., Rani, S., & Khangura, S. S. (Eds.). (2023). Machine Learning for Sustainable Manufacturing in Industry 4.0: Concept, Concerns and Applications. CRC Press.