Applying Deep Learning and Computer Vision Techniques for an e-Sport and Smart Coaching System Using a Multiview Dataset: Case of Shotokan Karate

https://doi.org/10.3991/ijoe.v18i12.30893

Fatima-Ezzahra Ait-Bennacer[1(✉)], Abdessadek Aaroud[1], Khalid Akodadi[1], Bouchaib Cherradi[1,2]
[1]LAROSERI Laboratory, Department of Computer Science, Faculty of Sciences, Chouaib Doukkali University, El Jadida, Morocco
[2]STIE Team, CRMEF Casablanca-Settat, Provincial Section of El Jadida, El Jadida, Morocco
aitbennacer.fati@gmail.com

Abstract—Smart coaching and e-sport platforms have shown a great interest in the recent research studies. Through this study, we aim to globalize the practice of sport, especially Shotokan Karate, by connecting participants and coaches on an international scale through the integration of Artificial Intelligence techniques such as computer vision and deep learning, to give the possibility of carrying out national and international virtual training courses without logistical constraints. The proposed work aims to apply the latest action detection, action recognition, and action classification methods for different Karate movements using the LSTM and the ST-GCN algorithms and proposes these movements in 3D using Video Inference for Human Body Pose and Shape Estimation (VIBE). Our proper Multiview Dataset contains pose estimations of a set of basic movements captured by a karate Shotokan expert (6th DAN Black Belt) from three views (Front view, Left view, and Right view) using OpenPose and FastPose for human body keypoint detection. The current study sets out to detect, recognize, classify and score different participants' movements. We achieved greater than 96% recognition accuracy of this dataset using the LSTM algorithm, and 91.01% using the ST-GCN algorithm.

Keywords—e-Sport, sport service continuity, deep learning, computer vision, smart coaching

1 Introduction

In Morocco, the Moroccan Games and Sports Ministry (MDJS) launches the mobile application "NT7arko f'dar" ("Let's move at home"), the first in Morocco and the region; this new mobile application that enriches the program "NT7arko" offers users more than 150 exercises to practice at home and without equipment[1]. Likewise, other proposals offer solutions for practicing sports at home, on web/mobile platforms, while

[1] https://www.nt7arko.com/

neglecting the interaction with the participant. Moreover, these solutions do not involve collective sports, such as Shotokan Karate, which will decrease the collective motivation of participants.

Karate is a martial art developed in the Ryukyu Kingdom. Karate is now a striking art using punching, kicking, knee strikes, elbow strikes and open-hand techniques such as knife-hands, spear-hands and palm-heel strikes. A karate practitioner is called a karateka (空手家) [1]. Figure 1 shows Shotokan Karate history and description.

WHAT	Shotokan (松濤館, Shōtōkan)
WHO	Gichin Funakoshi (1868–1957) and his son Yoshitaka Funakoshi (1906–1945)
WHEN	1938
WHERE	Okinawa, Japan
HOW	• Kihon: the practice of basics techniques • Kata: a set sequence of karate moves, it consists of kicks, punches, sweeps, strikes, and blocks • Kumite: the practice and application of Kihon and Kata to real opponents

Fig. 1. Shotokan Karate history and description

Today, the Royal Moroccan Federation of Karate and Associated Disciplines (FRMK & DA) has more than 40.000 members and 900 affiliated associations, with an estimated 350.000 karate practitioners in Morocco.[2]

Smart coaching in martial arts [2] is considered as one of the recent research areas in Human Motion Analysis, or frequently called human motion capture, that is growing rapidly due to the large number of potential applications [3], [4].

Recently, many applications of Computer Vision and Deep Learning techniques have been done in sports, like basketball, football, and Karate, [5], [6]. Besides, other studies have implemented Maching Learning algorithms to predict and classify sports activities [7], and others by integrating wearable sensors for activity recognition [8].

The primary goal of our approach is to propose a system that will automatically detect, recognize and classify movements, just with cameras, phones or a laptop. The new concept is based on Artificial Intelligence techniques for an e-Sport and smart coaching system. We aim to keep the same enthusiasm, a group of participants in real time or deferred time with an intelligent coaching that can detect, recognize and classify a given movement by involving emerging techniques such as action detection, action recognition using computer vision.

Through this work, we aim to globalize the practice of collective sports by connecting participants and coaches at national and international levels and even by offering

[2] https://www.frmk.ma/

the possibility of carrying out national and international virtual training sessions without logistical constraints.

The dataset of our study contains pose estimations of a set of basic movements captured by a karate Shotokan expert (6th DAN Black Belt) from three views (Front view, Left view and Right view) using OpenPose and FastPose for human body keypoint detection. We choose eight movements: Gidan Barei, Tsuki Chudan, Tsuki Jodan, Soto Uke, Shuto Uke, Mae Giri, Mawashi Giri and Yoko Giri from three views, face, right and left.

The main contributions of this work can be summarized as follows:

– We establish a novel Multiview Karate Dataset that contains pose estimation of basic Karate movements from three views.
– We apply two Deep Learning algorithms, the LSTM and the ST-GCN, to detect, classify and score the Karate movements. We achieve 96% recognition accuracy using the LSTM, and 91.01% using the ST-GCN.
– We propose a 3D Karate movement using fewer materials, by integrating the VIBE technique.

This paper is structured as follows: Section 2 discusses the related work and explains the selected methods. Section 3 describes the current approaches and technologies that are used. Section 4 shows the implementation and discusses results that each approach has reached. Finally, section 5 presents the conclusion and the future research perspectives.

2 Related works

2.1 Applying computer vision and deep learning techniques in human motion analysis

Human motion analysis has become one of the popular discussions in recent research [9]; thus it can be treated by many approaches. In this subsection, we discuss how the latest studies applied these techniques to understand human motion. The authors used Python and OpenCV (OpenPose: the real-time multiple-person detection library) to build the 2D real-time gesture grading system on the Linux platform. The users may record movies using any monocular camera.

Yan et al. proposed a novel model for skeleton sequences generic representation by extending graph neural networks to a spatial-temporal graph model, called Spatial-Temporal Graph Convolutional Networks (ST-GCN) [10]. The authors introduced the ST-GCN model as an extension of the neural networks graph to a spatial-temporal graph model. The main objective is to design a skeleton sequences generic representation based on the ST-GCN on two large-scale datasets (Deepmind Kinetics human action dataset and NTU-RGB+D dataset) for skeleton-based action recognition and to achieve high performance. Another recent study conducted by Noumeir et al. applied the new deep learning method ST-GCN that has the advantage of capturing spatial and temporal information simultaneously [11].

Existing works that applied computer vision technologies for behavior recognition are avoiding the importance of depth information to deal with the over-fitting and the poor performance issues. As addressed by Kai et al. who proposed a novel framework by including a target depth estimation algorithm for the skeleton and a network combining St-CNN and ATT-LSTM (Spatio-temporal convolution and attention-based LSTM) [12].

Zhe et al. presented an explicit nonparametric of keypoint association that encodes both the location and the orientation of human limbs. Besides, the authors have created an architecture that learns component detection and association at the same time [13].

The Table 1 summarizes some related works that we have cited above.

Table 1. Some related works in applying AI techniques in sport and smart coaching

Ref	Year	Overview
[10]	2018	The main objective is to design a skeleton sequences generic representation based on the ST-GCN for skeleton-based action recognition and to achieve high performance.
[12]	2020	The current work proposes a novel framework by including a target depth estimation algorithm for the skeleton and a network combining St-CNN and ATT-LSTM.
[13]	2021	The article shows that a greedy parsing method is suitable for producing high-quality body position parses. Then, the PAF refinement is considerably more relevant than PAF and body part location refinement when it comes to runtime performance and accuracy.
[11]	2021	The study applied the new deep learning method called ST-GCN proposed in [10]. They exploit skeleton data provided by the Kinect v2 camera and can detect up to two persons in the same scene by using three datasets: NTU-RGB+D, TST Fall detection v2, and Fallfree.

2.2 Applying AI techniques in sports context/smart coaching and karate

The sport and coaching context have increasingly developed by adopting trendy technologies. In our case, we focused on Artificial Intelligence techniques applied in sport and smart coaching, especially for Shotokan Karate. Some recent studies have treated this area, and they have proposed techniques to detect recognize and evaluate karate movements.

Starting with a traditional local martial art called Seni Silat Cekak Malaysia (SSCM), Idris et al. [14] presented a theoretical framework of extrinsic feedback using an automated evaluation system to evaluate the effectiveness of performing techniques of this martial art. The MoCap technology has been chosen to capture, track and record techniques. The proposed framework consists of 3 important modules: Motion Capture module, motion Recognition module, and motion evaluation module.

In addition for Karate context, Priya and Arulselvi in created a Multiview dataset containing new actions of karate martial arts and Bharatanatyam dance. The dataset contains defensive and counterattacking body movements and Bharatanatyam. These poses are captured and used to train the classifier, in an open environment and Bharatanatyam poses in the studio both with a cluttered background. The authors used Deep Convolutional neural networks to classify the poses without any feature extraction. Karate and Bharatanatyam data were achieved a classification accuracy of 68% and 62%, respectively. Also, Emad et al. in [15] presented a system using Kinects v2 sensors to capture Karate moves. In this work, the Fast-DTW (Fast Dynamic Time Warping) was implemented to manage the different speeds of the moves taken by the performers using the Kinect and to provide them a real-time feedback as accurately as possible. The authors used "Age-Uke", "Mae-Geri", "Gedan-Barei", "Soto-Uke" and other multiple Hein-Shodan moves. Another recent work realized by the same authors [5], fits with the smart coaching in Karate. The system captures and analyzes players' movements in real-time, generates an evaluation report that informs them how to develop their performance or warns them if they have executed a movement incorrectly. The equipment used in this methodology is the Kinect IR camera sensor, and the software was the Microsoft SDK. The best fitted supervised classification algorithm with the data-set and the segmentation method, the F-DTW has reached the highest accuracy; it is capable of detecting each movement with an overall accuracy of 91.07%.

As shown and discussed in these previous works, the current work will provide a new Multiview Dataset and propose an approach that combines recent applied AI techniques to recognize and classify basic Karate movements.

3 Proposed approach based on AI techniques

3.1 Preparing the dataset

The current study aims to detect and recognize automatically the basic karate movement sequences presented by participants, based on the expert coach movements. In addition, we try to adopt an approach with minimal materials and equipment, which means that the coach and the karateka will only need a smartphone or a laptop to capture the movements. The following subsections detail the steps.

Capturing video capsules of karate movements & movements splitting. The videos have been recorded in a Karate Club using 3 cameras from three views: front, right and left with a coach in the center wearing a kimono. Figure 2 presents the video capturing steps.

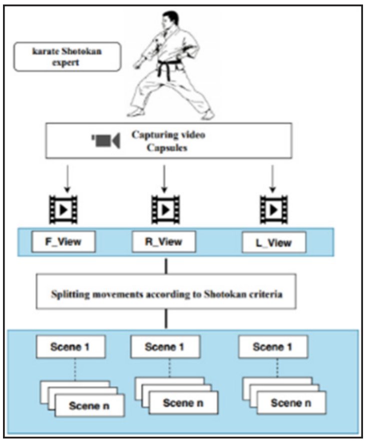

Fig. 2. Capturing video capsules and movements splitting

The movement split is done according to the Shotokan criteria; each movement is taken several times from 3 views and with a different speed, and each video starts at the beginning of the movement and finishes at the end of the movement. In this work we consider 8 basic movements as described as following (Figures 3 to 10):

Gedan Barei is a powerful blocking movement that deflects any kick or fist attack arriving at the belt level. The Figure 3 illustrates Gidan Barei performed from three views.

a) Face View b) Right View c) Left View

Fig. 3. Gidan Barei from three views

Tsuki Chudan is a punch at the medium level. The punch can be done between the belt and the neck as described in Figure 4.

a) Face View b) Right View c) Left View

Fig. 4. Tsuki Chudan from three views

Tsuki Jodan is a punch at the head level. The punch is usually done to the chin. Figure 5 depicts Tsuki Jodan performed from three views.

a) Face View b) Right View c) Left View

Fig. 5. Tsuki Jodan from three views

Soto Uke is a blocking technique used to block medium level incoming attacks as illustrated in Figure 6.

a) Face View b) Right View c) Left View

Fig. 6. Soto Uke from three views

Shuto Uke is a blocking technique by using the external edge of the hand. It can be performed by a circular motion from the inside to the outside to deflect attacks incoming from the sides. The Figure 7 shows this movement from three views.

a) Face View b) Right View c) Left View

Fig. 7. Shuto Uke from three views

Mae Giri is a direct kick that hits with the ball of the foot. Figure 8 illustrates Mae giri from three views.

a) Face View b) Right View c) Left View

Fig. 8. Mae Giri from three views

Mawashe Giri is a circular kick by rotating the hip, then the supporting foot, and using the instep of the foot. Figure 9 shows Mawashi Giri from three views.

a) Face View b) Right View c) Left View

Fig. 9. Mawashi Giri from three views

Yoko giri is a linear pushing kick that is normally used in Shotokan Karate. It is a powerful kick because it combines the thrusting motion of the leg with the power of the hips. The Figure 10 describes Yoko Giri movement from three views.

a) Face View b) Right View c) Left View

Fig. 10. Yoko Giri from three views

The next step is to estimate poses from all these captures to prepare our Dataset. We used the most well-known techniques in the state of the art that are OpenPose and FastPose as explained in the following subsections.

Pose estimation using OpenPose. As discussed in the related works, real-time multi-person 2D pose estimation is a key element that allows machines to understand humans in images and videos.

Using Human Pose Estimation in Computer Vision, we can detect the position and the orientation of objects; in our case, we detect the keypoints that describe the coach's body. Many human pose estimation approaches have been proposed over the last years. OpenPose provides an efficient and robust approach that allows applying pose estimation to images with crowded scenes. It is a real-time multi-person human pose detection library that has for the first time shown the capability to jointly detect the human body, foot, hand, and facial keypoints on single images. OpenPose is capable of detecting a total of 135 keypoints. Figure 11 illustrates the multi-stage CNN architecture used in OpenPose [13].

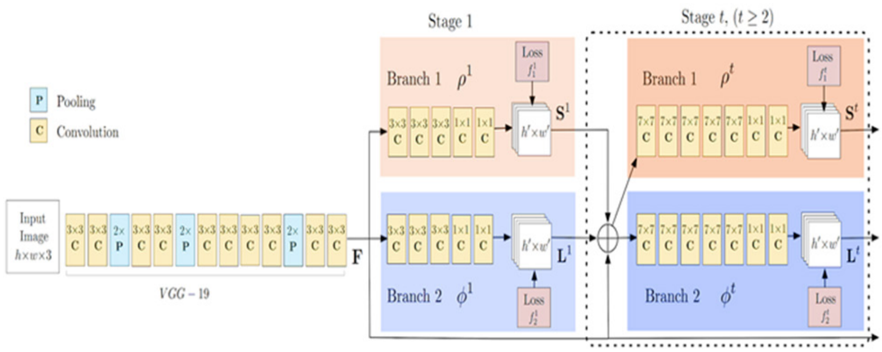

Fig. 11. Multi-person pose estimation model architecture

We used the COCO Keypoints challenge & MPII Human Pose Dataset to train the OpenPose. To create our proper dataset, we extracted the pose estimation of all the proposed coach's movements from 3 views as described in Figures 12 and 13.

Fig. 12. Pose estimation using OpenPose

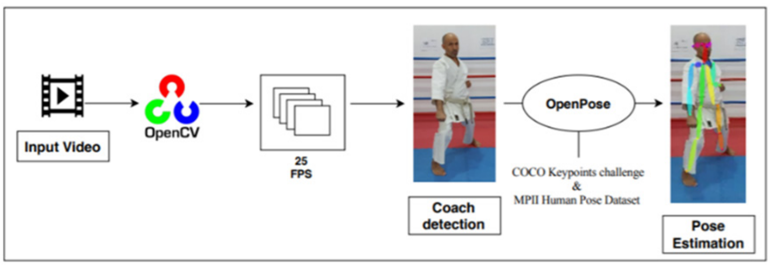

Fig. 13. Coach detection and pose estimation using OpenPose

Pose estimation using FastPose. In this part, the pose estimation will be extracted by using FastPose. It is an open source library that can perform a 2D/3D pose estimation in real time. In this work [16], authors mentioned that FastPose is 46% smaller and 47%

faster (forward time) than OpenPose. The feedforward network takes a three-channel picture as input and predicts detection confidence maps of the anatomical keypoints and the middle body part as output. Figure 14 gives the main elements of this architecture.

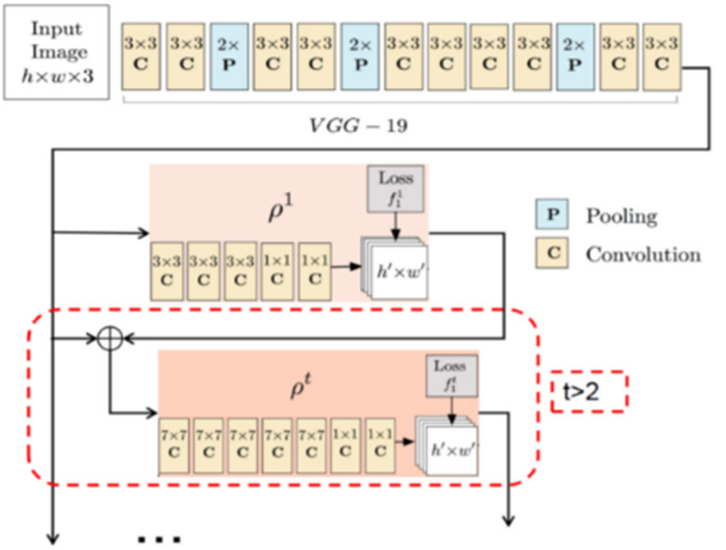

Fig. 14. Feedforward network architecture

After having the dataset from each technique, the next step is to recognize and classify the Karate movements. For this aim, we apply two Deep Learning algorithms, the LSTM and the ST-GCN as explained as follows.

3.2 Karate movement recognition and classification using the LSTM algorithm

Recently, deep learning methods such as recurrent neural networks like LSTMs have been shown to provide state-of-the-art results on challenging activity recognition tasks. The purpose of their application is to represent each movement as a single pose that comprises a set of body key points.

LSTM algorithm. One of the most used deep learning algorithms, the LSTM networks have been extensively used to process time series data as they can follow long-term dependencies in sequences [17]. The LSTM has been designed to cope with a variety of complex detection tasks. The Figure 15 describes the sequencing steps to classify a performed movement.

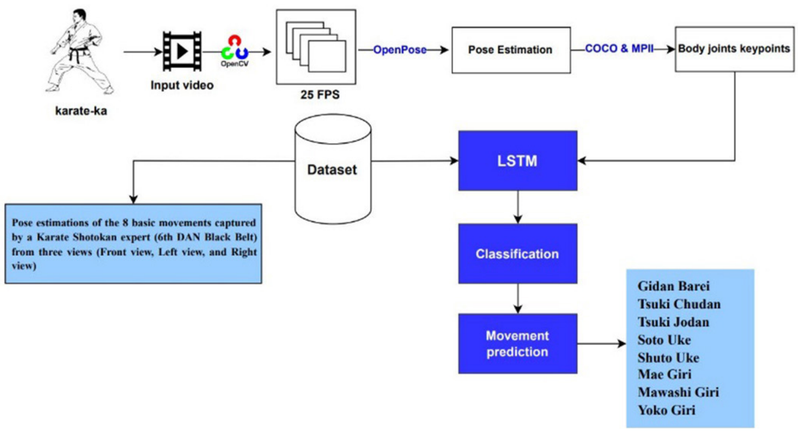

Fig. 15. Movement classification using LSTM algorithm

3.3 Karate movement recognition and classification using the ST-GCN algorithm

ST-GCN algorithm. This work [11] discussed a novel recognition method called Spatial-Temporal Graph Convolutional Networks (ST-GCN). It can be used for skeleton sequences generic representation by extending graph neural networks to a spatial-temporal graph model.

We aim to propose a system that provides a movement classification with a scoring which represents how the karateka has performed the chosen movement referring to coach movement. Figure 16 depicts the main stages of ST-GCN algorithm.

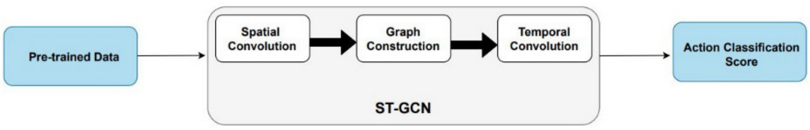

Fig. 16. ST-GCN approach for action classification

The pre-training steps can be described in Figure 17:

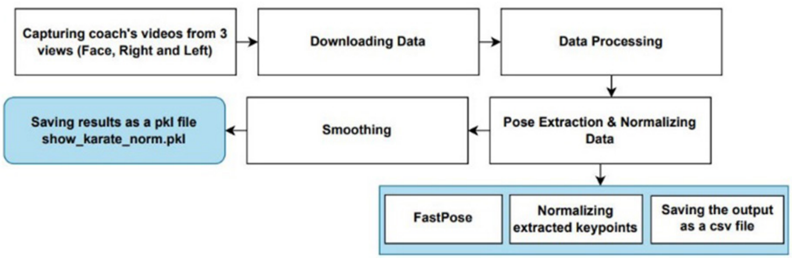

Fig. 17. Pre-training steps

3.4 A 3D Karate coach using VIBE technique

Muhammed et al. have mentioned in their work [18] that the existing video based state of the art methods fail to produce accurate and natural motion sequences due to a lack of ground-truth 3D motion data for training (Figure 18). The cited research defines a novel temporal network architecture with a self-attention mechanism and shows that adversarial training, at the sequence level, produces kinematically plausible motion sequences without in-the-wild ground-truth 3D labels. The authors have made the code and pre-trained models available at GitHub.[3]

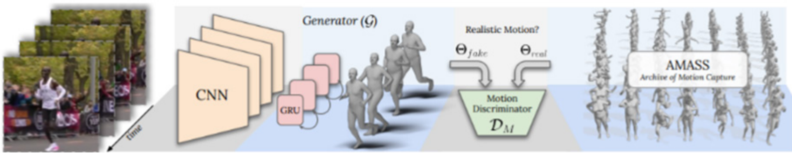

Fig. 18. VIBE architecture

For the training, VIBE takes in-the-wild images as input and predicts SMPL body model parameters using a convolutional neural network (CNN) pre-trained for single-image body pose and shape estimation, as shown in the figure above. In this study, the application of VIBE on our context can provide us with 3D Karate movements on minimizing materials as illustrated in Figure 19.

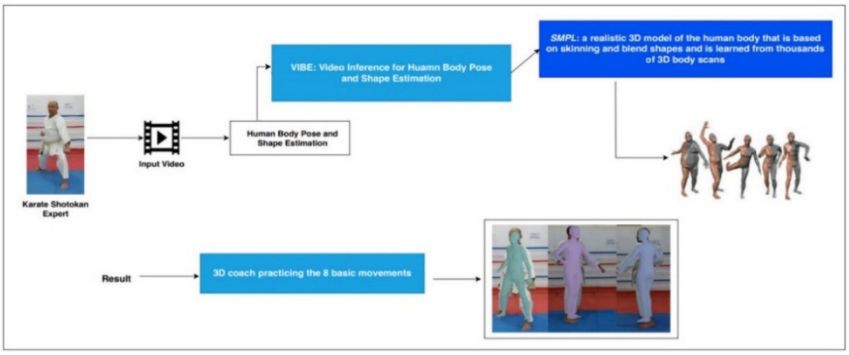

Fig. 19. A 3D Karate coach

4 Results and discussion

This section discusses the predictions obtained by both the LSTM and ST-GCN methods. The discussion will be done in three parts. The first part discusses the obtained

[3] https://github.com/mkocabas/VIBE

results using the LSTM algorithm. The second part presents the results using the ST-GCN algorithm. The last part concerns a comparison between these two approaches and other works in pose recognition based on their accuracy. We have to mention that our experiments were conducted on Acer computer AMD Ryzen 5 3500U with Radeon Vega Mobile Gfx 2.10 GHz, using Google Collab. Our program is developed in Python Programming Language, using Keras library.

4.1 Results using the LSTM algorithm

We trained our LSTM by multiple "Body joints keypoints" obtained from the pose-estimation method of the 8 basic movements captured by a Shotokan Karate expert; these movements were captured from three views (L_View, F_View, R_View). All parameters used are described in the Table 2 below.

Table 2. LSTM parameters

Optimizer	Learning_Rate	Error Function	Epochs	Batch_Size	Activation Function
SGD	0,0001	categorical_ crossentropy	800	256	softmax

The LSTM approach results are summarized in Figures 20 and 21. It can be observed that the LSTM method performed well with 96% overall accuracy. It can be seen also that accuracy increases while raising the number of epochs in the training/test set, which proves the high performances of LSTM.

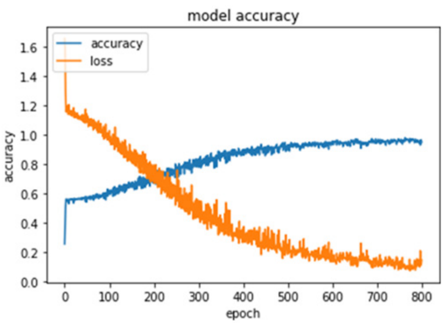

Fig. 20. LSTM model accuracy and loss

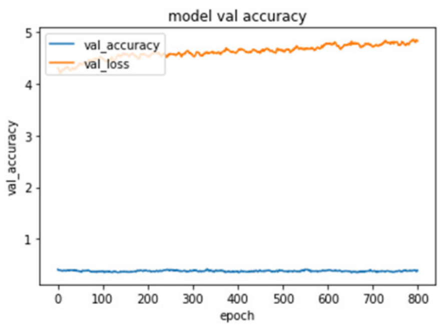

Fig. 21. LSTM model validation accuracy and validation loss

The Figure 22 represents the results achieved by using this technique. The Karateka performs a chosen movement, and the system will detect his body, and classify the movement.

Fig. 22. Karate movement recognition using LSTM

4.2 Results using the ST-GCN algorithm

Despite the high accuracy obtained by ST-GCN which achieved 91% overall, there are still a few limitations of this technique regarding its application in this domain. The results for the ST-GCN approach are summarized in Figure 23 (Confusion Matrix). It can be observed that the ST-GCN method performed well with 91% overall accuracy, and all the trained ST-GCN classifiers yielded satisfactory results for the test dataset :0,93% Gidan_barei, 0,94% Mae_giri, 100% Mawashi_giri, 0,85% Shuto_uki, 100% Suto_uki, 0,91% Seuki_jodan, 0,97% seuki_chudan and 0,73% yoko_giri. It can also be seen from Figures 24 and 25 (Accuracy and Loss) that accuracy increases while raising the number of epochs in the training/test set, which proves the high performances of ST-GCN.

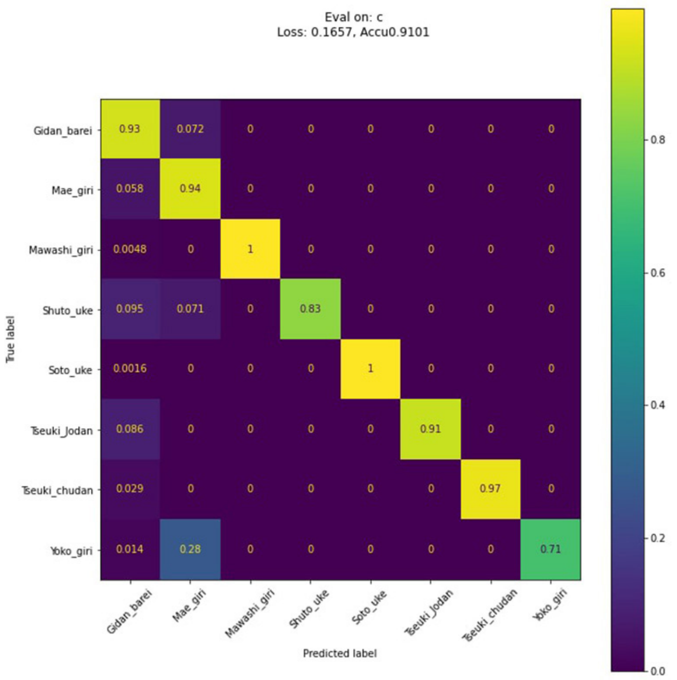

Fig. 23. Confusion Matrix – ST-GCN

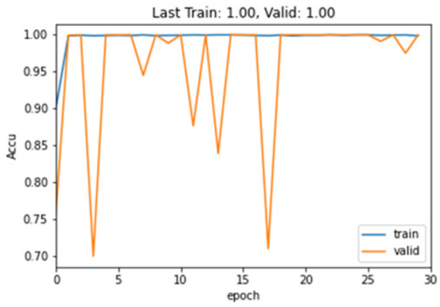

Fig. 24. Accuracy obtained using ST-GCN

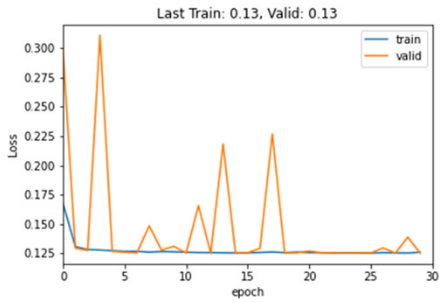

Fig. 25. Loss obtained using ST-GCN

We can notice that the valid test is fluctuating and there is a large gap between the test and valid test; this can be due probably to the number of instances (views) used in the project to train the model. The Figures 26, 27 and 28 represent the results achieved by using this technique. The Karateka performs a chosen movement, and the system will detect his body, classify the movement, and give a scoring which evaluates how the movement was performed referring to the coach's performance.

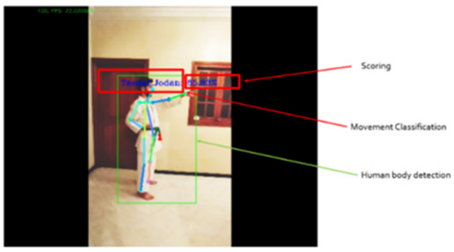

Fig. 26. Tseuki_Jodan movement classification and scoring

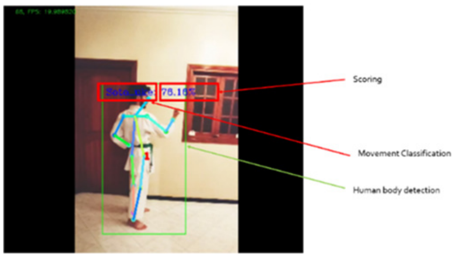

Fig. 27. Soto-Uke movement classification and scoring

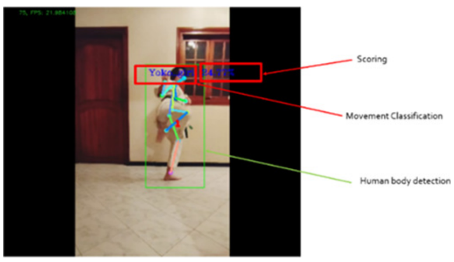

Fig. 28. Yoko Giri movement classification and scoring

4.3 Results comparison

Based on these results, the following conclusions can be obtained:

– The e-sport system can be implemented by several techniques such as LSTM, ST-GCN, and many others;

– When using LSTM, the efficiency of the built system achieved 96% overall with a stable system, which is a good precision for such problems;
– When using ST-GCN, the accuracy of the built system achieved 91% overall with some limitations;

The Table 3 compares our work to other works done in the same context but with different datasets:

Table 3. Results comparison

Ref	Year	Techniques	Metrics	Value
[19]	2017	DTW classfier with angle-based features	Overall recognition rate	94.2%
[20]	2018	Bidirectional & Unidirectional LSTM RNNs	Overall Accuracy (Uni)	86.7%
			Average Precision (Uni)	87.5%
			Average Recall (Uni)	86.7%
			F1 Score (Uni)	86.3%
			Overall Accuracy (Bi)	93.3%
			Average Precision (Bi)	95.0%
			Average Recall (Bi)	93.3%
			F1 Score (Bi)	93.1%
[15]	2020	Kinect v2 sensor and F-DTW	Accuracy	90%
[5]	2020	F-DTW	Accuracy	91.07%
Our work	2022	LSTM	Accuracy	**96%**
		ST-GCN	Accuracy	**91%**

4.4 3D coach with VIBE

As declared, we aim to represent the 8 Karate movements in 3D by using the VIBE technique. The Figure 29 shows Gidan_barei in 3D from 3 views. We noticed that this technique could not detect hand gesture, which is considered very important in Karate movements and makes difference between movements' evaluation.

Fig. 29. 3D Karate movements using VIBE

5 Conclusion

In this work, we have presented an automatic guided coach assistant for Shotokan Karate sport. The proposed approach is implemented based on LSTM and ST-GCN Deep Learning algorithms. The obtained results showed that these methods can estimate, classify, and score karate movements with high efficiency compared to other methods used in other research. The challenges faced in this work were to build a dataset with desired poses from different views (L_View, R_View, F_View) and to show the best approach to classify and score the movement. In addition, through this work, we propose a 3D Karate coach practicing the basic movements using the VIBE technique. As a perspective, we intend to generalize the program to be applicable on 3D images in real time to make a hybrid approach combining between LSTM and ST-GCN.

6 Data availability statement

The dataset and source code that support the findings of this study are available here: https://github.com/fatibennacer/e-Sport-and-Smart-Coaching-System-using-a-Multiview-Dataset.

7 Acknowledgements

This work was funded by the National Center for Scientific and Technical Research (CNRST-Rabat). We thank the expert coach, AIT BENNACER Abdellah, and all the karatekas from the Traditional Karate Shotokan Association in Beni-Mellal-Morocco who assisted in this work.

8 References

[1] M. Bishop, *Okinawan Karate: Teachers, Styles and Secret Techniques*. London, Eng.: A & C Black, 1993.
[2] D. Li and J. Zhang, "Computer Aided Teaching System Based on Artificial Intelligence in Football Teaching and Training," *Mob. Inf. Syst.*, vol. 2021, pp. 1–10, Nov. 2021, https://doi.org/10.1155/2021/9562782
[3] H. T. Sadeeq, T. H. Hameed, A. S. Abdi, and A. N. Abdulfatah, "Image Compression Using Neural Networks: A Review," *Int. J. Online Biomed. Eng. IJOE*, vol. 17, no. 14, Art. no. 14, Dec. 2021, https://doi.org/10.3991/ijoe.v17i14.26059
[4] M. E. Hoque and K. Kipli, "Deep Learning in Retinal Image Segmentation and Feature Extraction: A Review," *Int. J. Online Biomed. Eng. IJOE*, vol. 17, no. 14, Art. no. 14, Dec. 2021, https://doi.org/10.3991/ijoe.v17i14.24819
[5] B. Emad et al., "iKarate: Karate Kata Guidance System," *Procedia Comput. Sci.*, vol. 175, pp. 149–156, 2020, https://doi.org/10.1016/j.procs.2020.07.024
[6] S. J. Andajani, O. Wirawan, and P. Pamuji, "Development of Detector Vibrating Watch Model for 100-Meter Dash Race for Visually-impaired Athletes," *Int. J. Online Biomed. Eng. IJOE*, vol. 16, no. 10, Art. no. 10, Sep. 2020, https://doi.org/10.3991/ijoe.v16i10.15617

[7] S. Qiu et al., "Multi-Sensor Information Fusion Based on Machine Learning for Real Applications in Human Activity Recognition: State-of-the-Art and Research Challenges," *Inf. Fusion*, vol. 80, pp. 241–265, Apr. 2022, https://doi.org/10.1016/j.inffus.2021.11.006
[8] Y. Tu, L. Liu, M. Li, P. Chen, and Y. Mao, "A Review of Human Motion Monitoring Methods using Wearable Sensors," *Int. J. Online Biomed. Eng. IJOE*, vol. 14, no. 10, Art. no. 10, Oct. 2018, https://doi.org/10.3991/ijoe.v14i10.9313
[9] Y. Bai and Y. Chen, "Human Motion Analysis and Action Scoring Technology for Sports Training Based on Computer Vision Features," *J. Intell. Fuzzy Syst.*, no. Preprint, pp. 1–9, 2021, https://doi.org/10.3233/JIFS-219092
[10] S. Yan, Y. Xiong, and D. Lin, "Spatial Temporal Graph Convolutional Networks for Skeleton-Based Action Recognition," 2018, https://doi.org/10.1609/aaai.v32i1.12328
[11] O. Keskes and R. Noumeir, "Vision-Based Fall Detection Using ST-GCN," *IEEE Access*, vol. 9, pp. 28224–28236, 2021, https://doi.org/10.1109/ACCESS.2021.3058219
[12] K. Zhou, T. Wu, C. Wang, J. Wang, and C. Li, "Skeleton Based Abnormal Behavior Recognition Using Spatio-Temporal Convolution and Attention-Based LSTM," *Procedia Comput. Sci.*, vol. 174, pp. 424–432, 2020, https://doi.org/10.1016/j.procs.2020.06.110
[13] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 1, pp. 172–186, Jan. 2021, https://doi.org/10.1109/TPAMI.2019.2929257
[14] W. M. R. W. Idris, A. Rafi, A. Bidin, and A. A. Jamal, "A Theoretical Framework of Extrinsic Feedback Based-Automated Evaluation System for Martial Arts," *Int. J. Eng. Technol.*, vol. 7, no. 2.14, pp. 74–79, 2018, https://doi.org/10.14419/ijet.v7i2.14.11160
[15] B. Emad et al., "iKarate: Improving Karate Kata," *Procedia Comput. Sci.*, vol. 170, pp. 466–473, 2020, https://doi.org/10.1016/j.procs.2020.03.090
[16] F. Zhang, X. Zhu, and M. Ye, "Fast Human Pose Estimation," *ArXiv181105419 Cs*, Apr. 2019, Accessed: Feb. 26, 2022. [Online]. Available: http://arxiv.org/abs/1811.05419
[17] J. Shi, D. Peng, Z. Peng, Z. Zhang, K. Goebel, and D. Wu, "Planetary Gearbox Fault Diagnosis using Bidirectional-Convolutional LSTM Networks," *Mech. Syst. Signal Process.*, vol. 162, p. 107996, Jan. 2022, https://doi.org/10.1016/j.ymssp.2021.107996
[18] M. Kocabas, N. Athanasiou, and M. J. Black, "VIBE: Video Inference for Human Body Pose and Shape Estimation," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, Jun. 2020, pp. 5252–5262, https://doi.org/10.1109/CVPR42600.2020.00530
[19] T. Hachaj, M. Piekarczyk, and M. Ogiela, "Human Actions Analysis: Templates Generation, Matching and Visualization Applied to Motion Capture of Highly-Skilled Karate Athletes," *Sensors*, vol. 17, no. 11, p. 2590, Nov. 2017, https://doi.org/10.3390/s17112590
[20] S.-M. Lim, H.-C. Oh, J. Kim, J. Lee, and J. Park, "LSTM-Guided Coaching Assistant for Table Tennis Practice," *Sensors*, vol. 18, no. 12, p. 4112, Nov. 2018, https://doi.org/10.3390/s18124112

9 Authors

Fatima-Ezzahra Ait-Bennacer (corresponding author) is a Ph.D. candidate in the Computer Science Department of the University of Chouaib Doukkali, Faculty of sciences El Jadida in Morocco. She received a Master's degree in Techno-pedagogical engineering from Abdelmalek Essaâdi University Tétouan in 2017. Her research interests include Digital Transformation, Higher Education Institutions, Total Quality Management, and Artificial Intelligence. E-mail: aitbennacer.fati@gmail.com

Abdessadek Aaroud is a Professor at the Faculty of sciences El Jadida in Morocco, and member of LAROSERI Laboratory, and head of the computer sciences department. His research interests are the specification and verification of real-time systems, the use of a reactive agent approach for modeling, and real-time temporal logic as a formal method. E-mail: aaroud.a@ucd.ac.ma

Khalid Akodadi is a Ph.D. candidate in the Computer Science Department of the University of Chouaib Doukkali, Faculty of sciences El Jadida in Morocco. His research interests include Semantic middleware, Systems Analysis and Control, Artificial Intelligence, and neural networks. He currently works as a Technical Manager in 2M TV Channel. E-mail: akodadi.khalid@gmail.com

Bouchaib Cherradi is a Professor of computer science in CRMEF-El Jadida. He is associate member of Signals, Distributed Systems and Artificial Intelligence (SSDIA) Laboratory in ENSET Mohammedia, Hassan II University Casablanca (UH2C), and LaROSERI Laboratory on leave from the Faculty of Science El Jadida (Chouaib Doukali University), Morocco. His research interests include Massively Parallel Architectures, Cluster Analysis, Pattern Recognition, Image Processing and Fuzzy Logic systems. E-mail: bouchaib.cherradi@gmail.com

Article submitted 2022-03-15. Resubmitted 2022-05-20. Final acceptance 2022-06-01. Final version published as submitted by the authors.