Runners detection based on trajectory analysis

Jonay Sarmiento-Ramírez
Escuela de Ingeniería Informática
Universidad de Las Palmas de Gran Canaria
Las Palmas de G.C., Spain
jonay.sarmiento102@alu.ulpgc.es

Javier Lorenzo-Navarro*
Instituto Universitario SIANI
Universidad de Las Palmas de Gran Canaria
Las Palmas de G.C., Spain
javier.lorenzo@ulpgc.es (0000-0002-2834-2067)

Modesto Castrillón-Santana
Instituto Universitario SIANI
Universidad de Las Palmas de Gran Canaria
Las Palmas de G.C., Spain
modesto.castrillon@ulpgc.es (0000-0002-8673-2725)

Elena Sánchez-Nielsen
Dpto. Ingeniería Informática y de Sistemas
Universidad de La Laguna
La Laguna, Spain
enielsen@ull.edu.es (0000-0003-2114-4137)

***Abstract*—In some outdoor sports events, spectators and participants are not located in clearly different ubications. One example of such events is any ultratrail competition where both the spectators and runners share the space, with just some limitations in aid stations, where the organization staff may be present. An automatic visual-based runner detection could be of great help in mass events to support existing timing systems. For this aim, a challenging issue is to differentiate between participants and the rest in any recorded locations because sometimes the race bib number, which is a distinguishing cue, is not visible. In this work, the trajectories of the people in the scenario are explored as an element to solve the problem. Firstly, the trajectories are extracted combining YOLOV4 person detector and DeepSort tracker, and later manually annotated as runner or spectator. Then, a set of features are extracted from the trajectories and different classification methods are tested. For the experiments, 26 videos of a real ultra-trail race were used and a total of 756 trajectories were extracted. The best result was obtained with a two-level stacked LTSM that yields an accuracy of 98%.**

***Index Terms*—Computer Vision, Trajectory Classification, Sports analysis.**

I. Introduction

Computer Vision techniques have been widely applied in the sports context, as evidenced by the publication of recent surveys on the topic [1], [2], and the regular organization of workshops related such as CVsports and MMSports. Given the huge budget present in some highly demanded sports, there are many applications focused on the achievement of statistics related to team players [2], [3], but also in providing ways to improve the individual athlete performance [2], [4]–[6]. In both cases, the aim is to increase team or individual performance. Other Computer Vision techniques, such as tracking, have evidenced a large reliability improvement in the last years, and therefore introduced as an aid for assisting ball sports referees [7], [8]. However, biometric techniques as those related to people identification or re-identification, have not been exploited to aid the organizers of long-term sports, such as ultra-running, in detecting fraudulent actions as a mismatch between the runner and the registered person, or if a race bib number (RBN) is carried out by different runners during the course. A characteristic of long-term sports, which sometimes last more than 12 hours, is the fact that participants can change their clothes and wear different accessories like glasses or cap during the course. In this sense, biometrics can aid in runner identification but, as mentioned above, it has hardly been applied in the described scenario.

Adopting a supervised scenario, annotated data are necessary. Typical captures from ultra-running competitions may contain participants, staff, spectators, etc. Assuming video acquisition, the first problem to solve is to distinguish between runners from the rest as this kind of event takes place outdoors where the spectators can be very close to the runners (Figure 1). Detection of RBN can be considered as a solution because only runners wear it, but as stated before, in ultra-running events participants might wear outfits that occlude the RBN or after many hours of race the RBN can be folded or deteriorated (Figure 2). In this paper, we explore an approach based on the different behavior shown by runners and spectators when video footage is available. Runners normally exhibit a particular motion pattern, keeping a constant velocity/speed and moving in a defined path, unlike spectators that are static or wandering.

The contributions of the work are: 1) the annotation of people tracklets as runner/non-runner captured during a real ultra-running competition, 2) the exploration of features extracted from their extraction, and 3) the exploration of different classification methods to discriminate between both classes.

II. Related works

As previously stated, tracking players allows the analysis of tactics in team sports like basketball, soccer, or football. In [9], Chen et al. present a method based on player tracking with a

This work has been partially funded by the Spanish Ministry of Economy and Competitiveness (MINECO) under project RTI2018-093337-B-I00, and the Gobierno de Canarias and the FEDER funds under project ProID2020010024. We would also like to thank Arista Eventos S.L.U. to allow us to record images during Transgrancanaria 2020 and grant us their usage for research purposes.

*Corresponding author

Fig. 1. Spectators and runners in an ultra-marathon trail race. ©Alexis Martín (Creative Common License CC BY-NC 2.0)

Fig. 2. Two situations where RBN recognition will fail: RBN partially occluded (top) and folded RBN (bottom)

Kalman filter to recognize three of the most frequently used screen strategies in basketball using a set of rules. Also for basketball matches, Chen et al. [10] make use of a graph-based tracking method that can deal with player occlusions though only trajectories are obtained and no further analysis is carried out. Even if the players are not tracked to get tactics but the ball 3-D position estimated from 2-D, Chen et al. [11] propose a method to classify nine set types in volleyball games. In [12], Gómez et al. compare two tracking methods for both volleyball players and other two methods for ball tracking concluding that integral histogram and trajectory growth exhibit the best performance respectively. Deep learning approaches have been also introduced in this task. In [13], Schwenkreis maps the positions on the court (x,y) of the handball players during 15 seconds in a matrix of dimension 30 x 900, adopting a Convolutional Neural Network (CNN) to classify into 80 tactical moves. In Stein et al. [14], a technique for analysis of soccer is proposed where three types of relevant analysis are realized: region-based, event-based and player-based. Apart from sports, trajectory analysis is of interest in other fields as predicting the behavior of pedestrians in streets [15] or for assisted and autonomous driving [16]. In [17], Turchini et al. describe an unsupervised technique based on Landmark Based Spectral Clustering (LSC) of previously extracted trajectories with improved trajectories (IDT). Turchini's work aims to find a similarity measure between human activities in videos instead of classifying those activities as we propose in this work.

The application in ultra-running is limited, basically focused on identifying participants [18], [19]. The identification problem has certainly been already considered in the marathon scenario, even if the conditions are more controlled, compared to ultra-running, the literature is wider. In most cases, the Racing Bib Number (RBN) recognition strategy has been adopted for identification or re-identification [20], with rare attempts using facial or body appearance [19], [21]. Starting from the work described in [22], which includes the most used RBN benchmark, other approaches have recently been presented proposing different solutions using handcrafted features and deep learning, sometimes after a previous face, upper body, or body detection [23]–[25], to restrict the region of interest, or directly searching for RBNs [26]–[28]. Unlike marathon competitions where the spectators are placed separated from the runners, normally on the sidewalk, in the ultra-running scenario, those restrictions are in general more relaxed, as there are not big limitations because the sporting event takes place mainly along trails and paths, see Figure 1.

III. Methodology

A. Dataset

The experiments described below were carried out using some footage of the TGC2020 dataset [19], which was gathered during a real ultra-running competition held in March 2020. Runners taking part in Transgrancanaria (TGC) Classic 128KM in its 2020 edition left from the start line at 11 pm on March 6, to reach the finish line before its closure 30 hours later.

Participants were captured in six different recording points (RPs) (RP0-RP5) along the race track. For such purpose, a Sony Alpha ILCE 6400 (16-50mm lens and configured at 1920x1080@50fps) was employed at the different spots in continuous operation for a limited time, see Table I.

In locations RP1, RP4, and RP5, the spectators are close to the runners' track and in some cases, they walk on the race path (Figure 3 left) or even run near her/him (Figure 3 right). Additionally in RP1, the illumination conditions are not ideal because it was night (Figure 3 right) and the detection of RBN does not exhibit a reliable performance [18].

In total, 26 fragments of video from RP1, RP4, and RP5 were used to extract the trajectories of the people to be analyzed. YOLOv4 [29] and DeepSORT [30] were adopted for people detection and for tracking the detected person frame by frame, respectively. The number of trajectories obtained after this stage was 756, of which 655 correspond to the spectators and 101 to runners. For each frame, labels and bounding boxes for each detected person are obtained. The number of frames contained in the trajectories is 224,033, corresponding to 180,810 to spectators and 43,223 to runners. In Figure 4 it

TABLE I
RECORDING POINTS LOCATIONS AND RECORDING TIMETABLE.

Location	Timing Point	Km	Start Rec. Time	Footage (frames)
RP1	yes	16.5	00:06	140,616
RP2	yes	27.9	01:08	432,624
RP3	no	84.2	07:50	667,872
RP4	yes	110.5	10:20	1,001,208
RP5	yes	124.5	11:20	1,462,056

Fig. 3. Top: Spectators walking on the race path (RP4). Bottom: Spectator running close to the runner to record him (RP1).

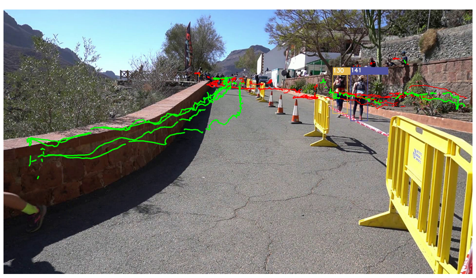

Fig. 4. Examples of labeled trajectories of runners (green) and spectators (red).

can be seen examples of runners (green) and spectators (red) trajectories for a clip recorded at RP4.

With a frame rate of 50 fps, those trajectories with less than 100 frames were ruled out because they correspond to trajectories that last less than 2 seconds which can be produced by artifacts or detection failures. After the previous preprocessing, 385 trajectories were obtained, corresponding to 301 to spectators and 84 to runners.

B. Trajectory features

As it was described in Section III-A, for each person the id and the bounding box in the frame are obtained with the DeepSORT method. From these inputs, the following handcrafted features are computed:

- Height: For each frame i-th, the height of the bounding box for each detected person id is obtained.
- Speed: For each frame i-th, the speed of each detected person id in the image is computed as the distance between the bounding box centroid of person id in the current frame, $centr_{bb}(id, fr_i)$, and the previous one divided by the inverse of the frame rate.

$$speed_{id}(i) = \frac{||centr_{bb}(id, fr_{i-1}), centr_{bb}(id, fr_i)||_2}{1/fps} \tag{1}$$

- Average speed: To avoid artifacts, the average speed of the five previous frames is also computed.
- Direction: For each frame i-th, the direction of the person id in the image is computed from the bounding box centroid of person id in the current frame, $centr_{bb}(id, fr_i)$, and the previous one.

C. Classical Approach

A first attempt to differentiate runners from spectators using trajectory features in the videos is performed with classical Machine Learning methods. These methods accept a descriptor of fixed length known as dimensionality. The number of frames of the trajectories differs depending on the duration, so it is necessary to condense the information of the whole trajectory into a fixed size descriptor. To this end, a Bag of Words approach is implemented and obtaining the codebook by discretizing the features described in Section III-B. Therefore, the speed and the height were discretized into 10 and 11 bins, respectively, using an equal-frequency strategy, and the direction feature was mapped into the eight main cardinal points (N, S, W, E, NW, NE, SW, SE). Making use of the computed codebook, each frame is mapped to a 29-dimension feature vector (codeword). Finally, the trajectory is represented by the histogram of the codewords of all the frames that made up the trajectory.

D. Deep Learning Approach

A trajectory by definition is a temporal signal that gives the position of the person along the time, so a method that takes into account this fact would be the logical approach. In this sense, Recurrent Neural Networks (RNNs) are designed to deal with temporal signals, therefore a Long Term Short Memory (LSTM) [31] is evaluated for classification. Thus, the dataset was converted into sequences that feed the LSTM, and those sequences must have the same length. The size of the sequence must be a tradeoff between being as long as possible to capture the different behaviors of runners and non-runners,

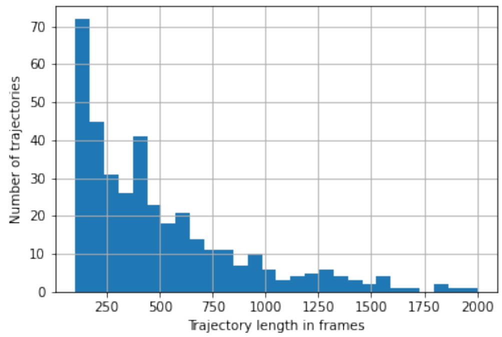

Fig. 5. Number of trajectories for different duration in frames

and not so long that the number of valid sequences in the dataset doesn't allow to train the network. Figure 5 shows the number of trajectories for different ranges of duration in frames, excluding those with less than 100 frames. A value of 400 frames was chosen so trajectories are restricted to have a length of 400 frames which implies sequences with a duration of 8 seconds. To compute these sequences, longer trajectories were divided into 400 frame sequences with a stride of 50 frames. Incomplete sequences were padded with zeros. In this case, the number of samples (sequences) is 3,218, of which 621 correspond to runners and 2,597 to spectators.

IV. Experiments and Results

The experimental setup for this problem was a holdout with 70% for training and 30% for testing. The classification methods used for the classical approach were: Decision Tree [32], Random Forest [33], Gradient Boosting [34], Naive Bayes [35] and Logistic Linear Regression. The hyperparameters for each method were: Gini criterion for the Decision Tree, 100 trees for the Random Forest, and logistic regression as loss function, and 100 stump trees as the base classifier for the Gradient Boosting.

As a reference, we perform a naive experiment to classify each frame as corresponding to spectator/runner. The frames are characterized with the features described in Section III-B. In this setup, the temporal component is not taken into account because each frame is considered independently of the rest. Therefore, in a trajectory, some frames may be classified as runner and the rest as spectator. It will serve as a baseline because no temporal information is utilized for classifying each person as runner/non-runner. In this setup, the training set is composed of 156,823 samples (126,584 spectators, 30,239 runners), and the test set is composed of 67,210 samples (54,226 spectators, 12,984 runners). Results are shown in Table II. It can be observed that even with no temporal information, an accuracy of 89% in classifying each frame as runner/spectator is achieved with a Random Forest classifier exhibiting a balanced performance in both classes according to precision and recall. The confusion matrix for the Random Forest is shown in Figure 6. It can be observed that the

TABLE II
Results of classical methods for frames (no trajectories).

Classifier	Accuracy	Precision	Recall
Decision Tree	0.84	0.83	0.84
Random Forest	**0.89**	**0.88**	**0.89**
Gradient Boosting	0.85	0.86	0.85
Naive Bayes	0.74	0.74	0.74
Logistic Regression	0.80	0.76	0.81

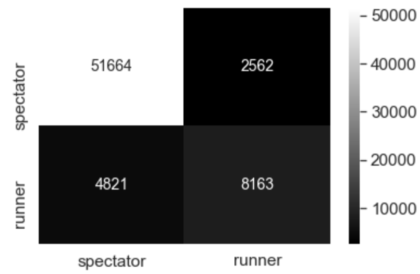

Fig. 6. Confusion matrix for Random Forest using frames.

TABLE III
Results of classical methods for trajectories using BoW approach.

Classifier	Accuracy	Precision	Recall
Decision Tree	0.75	0.80	0.75
Random Forest	**0.91**	**0.92**	**0.91**
Gradient Boosting	0.91	0.90	0.91
Naive Bayes	0.49	0.77	0.49
Logistic Regression	0.78	0.78	0.78

number of false positives for the runner class is almost the 40% of the runner frames as expected due that many frames that correspond to the runner trajectories are similar to the spectator frames. Though it has not been included in this work, the classification of the person can be obtained with a majority strategy, if the majority of the person's trajectory frames correspond to runner then is classified as a runner, otherwise is classified as a spectator.

When the trajectory information is condensed into a histogram following a BoW approach, two classifiers, Random Forest and Gradient Boosting, exhibit the highest performance with an accuracy of 91% (Table III), which implies an increase of a 2% over the baseline. Though both classifiers match in accuracy and recall, Random Forest improves a little in precision over Gradient Boosting. To note that in this case, the training and test set contains 269 samples (209 spectators, 60 runners) and 116 samples (92 spectators, 24 runners) respectively. Figure 7 shows the confusion matrix for the Random Forest classifier when the BoW approach is used for representing the trajectories. In this case, the percentage of false positives for the runner class is very similar to the previous setup which is due to the loss of temporal information because with the histogram the adjacency information among frames is lost.

Finally, to evaluate the use of the RNNs, the experimental

TABLE IV
RESULTS OF DIFFERENT LSTM TOPOLOGIES.

Topology	1st layer	2nd layer	FC layer	Accuracy	Precision	Recall
Single LSTM	32	-	32	0.92	0.87	0.68
Stacked LSTM	64	32	32	**0.98**	**0.95**	**0.93**
Stacked LSTM	64	64	64	0.96	0.88	0.87
Stacked LSTM	128	32	32	0.95	0.88	0.84

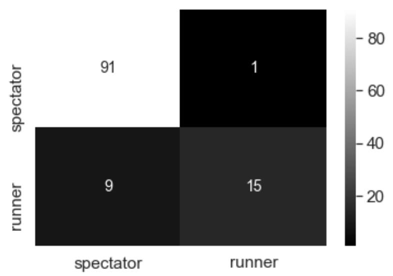

Fig. 7. Confusion matrix for Random Forest using BoW approach.

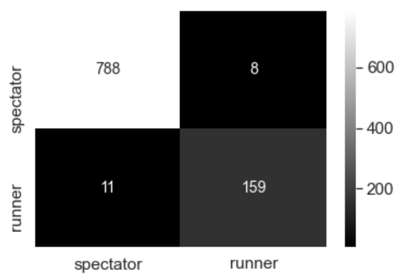

Fig. 8. Confusion matrix for Stacked LSTM (64-32 + 32)

setup was a holdout with 70% for training and 30% for testing, as with the classical approach. The training and test set has 2,252 samples (1,631 spectators, 621 runners) and 966 samples (796 spectators, 170 runners), respectively. The features considered as input to the neural network are the height, speed, average speed, and the eight main cardinal points. The three continuous features were normalized to the range $[0, 1]$. The chosen RNN architecture has been an LSTM and different topologies were considered. The upper layer of the LSTM feeds a fully connected layer with a sigmoid output function which is suitable for a bi-class problem as the one at hand. For all the configurations, the loss function was the binary cross-entropy and trained using the Adam optimizer with a learning rate of 0.001. To reduce the overfitting risk, a dropout of 0.5 was introduced in the training stage. As it is shown in Table IV, an accuracy of 98% was obtained with a two-level stacked LTSM with 64 cells for the first level, 32 for the second one, and 32 units in the fully connected layer. This result is a noticeable improvement over the best results obtained with classical approaches. As it can be observed in Figure 8, the confusion matrix for the best LSTM configuration shows a more balanced behavior in both classes as it was expected in comparison with the two previous experimental setups.

V. CONCLUSIONS

This work has presented the application of two different approaches to solving the problem of differentiating automatically runners from non-runners in ultra-marathon scenarios where the spectators may be close to participants, adding difficulties not present in other sports where participants and spectators occupy separated spaces. The two explored approaches are classical classifiers on the one hand, and RNNs on the other hand. In both cases, hand-crafted features obtained from the spectators and runners trajectories were used as input. As expected, the performance of RNNs, with an accuracy of 98%, was significantly superior to the classical approach because the temporal component is implicit in this kind of neural network. With the classical approach, only isolated frames can be considered or using the BoW approach to deal with variable length trajectories characterizing them with a fixed size descriptor. The experiments were carried out with real footage of a ultra-trail competition that covered both night and day scenarios.

REFERENCES

[1] T. B. Moeslund, G. Thomas, and A. Hilton, *Computer Vision in Sports*. Switzerland: Springer, 2014.

[2] G. Thomas, R. Gade, T. B. Moeslund, P. Carr, and A. Hilton, “Computer vision for sports: Current applications and research topics,” *Computer Vision and Image Understanding*, vol. 159, pp. 3 – 18, 2017.

[3] H. Liu and B. Bhanu, “Pose-guided r-cnn for jersey number recognition in sports,” in *Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Long Beach, CA, USA: IEEE, 2019, pp. 2457–2466.

[4] G. Li and C. Zhang, “Automatic detection technology of sports athletes based on image recognition technology,” *EURASIP Journal on Image and Video Processing*, vol. 2019, no. 1, p. 15, Jan 2019.

[5] C. Ning, “Design and research of motion video image analysis system in sports training,” *Multimedia Tools and Applications*, March 2019.

[6] A. Rajšp and I. Fister, “A systematic literature review of intelligent data analysis methods for smart sport training,” *Applied Sciences*, vol. 10, no. 9, 2020. [Online]. Available: https://www.mdpi.com/2076-3417/10/9/3013

[7] P. Kurowski, K. Szelag, W. Zaluski, and R. Sitnik, “Accurate ball tracking in volleyball actions to support referees,” *Opto-Electronics Review*, vol. 26, no. 4, pp. 296–306, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1230340218301045

[8] M. Labayen, I. G. Olaizola, N. Aginako, and J. Florez, “Accurate ball trajectory tracking and 3d visualization for computer-assisted sports broadcast,” *Multimedia Tools and Applications*, vol. 73, pp. 1819–1842, 2014.

[9] H.-T. Chen, C.-L. Chou, T.-S. Fu, S.-Y. Lee, and B.-S. P. Lin, “Recognizing tactic patterns in broadcast basketball video using player trajectory,” *Journal of Visual Communication and Image Representation*, vol. 23, no. 6, pp. 932–947, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1047320312000995

[10] L.-H. Chen, C.-W. Su, and H.-A. Hsiao, “Player trajectory reconstruction for tactical analysis,” *Multimedia Tools and Applications*, vol. 77, no. 23, p. 30475–30486, Dec. 2018.

[11] H.-T. Chen, W.-J. Tsai, S.-Y. Lee, and J.-Y. Yu, "Ball tracking and 3d trajectory approximation with applications to tactics analysis from single-camera volleyball sequences," *Multimedia Tools and Applications*, vol. 60, no. 3, p. 641–667, Oct. 2012.

[12] G. Gomez, P. H. López, D. Link, and B. Eskofier, "Tracking of ball and players in beach volleyball videos," *PLoS ONE*, vol. 9, 2014.

[13] F. Schwenkreis, "An approach to use deep learning to automatically recognize team tactics in team ball games," in *Proceedings of the 7th International Conference on Data Science, Technology and Applications (DATA 2018)*, Porto, Portugal, 2018, pp. 157–162.

[14] M. Stein, H. Janetzko, A. Lamprecht, T. Breitkreutz, P. Zimmermann, B. Goldlücke, T. Schreck, G. Andrienko, M. Grossniklaus, and D. A. Keim, "Bring it to the pitch: Combining video and movement data to enhance team sport analysis," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 1, pp. 13–22, 2018.

[15] J. Liang, L. Jiang, J. Carlos Niebles, A. Hauptmann, and L. Fei-Fei, "Peeking into the future: Predicting future person activities and locations in videos," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2019.

[16] A. Bhattacharyya, M. Fritz, and B. Schiele, "Long-term on-board prediction of people in traffic scenes under uncertainty," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[17] F. Turchini, L. Seidenari, and A. Del Bimbo, "Understanding and localizing activities from correspondences of clustered trajectories," *Computer Vision and Image Understanding*, vol. 159, pp. 128–142, 2017, computer Vision in Sports.

[18] P. Hernández-Carrascosa, A. Penate-Sanchez, J. F. O. Lorenzo-Navarro, and M. David; Castrillón-Santana, "TGCRBNW: A dataset for runner bib number detection (and recognition) in the wild," in *Proceedings International Conference on Pattern Recognition*, Milan, Italy, 2020.

[19] A. Penate-Sanchez, D. Freire-Obregón, A. Lorenzo-Melián, J. Lorenzo-Navarro, and M. Castrillón-Santana, "TGC20ReId: A dataset for sport event re-identification in the wild," *Pattern Recognition Letters*, vol. 138, pp. 355–361, 2020.

[20] P. X. Kamlesh, Y. Yang, and Y. Xu, "Person re-identification with end-to-end scene text recognition," in *Chinese Conference on Computer Vision*. Tianjin, China: Springer, 2017, pp. 363–374.

[21] A. Wrońska, K. Sarnacki, and K. Saeed, "Athlete number detection on the basis of their face images," in *Proceedings International Conference on Biometrics and Kansei Engineering*. Kyoto, Japan: IEEE, 2017, pp. 84–89.

[22] I. Ben-Ami, T. D. (Basha), and S. Avidan, "Racing bib number recognition," in *British Machine Vision Conference*. Surrey, UK: British Machine Vision Association, 2012, pp. 1–10.

[23] N. Boonsim, "Racing bib number localization on complex backgrounds," *WSEAS Transactions on Systems and Control*, vol. 13, pp. 226–231, 2018.

[24] S. Nag, R. Ramachandra, P. Shivakumara, U. Pal, T. Lu, and M. Kankanhall, "Crnn based jersey-bib number/text recognition in sports and marathon images," in *International Conference on Document Analysis and Recognition (ICDAR)*. Sydney, Australia: IEEE, 2019, pp. 1149–1156.

[25] P. Shivakumara, R. Raghavendra, L. Qin, K. B.Raja, T. Luc, and U. Pal, "A new multi-modal approach to bib number/text detection and recognition in marathon images," *Pattern Recognition*, vol. 61, pp. 479–491, 2017.

[26] A. Apap and D. Seychel, "Marathon bib number recognition using deep learning," in *11th International Symposium on Image and Signal Processing and Analysis (ISPA)*. Dubrovnik, Croatia: IEEE, 2019, pp. 21–26.

[27] W. M. de Jesús and D. L. Borges, "An improved stroke width transform to detect race bib numbers," in *Proceedings of the Mexican Conference on Pattern Recognition*. Puebla, Mexico: Springer, 2018, pp. 267–276.

[28] Y. C. Wong, L. J. Choi, R. S. S. Singh, H. Zhang, and A. R. Syafeeza, "Deep learning based racing bib number detection and recognition," *Jordanian Journal of Computers and Information Technology (JJCIT)*, vol. 5(3):, no. 3, pp. 181–194, 2019.

[29] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv*, vol. 2004.10934, 2020. [Online]. Available: https://arxiv.org/abs/2004.10934

[30] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," in *2017 IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 3645–3649.

[31] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[32] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, pp. 81–106, 1986.

[33] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct 2001.

[34] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics*, pp. 1189–1232, 2001.

[35] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification (2nd Edition)*. Wiley-Interscience, 2001.