An Evaluation of General-Purpose Optical Character Recognizers and Digit Detectors for Race Bib Number Recognition

Modesto Castrillón-Santana[1][a], David Freire-Obregón[1][b], Daniel Hernández-Sosa[1][c], Oliverio J. Santana[1][d], Francisco Ortega-Zamorano[2][e], José Isern-González[1][f] and Javier Lorenzo-Navarro[1][g]

[1]*SIANI, Universidad de Las Palmas de Gran Canaria, Spain*

[2]*Universidad de Málaga, Spain*

Keywords: Race Bib Number Recognition, OCR, General Object Detection.

Abstract: Bib numbers are used in mass competitions to identify participants, especially in long-distance races where runners commonly wear tags to verify that they pass mandatory checkpoints. In this paper, we delve deeper into the use of existing computer vision techniques for recognizing the digits present in bib numbers. Our analysis of bib recognition involves evaluating OCRs (Optical Character Recognition) techniques and a YOLOv7 digit detector on two public datasets: RBNR and TGCRBNW. The results reveal that the former scenario is solvable, while the latter presents extremely in-the-wild challenges. However, the findings suggest that more than relying solely on RBN for runner identification, other appearance-based cues, e.g., clothing and accessories, may be required due to various circumstances, such as occlusion or incomplete bib recognition. In any case, all those cues do not necessarily imply that the same person is wearing the RBN across the competition track, as they are not biometric traits.

1 INTRODUCTION

Since the early 90s, timing systems have been an essential tool for organizers of massive running events to keep track of runners at specific locations along the course track. However, these systems do not control the real presence of the runner, but rather the presence of the tag they carry, and strictly, they do not verify the participant's identity. This has led to incorrect classification results and insurance policy problems for organizers. The computer vision community has started to address the participant identification problem, mostly focusing on solving the Racing Bib Number (RBN) automatic recognition problem (Ben-Ami et al., 2012), i.e., recognizing the precise number exhibited in the bib. However, this solution suffers from the same drawbacks as tag-based systems. Biometrics could be an alternative, but it has yet to be applied more extensively in this challenging scenario.

In this paper, we do not make use of biometrics. Instead we explore using state-of-the-art computer vision techniques to recognize RBNs in running competitions, considering real-life datasets used in the literature to which we have been granted access. The scenario comprises significant challenges, such as variations in illumination conditions, runner's image resolution, body pose, and image sharpness, among others. RBN recognition is a valid modality in a flexible multimodal approach for runner recognition, which should also integrate biometric traits, such as face and body appearance or gait recognition. However, all those mentioned modalities present the risk of being unavailable at a particular moment. The mid-term expectations offer exciting possibilities for improving participant identification and can be a game-changer for the running event industry.

In summary, the main contribution of this paper is evaluating existing RBN recognition benchmarks, defining a new digit-based instead of RBN-based metric focus, and proposing relevant features for real-world or in-the-wild benchmarks.

[a] https://orcid.org/0000-0002-8673-2725
[b] https://orcid.org/0000-0003-2378-4277
[c] https://orcid.org/0000-0003-3022-7698
[d] https://orcid.org/0000-0001-7511-5783
[e] https://orcid.org/0000-0002-4397-2905
[f] https://orcid.org/0000-0001-5830-7732
[g] https://orcid.org/0000-0002-2834-2067

Castrillón-Santana, M., Freire-Obregón, D., Hernández-Sosa, D., Santana, O., Ortega-Zamorano, F., Isern-González, J. and Lorenzo-Navarro, J.
An Evaluation of General-Purpose Optical Character Recognizers and Digit Detectors for Race Bib Number Recognition.
DOI: 10.5220/0012562400003654

In *Proceedings of the 13th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2024)*, pages 910-917
ISBN: 978-989-758-684-2; ISSN: 2184-4313

2 RELATED WORK

In sports, computer vision is an active field tackled by recent surveys (Mendes-Neves et al., 2023) and well-established workshops such as CVsports and MM-Sports. Several studies have primarily focused on analyzing the individual athlete's movements to find ways of improvement (Li and Zhang, 2019; Thomas et al., 2017) or studying team sports to collect statistical data, increasing the team's performance (Liu and Bhanu, 2019; Thomas et al., 2017). Despite the active research in computer vision for sports, our focus is diverse, as in this paper we are just interested in identifying mass runner competition participants by their RBN. This section summarizes existing proposals for recognizing runners using computer vision.

2.1 RBN Detection and Recognition

As previously discussed, RBN recognition is the main approach for participant identification in running competitions, with most previous research primarily focusing on marathon-like scenarios (Kamlesh et al., 2017). Although such competitions are now organized everywhere, the amount of publicly available data is limited. The RBN dataset most widely used as a benchmark in the literature (Ben-Ami et al., 2012) includes 217 high-quality images captured by professional photographers, with 290 annotated RBNs. The proposed approach detects faces in the images and estimates the RBN Region of Interest (RoI) to detect text using the Stroke Width Transform (SWT) and Optical Character Recognition (OCR). Annotated RBNs correspond to those individuals whose face was correctly located by a face detector, being eligible to estimate an RBN RoI. This means that other RBNs may be present in the dataset images.

The approach described in (Boonsim, 2018) also adopts the initial face detection step but focuses on morphological operations applied to the torso area of the detected runners. In (Shivakumara et al., 2017), upper body detection is performed and the runner is segmented using GrabCut to restrict the text detection and recognition area. This approach is evaluated on their own dataset of 212 images captured in marathon competitions. The same research group has more recently described a unified method for detecting text from marathon runners and sports players in video, achieving high accuracy in text extraction in that scenario (Nag et al., 2020). In (de Jesús and Borges, 2018), the focus is on text detection, skipping the previous face/body detection, which may undoubtedly be affected by the runner's poses, reducing the number of possible RBNs.

In the past few years, the community has explored using deep learning for RBN recognition. The approach in (Wong et al., 2019) uses the convolutional network YOLO (You Only Look Once) (Bochkovskiy et al., 2020) for RBN detection and a convolutional recurrent neural network (CRNN) for text recognition. A similar approach is described in (Apap and Seychel, 2019), where a convolutional neural network (CNN) is used for RBN detection, followed by a CRNN for classification. Both approaches report results on the different literature datasets. The YOLO architecture is also adopted in (Carty et al., 2021) for RBN detection, applying a Tesseract OCR for the bib number recognition, which achieves promising results on a private dataset. In (Nag et al., 2019), runners are detected using a Single Shot Multibox Detector (SSD), and their body parts are extracted to locate the RBN. Finally, a CRNN is applied for classification. In (Kamlesh et al., 2017), a re-identification approach based on RBN recognition is applied, using TextBoxes (Minghui Liao and Bai, 2018) for RBN detection and a CRNN for classification. This approach is evaluated on their non-public dataset.

RBN digit recognition may be considered a subproblem of the overall problem of text recognition in natural scene images and video, which usually needs the previous step to spot the pieces of text to be located and later decode each separate character. The literature describes two main approaches for text recognition: 1) the combination of a recurrent neural network (RNN) on top of a CNN and a connectionist temporal classification (CRNN+CTC), and 2) the combination of a CNN and an attention RNN (Luo et al., 2019). Tesseract, one of the most widely used OCRs, adopted the first approach, i.e., the combination of RNN on top of CNN and CTC for RBN recognition, as mentioned earlier. The use of the two-steps method applied to RBN recognition is presented in (Ivarsson and Mueller, 2019), reporting the performance in datasets captured during marathon-like events. The FOTS (Fast Oriented Text Spotting) method proposed in (Liu et al., 2018) simultaneously detects and recognizes text, in contrast to the previous methods, which perform these tasks separately. This is done because there is a high correlation between the features used for detection and recognition.

2.2 Biometrics

Although body information has been used to determine the Region of Interest (RoI) for RBNs through face, upper body, or body part location. The use of facial and body recognition or clothing appearance, to determine the identity of runners has rarely been employed. One of the few references that combine facial appearance with RBN recognition for improving identification performance is (Wrońska et al., 2017). This multimodal strategy may be well-suited for the task at hand, as both face and RBN occlusions are frequently present, and the target pose may introduce difficulties for both face and RBN detection.

While the marathon-like scenario, with daylight conditions, reduces the presence of wild variations in the problem dataset, the re-identification in the ultra-running scenario proposed in (Penate-Sanchez et al., 2020) aims to define a new benchmark for state-of-the-art re-identification approaches based on body/clothing appearance. The final evaluation of top-ranked re-identification approaches suggests relevant challenges in the scenario. More recently, Choi et al. utilized gait traits, extracting arm swing features from the silhouette, to overcome issues such as RBN recognition, face occlusion, and clothing appearance similarity (Choi et al., 2021).

3 METHODOLOGY

As mentioned earlier, this paper focuses not on bib detection but on bib number recognition. Therefore, in the experimental evaluation, we assume that the bounding box of the bib number has already been obtained using one of the aforementioned detection approaches. The main emphasis of this study is on text recognition, precisely digit recognition. In this sense, the present study allows us to evaluate the ideal bib number recognition performance, which may be considered as an upper bound of the reachable performance in a real application scenario where firstly the bib detection step is required. Previous literature on bib recognition has utilized various OCR techniques or deep learning strategies for character classification.

General object detectors are, in principle, less specialized for text recognition tasks than OCRs, as they are designed as flexible solutions for any object detection task. However, two remarkable features of bib number text might make general object detectors suitable for the task: 1) the RBN text to recognize is composed of digits, i.e., just ten classes or categories are needed to be trained, and 2) the RBN text is not manuscript, i.e., not written by hand, and thus there is great homogeneity in the typography. In this sense, we are interested in exploring whether a fine-tuned general object detector, trained with non manuscript digits, may outperform OCRs in the task of RBN recognition.

In this sense, we will evaluate as baseline two freely available OCRs in the experimental evaluation below, focusing on recognizing only digits whenever possible. Specifically, we have selected two open-source alternatives: Tesseract[1], which has been used for this specific problem in the literature, and EasyOCR[2], which integrates CRAFT (Baek et al., 2019) for character region cropping and CRNN for text recognition.

Furthermore, to test the hypothesis that an object detector trained to detect digits improves OCR performance, we have adopted the YOLO (You Only Look Once) architecture, specifically YOLOv7 (Wang et al., 2023) using a digit dataset. For this aim, we have adopted the Street View House Numbers (SVHN) dataset (Netzer et al., 2011), which is composed of only street numbers captured in varying real-world conditions sharing similar features with RBNs, see Figure 1. The ground truth dataset contains bounding boxes per digit in the street number. This strategy is based on reported experiences by the community training a digit recognizer with SVHN[3] with RetinaNet (Lin et al., 2017) and YOLOv4[4] for digit spotting within the same scenario.

Even if the chosen scenario of application is different, we have adopted just the SVHN training subset for training, not fusing it with the rest of available samples in the SVHN data collection. We launched a 200-epoch training with a learning rate of 0.01 in a GTX 3080 with a batch size of 32 and an image width of 640. The default YOLOv7 data augmentation strategies were adopted, except for deactivating the possibility of augmenting data by flipping the image, since most digits are not symmetric.

The fine-tuned YOLOv7 detector will provide a collection of digit-bounding boxes. Assuming only the bib number area is processed, any detected digit will be considered belonging to the bib number. We, therefore, compound the recognized bib number by first sorting the digit bounding boxes along the x-axis; see Figure 3.

4 EXPERIMENTAL EVALUATION

The experimental evaluation focuses on the two previously mentioned approaches for recognizing RBNs. Firstly, we investigate the usage of general-purpose OCRs, such as Tesseract and EasyOCR. Secondly, we fine-tune the general object detection architecture YOLOv7 using SVHN annotated street number samples to create a dedicated digit detector and classifier.

[1] https://github.com/tesseract-ocr/tesseract
[2] https://github.com/JaidedAI/EasyOCR
[3] https://github.com/penny4860/retinanet-digit-detector
[4] https://github.com/Lwhieldon/BibObjectDetection

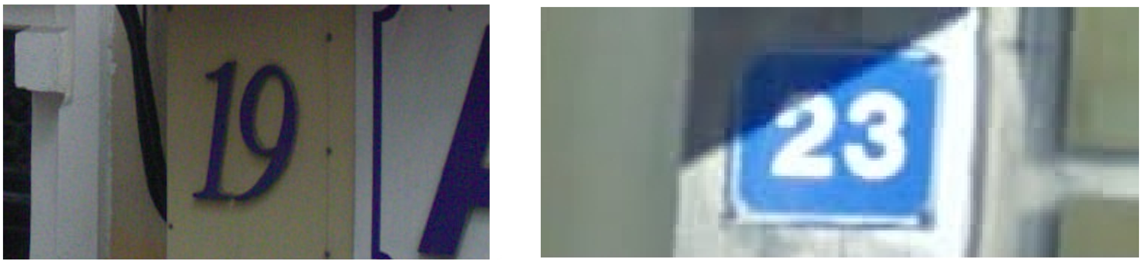

Figure 1: Samples from SVHN dataset (Netzer et al., 2011). Similarly to RBNs, both possibilities, i.e. clearer and darker fonts, are present in the data collection.

Figure 2: Upper row: samples from RBNR dataset (Ben-Ami et al., 2012). Bottom row: samples from TGCRBNW dataset (Hernández-Carrascosa et al., 2020).

Figure 3: Three digits are detected: 4, 2, and 8. Their respective x-coordinates are sorted to compose the recognized bib number: 428.

4.1 Datasets

Although various datasets are referred to in the literature for detecting and/or recognizing race bib numbers (RBNs) in running competitions, most are considered private, and we have not obtained permission to include them in the experimental evaluation. Therefore, we assess the different approaches using the only publicly available datasets accessible to us: RBNR (Ben-Ami et al., 2012) and TGCRBNW (Hernández-Carrascosa et al., 2020).

Both datasets consist of images of long-distance runners, but RBNR comprises photographs of marathon runners, while TGCRBNW contains frames extracted from videos of ultra-distance trail runners. Their main distinction is that TGCRBNW comprises images captured in low-light conditions. Furthermore, as shown in Figure 2, the RBNR dataset samples have larger and different fonts. The authors of TGCRBNW (Hernández-Carrascosa et al., 2020) claim that despite containing only a single font, their dataset is more extensive and more challenging for detecting and recognizing race bib numbers. In fact, the RBNs are annotated even if they are not easily recognized by humans in the image, as seen in the bottom row of Figure 2. We will evaluate this claim in the experiments below.

Figure 4: Upper row: cropped samples from RBNR dataset (Ben-Ami et al., 2012). Bottom row: cropped samples from TGCRBNW dataset (Hernández-Carrascosa et al., 2020).

4.2 RBN Recognition

As mentioned earlier, the experimental evaluation focuses on RBN recognition. It skips the RBN detection step, as both datasets provide ground truth in rectangular bounding boxes around the RBNs. Before locating and recognizing the text, we decided to enlarge the bounding boxes by 50% of their dimensions in width and height. This preprocessing was adopted as, in some cases, the RBN bounding box has been defined very tightly, and some bib digits do not appear whole in the cropped area. Examples of the resulting samples for both datasets are shown in Figure 4.

After enlarging the annotated bounding boxes, we first evaluate the performance of our approach on the RBNR dataset using the same metrics as the original dataset authors defined in (Ben-Ami et al., 2012). Those metrics are adopted to provide with the precision the ratio of correctly recognized RBNs out of the total number of recognized RBNs, and with the recall the ratio of correctly recognized RBNs out of the total number of annotated RBNs. The reader must observe that in their work, Ben-Ami et al. assume that detecting a face triggers the RBN RoI estimation. Therefore, the metrics do not consider the RBN detection but the full recognition of the RBN digits. This means that, for the first RBN present in Figure 4, it will be considered a correct recognition only if the system outputs 3628. The computation of those metrics required the number of correctly recognized bibs or true positives (TP) when all its digits were identified, the number of incorrectly recognized bibs or false positives (FP), when at least one but not all the digits were identified, and the number of non-classified bib numbers or false negatives (FN), those without any digit located. The recall (R), precision (P), and F-score (F) are defined respectively as:

$$P = \frac{TP}{TP+FP} \quad R = \frac{TP}{TP+FN} \quad F = 2 \cdot \frac{P \cdot R}{P+R} \tag{1}$$

The results achieved for RBNR dataset are summarized in Table 1. We compare the performance of our approach, which uses the fine-tuned YOLOv7-based digit detector, with the mentioned OCRs and the latest published result for this dataset that we are aware of (Nag et al., 2020), which refers to the work (Bartz et al., 2018).

Among the OCRs, Tesseract produced significantly poor results, but EasyOCR provided fairly competitive results compared to the latest reported state-of-the-art for RBNR (Bartz et al., 2018), although still lower. However, the results achieved by the specifically trained digit detector significantly outperformed any other approach applied to this dataset.

Table 1: Bibs recognition results for RBNR.

Approach	P	R	F
EasyOCR	0.67	0.54	0.60
Tesseract	0.33	0.05	0.08
See (Bartz et al., 2018)	0.76	0.77	0.77
YOLOv7	**0.91**	**0.91**	**0.91**

With those results, we decided to explore the errors of the YOLOv7 digit detector, focusing specifically on FPs, to understand why they occur. As mentioned above, FPs are outputs that do not match the annotated RBN completely. Upon observing the FPs in the evaluation, we noticed that in most cases, even when the entire bib number is not recognized, most detected digits are correctly identified, as shown in

Figure 5: Partially recognized bib numbers, just three digits of four were detected by the digit detector in both samples. The previously adopted metric considers both cases as FPs. A metric evaluating the digit recognition performance provides a better insight into the approach validity.

Figure 6: Images that reported a digit FP according to RBNR ground annotation. Upper row: the YOLOv7 digit detector confused a 3521 with 3527. Second row: bibs annotated as 80653, 891, 359, and 764 were recognized as 80635, 0891, 3594, and 3764, respectively. Indeed, those recognized bibs must be considered correct.

Figure 5. Fixing a detection confidence threshold of 0.5, we found that in both images, out of the eight digits present, only two of them were not located. As the bib number is not correctly matched, they are counted as FPs. That consideration is not fair to establish a real usability of the digit detector. Therefore, we propose a novel metric approach to evaluate the performance more precisely for the RBN recognition problem: considering digit recognition instead of bib number recognition. By adopting this evaluation strategy, the values for precision (P), recall (R), and F-score (F) for the YOLOv7 fine-tuned digit detector significantly increase to 0.992, 0.975, and 0.984, respectively.

These results suggest that the RBNR dataset is relatively simple for current general object detectors. An example of the type of FNs in the digit detector is visible in the previously mentioned Figure 5. As for the still remaining digit FPs, only five bib numbers were reported to have digit FPs. However, upon closer inspection, it was revealed that only one of them was a true digit FP, while the rest correspond to annotation errors, as seen in Figure 6. This means that the actual value of precision (P) would be indeed higher, indicating that the RBNR dataset is certainly trivial for current general object detectors.

Now that we have established that RBNR is a relatively simple dataset for existing tools, we evaluated the best approaches for bib number recognition on TGCRBNW. Before discussing the results, it is important to note that TGCRBNW contains 3223 annotated bibs captured in both low-light and daylight conditions, with a split of 69% and 31%, respectively. Those images correspond to video frames, thus there is no human photographer behind the camera adjusting the camera settings for each sample. Additionally, it is worth mentioning that the bib number is not necessarily visible in the images, as the human annotators had access to the whole runner tracklet to annotate the identity, i.e., his/her bib number. The results of bib number recognition are summarized in Table 2.

Indeed, those results are significantly worse than those achieved for the RBNR dataset, suggesting that the TGCRBNW dataset is considerably more challenging. A first impression suggests that the YOLOv7-based recognizer has a lower precision (P), i.e., it exhibits a higher number of false positives (FPs). However, if, similarly to the consideration made for RBNR, we adopt a digit-based metric instead of a bib number-based metric, the reported precision (P), recall (R), and F-score (F) for the YOLOv7 detector are 0.776, 0.779, and 0.778, respectively, which suggests promising performance also in severe conditions. We remind again that among the 3223 bib numbers contained in TGCRBNW, only 1000 were captured in daylight. The reported values, as seen in Table 3, suggest a slightly easier scenario with daylight illumination conditions; however, there is still much room for improvement.

Table 2: Bibs recognition results for TGCRBNW.

Approach	P	R	F
EasyOCR	**0.259**	0.002	0.004
YOLOv7	0.216	**0.047**	**0.077**

Table 3: Digits recognition results for TGCRBNW including overall and daylight and nightlight splits metrics.

TGCRBNW split	P	R	F
Complete	0.776	0.779	0.777
Nightlight	0.714	0.714	0.714
Daylight	0.825	0.831	0.827

5 CONCLUSIONS

In this paper, we analyzed RBN recognition using two available datasets: RBNR and TGCRBNW. To achieve this, we evaluated two well-known general-purpose OCRs and a specific digit detector trained with samples from a different scenario, specifically street numbers, using a general object detection architecture. After an initial evaluation of RBNR, we defined a new digit-based metric instead of an RBN-based one to get more precise evidence of the partially recognized RBNs. The results suggest that the scenario defined by datasets similar to RBNR is now solvable. However, the scenario presented by TGCRBNW poses a real in-the-wild benchmark, likely due to challenging features such as illumination conditions (nightlight, shadows, etc.), RBN resolution, and motion blur, among others.

In any case, relying solely on RBN for recognition is not enough due to the presence of occlusions, which are common in such scenarios. Integrating other cues, such as biometric traits, might help achieve coherent recognition of specific runners, especially if a video stream is available instead of single frames. However, this concept is beyond the scope of the present paper.

ACKNOWLEDGEMENTS

This work is partially supported by the the Spanish Ministry of Science and Innovation under project PID2021-122402OB-C22 and by the ACIISI-Gobierno de Canarias and European FEDER funds under projects ProID2021010012, ULPGC Facilities Net, and Grant EIS 2021 04

REFERENCES

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

Baek, Y., Lee, B., Han, D., Yun, S., and Lee, H. (2019). Character region awareness for text detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 9365–9374.

Bartz, C., Yang, H., and Meinel, C. (2018). See: Towards semi-supervised end-to-end scene text recognition. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1).

Ben-Ami, I., (Basha), T. D., and Avidan, S. (2012). Racing bib number recognition. In *British Machine Vision Conference*, pages 1–10, Surrey, UK. British Machine Vision Association.

Bochkovskiy, A., Wang, C.-Y., and Liao, H.-Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.

Boonsim, N. (2018). Racing bib number localization on complex backgrounds. *WSEAS Transactions on Systems and Control*, 13:226–231.

Carty, G., Raja, M. A., and Ryan, C. (2021). Running to get recognised. In Thampi, S. M., Krishnan, S., Hegde, R. M., Ciuonzo, D., Hanne, T., and Kannan R., J., editors, *Advances in Signal Processing and Intelligent Recognition Systems*, pages 3–17, Singapore. Springer Singapore.

Choi, Y., Napolean, Y., and van Gemert, J. C. (2021). The arm-swing is discriminative in video gait recognition for athlete re-identification. In *IEEE International Conference on Image Processing*.

de Jesús, W. M. and Borges, D. L. (2018). An improved stroke width transform to detect race bib numbers. In *Proceedings of the Mexican Conference on Pattern Recognition*, pages 267–276, Puebla, Mexico. Springer.

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

Ivarsson, E. and Mueller, R. M. (2019). Racing bib number recognition using deep learning. In *Proceedings of the 25th Americas Conference on Information Systems*

(AMCIS), Cancún, Mexico. Association for Information Systems.

Kamlesh, P. X., Yang, Y., and Xu, Y. (2017). Person reidentification with end-to-end scene text recognition. In *Chinese Conference on Computer Vision*, pages 363–374, Tianjin, China. Springer.

Li, G. and Zhang, C. (2019). Automatic detection technology of sports athletes based on image recognition technology. *EURASIP Journal on Image and Video Processing*, 2019(1):15.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017). Focal loss for dense object detection. In *In IEEE International Conference on Computer Vision (ICCV)*.

Liu, H. and Bhanu, B. (2019). Pose-guided R-CNN for jersey number recognition in sports. In *Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 2457–2466, Long Beach, CA, USA. IEEE.

Liu, X., Liang, D., Yan, S., Chen, D., Qiao, Y., and Yan, J. (2018). FOTS: Fast oriented text spotting with a unified network. In *IEEE Conf. on Computer Vision and Pattern Recognition*.

Luo, C., Jin, L., and Sun, Z. (2019). Moran: A multi-object rectified attention network for scene text recognition. *Pattern Recognition*, 90:109–118.

Mendes-Neves, T., Meireles, L., and Mendes-Moreira, J. (2023). A survey of advanced computer vision techniques for sports.

Minghui Liao, B. S. and Bai, X. (2018). TextBoxes++: A single-shot oriented scene text detector. *IEEE Transactions on Image Processing*, 27(8):3676–3690.

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

Nag, S., Shivakumara, P., Pal, U., Lu, T., and Blumenstein, M. (2020). A new unified method for detecting text from marathon runners and sports players in video (prd-19-01078r2). *Pattern Recognition*, 107:107476.

Netzer, Y., Wang, T., Coates, A., Bissacco, A., Wu, B., and Ng, A. Y. (2011). Reading digits in natural images with unsupervised feature learning. In *NIPS Workshop on Deep Learning and Unsupervised Feature Learning*.

Penate-Sanchez, A., Freire-Obregón, D., Lorenzo-Melián, A., Lorenzo-Navarro, J., and Castrillón-Santana, M. (2020). TGC20ReId: A dataset for sport event reidentification in the wild. *Pattern Recognition Letters*, 138:355–361.

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

Thomas, G., Gade, R., Moeslund, T. B., Carr, P., and Hilton, A. (2017). Computer vision for sports: Current applications and research topics. *Computer Vision and Image Understanding*, 159:3 – 18.

Wang, C.-Y., Liao, H.-Y. M., and Yeh, I.-H. (2023). Designing network design strategies through gradient path analysis. *Journal of Information Science and Engineering*.

Wong, Y. C., Choi, L. J., Singh, R. S. S., Zhang, H., and Syafeeza, A. R. (2019). Deep learning based racing bib number detection and recognition. *Jordanian Journal of Computers and Information Technology (JJCIT)*, 5(3):(3):181–194.

Wrońska, A., Sarnacki, K., and Saeed, K. (2017). Athlete number detection on the basis of their face images. In *Proceedings International Conference on Biometrics and Kansei Engineering*, pages 84–89, Kyoto, Japan. IEEE.