View Reviews

Paper ID
1789

Paper Title
CricXpert: A Hybrid Spatial Fusion Model For Enhanced Player Recognition

Track Name
Main Track

Reviewer #1

Questions

2. Relevance to IJCNN
Good

3. Technical quality
Poor

4. Novelty
Poor

5. Quality of presentation
Fair

9. Comments to Authors
The authors present a paper on player classification from bounding boxes extracted by an object detection model. While the problem may be interesting, I believe that the contribution is not significant enough for publication. Please find detailed comments in the following:
- Performance results is very high, which raises concerns about the actual difficulty of the dataset being used. A comparison with state-of-the-art methods is absent, leaving readers to question how these results measure up against existing benchmarks (which appear to exist, based on the mentioned related works).
- The methodology lacks clarity in several key aspects Specifically, what improvements does the proposed methodology offer over existing approaches utilizing SVM and kNN with features extracted by CNNs?
- Has the new dataset has been manually checked for quality? The dataset size is another concern; it appears insufficient to conclude that the task is challenging, especially given the reported high accuracy.
- The authors claim that training a ViT model would take too long, which is strange

to me, given the small size of the dataset, particularly if leveraging efficient fine-tuning methods like LoRA.
- The proposed architecture involving a stacking ensemble is not described in sufficient detail, making it impossible to reproduce the results. The paper does not clarify the input and output of each component nor their significance or rationale of inclusion.
- The motivation behind using CLIP-based dataset validation is unclear. The authors do not explain what this process aims to demonstrate.
- What does the evaluation procedure using the LLM do? It is not mentioned and it is not clear what its point is.
- What is the custom CNN?
- How were CNNs compared—were they fine-tuned or frozen during training?
- Given that ResNet already achieves high accuracy, the lack of statistical significance values makes it impossible to evaluate the contribution of the stacking ensemble effectively.
- The authors express concerns about overfitting in a model that reportedly achieves over 95% test accuracy. This contradiction raises questions about the reliability of the reported results.

Reviewer #2

Questions

2. Relevance to IJCNN
Good

3. Technical quality
Good

4. Novelty
Fair

5. Quality of presentation
Good

9. Comments to Authors
1.The authors should justify the novelty of the paper by highlighting the advantages of the proposed model.
2.The language should be refined for clarity and better comprehension.
3.The references to existing methods used in the article are not properly provided and should be appropriately cited.

Reviewer #3

Questions

2. Relevance to IJCNN

Very good

3. Technical quality

Good

4. Novelty

Fair

5. Quality of presentation

Fair

9. Comments to Authors

The paper titled: "CricXpert: A Hybrid Spatial Fusion Model
For Enhanced Player Recognition" is focused on vision-based player recognition using the hybrid ML and DL technique.
Here are a few observations and suggestions:
Abstract is written well.
Fig. 1. needs improvements. The font size of the text is very small. The block diagram of the architecture of the proposed study should be presented in fig. 1.
Fig. 2. has a lot of text and also the font size is very small. The details of the proposed model/technique should be provided for reproducibility.
The comparison with existing techniques is presented in the Table 1. The baseline techniques with mentioned results of dataset taken from literature should be cited in table.
The text font size in almost all figures is very small.
The performance comparison of techniques should be provided in bar chart/plot form to compare the evaluation metrics.
Conclusion should be concise.
Compute cost anaysis for inference should be provided for proposed hybrid model.
Why ViT and other DL techniques get very low performance in comparison to the proposed technique. The author should explain in the discussion section.

Reviewer #4

Questions

2. Relevance to IJCNN

Very good

3. Technical quality

Good

4. Novelty

Fair

5. Quality of presentation

Good

9. Comments to Authors

The paper has good technical quality with solid methodology and rigorous evaluation, but it could be strengthened with more dataset variety, efficiency analysis, and statistical validation. Areas for improvement: Larger dataset, analysis of computational efficiency, statistical tests to confirm performance gains and why did ViTs fail.

The paper is novel in its application to cricket analytics, but its methodological contributions are incremental rather than groundbreaking. A stronger contribution could have been made by:

Releasing the dataset publicly for wider adoption.
Proposing modifications to ResNet50 or the classifiers instead of using them in standard form.
Exploring transfer learning to extend the model to other sports.
The paper is innovative for sports analytics but does not introduce a fundamentally new ML approach.