Multi-Task Temporal and Spatial Networks for High-Precision Event Spotting in Volleyball Videos

Hoang Quoc Nguyen[1,2][a], Second Author Name[1][b] and Third Author Name[2][c]
[1]*Korea Institute of Science and Technology, Seoul, Republic of Korea*
[2]*University of Science and Technology, Daejeon, Republic of Korea*
523503, second_author}@kist.re.kr, third_author@dc.mu.edu

Keywords: Temporally Precise Spotting, Video Understanding, Spatial Temporal Event Spotting, Volleyball, Sport, Deep Learning

Abstract: Understanding the precise timing and location of events is crucial for analyzing sports videos, especially in fast-paced sports like volleyball. We introduce a new task: high-precision spatial-temporal event spotting, which aims to detect both when and where key actions occur. To support this, we present the KOVO Volleyball Event Dataset, featuring 947 rally videos, and 5,935 events, annotated for both temporal and spatial localization. Our best model achieves a combined mAP of 85.46 across various temporal and spatial thresholds. Notably, we find that incorporating spatial predictions enhances temporal mAP by 5.89, underscoring the synergy between spatial and temporal analysis. To the best of our knowledge, this is the first work addressing this task, establishing a strong baseline for future research in spatial-temporal event spotting.

1 INTRODUCTION

Video understanding has emerged as a cornerstone in computer vision, offering valuable insights into dynamic scenes for applications such as sports analytics, surveillance, and autonomous systems. Within this field, various tasks have been defined to interpret actions over time. *Temporal Action Detection (TAD)* focuses on pinpointing time intervals where specific actions occur within untrimmed videos, while *Temporal Action Segmentation (TAS)* aims to divide videos into continuous sequences of actions. Complementing these is the task of *Action Spotting*, which zeroes in on identifying the precise frames that capture key events, requiring models to discern subtle temporal differences and visually similar frames (Hong et al., 2022).

Recent advancements in action spotting, such as *T-DEED* (**?**) and *spot22* (Hong et al., 2022), have demonstrated the ability of models to achieve frame-level precision in fast-paced events by leveraging deep learning architectures. However, most of these efforts focus solely on detecting events in specific domains like figure skating and diving. While effective in their respective domains, they do not capture the unique challenges posed by team-based sports with rapid dynamics, such as volleyball.

Despite progress in action spotting, there is a gap in applying these techniques to sports like volleyball, where rapid play transitions occur within specific areas of the court. Addressing this requires a broader focus: *high-precision spatial-temporal event spotting*, a task designed to detect both the exact timing and spatial location of key events. Unlike conventional action spotting, this task provides richer insights into player positioning and movement patterns, crucial for analyzing volleyball gameplay.

In other sports, datasets like *SoccerNet* have pushed the boundaries of action spotting through rich temporal and spatial annotations, significantly advancing model capabilities. Yet, no equivalent dataset exists for volleyball, a sport characterized by its rapid exchanges and the need for precise localization of actions. To fill this gap, we introduce the *KOVO Event Dataset*, comprising 947 rally videos, 890,797 frames, and 5,935 annotated key actions. This dataset offers granular annotations for both temporal and spatial event localization, making it a valuable resource for developing models that capture the intricacies of volleyball.

Our contributions are threefold. First, we introduce the new task of high-precision spatial-temporal

[a] https://orcid.org/0009-0002-2004-9285
[b] https://orcid.org/0000-0000-0000-0000
[c] https://orcid.org/0000-0000-0000-0000

event spotting, specifically tailored for the dynamics of volleyball. Second, we present the *KOVO Event Dataset*, the first of its kind to include detailed temporal and spatial annotations for volleyball rallies. Third, we propose a multi-task deep learning model that jointly predicts event timing and spatial positions, leveraging this dual focus to achieve improved performance. Notably, incorporating spatial predictions into our model enhances temporal mAP by 5.89 points. Our best model achieves a temporal mAP of 90.59, a spatial mAP of 77.94, and a combined mAP of 85.46, providing a strong baseline for this new task. To the best of our knowledge, this work is the first to explore high-precision spatial-temporal event spotting in volleyball, setting the stage for future research in this area. The paper proceeds with a discussion of related work in Section 2, a detailed description of our approach in Section 3, the experimental setup in Section 4, results in Section 5, and conclusions in Section 6.

2 RELATED WORK

We strongly encourage authors to use this document for the preparation of the camera-ready. Please follow the instructions closely in order to make the volume look as uniform as possible (MooreX and Lopes, 1999).

Please remember that all the papers must be in English and without orthographic errors.

Do not add any text to the headers (do not set running heads) and footers, not even page numbers, because text will be added electronically.

For a best viewing experience the used font must be Times New Roman, except on special occasions, such as program code 2.4.8.

2.1 Manuscript Setup

The template is composed by a set of 7 files, in the following 2 groups:
Group 1. To format your paper you will need to copy into your working directory, but NOT edit, the following 4 files:

- `apalike.bst`
- `apalike.sty`
- `article.cls`
- `scitepress.sty`

Group 2. Additionally, you may wish to copy and edit the following 3 example files:

- `example.bib`
- `example.tex`
- `scitepress.eps`

2.2 Page Setup

The paper size must be set to A4 (210x297 mm). The document margins must be the following:

- Top: 3,3 cm;
- Bottom: 4,2 cm;
- Left: 2,6 cm;
- Right: 2,6 cm.

It is advisable to keep all the given values because any text or material outside the aforementioned margins will not be printed.

2.3 First Section

This section must be in one column.

2.3.1 Title and Subtitle

Use the command *title* and follow the given structure in ”example.tex”. The title and subtitle must be with initial letters capitalized (titlecased). The separation between the title and subtitle is done by adding a colon ”:” just before the subtitle beginning. In the title or subtitle, words like ”is”, ”or”, ”then”, etc. should not be capitalized unless they are the first word of the title or subtitle. No formulas or special characters of any form or language are allowed in the title or subtitle.

2.3.2 Authors and Affiliations

Use the command *author* and follow the given structure in ”example.tex”. Please note that the name of each author must start with its first name.

2.3.3 Keywords

Use the command *keywords* and follow the given structure in ”example.tex”. Each paper must have at least one keyword. If more than one is specified, please use a comma as a separator. The sentence must end with a period.

2.3.4 Abstract

Use the command *abstract* and follow the given structure in ”example.tex”. Each paper must have an abstract up to 200 words. The sentence must end with a period.

2.4 Second Section

Files ”example.tex” and ”example.bib” show how to create a paper with a corresponding list of references.

This section must be in two columns.

Each column must be 7,5-centimeter wide with a column spacing of 0,8-centimeter.

The section text must be set to 10-point.

Section, subsection and sub-subsection first paragraph should not have the first line indent.

To remove the paragraph indentation (only necessary for the sections), use the command $\backslash noindent$ before the paragraph first word.

If you use other style files (.sty) you MUST include them in the final manuscript zip file.

2.4.1 Section Titles

The heading of a section title should be in all-capitals.

Example: $\backslash section\{FIRST\ TITLE\}$

2.4.2 Subsection Titles

The heading of a subsection title must be with initial letters capitalized (titlecased).

Words like ”is”, ”or”, ”then”, etc. should not be capitalized unless they are the first word of the subsection title.

Example: $\backslash subsection\{First\ Subtitle\}$

2.4.3 Sub-Subsection Titles

The heading of a sub subsection title should be with initial letters capitalized (titlecased).

Words like ”is”, ”or”, ”then”, etc should not be capitalized unless they are the first word of the sub subsection title.

Example: $\backslash subsubsection\{First\ Subsubtitle\}$

2.4.4 Tables

Tables must appear inside the designated margins or they may span the two columns.

Tables in two columns must be positioned at the top or bottom of the page within the given margins. To span a table in two columns please add an asterisk (*) to the table *begin* and *end* command.

Example: $\backslash begin\{table*\}$
$\backslash end\{table*\}$

Tables should be centered and should always have a caption positioned above it. The font size to use is 9-point. No bold or italic font style should be used.

The final sentence of a caption should end with a period.

Please note that the word ”Table” is spelled out.

Table 1: This caption has one line so it is centered.

Example column 1	Example column 2
Example text 1	Example text 2

Table 2: This caption has more than one line so it has to be justified.

Example column 1	Example column 2
Example text 1	Example text 2

2.4.5 Figures

Please produce your figures electronically, and integrate them into your document and zip file.

Check that in line drawings, lines are not interrupted and have a constant width. Grids and details within the figures must be clearly readable and may not be written one on top of the other.

Figure resolution should be at least 300 dpi.

Figures must appear inside the designated margins or they may span the two columns.

Figures in two columns must be positioned at the top or bottom of the page within the given margins. To span a figure in two columns please add an asterisk (*) to the figure *begin* and *end* command.

Example: $\backslash begin\{figure*\}$
$\backslash end\{figure*\}$

Figures should be centered and should always have a caption positioned under it. The font size to use is 9-point. No bold or italic font style should be used.

Figure 1: This caption has one line so it is centered.

Figure 2: This caption has more than one line so it has to be justified.

The final sentence of a caption should end with a period.

Please note that the word ”Figure” is spelled out.

2.4.6 Equations

Equations should be placed on a separate line, numbered and centered.

The numbers accorded to equations should appear in consecutive order inside each section or within the contribution, with the number enclosed in brackets and justified to the right, starting with the number 1.

Example:

$$a = b + c \tag{1}$$

2.4.7 Algorithms and Listings

Algorithms and Listings captions should have font size 9-point, no bold or italic font style should be used and the final sentence of a caption should end with a period. The separator between the title of algorithms/listings and the name of the algorithm/listing must be a colon. Captions with one line should be centered and if it has more than one line it should be set to justified.

Data: this text
Result: how to write algorithm with LaTeX2e
initialization;
while *not at end of this document* **do**
  read current;
  if *understand* **then**
    go to next section;
    current section becomes this one;
  else
    go back to the beginning of current
     section;
  end
end

Algorithm 1: How to write algorithms.

2.4.8 Program Code

Program listing or program commands in text should be set in typewriter form such as Courier New.

Example of a Computer Program in Pascal:

```
Begin
    Writeln('Hello World!!');
End.
```

The text must be aligned to the left and in 9-point type.

2.4.9 Reference Text and Citations

References and citations should follow the APA (Author, date) System Convention (see the References section in the compiled manuscript). As example you may consider the citation (Smith, 1998). Besides that, all references should be cited in the text. No numbers with or without brackets should be used to list the references.

References should be set to 9-point. Citations should be 10-point font size.

You may check the structure of "example.bib" before constructing the references.

For more instructions about the references and citations usage please see the appropriate link at the conference website.

3 PROPOSED METHOD

For the mutual benefit and protection of Authors and Publishers, it is necessary that Authors provide formal written Consent to Publish and Transfer of Copyright before publication of the Book. The signed Consent ensures that the publisher has the Author's authorization to publish the Contribution.

The copyright form is located on the authors' reserved area.

The form should be completed and signed by one author on behalf of all the other authors.

4 EXPERIMENTS

Please note that ONLY the files required to compile your paper should be submitted. Previous versions or examples MUST be removed from the compilation directory before submission.

We hope you find the information in this template useful in the preparation of your submission.

5 CONCLUSIONS

Please note that ONLY the files required to compile your paper should be submitted. Previous versions or examples MUST be removed from the compilation directory before submission.

We hope you find the information in this template useful in the preparation of your submission.

ACKNOWLEDGEMENTS

If any, should be placed before the references section without numbering. To do so please use the following command:

REFERENCES

Hong, J., Zhang, H., Gharbi, M., Fisher, M., and Fatahalian, K. (2022). Spotting temporally precise, fine-grained events in video.

MooreX, R. and Lopes, J. (1999). Paper templates. In *TEMPLATE'06, 1st International Conference on Template Production*. SCITEPRESS.

Smith, J. (1998). *The Book*. The publishing company, London, 2nd edition.

APPENDIX

If any, the appendix should appear directly after the references without numbering, and not on a new page. To do so please use the following command: