Dive Performance Analytics: A Pose Estimation Pipeline Using RVM and RTMPose

Prabhu Nithin Gollapudi, Matrikelnummer: 23269243

Friedrich-Alexander-Universität Erlangen-Nürnberg
Machine Learning and Data Analytics Lab
https://www.mad.tf.fau.de/
Carl-Thiersch-Straße 2b, 91052 Erlangen, Germany
Email: prabhu.nithin.gollapudi@fau.de

***Abstract*—Accurate human pose estimation is essential for analyzing athletic performance, optimizing training, and supporting technical assessment in sports. In dynamic and high-speed disciplines such as diving, pose estimation systems must overcome challenges related to rapid motion, complex body orientations, and real-time processing demands. This study presents a pose estimation pipeline designed specifically for diving analysis. We used the RVM segmentation technique to accurately extract the diver from the background. Subsequently, we employ the RTMPose and RTMDet algorithms to track and extract keypoints from the diver's body and generate KPIs including total rotation angle, joint angles, and maximum height at various stages. Experimental evaluations on real-world diving videos demonstrate that our approach achieves a 3x improvement in FPS processing speed without sacrificing accuracy. The idea of a pose estimation pipeline for dive analysis can be seen as a first step towards significantly improving real-time analytics in the sports industry with the help of AI and data-driven approaches.**

I. Introduction

In recent years, the use of artificial intelligence(AI) to evaluate athletic performance has gained momentum in multiple sports. Among these techniques, **human pose estimation** stands out because it provides deep insight into the movements of an athlete, allowing coaches and trainers to optimize their training. In fast-paced sports such as **diving**, where precise posture and body control are critical, pose estimation faces unique challenges, particularly due to rapid movements, complex rotations, and occlusions.

Traditional evaluation methods for diving are heavily based on manual annotation and subjective judgment, which makes them time consuming and prone to human error and bias. Although recent advances in pose estimation - **MediaPipe** [1] and **Movenet** [2] have improved speed and accuracy, these models often face limitations with the specific demands of diving. MediaPipe, for instance, provides real-time performance, but lacks robustness under rapid, dynamic movements. Movenet provides high accuracy at the cost of computational efficiency.

Similarly, segmentation of divers from the background presents its own challenges. Existing tools like **Segment Anything** [3] and **Mediapipe Selfie Segmentation** [4] and **YOLOv5** [5] have shown potential in general segmentation tasks but are not optimized for high-speed sports footage. Their limitations in precision and latency make them less suitable for real-time analysis in diving.

A more domain-specific approach, **DiveNet** [6], focuses on diving action localization and physical parameter extraction. Although DiveNet successfully detects the start and end of dive sequences and extracts certain motion attributes, it lacks continuous full-body pose tracking throughout the dive. As a result, it cannot provide detailed frame-level insights such as joint angles, total rotation, or precise trajectory tracking.

To address these challenges, we propose a novel pipeline tailored for dive pose estimation and performance analysis. Our system integrates **Robust Video Matting (RVM)** [7] for high-quality segmentation and combines **RTMPose** [8] with **RTMDet** [9] for efficient and continuous extraction of keypoints. This architecture achieves high accuracy while maintaining near-real-time frame rates.

From the extracted pose data, we compute a set of **key performance indicators (KPIs)** essential for the analysis of dives: joint angles, total rotation, diver height, trajectory, and identification of dive phases. These metrics offer actionable insights for coaches and athletes, improving both training quality and performance assessment.

Our main contributions are as follows.

- **Segmentation** : A robust RVM-based approach that isolates the diver even under complex motion and cluttered backgrounds.
- **Pose Estimation**: A lightweight RTMPose+RTMDet-based module that provides near real-time, continuous keypoint tracking for each frame.
- **KPI Generation**: A framework for extracting key performance metrics, such as joint angles, height, and rotation, that can aid in training optimization.

This work demonstrates the potential of AI-based pose estimation to advance into real-time analysis in diving. The proposed system serves as a step towards the practical deployment of intelligent feedback tools in competitive sports environments.

II. Methods

This section includes everything the reader needs to understand your system and results. Things you just used only have to be mentioned and cited. For data recording, we used

the mobile system developed by Kugler et al. [11]. Make sure to describe your contribution as short and precise as possible and as detailed as needed to reimplement it. You have to include a description of data, hardware, math, algorithms, system structure and evaluation methods and everything that was important to solve the defined problem. Make sure to order the methods so that the reader can follow your description. First things first! Make sure to structure the methods in a proper way. Use a dry description and try not to teach the reader.
The methods part is written in simple past, even if you created a live system that can still do things. It is only important what the system did on the data you presented. Only describe the best and final version of the system. The reader is not interested what you did wrong in developing the final system. Things that did not work don't have to be mentioned.

A. *Study-Design*

A rule of thumb is that you describe one fact in each paragraph. That means, if you go to the next aspect, you start with a new paragraph. Try to maintain a logical structure throughout the paper where paragraphs build upon each other. A good way to structure the paper are subsections. You can also use enumerations to structure the paper. Use an unordered list if the items don't have a structure.

- Apple
- Peach
- Melon
- Grape

Use an ordered list if there is a hierarchical structure in the list.

1) Erlangen
2) Bavaria
3) Germany
4) Europe

You should always have more than one subsection. Otherwise the structure does not make sense.

B. *Data Collection*

Add figures and cite the figures at the end of a sentence (Fig. 1). Good figures make a paper a lot better. Don't use figures if they are of bad quality or not exactly what you want to show with it. Latex puts figures where they fit best regarding the document structure. Often, figures are not at the same place as they are defined in the source code. Don't worry too much about that. Often, they are place on top of a page but the desired (not defined) positioning can be specified in the brackets after begin{figure}. Make sure the figure is readable when printed in black and white. This means that the lines in a plot have to be of different shape (dotted, dashed,...) and the coloring has to be adapted if used. Print the paper in black and white if unsure.
Add important equations that a relevant for understanding your system and cite them (Eq. 1). Every variable has to be explained. The variable f denotes the force, m the mass and c a constant. Make sure to reference the equation or provide an explanation. When writing about value make sure to insert a safe blank between value and unit like 9.81 g or 67 % to avoid a line break between value and unit. The safe blank is also needed when referencing figures.

Fig. 1. The caption of a figure explains the complete figure. The reader must be able to completely understand what he/she sees with the information in the caption. Do not expect the reader to read the text before trying to understand the figure.

$$f = m \times c^2 \tag{1}$$

C. *Analysis*

Make sure to use proper English in your papers. Get the paper reviewed by another person to avoid stupid typos and check the language for common mistakes. If the paper is important, try to find a native speaker for review. Some common language related problems:

- Make sure to use lower case for nouns.
- There are no strict rules for punctuation in English and, in some cases, the rules are different to German rules. One common problem is the use of a comma in a relative clause. There is no comma after "that" and "which" if it is a restrictive relative clause like "The book which I read is well written.". There is a comma if it is a non-restrictive relative clause like "That book, which incidentally I just finished reading, is well written.". A rule of thumb is that you use a comma if the relative sentence is not needed to make it a proper sentence.
- Use short sentence in paratactic form. This is easier to read. Long sentences are hard to understand. You don't write poems that are supposed to be of nice language.
- You can alternate between active and passive. Contrary to German, the use of active is also elegant. When talking about yourself you should write "the author" or "we".
- Do not abbreviate terms like "do not", "cannot" or "it is".
- Find scientific synonyms for colloquial language. You can start colloquial and then transform into scientific language afterwards. If you are unsure how to start you can use a phrasebank [14] to copy nice formulations.
- If you are not sure if you found the right word in a dictionary use a single language dictionary like [13].
- Numbers: write words when referring to numerals below ten, referring to fractions, referring to approximate numbers and at the beginning of a sentence. Use figures when referring to sets of numbers, numbers

TABLE I. THE CAPTION IN TABLES HAS TO BE SELF EXPLAINING LIKE IN FIGURES (SEE FIG. 1). MAKE SURE TO EXPLAIN ALL ABBREVIATIONS THAT YOU USE IN THE CAPTION.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
RULE	0.5	0.7	0.3	0.7	0.8	0.2	0.3	0.5	0.7	0.9
EPIC	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

including decimal points and when referring to pages. Make sure to use a decimal point instead of a comma as in German.

- Get yourself inspired from the literature. Use nice formulations from other papers. Never copy words but copy style.

It is very important to be consistently in wording and punctuation. Call the same things with the same name even if you think that it is boring. Use the same punctuation rule throughout the complete paper.

D. Evaluation

You should always avoid to change the template. It was created for a good reason. If your manuscript is too long you should change the content and not the template. Concentrate on the purpose of your project and get rid of side aspects. Use references and keep work of others short. You HAVE to learn to describe your work in a given length. A rule of thumb is a partitioning of 15 % (Introduction), 40 % (Methods), 10 % (Results), 25 % (Discussion) and 10 % (Summary and outlook).

E. Subsection

Normally, you don't present source code. The reader is more interested in a flow chart to understand how your algorithm works. Describe the most important steps of algorithms in detail and briefly describe other parts. Always provide a version when referring to other software packages or libraries. If you talk devices like smart phones or treadmills provide the manufacturer and the companies headwater details like (Google Inc., Mountain View, USA).

III. RESULTS

You describe the pure results in this section. Performance measures, reference system and experimental setup have to be explained in the methods part. Don't provide an interpretation of the results in this section, just talk about the results. Don't provide reasons why your system didn't work in this or that condition. Just mention that this was the case. Quantitative results are most striking and you might want to sum them up in a table (Tab. I). If you have a lot of different results, pick the best and most important ones. Focus on the purpose of your system! You can use the same structure as in the methods section if appropriate.
The result section is written in simple past. Even if you have a live system, you describe the results as achieved in the evaluation.

IV. DISCUSSION

Begin with a short problem specific summary. Discuss every aspect from the result section and give an interpretation of the results. You can use the same structure as in the results. Why is the algorithm so good/bad? What are limitations or assumptions? What was remarkable in the project and can be seen in the data. Make sure to cite the literature when comparing to existing systems or algorithms. Always mention pros and cons! Discuss global strength and weakness and concentrate on major aspects.
The discussion is mainly written in simple past. You can use present tense when talking about facts that you are absolutely sure about. If you are very self confident, you can deduce facts from your findings but often conditional formulations are more polite. Sometimes it is better to combine the discussion with an outlook so you might want to use future as well.

V. SUMMARY AND OUTLOOK

Sum up the most important project findings and mention what the next steps, improvements, enhancements could be. The outlook can consist of ideas that will be hard to realize. Don't just think about tomorrow, think about the next ten years. You can finish with advertising your work and point out the most important finding or development again.
The summary part is written in simple past. The outlook is written in present tense or future.

ACKNOWLEDGMENT

The authors would like to thank...

REFERENCES

[1] Google AI, MediaPipe Repository. https://github.com/google-ai-edge/mediapipe/tree/master. Last visited: 28.03.2025 (2025)

[2] Google AI, MoveNet - SinglePose Thunder. https://www.kaggle.com/models/google/movenet/tensorFlow2/singlepose-thunder. Last visited: 28.03.2025 (2025)

[3] Kirillov, A. et al.: *Segment Anything.* https://arxiv.org/abs/2304.02643. Last visited: 28.03.2025 (2023)

[4] Google AI, MediaPipe Selfie Segmentation. https://ai.google.dev/edge/mediapipe/solutions/vision. Last visited: 28.03.2025 (2025)

[5] Glenn Jocher, YOLOv5. https://github.com/ultralytics/yolov5. Last visited: 28.03.2025 (2020)

[6] Murthy, P., Taetz, B., Lekhra, A., and Stricker, D.: *DiveNet: Dive Action Localization and Physical Pose Parameter Extraction for High Performance Training.* IEEE Access 11, 37749–37767 (2023), doi: 10.1109/ACCESS.2023.3265595

[7] Lin, S., Yang, L., Saleemi, I., and Sengupta, S.: *Robust High-Resolution Video Matting with Temporal Guidance.* https://arxiv.org/abs/2108.11515

[8] Jiang, T., Lu, P., Zhang, L., Ma, N., Han, R., Lyu, C., Li, Y., and Chen, K.: *RTMPose: Real-Time Multi-Person Pose Estimation based on MMPose.* https://arxiv.org/abs/2303.07399

[9] Lyu, C., Zhang, W., Huang, H., Zhou, Y., Wang, Y., Liu, Y., Zhang, S., and Chen, K.: *RTMDet: An Empirical Study of Designing Real-Time Object Detectors.* https://arxiv.org/abs/2212.07784

[10] Cook, G. et al.: *Pre-Participation Screening: The Use of Fundamental Movements as an Assessment of Function - Part 1.* North American Journal of Sports Physical Therapy 1(2), 62–72 (2006)

[11] Kugler, P. et al.: *Mobile Recording System for Sport Applications.* In: Proc. of the 8th International Symposium on Computer Science in Sport (IACSS2011) , pp. 67–70, Shanghai (2011)

[12] Theodoridis, S. and Koutroumbas, K.: Pattern Recognition. Academic Press (2008)

[13] Longman Dictionary http://www.ldoceonline.com/. Last visited: 16.09.2013 (2013)

[14] Academic Phrasebank http://www.phrasebank.manchester.ac.uk//. Last visited: 16.09.2013 (2013)