The Shot Predictor

Laith Altarabishi
University of Texas at Austin
laithaustin@utexas.edu

Sidharth Babu
University of Texas at Austin
<insert email>

Afnan Mir
University of Texas at Austin
afnanmir@utexas.edu

Zayam Tariq
University of Texas at Austin
<insert email>

1. Problem Description

In basketball, the ability of a player to effectively shoot the basketball typically comes down to the player's shooting form. While the form of the best shooters tend to look different, they all typically use the same fundamentals. In our project, we will attempt to capture these fundamental aspects of a player's shooting form and attempt to predict the outcome of a shot using these features. Research has been done on extracting features from a player's movement to classify the action a player is performing (shooting, dribbling, etc.) [1], but we would like to focus our energy on feature extraction from the shooting motion using pose estimation [2], object detection [3], and possibly other methods to extract feature descriptors of a shot and attempt to identify it as a make or a miss.

This problem is a particularly nontrivial application of pose estimation for two main reasons. The first being that there are multiple stages to a basketball shot that need to be taken into account. From dipping the ball to waist level, to the motion of bringing the ball to eye level, to releasing the ball, each plays a significant role in the outcome of a shot, so each stage needs to be taken into account. The second reason is the variability of the average shot. It can be argued that no two shots are identitical, when taking into account the speed of the shot and different shooting forms. This will require us to develop a method of capturing all the motions of a shot in a way that is invariant to the shot's speed and overall form.

2. Related Works

2.1. Pose Estimation

Pose Estimation is a critical topic in computer vision that will intersect with our goal of trying to accurately capture the motion and actions of a person taking a shot in basketball. Pose estimation, in the context of 2D videos of humans, is the problem of localizing anatomical keypoints or joints in a frame by frame video or image [2]. To fulfill our goal of predicting the outcome of a basketball shot, it will be critical to assess the form of a player who's taking a shot - where form can be decomposed into various classifications of joints in space. Pose estimation methods can be categorized into bottom-up or top-down methodologies. Bottom-up methodologies start by estimating keypoints and body joints first, and then these points are clustered to form poses. In contrast, top-down methodologies of pose estimation first run a person detector before decomposing each person into their respective body joints within detected bounding boxes [10]. Computational complexity is a major consideration for landmark pose estimation algorithms, and modern SOTA pose estimation algorithms deploy deep learning and CNNs to improve compuational overhead and speed [2]. We list some examples of prevalent and SOTA pose estimation models that have been employed and researched below.

> OpenPose: The first multi-person realtime 2D pose estimation system that uses a bottom-up approach that implements nonparametric representation to associate human keypoints and body parts with an individual in an image [2].
>
> DeepPose: SOTA pose estimation method that uses DNNs to classify human body joints through the usage of cascading DNN regressors that produce high precision pose estimates [4].
>
> AlphaPose: Multi-person SOTA realtime pose estimation system that outperforms OpenPose in AP score and has a high mAP score [9].
>
> DeepCut: Proposes an approach to solving issues in both pose estimation and detection by using a partitioning and labeling formulation of a set of CNN part detectors [11].

2.2. Type style and fonts

Wherever Times is specified, Times Roman may also be used. If neither is available on your word processor, please use the font closest in appearance to Times to which you have access.

MAIN TITLE. Center the title $1\frac{3}{8}$ inches (3.49 cm) from the top edge of the first page. The title should be in Times 14-point, boldface type. Capitalize the first letter of nouns, pronouns, verbs, adjectives, and adverbs; do not capitalize articles, coordinate conjunctions, or prepositions (unless the title begins with such a word). Leave two blank lines after the title.

AUTHOR NAME(s) and AFFILIATION(s) are to be centered beneath the title and printed in Times 12-point, non-boldface type. This information is to be followed by two blank lines.

The ABSTRACT and MAIN TEXT are to be in a two-column format.

MAIN TEXT. Type main text in 10-point Times, single-spaced. Do NOT use double-spacing. All paragraphs should be indented 1 pica (approx. $\frac{1}{6}$ inch or 0.422 cm). Make sure your text is fully justified—that is, flush left and flush right. Please do not place any additional blank lines between paragraphs.

Figure and table captions should be 9-point Roman type as in **????**. Short captions should be centred.
Callouts should be 9-point Helvetica, non-boldface type. Initially capitalize only the first word of section titles and first-, second-, and third-order headings.

FIRST-ORDER HEADINGS. (For example, **1. Introduction**) should be Times 12-point boldface, initially capitalized, flush left, with one blank line before, and one blank line after.

SECOND-ORDER HEADINGS. (For example, **1.1. Database elements**) should be Times 11-point boldface, initially capitalized, flush left, with one blank line before, and one after. If you require a third-order heading (we discourage it), use 10-point Times, boldface, initially capitalized, flush left, preceded by one blank line, followed by a period and your text on the same line.

2.3. Footnotes

Please use footnotes[1] sparingly. Indeed, try to avoid footnotes altogether and include necessary peripheral observations in the text (within parentheses, if you prefer, as in this sentence). If you wish to use a footnote, place it at the bottom of the column on the page on which it is referenced. Use Times 8-point type, single-spaced.

2.4. Cross-references

For the benefit of author(s) and readers, please use the

```
\cref{...}
```

command for cross-referencing to figures, tables, equations, or sections. This will automatically insert the appropriate label alongside the cross-reference as in this example:

> To see how our method outperforms previous work, please see **??** and Tab. 1. It is also possible to refer to multiple targets as once, *e.g.* to **????**. You may also return to Sec. 2 or look at **??**.

If you do not wish to abbreviate the label, for example at the beginning of the sentence, you can use the

```
\Cref{...}
```

command. Here is an example:

> **??** is also quite important.

2.5. References

List and number all bibliographical references in 9-point Times, single-spaced, at the end of your paper. When referenced in the text, enclose the citation number in square brackets, for example [4]. Where appropriate, include page numbers and the name(s) of editors of referenced books. When you cite multiple papers at once, please make sure that you cite them in numerical order like this [1–5]. If you use the template as advised, this will be taken care of automatically.

2.6. Illustrations, graphs, and photographs

All graphics should be centered. In LaTeX, avoid using the `center` environment for this purpose, as this adds potentially unwanted whitespace. Instead use

```
\centering
```

at the beginning of your figure. Please ensure that any point you wish to make is resolvable in a printed copy of the paper. Resize fonts in figures to match the font in the body text, and choose line widths that render effectively in print. Readers (and reviewers), even of an electronic copy, may choose to print your paper in order to read it. You cannot insist that they do otherwise, and therefore must not assume that they can zoom in to see tiny details on a graphic.

When placing figures in LaTeX, it's almost always best to use `\includegraphics`, and to specify the figure width as a multiple of the line width as in the example below

Method	Frobnability
Theirs	Frumpy
Yours	Frobbly
Ours	Makes one's heart Frob

Table 1. Results. Ours is better.

[1] This is what a footnote looks like. It often distracts the reader from the main flow of the argument.

```
\usepackage{graphicx} ...
\includegraphics[width=0.8\linewidth]
                {myfile.pdf}
```

2.7. Color

Please refer to the author guidelines on the CVPR 2022 web page for a discussion of the use of color in your document.

If you use color in your plots, please keep in mind that a significant subset of reviewers and readers may have a color vision deficiency; red-green blindness is the most frequent kind. Hence avoid relying only on color as the discriminative feature in plots (such as red *vs*. green lines), but add a second discriminative feature to ease disambiguation.

3. Methodology

3.1. General Overview

We will have a camera taking in video input of a player shooting the basketball, potentially from two different angles. One angle would be a side view of the shooter that can capture the relative positions of the shooter and the hoop, and the other would be a head-on view of the of the shooter to capture the full pose estimation of the shooter.

In the background, we will have our pose estimation algorithm running to capture the pose estimation of the shooter which will be used to make our predictions. Additionally, we will potentially be retrieving other data to create our feature vector, including the angle relationship between the shooter and the hoop, and the speed of the shot. We will be attempting to capture the pose estimation at multiple stages of the shot, as each stage can have a significant impact on the outcome of the shot. We will then take this feature vector and put it through a classification model. We will perform binary classification and predict whether a shot will go in or not.

References

[1] FirstName Alpher. Frobnication. *IEEE TPAMI*, 12(1):234–778, 2002. 2

[2] FirstName Alpher and FirstName Fotheringham-Smythe. Frobnication revisited. *Journal of Foo*, 13(1):234–778, 2003. 2

[3] FirstName Alpher and FirstName Gamow. Can a computer frobnicate? In *CVPR*, pages 234–778, 2005. 2

[4] FirstName LastName. The frobnicatable foo filter, 2014. Face and Gesture submission ID 324. Supplied as supplemental material `fg324.pdf`. 2

[5] FirstName LastName. Frobnication tutorial, 2014. Supplied as supplemental material `tr.pdf`. 2