

バスケットボールのシュート練習における シュート数と成否のリアルタイム可視化

坂口 泰章^{1,a)} 井尻 敬¹

概要：バスケットボールのシュート練習は、一般的に複数人が1つのリングを共有して実施される。本研究では、このシュート練習の振り返り支援とモチベーション向上を目的とし、練習中のシュートを自動検出し、その成否をリアルタイムに可視化する手法を提案する。具体的には、バックボード上部に設置した広角カメラで練習の様子を撮影し、映像中のボールを追跡することで「シュート判定」と「シュート成否判定」を行う。ボール追跡では、まずYOLOを用いて複数のボールを検出し、映像中のボールの予測位置およびカメラ3次元座標系におけるボール位置を利用して追跡を行う。さらに、追跡中のボールとリングの位置関係を用いてシュートの成否を推定する。各シュートの追跡後、シュート時の選手のサムネイル画像を作成し、シュート成否に応じて色付けして、コート横に配置したディスプレイ上に提示する。約5分間の練習動画を用いて「シュート判定」と「シュート成否判定」の精度を計測したところ、シュート判定のF値は0.99、シュート成否判定のF値は0.93と高い値を確認した。提案手法の有用性を検証するためにユーザスタディを実施した結果、提案手法によりシュート成否をリアルタイム可視化することで雑なシュートが統計的に有意に減ることが確認された。



図1：提案手法を用いたシュート練習の様子（左）と提案手法の可視化画面（右）。

1. はじめに

バスケットボールのシュート練習は、一般的に複数人が1つのリングを共有して実施される。このシュート練習において、シュート数や成功数を計測することは、各選手の実力把握や成長の記録のために重要である。しかし、複数人が同時に練習を行うため、各選手のシュート結果を目視で計測することは困難である。さらに、各個人の取り組みが見えにくいため、練習への集中力やモチベーションを維持しづらいという課題もある。

シュート練習における計測の支援やモチベーション向上を目的として、これまでに様々な研究やアプリケーションが開発されている。例えば、カメラでシュート練習の様子を撮影することで、自動でシュート結果を計測し、シュートのクリップを生成するスマートフォン用のアプリケーションがリリースされている[1], [2]。しかし、これらのアプリケーションは1人で実施する練習を想定しており、複数人で同時に行う状況には対応していない。

本研究では、複数人が同時に実施するシュート練習のモチベーション向上と振り返り支援を目的とし、練習中のシュートを自動検出し、その成否をリアルタイムに可視化する手法を提案する（図1左）。具体的には、バックボーダー

¹ 芝浦工業大学

a) al20101@shibaura-it.ac.jp

ド上部に設置した広角カメラで練習の様子を撮影し、映像中のボールを追跡することでシュート判定およびシュート成否判定を行う。各シュートの追跡後、シュート時のユーザのサムネイル画像を作成し、シュート成否に応じて色付けしたのち、コート脇に配置したディスプレイに表示する(図1右)。複数人で行う練習であってもシュートの結果がサムネイルとして可視化されるため、ユーザはより丁寧なシュートを心がけるなど、練習へのモチベーション向上が期待される。

提案手法の計測性能と有用性を評価するため、精度評価およびユーザスタディを実施した。まず精度評価では、事前に用意した約5分間のシュート練習動画中のシュート数とその成否を提案手法により計測し、この結果を著者が目視で確認した実際の結果と比較した。その結果、シュート判定のF値は0.99、シュート成否判定のF値は0.93であり、提案手法は高い精度でシュート判定を行えることを確認した。次にユーザスタディでは、6名の実験参加者に「アップ」「可視化なし」「可視化あり」という3つの条件で5分間のシュート練習を行ってもらい、各条件における雑なシュート(意図が伴わない、または、試合では打たない非実践的なシュート)を目視で計測した。計測の結果「可視化あり」は「可視化なし」に比べ、雑なシュートの割合が統計的に有意に減ることが確認された。

2. 関連研究

バスケットボールの試合や練習中の映像を活用し、プレイを分析・可視化する手法が研究されている。ここでは、試合の分析支援手法とシュート練習支援手法に分けて紹介する。

2.1 試合の分析支援手法

試合映像中のボールを追跡し、自動で得点計測を行う研究がなされている。Wenら[3]は、物体検出器であるYOLOv8[4]を拡張したDWC-YOLOv8を用いて、試合映像中のシュート数と得点数をリアルタイムに計測する手法を提案した。Fuら[5]は、Convolutional Neural Network(CNN)と背景差分法を組み合わせ、固定カメラで撮影された試合映像中の得点数を計測する手法を提案した。

試合映像からプレイの分析を行う手法も提案されている。Jia-Shingら[6]は、軽量化したYOLOv5を用いて試合映像中の選手をリアルタイムに検出し、コート俯瞰図上に動きを可視化するシステムを提案した。Yoonら[7]は、YOLOv5と画像上の距離に基づく追跡手法によって、試合映像中の選手とボールの動きを自動解析するシステムを提案した。Khobdehら[8]は、YOLOv4とLSTM[9]とFuzzy推論を組み合わせ、試合映像中の選手の動作を認識する手法を提案した。Fuら[10]は、NBAの試合のデータを活用し、様々な条件において選手やチームのデータを比

較できるシステムHoopInSightを提案した。しかし、これらの手法はいずれも試合中の得点計測やプレイ分析を目的としており、複数のボールが同時に存在するシュート練習のためのものではない。

2.2 シュート練習の支援手法

バスケットボールのシュート練習支援のための手法が研究されている。Linら[11]は、HoloLensを用いてユーザがAugmented Reality(AR)空間に入り、理想的なシュート軌道と自身のシュート軌道を比較・可視化することで、シュート改善を支援する手法を提案した。Liuら[1]は、シュート軌道と時間情報を1枚の画像にまとめたBasketball Energy Imageを用いることで、軽量なCNNによる成否判定を可能にし、スマートフォンで手軽にシュート結果の計測やハイライト生成、結果の共有を行うアプリケーションを開発した。他にも、シュートを自動検出し、シュート数や成功数の計測とクリップの生成が可能な商用アプリケーションもリリースされている[2]。しかし、これらのシステムは1人で練習する状況を想定しており、我々が着目する複数人で実施するシュート練習には対応していない。

我々の研究グループでは、本研究の前身となる手法をエンタテインメントコンピューティング研究発表会にて発表した[12]。この手法にはボールの追跡やシュート成否判定の精度をさらに向上させる余地があった。そこで本研究では、3次元カメラ座標を利用したボール追跡法、および、リング内におけるボールサイズの変化を考慮した成否判定アルゴリズムを新たに導入する。加えて本研究では、提案手法の有用性を評価するためのユーザスタディも実施する。

3. 提案手法

3.1 提案手法の全体像

提案手法は、魚眼レンズ付きのカメラ・PC・モニタから構成される(図2a)。バックボード上部に固定したカメラにより練習の様子を撮影し(図2b, c)、USB接続によってその映像をリアルタイムにPCへ転送する。PCでは、まず入力された映像中のボールを検出・追跡(3.2節参照)、次にそのボールがシュートであるかどうかを判定(シュート判定)する。ボールがシュートと判定された場合、ボールとリングとの位置関係とサイズ比を利用してシュート成否を判定し(3.3節参照)、最後にシュート時のユーザのサムネイルを作成して成否に応じて色付けを行い、モニタに表示する(3.4節参照)。

提案手法の可視化例を図2dに示す。画面左上にはリアルタイムの映像が、画面左下にはシュートを打った範囲を示すシュートマップが表示される。さらに、生成されたサムネイルが画面右側に一覧表示され、シュート成功の際は赤枠が、失敗の際は青枠が付与される。各個人のシュート結果がひとめで分かる形で可視化されるため、より丁寧な

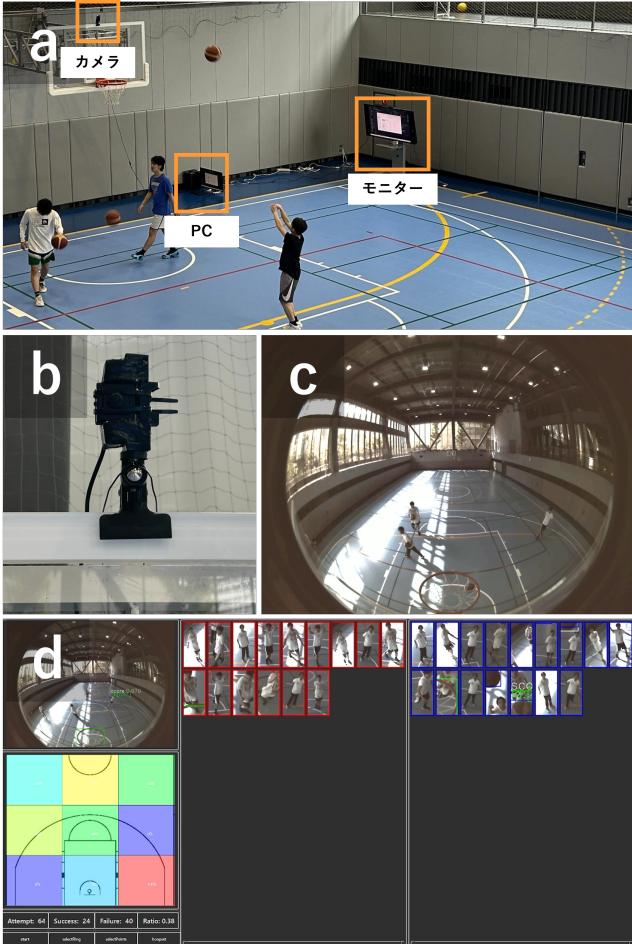


図 2: (a) 提案手法の構成. (b) バックボード上部に固定された広角カメラ. (c) 撮影される映像. (d) ユーザに提示される可視化画面.

シュート練習への動機づけにつながることが期待される.

3.2 ボールの検出と追跡

提案手法では映像中のボールを検出するため、CNNベースの物体検出モデルである YOLOv10[4] を追加学習して使用する。このため、練習時と同じ環境にて撮影された約 2400 枚に対して、手作業でボール位置をアノテーションしたデータセットを用意する。さらにこの自作データセットにオープンソースのデータセット [13] を統合し、全データを訓練データ 5839 枚、検証用データ 326 枚、テストデータ 214 枚に分割する。このデータを利用し、バッチサイズ 16、エポック数 100 として追加学習を行う。データセットの画像例を図 3 に示す。

提案手法は、連続して入力されるカメラ映像を利用し、リアルタイムでボールを追跡する。まず、入力映像の現在フレームに対して YOLOv10 を適用し、画像中のボールの矩形領域を検出する。検出した矩形領域が、追跡中のボールの移動予測位置の矩形領域と十分大きく重なる場合 (IoU が閾値 T_{IoU} 以上)，検出されたボールを追跡中のボールに



図 3: 学習に用いたデータセットの画像例。オープンソースデータセット（左）と自作データセット（右）。

対応付ける。一方、検出した矩形領域が追跡中のボールと対応付けされない場合、検出の信頼度が T_n 以上であれば新規に現れたボールとして新しい追跡を開始する。本研究ではこの処理に、カルマンフィルタによる移動予測を組み込んだ ByteTrack[14] を利用する。

ここでボールがカメラに近い場合には、映像中におけるボールの移動距離が相対的に大きくなり、かつ、リングやボードとの接触によりボールが不規則な動きをするため、前述の手法 [14] を単純に適用するだけでは、正確な追跡が困難である。そこで本研究では、追跡中のボールのサイズが閾値 T_s を超えた場合、カメラ座標系におけるボールの 3 次元位置を推定し、この位置を利用して追跡中のボールとの対応付けを行う。具体的には、追跡中のボールのうちサイズが閾値 T_s を超えたものと、現在フレームで検出されたボールのうちサイズが T_s を超えたものとの間で、カメラ座標系における距離による対応付けを行う。

魚眼カメラにより撮影された画像において、中心 (u, v) ・高さ r のボール矩形領域が検出されたとする。等距離射影モデルを仮定すると、このボールのカメラ座標系における 3 次元位置は下式の通り計算できる、

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = d \begin{bmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{bmatrix}. \quad (1)$$

ここで、 θ および ϕ は、それぞれ、カメラ座標系における入射角と方位角であり、 d はカメラから物体までの距離を表す（図 4）。これらの値は下式の通り計算できる、

$$\theta = \frac{\rho}{f} = \frac{\sqrt{(u - c_x)^2 + (v - c_y)^2}}{f}, \quad (2)$$

$$\phi = \tan^{-1} \left(\frac{v - c_y}{u - c_x} \right), \quad (3)$$

$$d = \frac{R}{\tan(r/f)}. \quad (4)$$

式 (2) は等距離射影モデル ($\rho = f\theta$) より計算され、 ρ は画像中心からの距離である。また、 (c_x, c_y) は画像の中心位置、 f は魚眼レンズのスケール係数、 R は実際のボールの半径である。

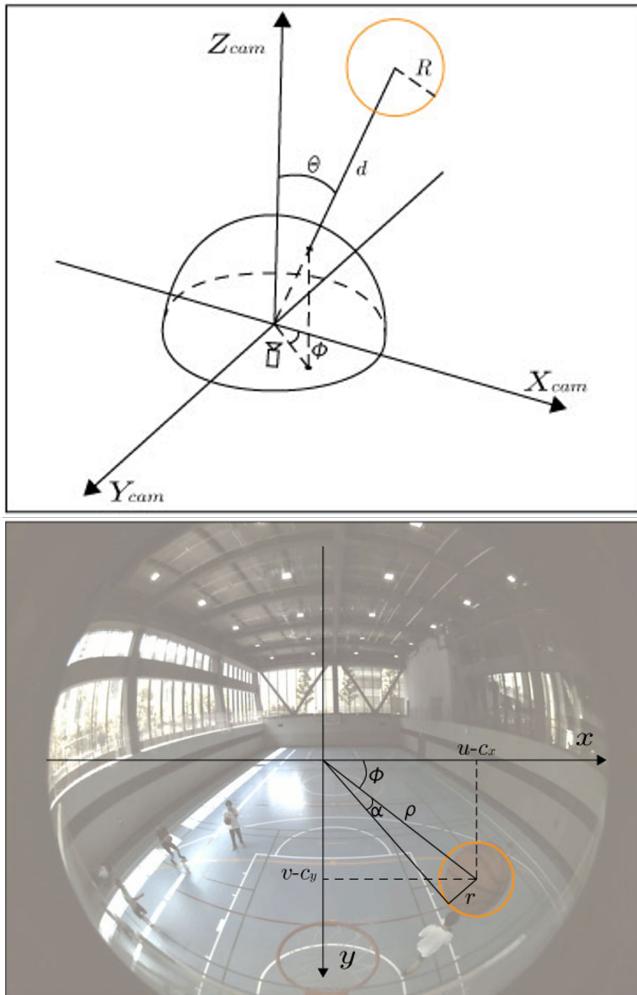


図 4: 提案手法で行った 3 次元位置推定におけるカメラ座標系（上）と画像平面上（下）の各パラメータ。

3.3 シュート判定とシュート成否判定

提案手法では、カメラがバックボードに固定されているため、シュートされたボールはリング付近で大きく撮影される。そこで本研究では、追跡中のボールのサイズが T_s を超えた場合に追跡中のボールを「シュート」と判定する。シュート判定を受けたボールについて、そのシュートの成否を判定する。シュート成否判定には、「ボールとリングの位置関係」と「リング内におけるボールの大きさ」を利用する。具体的には、シュートと判定されたボールの中心座標が N_g フレーム以上連続でリングのバウンディングボックス内に存在し、かつ、ボールのサイズが T_g 以下となった場合に「シュート成功」と判定する。一方、シュート成功判定を受けずに、ボールサイズが T_e 以下になった場合、「シュート失敗」と判定する。

前述のスケール係数 f については、利用したカメラ(ELP-USB16MP01-BL180)を用いて実際にボールを異なる距離で計測することで推定し、 $f = 600$ が得られた。また、本研究では、 $T_{IoU} = 0.1$ 、 $T_n = 0.3$ 、 $T_s = 0.39 \times L$ 、 $T_e = 0.3 \times L$ 、 $T_g = 0.4 \times L$ 、 $N_g = 3$ とした。なお、 L は

手作業でアノテーションしたリングのバウンディングボックスの対角線の長さである。今回はこれらのパラメータを経験的に設定したが、収集データを利用し機械学習によりパラメータを自動決定できる手法の確立が今後の課題である。

3.4 サムネイル生成

提案手法は、ユーザが各シュートの結果を把握できるようにシュート時のサムネイルを自動生成し、成否によって色分けして可視化する。追跡中のボールがシュートであると判定されたフレームから、過去 5 フレーム分に着目し、以下の処理を行う。まず、着目フレームにおけるボール位置の下方 550×550 ピクセルの矩形領域をおおまかに切り出し、この領域に対して MediaPipe を適用して人物の骨格推定を行う。推定に成功した場合、推定された腰位置を中心にして、 100×200 ピクセルの矩形領域を切り出し、サムネイル画像として記録する。その後、このサムネイルに対して、シュートが成功した場合は赤色、失敗した場合は青色の枠を付与して可視化する。なお骨格推定に失敗した場合はサムネイル生成は行わない。

4. 評価実験

4.1 精度評価

提案手法のシュート判定およびシュート成否判定の精度を評価するために実験を行う。芝浦工業大学豊洲キャンパスの体育館にて、本学のバスケットボール部員 3 名にシュート練習を行ってもらい、その様子を撮影した 5 分間の動画を用意する。この動画中の「シュート数」「シュートの成功数」「シュートの失敗数」を提案手法により自動で計測し、この結果を目視にて確認した実際の数と比較する。

用意した練習動画に対して提案手法によりシュート判定を行った際の混合行列を表 1 に示す。この練習動画には 94 本のシュートが含まれていた。提案手法は、追跡したボールが「シュートである」場合に検出をするものであるため、「実際には非シュート」かつ「シュート判定なし」に対応するセルに値は存在しない。このデータより、提案手法のシュート判定に関する適合率は 0.98、再現率は 1.00、F 値は 0.99 であり、提案手法は非常に高い精度でシュート判定をできたことがわかる。提案手法は、ボールがカメラに近い場合に、追跡アルゴリズムを切り替える。これが、移動量が大きくかつ衝突により不規則な動きをするボールの正確な追跡に寄与したと考えられる。一方で、2 度の誤判定が確認された。詳細を確認したところ、YOLO がシュート判定対象となる 1 つのボールに対して 2 つの領域を検出したためであった。

用意したデータに含まれる 94 本のシュートに対する提案手法のシュート成否判定に関する混合行列を表 2 に示す。シュート成否判定に関する正答率は 0.95、適合率は 0.94、

再現率は 0.92, 特異率は 0.97, F 値は 0.93 であり, 非常に高い精度で判定できた. 提案手法はリングとボールの位置関係だけでなく, リング内のボールサイズも考慮し成否の判定を行っているため, シュートがリング付近で不規則な動きをした場合でも正しく成否判定できたと考えられる.

表 1: 提案手法の自動シュート判定結果と実際のシュートの有無の比較.

	実際にシュート	実際には非シュート
シュート判定した	94 本	2 本
シュート判定なし	0 本	-

表 2: 提案手法により自動判定したシュート成否と実際の成否の比較.

	成功したシュート	失敗したシュート
成功と判定	33 本	2 本
失敗と判定	3 本	56 本

4.2 ユーザスタディ

提案手法による可視化がシュート練習に与える影響を調査することを目的とし, ユーザスタディを実施する(図 5). 実験参加者は 3 名 1 組のグループとなり, 「アップ」「可視化なし」「可視化あり」の 3 条件でシュート練習を実施する. 「アップ」は, 通常のウォーミングアップに相当するもので, 必ず最初に実施する. 「可視化なし」では提案手法を用いずに練習を行い, 「可視化あり」では提案手法により結果を可視化しながら練習を行う. 3 つのそれぞれの条件において, 5 分間シュート練習を行い, 条件間には 1 分間の休憩を挟む. また, 条件の実施順序による影響を軽減するために, 参加者を 2 グループに分けて, グループ間で「可視化なし」「可視化あり」の順序を入れ替えて実施する.

本研究では, 評価指標として『雑なシュート』の回数を計測した. ここで雑なシュートとは, 意図が伴わないシュートや試合では打たない非実践的なシュートと定義する. 例えば, 周囲との会話に気を取られ, リングを確認せずに打ったシュートや, 不適切なフォームでボールを投げているものが該当する. この雑なシュートの回数が少ないほど, より真面目に練習に取り組んでいるとする. 計測のため, ギャラリーに設置したビデオカメラで実験の様子を撮影し, 筆者が映像を確認して雑なシュートの判定を行った.

現役のバスケットボール部員 3 名および引退 1 年後の部員 3 名, 計 6 名の協力のもとユーザスタディを実施した. 引退した 3 名をグループ A, 現役の 3 名をグループ B として実験を行った. 各条件における雑なシュートの割合を表 3 に示す. 2 グループ合計した雑なシュートの割合について, 2 群間の母比率の差の検定(χ^2 検定)を行ったところ



図 5: ユーザスタディの様子.

ろ, 「可視化なし」と「可視化あり」の間において有意な差 ($p < 0.01$) が確認された. このことから, 提案手法による可視化が参加者の意識に影響を与え, より丁寧にシュートを行うように促した可能性が示唆される.

続いて, 3 条件それぞれにおける, シュート成功率を表 4 に, シュートの種類の内訳を図 6 に示す. なお図 6 は, 雜なシュートを除外した結果である. 2 グループの合計シュート成功率について同様に χ^2 検定を行ったところ, 「可視化なし」と「可視化あり」の間において有意な差 ($p < 0.05$) が確認され, 「可視化あり」のほうが成功率は低かった. これは「可視化なし」ではレイアップやリングから近い位置のシュートが多かった一方, 「可視化あり」では 3 ポイントなどの距離のあるシュートが増加したためと考えられる(図 6). 今回着目したシュート練習においては, 本来 3 ポイントや距離のあるミドルレンジからのシュートタッチを養うことが重視されるため, 可視化によって参加者がより実践的な距離からシュートに取り組んだ結果と考えられる.

表 3: ユーザスタディにおける雑なシュートの計測結果.

	アップ	可視化なし	可視化あり
グループ A	14%(9/65)	17%(14/81)	12%(10/81)
グループ B	7%(7/94)	21%(17/80)	7%(6/91)
合計	10%(16/159)	19%(31/161)	9%(16/172)

表 4: ユーザスタディにおけるシュート成功率の計測結果.

	アップ	可視化なし	可視化あり
グループ A	55%(36/65)	47%(38/81)	36%(29/81)
グループ B	40%(38/94)	53%(42/80)	37%(34/91)
合計	47%(74/159)	50%(80/161)	37%(63/172)

5. まとめと展望

本研究では, 複数人で実施するバスケットボールのシュート練習のモチベーション向上を目的とし, 練習中のシュート成否をリアルタイムに可視化する手法を提案した. 提案手法の精度評価のために 5 分程度の練習動画を用意し, 提

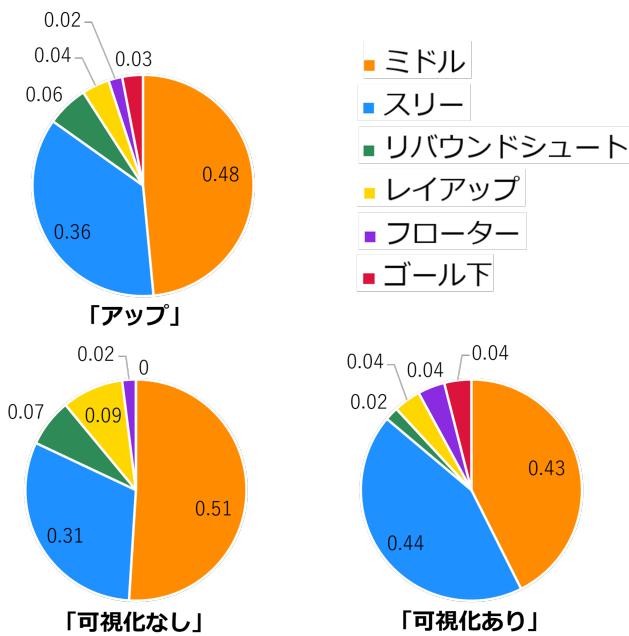


図 6: 雜なシュートを除外した各条件の合計シュート種別分布。

案手法による自動計測結果と実際の結果を比較した。その結果、シュート判定における F 値は 0.99、シュート成否判定における F 値は 0.93 と、シュート練習のリアルタイム可視化という目的においては十分高い精度を確認した。さらに、提案手法の有用性を検証するために、ユーザスタディを行った。その結果、提案手法を用いることで、雑なシュートの割合が減少する傾向や、提案手法がユーザの練習への意識に影響を与え、より丁寧にシュートを行うように促す可能性が示唆された。

提案手法の将来課題として、サムネイル生成におけるエラーが挙げられる。現在は、シュート成否までは識別できるものの、サムネイル生成時に画像中の人の検出できず、サムネイルが生成されないケースが少なくない。より高画質なカメラの導入や、高性能な人物検出器の導入を検討したい。さらに、将来展望として雑なシュートの自動検知も挙げられる。成功・失敗のサムネイルだけでなく、雑なシュートを行った際のサムネイルもリアルタイムに提示することで、ユーザのモチベーション維持に寄与できると考えられる。骨格推定により抽出した動作データを用い、機械学習によって通常のシュートと雑なシュートを分類する手法の検討を進めたい。

参考文献

- [1] Liu, W., Yan, C. C., Liu, J. and Ma, H.: Deep learning based basketball video analysis for intelligent arena application, *Multimedia Tools and Applications*, Vol. 76, pp. 24983–25001 (2017).
- [2] Nex Team Inc.: HomeCourt, <https://www.homecourt.ai/> (2018).
- [3] Wen, Z., Zou, J. and Guang, L.: Basketball Shooting and Goalie Detection Based on DWC-YOLOv8, *Computer Networks and IoT* (Jin, H., Pan, Y. and Lu, J., eds.), Singapore, Springer Nature Singapore, pp. 98–107 (2024).
- [4] Redmon, J., Divvala, S. K., Girshick, R. B. and Farhadi, A.: You Only Look Once: Unified, Real-Time Object Detection, *CoRR*, Vol. abs/1506.02640 (2015).
- [5] Fu, X.-B., Yue, S.-L. and Pan, D.-Y.: Camera-based basketball scoring detection using convolutional neural network, *International Journal of Automation and Computing*, Vol. 18, No. 2, pp. 266–276 (2021).
- [6] Sheu, J.-S. and Lin, S.-J.: An Efficient Application of Modified YOLOv5 in Basketball Player Detection and Analysis, *Advances in Technology Innovation*, Vol. 9, No. 3, p. 157–171 (2024).
- [7] Yoon, Y., Hwang, H., Choi, Y., Joo, M., Oh, H., Park, I., Lee, K.-H. and Hwang, J.-H.: Analyzing Basketball Movements and Pass Relationships Using Realtime Object Tracking Techniques Based on Deep Learning, *IEEE Access*, Vol. 7, pp. 56564–56576 (2019).
- [8] Khobdeh, S. B., Yamaghani, M. R. and Sareshkeh, S. K.: Basketball action recognition based on the combination of YOLO and a deep fuzzy LSTM network, *The Journal of Supercomputing*, Vol. 80, No. 3, pp. 3528–3553 (2024).
- [9] Donahue, J., Anne Hendricks, L., Guadarrama, S., Rohrbach, M., Venugopalan, S., Saenko, K. and Darrell, T.: Long-Term Recurrent Convolutional Networks for Visual Recognition and Description, *CVPR* (2015).
- [10] Fu, Y. and Stasko, J.: HoopInSight: Analyzing and Comparing Basketball Shooting Performance Through Visualization, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 30, No. 1, pp. 858–868 (2024).
- [11] Lin, T., Singh, R., Yang, Y., Nobre, C., Beyer, J., Smith, M. A. and Pfister, H.: Towards an Understanding of Situated AR Visualization for Basketball Free-Throw Training, *CHI '21*, New York, NY, USA, Association for Computing Machinery (2021).
- [12] 坂口泰章, 井尻 敏: バスケットボールのシュート練習におけるシュート成否のリアルタイム可視化, 技術報告 (2024).
- [13] Physics, T.: Basketball Dataset, <https://universe.roboflow.com/thesis-physics/basketball-1zkue> (2024).
- [14] Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., Luo, P., Liu, W. and Wang, X.: ByteTrack: Multi-object Tracking by Associating Every Detection Box, *Computer Vision – ECCV 2022* (Avidan, S., Brostow, G., Cissé, M., Farinella, G. M. and Hassner, T., eds.), Cham, Springer Nature Switzerland, pp. 1–21 (2022).