

畳み込みニューラルネットワークを用いた 複数カメラ環境における背番号認識手法

渡邊 良亮^{1,a)} 野中 敬介¹ 三功 浩嗣¹ 内藤 整¹

概要: 近年、スポーツ映像を利用して戦術の分析を行うことや、スポーツ映像内での同一選手の追跡を目的として、選手の背番号を認識し、選手の同定を行う技術について研究が行われている。スタジアム等に複数のカメラを配置して背番号認識を行う場合、いずれかのカメラが特定選手の背番号を捉える可能性が高くなるが、背番号が映っていない方向からの画像も大量に抽出されるため、それが誤認識を招く原因となる。そこで本研究では、背番号が映っている可能性の低い画像を選手の移動方向に基づき除外することに加え、ニューラルネットワークの出力関数から得られる確率を基に誤認識を抑制することが可能な畳み込みニューラルネットワークを用いた背番号認識手法について提案する。

Jersey Number Recognition Method Using Convolutional Neural Network with Multiple Cameras

RYOSUKE WATANABE^{1,a)} KEISUKE NONAKA¹ HIROSHI SANKOH¹ SEI NAITO¹

1. はじめに

近年、スポーツ映像を利用して戦術の分析を行うことや、スポーツ映像内での同一選手の追跡を目的として、選手の背番号を認識し、選手の同定を行う技術についての研究が行われている [1–3]。選手を同定する上で背番号は有力な情報になるものの、単一のカメラを用いて認識を行う場合には、選手の立つ角度や重なりなどによって、背番号が見えないシーンが多く発生し、動画の中で高い認識精度で正しい背番号を取得し続けることは難しかった。

そこで、本研究ではスポーツ映像内での選手の同定のために、複数カメラを用いた背番号認識手法の提案を行う。競技場に設置された複数のカメラを用いて背番号認識を行う場合、いずれかのカメラが特定選手の背番号を捉える可能性が高くなるが、背番号が映っていない画像も大量に抽出されるため、誤認識を招く原因となる。よって、本研究では選手の追跡による移動ベクトルを基に背面を映し出す可能性の高いカメラを推定することで、背番号が映ってい

る可能性の低い画像を効率的に排除し、誤認識を抑制することのできる複数カメラからの背番号認識手法について提案を行う。

2. 関連研究

選手の背番号を基に選手の同定を行う試みは、過去多く提案されてきた。そのアプローチとして、既存の光学文字認識を基に背番号の認識を行う手法 [1] や、SIFT に代表される特徴点を基にマッチングを行い背番号の認識を行う手法 [2] が提案されてきた。これらの手法は、用意したテンプレートとのずれが大きくなると認識精度が低下するという懸念があり、様々な角度や姿勢の画像が得られるスポーツ競技において、高い精度を出すことは難しかった。

そこで、ニューラルネットワークに代表される深層学習手法を適用することが考えられてきた。背番号認識への深層学習を適用した研究として、畳み込み層 3 層、全結合層 3 層から成る畳み込みニューラルネットワーク (Convolutional Neural Network : CNN) を用いることで約 83% という高い精度を実現したことが報告されている [3]。しかし、この約 83% という精度は、背番号が見える画像を手動でピッ

¹ KDDI 総合研究所
KDDI Research, Inc.

^{a)} ru-watanabe@kddi-research.jp



図 1: 複数カメラからの映像

クアアップして精度を検証しているに過ぎず、背番号が見えない画像が抽出された場合にどのような結果を引き起こすかに関しては検討が成されていない。特に今回のように複数カメラを用いると、いずれかのカメラが背番号を捉える可能性は高くなるものの、他のカメラは背番号領域を捉えることができず、誤認識を招いてしまう。そこで、本研究では複数カメラにて効率的に背番号認識の精度を高める手法について提案を行う。

3. 提案手法

本章では提案手法の詳細について述べる。本提案手法では図 1 のような時刻同期が取れた複数の固定カメラを競技場に設置して用いることを想定している。

提案手法では初めに背景差分法に基づいて得たシルエット画像に視体積交差法を適用することで人物位置の推定を行う。次に、得た人物位置の人物の画像の色情報を基に選手のチーム分類を行う。その後、背番号部分の抽出を行い、抽出された背番号に対し CNN を用いた認識を行うが、背番号部分の抽出を行う際に追跡で得られる移動ベクトルを基に、選手の背面となりやすい方向を推定し、背面を映し出している可能性が高いカメラからの画像を抽出し、背番号認識を行う。最後に、認識結果と、過去フレームからの追跡によって継承された認識結果を統合し、最終的な認識結果を算出する。本提案手法の処理手順を図 2 に示す。

3.1 選手位置推定手法

背番号の認識による選手の同定を行うためには、まず選手の位置と、その位置に存在する選手の背番号部分を抽出する必要がある。今回は複数カメラからの人物位置を推定する方法として、視体積交差法 [4] に基づく位置推定方法を適用した。視体積交差法は複数のカメラから撮影された対象領域のシルエット画像を基に、各シルエットを断面とする錐体の積集合を計算することで 3D モデルを得る手法である。視体積交差法で位置推定を行うためにはシルエット画像を得る必要があるが、今回は背景差分法 [5] を基にシルエット画像を取得することとした。視体積交差法で得られた 3D モデルに関し、一定以上の大きさを持つボクセルの塊を選手であると判断することで、選手位置の推定を

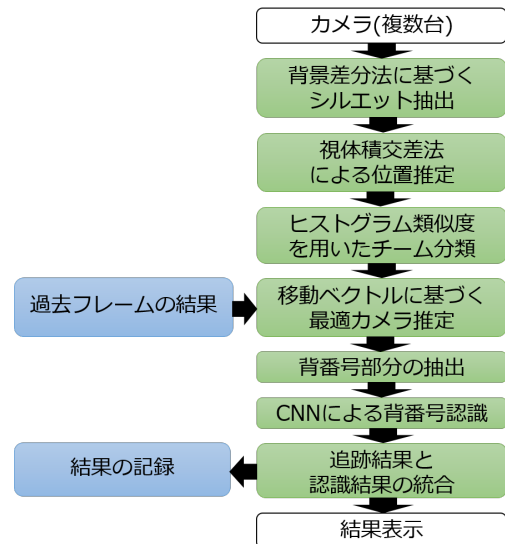


図 2: 提案手法の処理手順

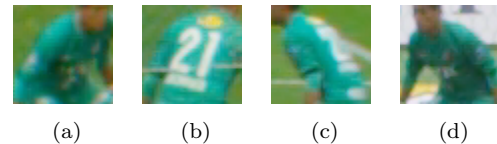


図 3: 各カメラの抽出画像の例

行った。

次に背番号部分を抽出するため、前記の計算で得られた選手の視体積部分をカメラ方向に逆投影した際にできる選手部分のシルエット画像に対し、重心位置を計算する。重心位置の計算法としては、対象画像 $f(x, y)$ の $(p + q)$ 次モーメントが

$$M_{pq} = \sum_x \sum_y x^p y^q f(x, y) \quad (1)$$

で定義されるとき、重心座標 (G_x, G_y) は次の式で求められる。

$$(G_x, G_y) = \left(\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}} \right) \quad (2)$$

こうして得られた重心位置に対して、その周囲を 50×50 [pixels] のサイズで切り取ることで背番号画像とした。抽出された画像の例を図 3 に示す。今回の実験では複数台のカメラを用いていることから、多くの背番号が映っていない画像も出力されるが、この除去は後段の背番号認識部分等で行うこととした。

3.2 チーム分類手法

選手のユニフォームの色に基づいて所属チームの分類を行う。文献 [1] と同様に、Bhattacharyya 距離に基づいて、ヒストグラム類似度を計算することでチームの分類を行った。選手の視体積部分をカメラ方向に逆投影した際にできる選手部分のシルエット画像を基に、元のカラー画像をマスクした選手画像に対し、YCrCb 色空間にてヒストグラ

表 1: ネットワーク構成表

Type	Filter	Stride	Output	Option
conv1	7×7	1×1	$50 \times 50 \times 32$	ReLU LRN
pool1	3×3	2×2	$25 \times 25 \times 32$	maxpool
conv2	5×5	1×1	$25 \times 25 \times 64$	ReLU LRN
pool2	2×2	2×2	$13 \times 13 \times 64$	maxpool
conv3	3×3	1×1	$13 \times 13 \times 96$	ReLU
conv4	3×3	1×1	$13 \times 13 \times 96$	ReLU
conv5	3×3	1×1	$13 \times 13 \times 64$	ReLU
pool3	2×2	2×2	$7 \times 7 \times 64$	maxpool
fc1	-	-	800	dropout(0.5) [7]
fc2	-	-	100	dropout(0.5)
fc3	-	-	50	softmax

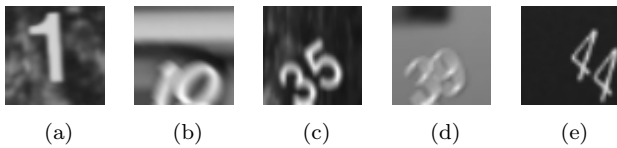


図 4: 実験に用いた訓練画像の例

ム類似度を計算することで所属チームの分類を行った。また、実験にて後述するが、今回は 5 種類のユニフォームについて判別できればよいことから、各ユニフォームの色の分布は事前に既知で得られるものとして、事前に用意したヒストグラムと、前述の選手画像から得られるヒストグラムに対し類似度を計算し、複数台のカメラからの類似度スコアの中で最小となるスコアを最終結果のチーム分類として採用することとした。

3.3 背番号認識手法

本手法では CNN を用いて、抽出された背番号画像の認識を行う。その際のモデルとして AlexNet [6] を参考に、畳み込み層 5 層と全結合層 3 層から成るニューラルネットワークを構築し、出力層の活性化関数には softmax 関数を用いた。表 1 に、用いたモデルの構成とパラメータの詳細を示す。今回の実験で対象としたシーンに登場する選手の背番号が最大で 49 だったことから、分類は 0~49 までの 50 クラスとした。本ネットワークの学習には、図 4 のようなフォントを重畳した人工的な画像を用いた。多数の任意の背景画像を用意し、その背景画像上に様々なフォントの数字をランダムに回転させて重畳することで、大量の人工的な学習画像を生成した。また背番号画像特有の動きのぶれなどを表現するために、ガウシアンフィルタやブラーをランダムな比率で重ねることで様々な背番号画像に対して適用できる学習画像を生成した。重ねる番号がわかっていることから、正解ラベルを手動で振る必要がなく、モデルの作成を全自動で行うことができる。今回はこのような人工画像から作成したモデルを予め用意しておくことにより、認識を行うこととした。

また、本手法においては複数のカメラを用いていることから、実際には背番号が映っていない画像が多く抽出されてしまう。この画像をそのまま認識させてしまうと誤認識を招くことから、出力関数として用いた softmax 関数にて i 番目のクラスの確率を計算する際に、

$$y_i = \frac{\exp(a_i)}{\sum_j \exp(a_j)} > p_{th} \quad (3)$$

上記の式 (3) のように閾値 p_{th} を上回ったもののみ認識成功とし、認識失敗に分類されたものの番号認識結果は採用しない。ただし、(3) 式の a_i は i 番目のクラスの CNN の出力を示し、 y_i は i 番目のクラスの確率を示している。その結果、複数のカメラで認識成功となった場合には認識結果の多数決を行う。これは、例えばカメラが 4 台の際に認識結果がそれぞれ「21, 27, 21, 認識失敗」であったときに、多数決で 21 を採用することを意味している。さらに「21, 27, 認識失敗, 認識失敗」のように同数で結果が割れた場合には、21 と 27 それぞれの y_i の値を比較し、大きい方を採用することとした。

3.4 追跡手法

背番号認識を行う上では、各フレームで全ての選手に対し番号認識を行う必要はなく、例え認識に失敗したとしても以前のフレームから追跡ができていのであれば、過去フレームの認識結果を採用すればよい。また、追跡の中で数フレーム連続で同じ番号が認識されれば、その番号は信頼性が高いと言える。よって本提案手法では選手の追跡と、追跡成功時に各認識結果の信頼度の計算を行い、信頼度の高い番号を採用する。

まず、本研究では簡易な追跡法として、現フレームの人物位置ベクトル $\mathbf{P}$ に対し、1 つ前のフレームで同一チームに分類されている全ての人物の位置ベクトル $\mathbf{B}_1$ との距離 d を以下のように計算する。

$$d = |\mathbf{P} - \mathbf{B}_1| < d_{min} \quad (4)$$

このときに距離の閾値 d_{min} よりも小さい距離の人物が存在する場合には、その中で最小の距離の人物から追跡が行われていることとした。1 フレーム間で移動する距離は非常に小さいが、位置推定の誤差等も考慮して本実験では $d_{min} = 1[m]$ とした。

次に信頼度の計算法を説明する。3.3 節の認識手法において、ある選手に対して番号 n が認識されたときに、本選手は n を候補番号として取得し、信頼度の計算を開始する。 t フレーム目での番号 n の信頼度を $R(t, n)$ としたとき、

$$R(t, n) = R_{rec}(n) + rR(t-1, n) \quad (5)$$

として計算される。 $R_{rec}(n)$ は現フレームの認識結果が n であったときに 1 となり、認識結果が n でなければ 0 となる。また、 r は時間経過と共に信頼度を落とすための信頼

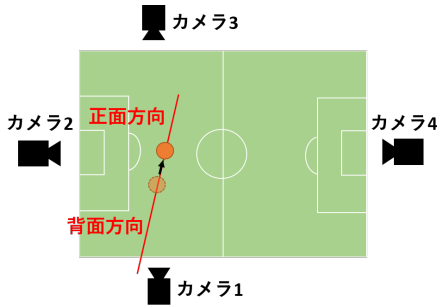


図 5: 移動に基づく背面方向推定

度の減少率を示している。ただし、 $t = 0$ の場合には追跡の結果が存在しないため、 $R(t-1, n)$ は 0 とする。本研究では $r = 0.9$ として信頼度計算を行い、各フレームで $R(t, n)$ が最大となる n を認識結果として表示することとした。

3.5 追跡ベクトルを用いた最適カメラ推定

本手法では複数カメラの映像を用いることから、図 3 のように 1 人の選手に対してカメラ台数分の画像が抽出されてしまう。これに対し、式 (3) のように閾値 p_{th} を下回る結果を認識失敗と判断することで、背番号の映っていない画像をある程度は排除可能であるが、完全に除去できるわけではない。このことから、本手法では選手の移動ベクトルを用いて最適なカメラの推定を行う。一般にサッカーなどの競技において選手が移動する場合、図 5 のように移動方向から選手の背面の可能性が高い方向を推定することが可能であり、選手の移動していく方向の反対にあるカメラが背番号を捉える可能性が高くなる。本手法では、追跡が成功している場合においては、選手の移動方向を知ることが可能であるため、これを利用して背面の可能性が高いカメラを推定する。

ここで、ある選手に対して s フレームに渡り、式 (4) に基づいて追跡が成功している場合において、 s フレーム前のフィールド平面上の選手位置ベクトルを B_s とし、現フレームの選手位置ベクトルを P としたときに得られる単位移動ベクトル

$$T = \frac{P - B_s}{|P - B_s|} \quad (6)$$

を計算する。ただし、比較的直近の選手位置を基に移動ベクトルを割り出すため、本実験では s が 30 を上回る場合には全て $s = 30$ とし、30 フレーム前の位置との移動ベクトルを割り出すこととした。また、1 フレームの移動量は微小で誤差も大きいため、 $s < 10$ のケースでは、最適カメラ推定を行わず全てのカメラからの画像を認識に掛けることとした。次に、この T と、カメラ番号 i のカメラの単位方向ベクトル C_i との内積を以下のように計算する。

$$d_i = T \cdot C_i \quad (7)$$

このとき内積 d_i が最大となるカメラ i からの画像のみを抽

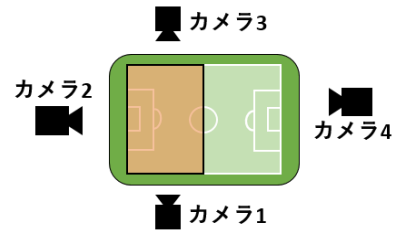


図 6: カメラ配置



図 7: 実験に用いた抽出画像の例

出し、背番号認識を行うこととした。今回の実験ではカメラ 4 台という少ない台数を用いているため、内積が最大となるカメラ 1 台を算出し最適カメラとしたが、実際にさらに多くのカメラがある場合には、内積値が大きいものを数台選択するような実装も考えられる。

また、今回の実験では移動ベクトルと同時に s フレーム前との移動量を計算し、一般的な歩行の速度とされる $1[m/s]$ 以下の速度で移動している場合には、選手がゆっくりと後ずさりをしたり、方向転換をしている可能性を考慮し、最適カメラ推定を行わず全てのカメラからの画像を認識に掛けることとした。

4. 実験

本手法の有用性を示すために 3 つの実験を行った。また今回の実験に際し、図 1 のような固定の 4 台の 8K カメラにて撮影した約 15 秒のサッカーの映像を用い、図 6 のようにハーフピッチを対象として実験を行った。

4.1 実験 1: 作成モデルの精度の確認

まず 3.3 節で紹介した背番号認識手法の精度を確認するために、背番号が見える画像に絞って認識精度を確認した。前述の固定の 4 台の 8K カメラにて撮影した 538 フレームから成るシーンについて、30 フレームごとの頻度でフレームを選択した $30n(n = 0, 1, \dots, 17)$ フレーム目の画像に対し、3.1 節の背番号抽出手法を用いて背番号部分の画像を抽出した。本手法では背番号ではない画像も大量に抽出されるが、精度確認のために目視で正しい背番号を判定可能なものだけを手動でピックアップした。その結果として 178 枚の背番号画像から成る検証用のデータセットを作成した。実際に抽出された画像の例を図 7 に示す。今回は、目視で正しい背番号を判定可能なものを対象としているため、図 7(a)(b) のような正面から比較的に鮮明に背番号が見える画像もあれば、図 7(c)(d) のような背番号の一部が欠けたものや、図 7(e) のような斜めから背番号を見た画像も

表 2: 閾値 p_{th} による正解率の変化

p_{th}	総枚数	正解枚数	正解率 [%]
0	178	124	69.7
0.5	153	118	77.1
0.55	149	116	77.8
0.6	144	114	79.2
0.65	139	113	81.3
0.7	134	112	83.6
0.75	131	109	83.2
0.8	130	108	83.1
0.85	118	102	86.4
0.9	113	101	89.3
0.95	106	97	91.5

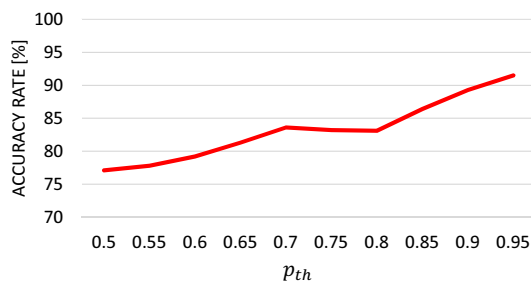


図 8: 閾値 p_{th} による正解率の変化のグラフ

混在するデータセットとなっている。これらの画像から背番号を認識するために、3.3 節で紹介した CNN を用いたモデルを 50 クラスで作成した。学習画像の枚数は 60000 枚 $\times$ 50 クラスで 300 万枚とした。実験の結果、178 枚の背番号画像に対して、後処理なしで認識を行った場合、124 枚 (69.7%) において正しい背番号を認識することができた。

次に、式 (3) で示される閾値 p_{th} の値を変更し、それに基づいて認識結果の精度にどのような変化が及ぶのかについて確認を行った。 p_{th} を 0.5 から 0.05 ずつ変化させ、確認した結果が表 2 と図 8 である。 p_{th} の値を高くすることで、正しく映っている背番号画像も除外されてしまうものの、 p_{th} を高めるにつれて誤認識が減っていることがわかる。本来、サッカーのようなコンテンツでは追跡により過去フレームから認識結果を引き継ぐことが可能なため、例えば認識が失敗に終わっても追跡に基づいて認識結果を得ることが可能であり、極力誤認識を減らすことが重要であると考えられる。よって p_{th} を高く設定することは有効であることが確認できた。

4.2 実験 2: 移動ベクトルによるカメラ選択の有効性確認

実験 2 として、移動ベクトルを用いた最適カメラ推定の有効性を確認するために、移動ベクトルを用いた場合と、用いない場合について比較を行う。実験 1 と同様に 538 フレームから成るシーンから、30 フレームごとの頻度でフレームを選択し、3.1 節の背番号抽出手法で自動抽出される画像全てを集めたところ 1296 枚となった。これを移動ベク

表 3: 移動ベクトルを適用した場合の正解率

dataset	総枚数	有効枚数	正解数	正解率 [%]
NO_VECTOR	1296	182	97	53.3
VECTOR	546	108	75	69.4

トルを用いていない場合のデータセット (NO_VECTOR) とした。次に、この 1296 枚の抽出画像に対し、移動ベクトルを用いて最適カメラを推定した場合に 546 枚の画像を得ることができた。よって、この 546 枚を移動ベクトルを用いた場合のデータセット (VECTOR) とした。このとき、546 枚の画像の中で、実験 1 で手動抽出したのと同様に背番号が見える画像を手動で確認したところ 119 枚であった。総枚数は 1296 枚から 546 枚と 42.1% の画像が残されているのに対し、背番号が見えている画像に限定すると、178 枚から 119 枚と減少はしているものの 66.9% が残されている。このことから、移動ベクトルを用いることで背番号が映っていない画像を効率的に除去できることがわかる。

次に、移動ベクトルを用いていない 1296 枚のデータセットと、移動ベクトルを用いた 546 枚のデータセットそれぞれに対し、実験 1 と同様のネットワーク ($p_{th} = 0.9$) で処理を行ったところ、正解率は表 3 のようになった。ただし、 $p_{th} = 0.9$ として一定確率以下の画像は全て除去することで有効枚数としている。この有効枚数に対し、正解枚数を確認することで正解率とした。移動ベクトルを用いた場合には、背番号が映っている画像と映っていない画像が混在している状況でも、69.4% で正解することができていることがわかる。以上の実験により、移動ベクトルを用いた場合、正しく抽出されている背番号画像も除去されてしまうという懸念はあるものの、全体として誤認識を減らすことが可能であるということが示された。

4.3 実験 3: 平面マップの作成

本研究の成果物として、フィールドのハーフピッチに対し図 10 のような平面マップを作成し、表示を確認した。平面マップ上では人物位置に円形のマーカーを表示し、マーカー上に背番号認識結果を表示した。また人物について白のユニフォームのチームを A チーム、赤と青の混在したユニフォームのチームを B チームとし、A チームの選手 (赤)、A チームのキーパー (桃)、B チームの選手 (青)、B チームのキーパー (黄)、審判 (緑) の 5 色にチーム分類し、その色に応じたマーカーを表示している。ただし、今回は平面マップに重ねる上で、両チームの出場選手の番号は既知であるものとし、出場していないはずの番号が認識された場合には、その認識結果は破棄するものとした。

平面マップの比較を行うために、確率の閾値 p_{th} に基づく画像除去と移動ベクトルに基づく最適カメラ推定を用い



図 9: カメラからの映像 (350 フレーム目)

ない条件で作成した動画 (動画 1) と, $p_{th} = 0.9$ とし, 移動ベクトルに基づく最適カメラ推定を行っている動画 (動画 2) を作成し, $50n(n = 1, 2, \dots, 10)$ フレーム目として示される計 10 フレームを抜き出して精度を確認した. 確認方法としては, 図 9 のようにカメラ映像から目視で正解を確認し, チームと番号が共に正しいマーカーを正解とした. 前述の 10 フレームを通して, 延べ 192 名の人物が存在していたが, 動画 1 では 128 名 (66.7%), 動画 2 では 149 名 (77.6%) が正しく表示されており, 動画 2 の方が 10%ほど高い精度を記録した. これは提案手法によって, 正しく背番号が映り込んだ画像も一部除去されてしまうという懸念に対し, 本実験の条件において問題は発生せず, 精度向上を実現できたことを意味している.

しかしながら, 戦術の分析用途等に応用することを考えれば 77.6%という精度は十分とは言えない. そこで, 精度が上がらない原因について考察する. 動画 2 の結果の一部を図 10 に示す. ただし, 図 10 の黄色の枠は, 正解していることを表しており, この枠は後から手動で付与したものである. 図 10 のような比較的選手が離れて存在しているシーンでは高い精度を発揮できることがわかる. 一方, 精度が低下しているシーンの例としては図 11 のように選手が密集している部分で精度が落ちているケースがある. このフレームでは, 平面マップ上の赤枠の部分以外では全て正解できているものの, 赤枠部分は選手が密集しており, 多くの誤りが生じている. これは, 選手が密集する場面では各選手の視体積を綺麗に分割して生成できないことから位置推定や追跡の精度が落ちているためであると考えられる. このような密集部分では背番号の見づらさなどから番号認識の精度も低下することから, オクルージョン領域では精度面に課題がある. よって, 今後はこのような領域の精度を向上させる新たなアプローチの提案が求められる.

5. 結論

複数のカメラ環境において CNN を用いて背番号を認識することで, 対象の同定を一定の精度で実現できることを確認した. また, 複数カメラ環境で多く発生する誤認識という問題に対し, 移動ベクトルに基づいて背面方向を推定することの有効性を確認できた. 今後の課題としては, 使用するネットワークや追跡手法について検討を行うことで,

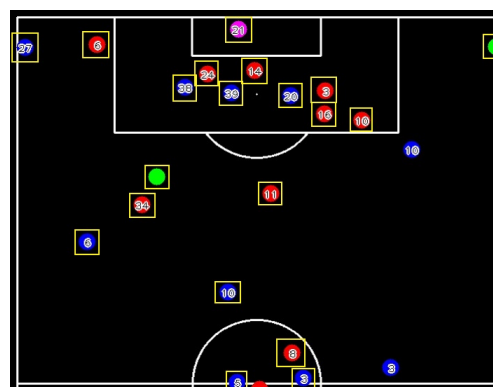


図 10: 平面マップ (350 フレーム目)

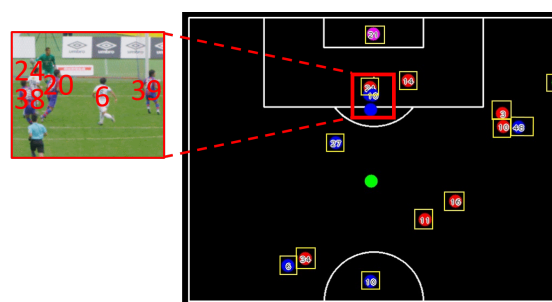


図 11: オクルージョン部分の認識結果 (200 フレーム目)

さらに認識精度を高めることが必要である. 加えて, オクルージョンが発生する領域では誤認識が増加することから, これを解決する新たなアプローチの提案が求められる.

参考文献

- [1] 山崎綾哉, 増田哲朗, 深井寛修, 徐剛, "フィールドスポーツにおける同期多カメラを用いた複数選手のチーム識別と背番号認識による実時間個人識別", 映像情報メディア学会技術報告, vol.36.34, pp.59–62, 2012.
- [2] T. Yamamoto, H. Kataoka, M. Hayashi, Y. Aoki, K. Oshima, and M. Tanabiki, "Multiple players tracking and identification using group detection and player number recognition in Sports Video", 41st Annual Conference of IEEE Industrial Electronics Society, pp.2442–2446, 2013.
- [3] S. Gerke, K. Müller, R. Schäfer, "Soccer jersey number recognition using convolutional neural networks", 2015 IEEE International Conference on Computer Vision Workshop, pp.734–741, 2015.
- [4] A. Laurentini, "The Visual Hull Concept for Silhouette Based Image Understanding", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.16, no.2, pp.150–162, 1994.
- [5] Q. Yao, H. Sankoh, H. Sabirin, and S. Naito, "Accurate silhouette extraction of multiple moving objects for free viewpoint sports video synthesis", Proc. Multimedia Signal Processing, pp.1–6, 2015.
- [6] A. Krizhevsky, I. Sutskever, G. Hinton, "ImageNet classification with deep convolutional neural networks", Advances in Neural Information Processing Systems, vol.25, pp.1106–1114, 2012.
- [7] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting", The Journal of Machine Learning Research, vol.15, pp.1929–1958, 2014.