

## 畳み込みニューラルネットワークによるボールの軌道の予測

日大生産工(院) ○安藤 景星

日大生産工 伊藤 浩

として表現でき、直球は円の半径が十分大きいときの円軌道とみなすことができる。

## 1. はじめに

現代のスポーツはデータが重視され、サッカーや野球等の球技はデータなしに勝つことは難しい。また近年、AI(Artificial Intelligence)は、その学習能力の向上に伴って、スポーツの分野でも様々に利用されている。例えば、データ収集を利用し、AIで勝敗予測、データ解析を行うサービスがある[1]。また、卓球のボールの軌道をエンコーダとデコーダを用いたディープネットワークに推論させた例がある[2]。文献[2]では多くのカメラを使用し、様々な角度からボールを追跡し、莫大なデータを用いて、予測を行っている。しかし、人間や動物は視覚のみから軌道を予測しており、同じことがニューラルネットワークでどこまで実現できるかを知ることは重要である。

本研究では、人間が実際に見ている映像から野球の投手が投げたボールの軌道を予測させる。ボールがホームベースに達したときの位置を審判の視点から推論するようにした。

## 2. ボールの見え方

ここではボールの動きが審判からどのように見えるかを説明する。Fig.1のように投手の位置を原点とし、投手から捕手の方向にx軸をとる。x軸に垂直に一塁方向にy軸をとり、投手から上方にz軸をとる。ボールの速度と回転は変化しないものとする。

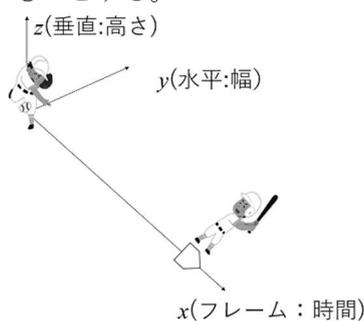


Fig.1 座標の設定

## 2.1 水平の軌道

水平の軌道はxy平面に投影したボールの軌道である。ボールは回転がなければ、直線的に進み、ボールに回転がかかれば、円軌道を描く。よって、水平の軌道は、カーブの場合は一塁側、シュートの場合は三塁側に中心を持つ円軌道

## 2.2 垂直の軌道

垂直の軌道はxz平面に投影したボールの軌道である。垂直方向にはボールにかかる重力が働く。そのため、ボールの軌道は回転による円軌道に放物線の軌道を足し合わせたものである。ストレートの場合は、バックスピがかかるため、直線に近い軌道になり、カーブの場合はトップスピがかかるため、ボールが大きく落ちる。フォークは比較的、回転が少ないため、重力によってボールが落ちる。

## 2.3 画像面への投影

Fig.2は水平方向のボールの位置が画像面にどのように投影されるかを示す図である。審判の目の位置( $x=l$ )にピンホール、そこから焦点距離 $f$ の位置にスクリーンがあり、スクリーンの中心からy軸と逆方向にY軸を取る。水平面上で( $x,y$ )の位置にあるボールのY座標は次式のようになる。

$$Y = y / (l-x) \quad (1)$$

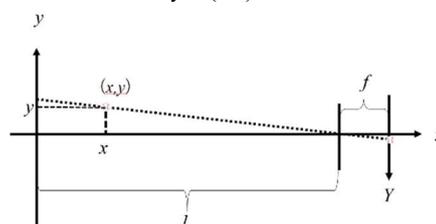


Fig.2 画像面への投影

垂直面についても、同様にZ軸をとって計算できる。以下、スクリーンに写った画像から既にボールの位置が特定されたものとして、( $x,Y$ )と( $x,Z$ )の時間的な変化からボールの軌道を推定するものとする。

## 3. ネットワークの設計

## 3.1 回路構成

本研究で使用するネットワークのモデルをFig.3に示す。入力信号は $32 \times 32$ の軌道データとし、これを畳み込み層とプリーング層で5回繰り返して処理する。各層で解像度は1/2に縮小し、チャンネル数は2倍に増やす。よって、第5層の出力は512チャンネルの一次元データにな

る。この後、全結合層を4回繰り返す。この各層でデータサイズを1/2ずつ縮小していき、最後の全結合層のサイズは32となる。最後に出力層として、再度全結合層を設け、一つの実数値を出力するようにした。この値はボールの予測位置を表す。

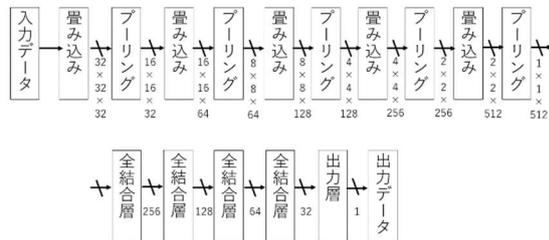


Fig.3 使用したモデルの構成

### 3.2 入力データ

入力データは32×32のグレースケール画像として生成した。この画像は背景を黒、ボールの位置を白で表示したものである。横軸はボールのx座標を表し、投手からホームベースまでを32等分したものである。縦軸はY座標(幅)またはZ座標(高さ)である。

縦軸の1画素はホームベース上で水平方向は約2.4cm、垂直方向は約5.5cmに相当する。ストライクゾーンの大きさは、水平方向は17画素、垂直方向は11画素である。Fig.4は水平方向の入力データの例であり、ボールがx=1/4, 1/2, 3/4, 1まで進んだ時の軌道を時系列に並べたものである。

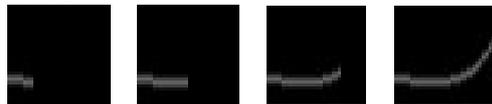


Fig.4 水平方向の入力データ

## 4. 評価

### 4.1 学習

入力データは訓練用として64000セット、検証用として16000セットをランダムに生成した。この中には、軌道が途中までのものを含む。正解データは、軌道が全て見えている入力画像の右端のボールの座標値をそのまま用いる。最適化アルゴリズムはAdamを使用し、Batch sizeを200、損失関数をMean Squared Errorに設定した。検証用データの損失が水平方向で25.2以下、垂直方向で1.7以下になれば、学習をやめ、そのときのモデルを保存した。学習回数とともに損失は減少し、水平方向の場合、180回目で損失は25.18になり、垂直方向の場合、106回目で1.664になった。このとき訓練用データの損失は水平、垂直でそれぞれ25.02、1.667であった。

### 4.2 テスト

訓練用と別に、ランダムな軌道を2000セット生成し、テストデータとした。水平、垂直方向の予測誤差をFig.5に示す。横軸はx軸上のボールの位置を示し、この地点までボールが到達したことを示している。1は投手の位置、32はホームベースの位置である。

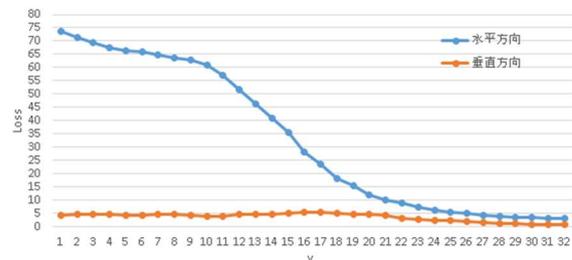


Fig.5 テストデータの予測誤差

水平方向はボールの進行とともに予測精度が向上し、特に中間地点の前後で急激に下がっている。ボールが打者から1/3の地点まで近づければ誤差は約3画素であり、これはホームベース幅の約20%に相当する。通常、人間がバットを振り始めるのはこの後であるため、この精度は人間がボールの位置を見極める精度とおよそ一致している。垂直方向については、ほぼすべての地点で予測精度が高い結果になった。これは垂直方向の視点の位置を高く設定してしまったためであり、ボールの最終的な位置の範囲が狭く、正解が推定しやすかったことによる。

## 5. まとめ

審判から見た画像を入力として、投手が投げたボールの軌道をCNNに予測させた。水平方向については、人間の感覚に合致する精度が得られた。

垂直方向は視点の位置を検討し直し、再評価する必要がある。また、本研究では既にボールの位置が特定されたものとしていたが、今後は画像からボールの検出を行った上で、予測を行わせる必要がある。本研究の成果は、ARグラスなどを使用した審判の補助装置などとして応用が期待できる。

### 参考文献

- [1] スパイア, <https://spaia.jp/> (参照 2023-10-01)
- [2] S.G. Gonzalez et.al. "Real time trajectory prediction using deep conditional generative models," arXiv:1909.03895, 2020.