

深層学習を用いた投球動作の予測手法の検討 Study of Prediction Methods for Pitching Motion Using Deep Learning

小林 将也[†] 田村 仁[‡]
Masaya Kobayashi Hitoshi Tamura

1.はじめに

近年、対人ロボットがたくさん開発されている。対人ロボットはカメラで取得した基本的にカメラ等で取得した現在の人の状況などのデータからロボットの制御を行っている、何かしらのアクシデントが発生した際には、その都度制御を行い対応する。しかしこの方法ではカメラから取得した情報を元にリアルタイムでロボットの制御を行うため動作が間に合わない恐れがある。カメラから取得した人の情報から動作の予測をし、その予測した情報からロボットを制御すれば、時間的に余裕を持った動作の制御が行える。

人の動作予測には、姿勢推定が必要である。このための手法の一つに `openpose`[1]が存在する。これは深層学習によって単眼カメラで人の姿勢推定が行えるものである。姿勢推定によって得られた関節座標の時系列データ、つまりモーションデータを用いて動作予測する。

深層学習によって人の動作を予測するには膨大な量のモーションデータが必要である、本研究では野球のスポーツ中継の映像を姿勢推定する元データとした。

以上のことから本研究の目的は野球の投手の投球の動作からその球種、コースを予測することである。

2.関連研究

深層学習を用いて未来を予測する研究には様々なものがある。対象物自体の情報を用いて対象物自身の動作を予測する、例としてビリヤードのボールの軌道から未来のボールの軌道を予測する研究がある[2]。移動体やロボットの制御に用いようとしている研究もあり、例として車載カメラの映像から危険度予測する研究がある[3]。人間の動きを測定し 0.5 秒後の人間の動きを推定する研究がある[4]。また本研究と同様に身体的特徴からボールの軌道を予測研究があり、例としてバレーのトスの軌道を予測する研究がある[5]。これらの研究は物体などの動作の軌道自体を予測しているもので本研究では動作を予測し、その後の結果を予測することが目的である。本研究では投球動作からコースと球種の予測を行う。

3.提案手法

目的を達成するために以下のように提案する。

- (1)野球のスポーツ中継の映像から投手が球をリリースする瞬間から 30 フレーム遡った画像をデータとする
- (2)各フレームを `openpose` で全身の関節位置を読み出す
- (3)関節の時系列データとスコアデータから球種とコースを深層学習の LSTM モデルで学習させる
- (4)未知の投球データからコースと球種を予測

LSTM とは再帰型ニューラルネットワークアーキテク

[†] 日本工業大学 機械システム専攻,
Nippon Institute of Technology Department of Mechanical Systems Engineering

[‡] 日本工業大学 先進工学部 ロボティクス学科
Nippon Institute of Technology Department of Advanced Engineering Faculty of Robotics

チャの一つであり、時系列データに強く学習できるモデルである。

また `openpose` を使用して投球の関節座標を解析した画像を図 1 に示す。



図 1 `openpose` で解析した投球画像

図 1 のように `openpose` を使用して投手の関節座標が取れる。

4.予備実験方法

この方法によって投球を予測できるのかどうか少数のデータから検証を行った、実験条件は以下の通りである。

- ①投手は読売巨人軍の菅野投手
- ②2020年9月15日の巨人対阪神の試合
- ③計 630 枚のデータから学習させる
- ④投球データはスポーツ報知プロ野球一球速報 (https://hochi.news/npb/main/?bw_pageId=oneball&gameId=2020091501)から取得する
- ⑤投球の右肩,右肘,右手首の座標データを用いて学習させる
- ⑥球種のみで学習を行い,ストレートを 0, 変化球を 1 とする
- ⑦ Sony が提供している `neural network console` の 'LSTM' モデルで学習させる

学習に使用している関節座標は `csv` ファイルにまとめている。またその関節座標のデータが入った `csv` ファイルを違う `csv` ファイルからすべてを読み込みするような形となっている。関節座標のデータが入っている `csv` ファイルを図 2 に、読み込んだ `csv` ファイルを図 3 に示す。

	A	B	C	D	E	F	G
1	0.362222	0.162222	0.335556	0.16	0.388889	0.228889	
2	0.364444	0.164444	0.337778	0.164444	0.391111	0.24	
3	0.366667	0.166667	0.337778	0.164444	0.391111	0.24	
4	0.368889	0.164444	0.344444	0.162222	0.393333	0.24	
5	0.368889	0.164444	0.344444	0.164444	0.393333	0.242222	
6	0.373333	0.166667	0.351111	0.166667	0.388889	0.248889	
7	0.375556	0.173333	0.348889	0.166667	0.402222	0.251111	
8	0.38	0.175556	0.353333	0.168889	0.397778	0.262222	
9	0.382222	0.177778	0.353333	0.168889	0.391111	0.264444	
10	0.384444	0.177778	0.355556	0.168889	0.391111	0.266667	
11	0.384444	0.18	0.357778	0.175556	0.395556	0.275556	
12	0.386667	0.184444	0.362222	0.18	0.397778	0.28	
13	0.388889	0.184444	0.366667	0.182222	0.393333	0.28	
14	0.391111	0.188889	0.368889	0.184444	0.391111	0.282222	
15	0.4	0.208889	0.38	0.206667	0.388889	0.313333	

図 2 関節座標が入った `csv` ファイルの一部

Copyright © 2021 by
The Institute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.

	A	B	C	D	E
1	x	y			
2	./csv/0.csv	0			
3	./csv/1.csv	1			
4	./csv/2.csv	1			
5	./csv/3.csv	0			
6	./csv/4.csv	1			
7	./csv/5.csv	1			
8	./csv/6.csv	0			
9	./csv/7.csv	0			
10	./csv/8.csv	0			
11	./csv/9.csv	1			
12	./csv/10.csv	0			
13	./csv/11.csv	1			
14	./csv/12.csv	1			
15	./csv/13.csv	1			

図 3 読み込みしている csv ファイルの一部

図 3 の A 列にて図 2 の csv ファイルを読み込んでいる。また B 列でその投球が変化球なのかストレートなのかを分けている。

5. 予備実験結果

実験から得た学習曲線を図 4 に示す。

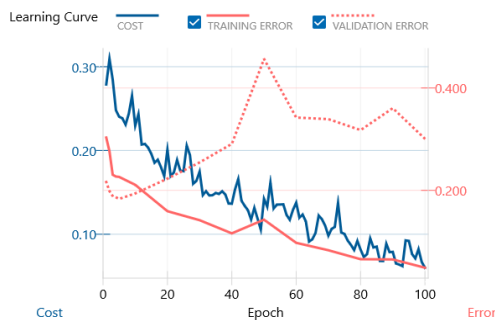


図 4 予備実験から得た学習曲線

図 4 の点線から約 8 エポックから過学習に陥ることがわかる。また過学習に陥る前に約 85 パーセントの判別率があることがわかる。

6 本実験

6.1 データを増やした学習

予備実験の結果を踏まえて今度はより多い投球データを用いて学習を行った。投手や試合日、投球データ、関節座標などの条件は同様で今回は計 3000 枚のデータで行った。実験から得た学習曲線を図 5 に示す。

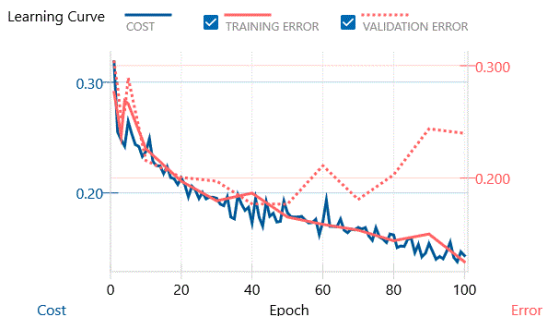


図 5 データを増やし学習させ得た学習曲線

図 5 の点線より 40 エポックから過学習に陥ることがわかる。また過学習に陥る前に約 80 パーセントの判別率があることが分かる。

6.2 学習反復世代数を半分にした学習

図 4 や図 5 では学習反復世代 100 で学習を行った、なので図 5 と同様の条件下で学習反復世代を半分の 50 にして学習を行った。この実験で得た学習曲線を図 6 に示す。

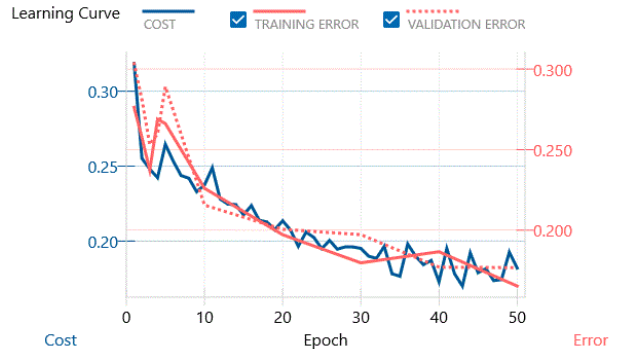


図 6 学習反復世代数 50 での実験から得た学習曲線

図 6 の点線より図 1, 図 2 と違い過学習に陥らず学習が行えていることが分かる。また判別率が約 85 パーセントあることが分かる。

6.3 球種をストレートに限定した学習

球種をストレートのみにしたデータで学習を行った、また他条件は予備実験と同じである。この実験で得た学習曲線を図 7 に示す。

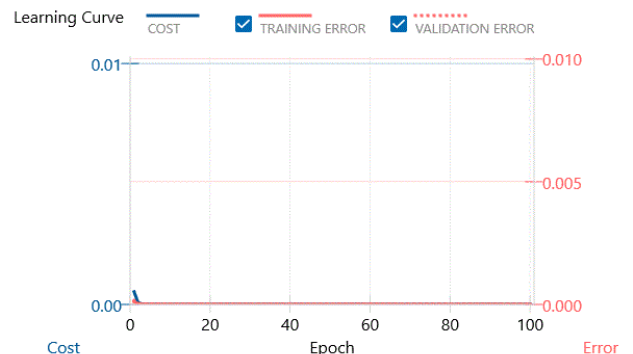


図 7 ストレートのみで得た学習曲線

図 7 より学習がしっかりと行われていないことが分かる。

6.4 球種を変化球に限定した学習

球種を変化球のみにしたデータで学習を行った、また他条件は予備実験と同じである。この実験で得た学習曲線を図 8 に示す。

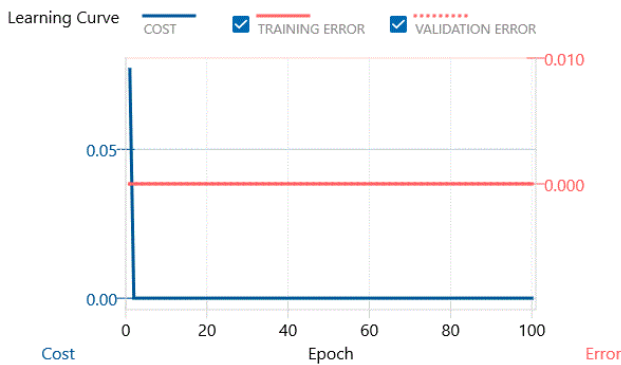


図 8 変化球のみのデータで得た学習曲線

図 8 より学習がしっかりと行われていないことが分かる。

6.5 試合前半と後半に分けて学習

データを試合前半と後半に分けてそれぞれ学習させた。他条件は予備実験と同様である。試合前半のデータを学習させた学習曲線を図 9、試合後半のデータを学習させた学習曲線を図 10 に示す。

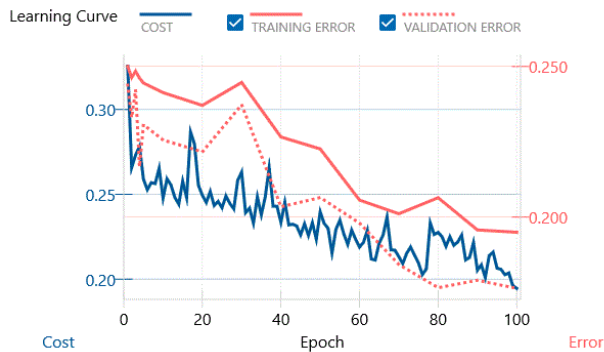


図 9 試合前半のデータで得た学習曲線

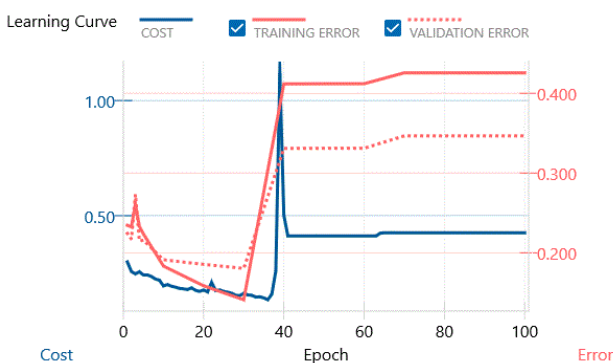


図 10 試合後半のデータで得た学習曲線

図 9 より過学習に陥らずに学習ができていることが分かる。また判別率は約 82 パーセントあることが分かる。

図 10 より 30 エポックから過学習に陥ってしまっていることが分かる。また判別率が約 75 パーセントあることが分かる。

7 考察

図 4, 5, 6 より球種が約 80 パーセント程度の判別率で予測できることが分かった。また図 1, 図 2 より学習データを増やした結果、過学習に陥るまでの時間が倍以上伸びることが分かった、しかしまだ過学習に陥ってしまっている。これは図 5 で用いたデータ数がまだまだ少ないことが原因と考えられる。また図 4, 5 を比べると判別率が 5 パーセント程度低くなってしまっていることが分かる。この原因としてはデータを増えたかつ過学習に陥ってしまったことが原因と考えられる。

図 6 より学習反復世代数を半分の 50 にしたことによって過学習に陥らずに学習が終わっていることが分かった。このことにより過学習に陥らずに学習ができる必要最低データ数は図 2, 3 での学習に用いた 3000 枚の二倍の 6000 枚程度必要なものと考えられる。

図 7 より学習が行えてないことが分かった。原因としてデータ数が少ないこと、何かしらのエラーが発生してしまったことが考えられる。また図 8 も図 7 と同様に学習が行えてないことが分かった。原因としては図 7 と同様にデータ数の少なさ、何かしらのエラーだと考えられる。

図 7, 8 より学習曲線の形に違いがあることが分かる。そのためストレートと変化球混ぜた状態よりもストレートと変化球分けて学習させた方がより良い結果が出る可能性があると考えられる。また学習結果の違いからストレートと変化球どちらかに絞って学習させた方がより良い判別率が出ると思われる。

図 9 より試合前半のデータを使用して学習を行った場合、試合全体のデータを使用した時と判別率も約 82 パーセントとさほど下がらずに学習が行えることが分かった。これは試合前半では投手自身が疲れておらず、関節座標のデータにばらつきが少ないためと考えられる。

図 10 より試合後半のデータを使用して学習を行った場合、今まで行った学習で得た学習曲線と比べて全く異なる形を示した。また判別率もどの学習曲線よりも低い約 75 パーセントを示した。これは試合後半ということで選手自身が疲れて関節座標のデータにばらつきが生じたためと考えられる。また図 9, 10 より試合前半の投球データを使用して学習した方が判別率も高く学習できることが分かった。これは上記のように選手自身がまだ疲れていないためと考えられる。また前半のデータのみを使用すれば判別率が高くなるのではないかと考えられる。

様々な条件で学習した結果判別率はほとんどの条件で約 80 パーセント前後であった。

9.終わりに

本研究では投手の投球モーションから球種とコースを予測することを目的とした。またその方法として openpose を使用して関節座標を求め、そのデータを csv ファイルにまとめ、その csv ファイルを読み込み実験では変化球かストレートかを分けて機械学習の LSTM モデルで学習させることで予測した。

実験を行い下記のようなことが分かった。

- (1) データ数に限らずに約 80 パーセント程度の判別率が出る。
- (2) データを増やし過学習が起こった際に判別率が下がる
- (3) 最低でも 6000 枚のデータがなければ過学習に陥ってしまう

- (4) ストレートと変化球のみにして学習を行うと異なる学習結果が出るためどちらかに絞れば結果が良くなる可能性がある
- (5) 試合前半と後半に分けてそれぞれ学習した際に試合前半よりも後半の方が良くない結果が出た
- (6) 試合前半の方が結果が良かったためデータを前半に絞って学習させればより良い結果が出るのではないか
- (7) 試合後半に絞った際に学習曲線が今までとは異なる形を示した

今後はもっとデータを増やし、様々な条件で学習しどのような結果が出るのかを確認、また判別率の変化がほとんどなかったためその向上を目指していく。

参考文献

- [1] Zhe.Cao, Tomas.Simon, Shin-En. Wei, Yaser. Shekh
“Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields”,
The Robotics Institute, Carnegie Mellon University.
arXiv:1611.08050v2[cs.CV]14Apr2017
- [2] Katerina.Fragkiadaki,Pulkit.Agrawal,Sergey.Levine,Jitendra.Malik,
” Learning Visual Predictive Models of Physics for Playing Billiards”
In Proc ICLR2016,2016.arXiv:1511.07404
- [3] 松田賢太,村田祐樹,後藤紳一郎,渥美雅保,
“領域ベース深層学習を用いた車載カメラ映像からの危険予測”,
第 80 回全国大会講演論文集 vol.80(1),pp507-508,2018
- [4] Yuuki.Horiuchi,Yasutoshi.Makino,Hiroyuki.Shinoda.Authors Info & Affiliations
“Computational Foresight: Forecasting Human Body Motion in Real-time for Reducing Delays in Interactive System”
ISS '17: Proceedings of the 2017 ACM International Conference on Interactive Surfaces and SpacesOctober 2017 Pages 312–317
- [5]Shuya.suda,Yasutoshi.Makino,Hiroyuki.Shinoda,
” Prediction of Volleyball Trajectory Using Skeletal Motions of Setter Player”
AH2019: Proceedings of the 10th Augmented Human International Conference 2019March 2019 Article No.: 16Pages 1–8