

将棋盤面の難易度評価のための 畳み込みニューラルネットワークによる人間の思考時間分類

池田英寿^{1,a)} 一ノ瀬元喜^{1,b)}

概要: 近年の将棋 AI の発展は目覚ましく、初学者への学習支援ツールとして活用させることを目的とした研究も存在する。しかし、学習支援を行う上で重要な「盤面ごとに人間が感じる難しさの違い」を定量化している研究は少ない。本研究では、人間の思考時間を盤面の難易度の指標として用いて、駒の配置と持ち駒の情報を入力、人間の思考時間を出力とした畳み込みニューラルネットワークによって、盤面の情報から人間の思考時間を予測した。その結果、合法手数を用いたランダムフォレストによる予測と比べて、畳み込みニューラルネットワークでは、駒のぶつかりといった特徴を捉えることで、僅かな違いの盤面でも正しい方向へ予測値を変えていることが確認でき、盤面の難解さの定量化に有効なモデルであることが分かった。

Classification of human thinking time using convolutional neural network for evaluation of Shogi-board difficulty

HIDETOSHI IKEDA^{1,a)} GENKI ICHINOSE^{1,b)}

Abstract: Shogi AIs have made remarkable progress in recent years, and some studies aim to use them as an educational support tool for beginners. However, few studies have quantified the “difference in difficulty perceived by humans to each Shogi board”, which is important for educational support. In this study, we regarded human thinking time as the index of board difficulty and predicted human thinking time by convolutional neural networks where piece positions and pieces in hand were used as input to the neural networks. As a result, compared to the prediction by random forest using the number of legal moves, the convolutional neural networks well predicted human thinking time by getting characteristics such as the state that pieces being captured, even for a slightly different situation. Thus, we find that this model effectively quantifies the boards’ difficulty.

1. はじめに

将棋 AI は、チェス AI で使用されたアルゴリズムを応用し、加えて機械学習や強化学習を取り入れたことで、2010年代にトッププレイヤーを何度も打ち破り世間を騒がせた。近年では、将棋 AI を強くすることを目的とした研究だけでなく、将棋 AI を初学者への学習支援ツールとして活用

することを目的とした研究も取り組まれている。先行研究では、どうぶつ将棋において逆転の余地を考慮した AI を作成し、人間を有利に導く手を提案する研究 [1] や、初学者の着手を参考に詰将棋の解説文を生成する研究が行われている [2]。

このように、ボードゲームをプレイする際の人間の思考をくみ取り、手助けを行う学習支援ツールの実現を目指している研究が存在するが、将棋において手助けを行う上で重要な指標である「人間が感じる将棋盤面の難解さ」の完全な定量化は成功していない。この指標の定量化が成功すれば、次の一手問題の難易度を自動で点数付けすることや、定跡の各盤面の難解さを調べることで戦法ごとの難易度を

¹ 静岡大学大学院総合科学技術研究科工学専攻数理システム工学コース
Department of Mathematical and Systems Engineering,
Shizuoka University, Hamamatsu 432-8561, Japan

a) ikeda.hidetoshi.18@shizuoka.ac.jp

b) ichinose.genki@shizuoka.ac.jp

測ることが可能となる。

将棋盤面の難解さの特徴量としては、例えば「人間の次の一手と AI の最善手の相違」など、さまざまな指標が考えられるが、本研究では「人間がその盤面で考えていた時間の長さ」が簡潔かつ有効な指標であると考えた。そこで本研究では、畳み込みニューラルネットワークを用いて将棋盤面から人間の思考時間を予測した結果、難易度の定量化のための有効なモデルが得られた。

2. 思考時間予測の方針

思考時間の予測に役立つであろう特徴量としては、盤面の合法手の数や駒の効きの数などが挙げられる。実際に先行研究では、ターン制戦略ゲームの局面難易度を合法手の数で定義している [3]。しかし、予測に役立つ良い特徴を人間の手で見つけていくことは労力のいる作業であるため、コンピュータで自動的に特徴を生成させる方法を考える。

先行研究では、畳み込みニューラルネットワーク（以下 CNN）を用いることで、任意のチェスの盤面に対して、駒の配置から人間の次の一手を良い精度で予測している [4]。言い換えると、駒の配置という複雑な情報から良い特徴量を取り出して、人間の次の一手を出力する近似関数を構築している。人間の次の一手の予測が可能ということは、CNN は盤面に対し人間がどのように思考をしているかを推論できているということである。CNN の出力対象を変えることによって、「次の一手予測」だけではなく、CNN を用いて駒の配置から人間の思考時間の特徴を抽出し、思考時間を予測することも可能だと考えられる。

将棋はチェスと異なり、盤面情報として駒の配置の他に持ち駒が必要のため、本研究では、駒の配置に加えて持ち駒の情報も入力とした CNN によって思考時間の予測を行った。

3. データの収集

本章では、CNN で用いるデータの詳細および収集方法について述べる。今回、人間の思考時間というデータが必要であるため、実際に行われた対局データを対象にすることが適切であると考えられる。対局データは、インターネット上で気軽に他人と対戦できる「将棋ウォーズ」というアプリケーションから収集した [5]。データを収集する際、四段以上同士が 10 分切れ負けというルールで対局したものを対象とし、Apery などのボットも取り除いた。四段以上のみとした理由としては、将棋に長い間触れているプレイヤーはある程度思考が似ているはずであり、外れ値が少なくなると考えたためである。

また、将棋における序盤は同一の盤面になってしまうことが多く、終盤は持ち時間が少ないため早指しすることが多い。したがって序盤と終盤は学習データとしては不適切と考えられるため、30 手目～80 手目の盤面のみを学習デー

タとした。これらの処理により、データ数は計 316,957 盤面となった。

4. 思考時間予測の手順

本章では、CNN の構造および学習と予測の手順を述べる。将棋の盤面情報のデータ加工や CNN の構造は、山岡の書籍を参考にした [6]。CNN に入力する盤面情報は、盤面上の駒の配置、先手および後手の持ち駒の種類と数、の 2 つの特徴から成る。盤面上の駒の情報は、盤面の升の数 $(9 \times 9) \times$ 駒・成り駒の種類 $(14) \times$ 先手後手 $(2) = 2,268$ 個で表現され、両者の持ち駒の情報は、盤面の升の数 \times 駒の数 $(38) \times$ 先手後手 $= 6,156$ 個で表現される。

特徴抽出を行う中間層には、残差ブロックを 5 個接続したネットワークを用いる。残差ブロック内では活性化関数に ReLU を使用した畳み込みを 2 回行い、活性化関数の前にバッチ正規化を行う。

また、人間の思考時間を学習するタスクだけでなく、その盤面で人間が指した手と、その対局の勝敗結果も同時に学習するマルチタスク学習を行う。思考時間は整数値ではなく、「0～2 秒（早指し）」、「3～9 秒（少し考慮）」、「10 秒～（長考）」の 3 カテゴリーを確率で出力するように設定した。閾値をこのようにした理由として、3 つのカテゴリーの分布が可能な限り平坦になるよう設定したためである。勝敗結果は、勝率として 0～1 の実数値として出力する。指し手は、駒の移動の種類 $(27) \times$ 移動先の座標 (9×9) で表現でき、計 2187 個の分類を確率で出力する。

以下に CNN の詳細を示す。

入力層 駒配置の情報 + 両者の持ち駒の情報 $(9 \times 9$ の 2 値画像 $\times 104$ チャンネル)

中間層 畳み込みとバッチ正規化が 2 回行われる残差ブロックを 5 つ接続する（畳み込み層：フィルターサイズ 3×3 、フィルター数 192、パディング 1、ストライド 1、活性化関数 LeRU）。

出力層

- 指し手：フィルターサイズ 1×1 、フィルター数 27 の畳み込みを行い、ソフトマックス関数で各指し手の確率を出力する。
- 勝敗結果：指し手と同様の畳み込みを行い、その後 256 ノードの全結合層を接続し、最後に 1 ノードの全結合層に接続する。畳み込み層と 1 つ目の全結合層の活性化関数は LeRU を用い、最後の全結合層はシグモイド関数で盤面の勝率を出力する。
- 思考時間：勝敗結果とほぼ同様のネットワークであり、最後の全結合層を 3 ノードに変更したものをを用いる。最後の全結合層はソフトマックス関数で 3 カテゴリー「0～2 秒（早指し）」、「3～9 秒（少し考慮）」、「10 秒～（長考）」の思考時間の確率を出力する。

以上のような構造を持つ CNN を用いて教師あり機械学習を行う。学習と予測の手順としては、まず収集したデータを 9:1 で学習データとテストデータに分割し、学習データで 9 分割の交差検証を行う。そして、交差検証によって得られた 9 つのモデルを用いて、テストデータに対して多数決による予測の精度を算出する。さらに、予測を行った盤面の中で、「早指しする可能性が高い盤面」および「長考する可能性が高い盤面」を抽出し、モデルの予測の特徴を考察する。

また、CNN との予測精度の比較として、盤面の合法手数を説明変数、人間の思考時間を目的変数としたランダムフォレストによる予測も行う。予測の手順としては CNN と同様、学習データで交差検証を行ったのち、テストデータに対して予測を行う。

5. 結果

はじめに、テストデータにおける 3 カテゴリーの実際の分布を図 1 に示す。

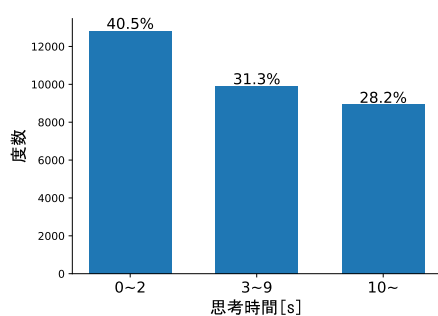


図 1 テストデータにおける思考時間の実際の分布。

思考時間の割合は、0~2 秒（早指し）が 40.5%、3~9 秒（少し考慮）が 31.3%、10 秒～（長考）が 28.2% となり、30 手目までのデータを取り除いたとは言え、早指しした盤面が多いことがわかる。この実際の思考時間のカテゴリーデータに対する予測結果を次節から示す。

5.1 合法手数を用いたランダムフォレストによる思考時間予測

始めに CNN と比較する「合法手数を用いたランダムフォレストによる思考時間の予測結果」を示す。まず、思考時間のカテゴリーごとの合法手数がどのような分布であるかを確認する。

図 2 に思考時間の各カテゴリーに対する合法手数のバイオリンプロットを示す。各カテゴリーで分布の山が 2 つ現れている、いわゆる二峰性となっている原因としては、歩以外の持ち駒を持っているか否かで分かれていると考えられる。歩は 2 歩というルールがあるため持ち駒にあったとしても合法手数は増えにくい、歩以外の持ち駒はほぼ自由に置けるため合法手数が急激に増加するとみられる。

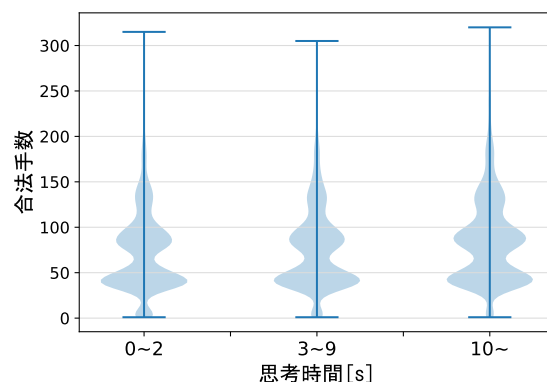


図 2 思考時間の各カテゴリーに対する合法手数のバイオリンプロット。薄い青色の箇所は人口ピラミッドのように縦軸に対する分布を表す。

次にランダムフォレストによって合法手数から人間の思考時間を学習し、予測を行った結果を示す。

表 1 合法手数による思考時間予測結果の混同行列。

	予測値			正解率
	0~2 秒	3~9 秒	10 秒~	
真値 0~2 秒	12,166	0	663	94.9%
真値 3~9 秒	9,266	0	641	0%
真値 10 秒~	8,241	0	684	7.7%

表 1 はテストデータに対する予測結果の混同行列であり、全体の正解率は 40.6% となった。このモデルでは、3~9 秒の予測を全く行わず、ほとんどの盤面に対し 0~2 秒と予測していた。このような予測結果となった理由は、図 2 の 0~2 秒の分布と 3~9 秒の分布が非常に似通っており、かつ 0~2 秒の真値の割合が多いことから、3~9 秒と予測するのであれば同じ分布で割合が多い 0~2 秒と予測した方が精度が向上するためであると考えられる。

5.2 駒配置と持ち駒情報を用いた CNN による思考時間予測

本節では、CNN によって駒の配置と両者の持ち駒という情報から人間の思考時間を学習し、予測を行った結果を示す。

表 2 駒配置と持ち駒情報による思考時間予測結果の混同行列。

	予測値			正解率
	0~2 秒	3~9 秒	10 秒~	
真値 0~2 秒	9,267	626	2,936	72.2%
真値 3~9 秒	5,678	848	3,381	8.6%
真値 10 秒~	3,972	713	4,240	47.5%

表 2 は予測結果の混同行列であり、全体の正解率は 45.3% となった。このモデルにおいても 0~2 秒の予測が多く、3~9 秒の予測が少ないという傾向が見られた。

以下では、テストデータの中で特に予測精度が高かった

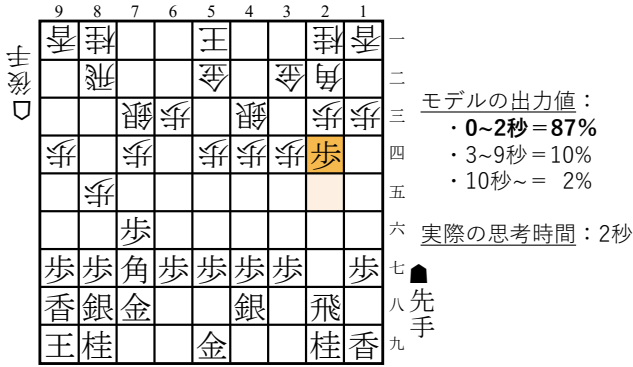


図3 テストデータの中で早指しの可能性が最も高いと予測した盤面。

盤面を示す。

図3は、テストデータの中で、0~2秒で指す可能性が最も高いと予測した盤面(31手目)である。実際、この盤面は歩を取る手以外を考へても意味がないため、早指しの可能性が高いと予測することは妥当であると言える。

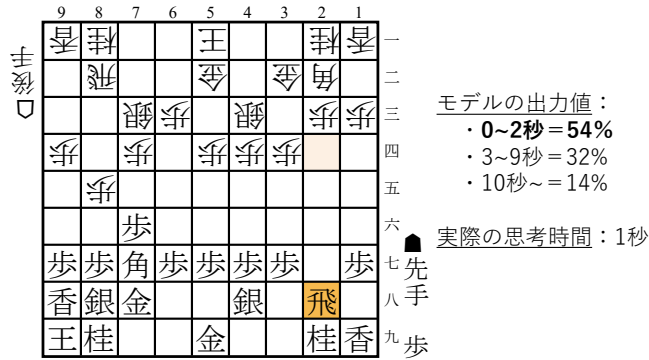


図4 図3から△同歩▲同飛△2三歩▲2八飛まで進んだ盤面。モデルは早指しする可能性が比較的高いと予測している。

図4は、図3から△同歩▲同飛△2三歩▲2八飛まで進み、飛車先の歩交換の流れを終えた盤面である。図3におけるモデルの出力値と比較すると「0~2秒:-33%, 3~9秒:+22%, 10秒~:+12%」に変化し、早指しの可能性が低くなると予測していた。図3は△同歩以外の手を指すと敗勢になる盤面であり、図4はさまざまな候補手が考えられるため、2四の歩が先手の持ち駒に変わったただけではあるが予測結果は良い方向へ変化していた。

次に、早指しの可能性が二番目、三番目に高いと予測した盤面を示す。

図5、図6はそれぞれ、早指しの可能性が二番目、三番目に高いと予測した盤面である。図3と共通して言えることとして、いずれも序盤における飛車先の歩交換の場面であり、指す手が1つに限定されるような盤面であることがわかった。序盤において「2三歩と2四歩」、「8六歩と8七歩」がぶつかる盤面は頻出であり、そのような盤面は思考時間が短い場合が多いと考えられるため、予測精度が高かったと考察できる。

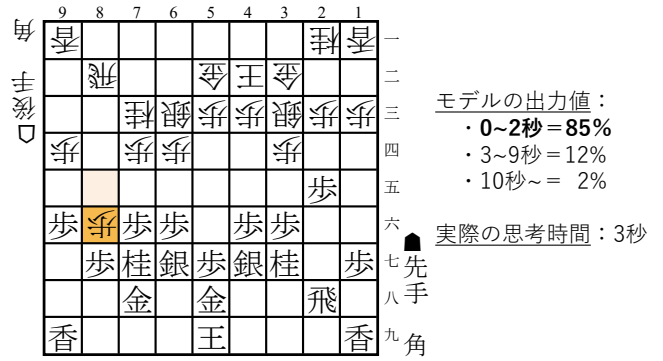


図5 テストデータの中で早指しの可能性が二番目に高いと予測した盤面(32手目)。

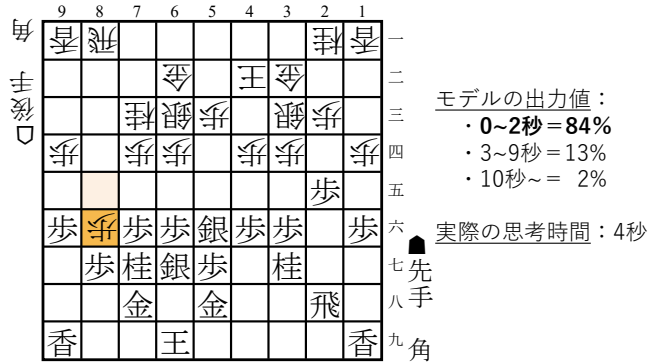


図6 テストデータの中で早指しの可能性が三番目に高いと予測した盤面(38手目)。

次に、長考の可能性が最も高いと予測した盤面を示す。

図7は、テストデータの中で、10~秒で指す可能性が最も

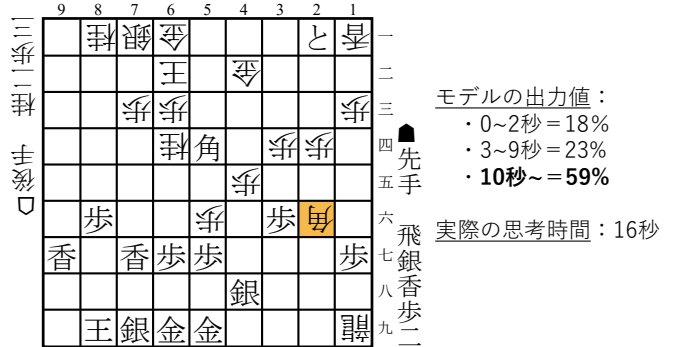


図7 後手が2六に攻防の角を打った盤面。テストデータの中で最も長考の可能性が最も高いと予測した。

高いと予測した盤面(78手目)である。先手番は大きく分けて攻め合いと受けの二通りの展開が考えられるので長考しても違和感がない盤面と言える。

ここでCNNによる全体の予測精度を表したグラフを示す。図8は「現在何手目の盤面なのかという手数」に対するCNNの予測精度を表したグラフである。序盤は精度が高く、終盤は精度が低い傾向にあり、30手目~34手目の精度の平均は49.5%、76手目~80手目の精度の平均は44.6%

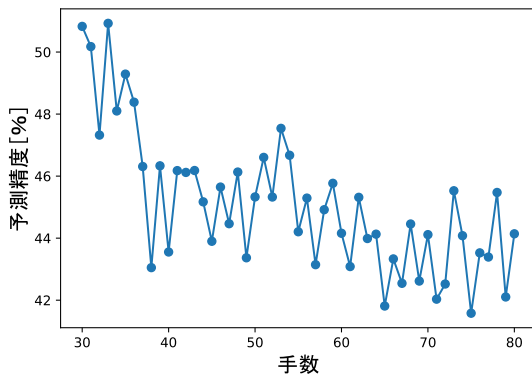


図 8 手数ごとの予測精度を表したグラフ。序盤は精度が高く、手数が増えていくと精度が低くなっている。

となった。序盤に関しては似た盤面になることが多いため、学習効率が良くなり予測精度が高くなると考察できる。さらに、似た盤面が多いということは人間にとって経験が多い盤面でもあるため、思考時間の外れ値が減ることにつながり、相乗効果で学習効率が上がりやすくなる。一方で終盤は似た盤面になることが非常に珍しいため、予測精度が低くなったと考察する。予測精度が低くなる終盤の場面で実際に CNN でうまく予測できなかった例を最後に示す。



図 9 ある対局の 77 手目の盤面。一見簡単な盤面に見えるがモデルの出力は平坦な確率分布を示した。

図 9 は、後手が 54 の桂馬を飛車で取った盤面である。手の流れで言えば▲5 四角と早指しする可能性が高い盤面のように見えるが、モデルの予測値はほぼ平坦な確率分布となっており、似た盤面が少ない終盤は学習が困難であったことがわかった。

6. おわりに

本研究では、畳み込みニューラルネットワークを用いて、駒の配置や持ち駒といった将棋盤面の情報のみから、その盤面で人間が考える時間の長さを予測し、思考時間の分類を行った。その結果、合法手数を用いたランダムフォレストによる予測と比べ、CNN による予測は、盤面の僅かな違いでも特徴を捉え、正しい方向へ予測値を変えていることが確認でき、盤面の難解さの定量化に有効なモデルであ

ることがわかった。高い確率で早指しすると予測していた盤面の共通の特徴としては、指す手が 1 つに限定されるような簡単な盤面であることがわかった。また、一例ではあるが、長考の可能性が高い盤面の特徴として、攻め合いか受けかといった複数の手が考えられる難解な盤面であることがわかった。

また、CNN による予測は、序盤における頻出する盤面に対しては予測精度が比較的高くなる傾向が見られたが、終盤における複雑な盤面に対しては簡単な盤面に見えても予測が低下する傾向が確認された。

今回、次の一手問題のようなものではなく、棋譜という連続的にイベントが起こったデータを扱ったが、現状では手の流れや残り時間の情報は一切考慮していないため、今後の方針としては扱うデータや説明変数の見直しを行い、人間の思考時間に対する予測精度の向上を目指す。

参考文献

- [1] 中屋敷太一, 金子知適, 逆転の余地を考慮した評価関数の設計とどうぶつしょうぎによる評価, The 25th Game Programming Workshop 2020, 22-29, 2020.
- [2] 石脇滉己, 小川直希, 荒川達也, 将棋初心者の着手を予測するための評価関数の検討, The 20th Game Programming Workshop 2015, 100-106, 2015.
- [3] 提橋凜, 武藤孝輔, 西野順二, ターン制戦略ゲームにおける合法手数応答型戦略切替の効果, 研究報告ゲーム情報学 (GI), 2017-GI-37, 1-4, 2017.
- [4] Reid, M-Y., et al., Aligning Superhuman with Human Behavior: Chess as a Model System, *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 1677-1687, 2020.
- [5] 将棋ウォーズ棋譜検索, 2022/010/03 閲覧. <https://www.shogi-extend.com/swars/search>
- [6] 山岡忠夫, 加納邦彦, 強い将棋ソフトの創りかた, マイナビ出版, 2021.