

# TUBSTAP におけるグラフ表現と画像表現の併用

高橋 光<sup>1,a)</sup> 野口 渉<sup>2</sup> 山本 雅人<sup>3</sup>

**概要：**ターン制戦略ゲームは人気のゲームジャンルであるが、未だ人間よりも強い AI は実現されていない。これは合法手の多さや地形の要素など、幾つかの扱いにくい性質があるためである。本稿ではターン制戦略ゲームの研究用プラットフォームである TUBSTAP を題材として、地形の要素を考慮できるグラフ表現を提案し、地形を含むマップで盤面評価を行う Graph Convolutional Networks(GCN) の学習を行った。さらに画像表現とグラフ表現を併用した統合モデルを複数提案し学習を行った。その結果、提案したグラフ表現を用いた GCN は、高精度な盤面評価が可能であることを確認した。提案した統合モデルはいずれも過学習の傾向があり GCN の性能に劣る結果となったが、複数の統合方法で過学習の大きさの差異を示すことができた。

**キーワード：**ターン制戦略ゲーム, グラフ畳み込みネットワーク, TUBSTAP

## Combined Graph Representation and Image Representation in TUBSTAP

HIKARU TAKAHASHI<sup>1,a)</sup> WATARU NOGUCHI<sup>2</sup> MASAHIITO YAMAMOTO<sup>3</sup>

**Abstract:** While turn-based strategy (TBS) games are a popular game genre, AI that is stronger than humans has yet to be realized. This is due to some unwieldy characteristics of TBS games, such as the large number of legal moves and terrain elements. In this paper, using TUBSTAP, a research platform for TBS games, we propose a graph representation that can account for terrain elements and train Graph Convolutional Networks (GCNs) to evaluate the board on a map including terrain. In addition, we propose several integration models that combine image and graph representations and train them. As a result, we confirmed that the proposed GCNs with graph representations are capable of highly accurate board evaluation. Although all the proposed integration models showed a tendency to overfit and were inferior to the performance of GCN, we show differences in the magnitude of overfitting among the different integration methods.

**Keywords:** Turn-Based Strategy(TBS) Game, Graph Convolutional Networks, TUBSTAP

### 1. 研究背景

近年のターン制ゲームの AI の発展はめざましく、囲碁や将棋においては人間のトッププレイヤーを打ち負かすほどの強さのものが登場している。ターン制戦略ゲームはターン制ゲームの 1 ジャンルであり、数多くのゲームが世界中で遊ばれている。その一方で、ルールの差異がある複数のゲームに人気分散していることなどから、AI の研究はあまり進んではいなかった。

このような背景のもと、ターン制戦略ゲームの研究用プラットフォームの TUBSTAP[1] が提唱され、統一的で再現性のある研究が可能になった。しかしながら、TUBSTAP の登場後もターン制戦略ゲームは幾つかの扱いにくい性質があることから、人間を打ち負かすほどの AI は実現されていないのが現状である。

TUBSTAP はターン制のゲームであり、お互いの駒を動かし取り合うことで勝敗を決するという点では将棋に似たゲームである。その一方で、よりゲームを複雑にする独特なルールも複数もっているため、将棋の手法をそのまま適用することはできない。一つの特徴としては、各プレイヤーの手番ではすべての自駒を任意の順番で動かせる点が挙げられる。そのため駒数の増加に伴って合法手数が爆

<sup>1</sup> 北海道大学大学院情報科学院

<sup>2</sup> 北海道大学 数理・データサイエンス教育研究センター

<sup>3</sup> 北海道大学大学院情報科学研究院

a) morinonusi421@eis.hokudai.ac.jp

発的に増え、ゲーム木の探索が困難になる。また、ゲーム開始時のユニットの配置が決まっていないという特徴や、フィールドも一定ではなく「平原」や「山」などの地形が存在するという特徴がある。このため生じうる可能な盤面が多様性に富み、駒の種類ごとの価値の重さも状況によってまったく異なる。このことから、AIが盤面の価値を適切に評価することが困難になっている。

## 2. TUBSTAP のルール

ここでは、本論文に関係する部分について簡単に説明する。

### 2.1 概要

TUBSTAP は 2 人のプレイヤーが、自分のもつユニットの移動や攻撃を行うことで、相手のもつユニットを先に全滅させることが目的のターン制ゲームである。基本的には相手を全滅させたプレイヤーの勝利となるが、既定のターンまで決着がつかなかった場合は残りユニットの合計 HP の多寡によって勝敗が決められる。

### 2.2 開始

開始時点の盤面は一定ではなく、「マップ」と呼ばれる初期盤面を選んでから試合を開始する。マップに含まれる要素、つまり試合ごとに決定される要素は以下の通りである。

- ユニットの数と種類
- 各ユニットの初期配置
- 各ユニットの初期 HP
- 地形の配置
- 最大ターン数
- 引き分け判定基準（総 HP 量の差の閾値）

#### 2.2.1 進行

各ターンでプレイヤーは、自身のもつユニットをすべて任意の順で行動させることができる。ユニットは「移動」、「移動後に隣接する敵を攻撃」、「移動せず攻撃」という行動が可能である。

#### 2.2.2 ユニット

ユニットは歩兵、戦車など全部で 6 種類が存在する。HP は 1 から 10 の範囲で、0 になるとゲームから取り除かれる。ユニット同士には有利不利の相性が存在し、戦闘のダメージ計算に用いられる。ユニットは移動力と等しいマンハッタン距離を移動することができる。ただし、山や海などの一部の地形では移動が阻まれる。本研究では以下のユニットを使用する。

- 歩兵：戦闘に弱い。移動力は 3。
- 戦車：戦闘に強い。移動力は 6。
- 自走砲：後述する特殊ユニット。移動力は 6。

#### 2.2.3 地形

地形は平原、林など全部で 7 種類が存在する。防御効果が高い地形にいたり、戦闘で受けるダメージを減らすことができる。本研究では以下の地形を使用する。

- 平原：防御効果は低い。移動しやすい。
- 林：防御効果が高い。戦車と自走砲は平原の 2 倍の移動コストがかかる。
- 山：防御効果がとても高い。戦車と自走砲は侵入ができない。歩兵は平地の 2 倍の移動コストがかかる。

#### 2.2.4 戦闘

ユニットは移動後に隣接する敵を攻撃することができる。攻撃でダメージを与えた後に敵の HP が残っていた場合、自動的に反撃が行われる。

ダメージは以下の要素から決まる。

- 攻撃側の HP が多いほど、与えるダメージが増える。
- 防御側の HP が多いほど、与えるダメージが減る。
- ユニット種の相性により、与えるダメージが増減する。
- 防御効果の大きい地形にいる相手には、与えるダメージが減る。

#### 2.2.5 自走砲

自走砲は特殊なユニット種であり、移動や戦闘のルールが他と異なる。唯一遠距離攻撃が可能であり、マンハッタン距離 2 と 3 の相手に対して反撃を受けずに攻撃することができる。ただし、移動と攻撃が同時にできない・近距離攻撃ができない・相手の攻撃に反撃できない、というデメリットを持っている。

自走砲は相手を一方的に攻撃できる強力なユニットだが、ユニットや地形の配置によっては有効に働かない場合がある。例えば、図 1 の右下、青チームの自走砲は一方的に敵戦車を攻撃可能な位置であり、山と林が邪魔で敵戦車はこちらにすぐは来られない。非常に有利な状況である。一方で同図の左上、赤チームの自走砲は周囲を山と敵に囲まれてしまっていて、移動も攻撃も行うことができない。このように自走砲はユニットと地形の配置によって全く強さが異なるため、よりゲームを複雑にする要素であるといえる。

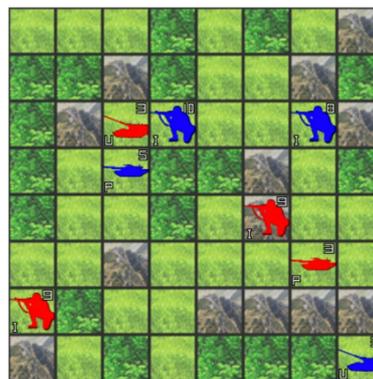


図 1 自走砲の有効性が顕著に異なるマップ

### 3. グラフによる表現

Wanxiang らは盤面の評価の難しさに対して、TUBSTAP におけるグラフを用いた盤面の表現と、Graph Convolutional Networks[2] を用いた盤面価値の学習法を提案した [3]. 囲碁などのゲームでは画像によって盤面を表現し、CNN によって盤面特徴の抽出を行うという手法が一般的であり、TUBSTAP の機械学習においてもこの方法が用いられてきた。Wanxiang らは TUBSTAP においてユニット同士の関係性が非常に重要であることから、それをグラフによって表現することで、より盤面の情報を適切に捉えることができると考えた。Wanxiang らの考案したグラフでは、敵味方合わせた全ユニットをグラフの頂点とし、攻撃可能な位置にいるユニット同士を辺で結んでいる。画像による表現よりも情報の量としては少なくなるが、このグラフによる表現はその分重要な情報のみに着目できる。

論文内では画像表現と CNN を用いたモデルと、グラフ表現と GCN を用いたモデルのそれぞれで盤面価値の学習を行った。その後、学習した盤面価値を利用した Mini-Max 法による AI の作成を行なった。結果として、グラフ表現と GCN を用いた AI が画像表現と CNN を用いた AI よりも高い勝率になることが示された。

## 4. 提案手法

### 4.1 目的

Wanxing らの実験においては、TUBSTAP の複雑さの要因の一つとなる地形の要素は簡略化され、すべて平原のみのマップで実験を行っていた。そのため、グラフ表現も地形の種類を考慮しないものとなっていた。

また、Wanxing らの用いたグラフ表現は有効ではあるものの、画像による表現には存在していた情報が抜け落ちてしまうという問題があった。例えば、ユニット同士の相対的な位置は表現されているものの、盤面の中での絶対的な位置は表現されていない。そのため、「盤面の角に相手を追い詰める」といった戦略を生み出すことはできない。ユニット同士の関係を表すことに長けたグラフ表現に加えて、位置の情報を表すことに長けた画像表現も組みあわせることで互いの欠点を補い、より高い性能をもつ手法となることが期待される。

本研究では、まず TUBSTAP において地形の要素を考慮できる新たなグラフ表現を提案し、そのグラフ表現を用いる GCN のモデルと、画像表現を用いる CNN のモデルについて述べる。次にグラフ表現と画像表現を併用する統合モデルを 3 種類提案する。最後に 5 種類のモデルそれぞれ盤面評価の学習を行い、その性能を比較する。

## 4.2 GCN

### 4.2.1 盤面のグラフ表現

敵味方合わせた全ユニットの各々に対してノードを設け、ユニット同士の関係性をエッジとする。すべてのユニット同士は 2 本の有向エッジで相互に結ばれる。

ノードの特徴量は 13 次元のベクトルとする。最初の 12 次元は (ユニット全 6 種類) × (手番と非手番どちらの所属か) に対応する one-hot ベクトルである。ただし該当部分は 1 ではなく (ユニットの HP / 10) にすることにより、ユニットの HP も表現に加える。これにより HP が多く影響の大きいユニットほどより大きな値となる。最後の 1 次元は行動可能フラグであり、手番プレイヤー所属で未行動のユニットは 1 に、それ以外は 0 になる。

ユニット  $i$  から  $j$  へのエッジの特徴量は以下に示す 4 次元ベクトルである。

- 味方フラグ:  $x_1 \in \{0, 1\}$
- 敵フラグ:  $x_2 \in \{0, 1\}$
- $i$  から  $j$  への攻撃ダメージ/10:  $x_3 \in \{0, 0.1, \dots, 0.9, 1\}$
- $j$  から  $i$  への最小反撃ダメージ/10  
:  $x_4 \in \{0, 0.1, \dots, 0.9, 1\}$

攻撃ダメージと最小反撃ダメージは  $i$  から  $j$  が攻撃できない場合は 0 になる。最小反撃ダメージを計算する際にはユニット  $i$  が  $j$  に攻撃できるすべての位置を探索し、最も防御効果の高い地形から攻撃した時の値を用いる。これらの特徴量はたとえユニットの配置が同じでも、地形の配置によっては異なる値をもつ。そのためこのエッジ特徴量を用いることで、GCN は地形の状況を考慮することができる。

### 4.2.2 畳み込み式

畳み込みには Gilmer らの提案を元にした、NNConv と呼ばれる以下の畳み込み式を用いる [4].

$$x'_i = Wx_i + \frac{1}{|N(i)|} \sum_{j \in N(i)} x_j \cdot h_{\Theta}(e_{i,j})$$

- $x_i$ : ユニット  $i$  のノード特徴量
- $e_{i,j}$ : ユニット  $i$  から  $j$  へのエッジ特徴量
- $N(i)$ :  $i$  を除くすべてのユニット
- $W$ : 結合重み
- $h_{\Theta}$ : 全結合層

この式では、ノード特徴量とエッジ特徴量の両方を考慮して新たなノード特徴量を得ることができる。また、全結合層の出力次元数を調整することで畳み込み後にノード特徴量を任意の次元に変えることができる。エッジ特徴量は計算に使われるのみで畳み込みの前後で変化しない。

### 4.2.3 モデル

GCN のモデルを図 2(左) に示す。入力ノード特徴量  $13 \times N$  ( $N =$  ユニットの数) 次元は 4 回の畳み込みを経て  $64 \times N$  次元となる。⑤では各特徴量について全ノードでの最大値を取り 64 次元の出力を得ている。その後全結合

層と Softmax 層を経て 3 次元の予想出力を得る。これは (手番プレイヤーの勝ち, 手番プレイヤーの負け, 引き分け) となる確率の予測である。

### 4.3 CNN

盤面を画像で表現する方法と, 画像表現から盤面評価を行う CNN モデルについて記述する。

#### 4.3.1 盤面の画像表現

画像表現ではユニット情報を表す 12 枚, 地形情報を表す 7 枚, 着手可能フラグを表す 1 枚, の計 20 枚の画像を用いる。各画像は盤面の大きさと等しい  $8 \times 8$  のサイズをもつ。ユニット情報の画像 12 枚は (ユニット全 6 種類)  $\times$  (手番と非手番どちらの所属か) に対応しており, 各地形に該当するマス目は (ユニットの  $HP/10$ ) に, その他のマスは 0 にする。地形情報の画像 7 枚は地形全 7 種類に対応しており, 該当するマス目は 1 に, その他のマスは 0 にする。行動可能フラグの画像は, 行動可能なユニットがいるマスのみを 1 に, その他のマスは 0 にする。

#### 4.3.2 モデル

CNN のモデルを図 2(右) に示す。入力サイズは  $20 \times 8 \times 8$  であり, 4 回の畳み込みと 2 回の全結合層を経て 3 次元の特徴量を得る。

### 4.4 統合ネットワーク

GCN と CNN を統合するネットワークはその組み合わせ方にいくつかの方法が考えられるため, 3 種類のモデルを提案する。

#### 4.4.1 統合モデル A

統合モデル A は, GCN と CNN それぞれから得た特徴量を組み合わせて全結合層で処理するモデルである。モデル図を図 3 に示す。まずグラフ表現が GCN(図 2) の⑤まで処理され 64 次元の特徴量を得る。また, 画像表現が CNN(図 2) の⑥まで処理され 64 次元の特徴量を得る。その後, GCN と CNN 2 つの特徴量を連結して 128 次元にし, 2 層の全結合層を通して 3 次元の予想出力を得る。

#### 4.4.2 統合モデル B

統合モデル B は, GCN と CNN それぞれが出力した勝敗予想を単純に足し合わせ, 最終的な勝敗予想とするモデルである。モデル図を図 4 に示す。まずグラフ表現が GCN(図 2) の⑥まで処理され 3 次元の特徴量を得る。また, 画像表現が CNN(図 2) の⑦まで処理され 3 次元の特徴量を得る。その後, GCN と CNN 2 つの特徴量を各次元ごとに足し合わせて 3 次元の予想出力を得る。

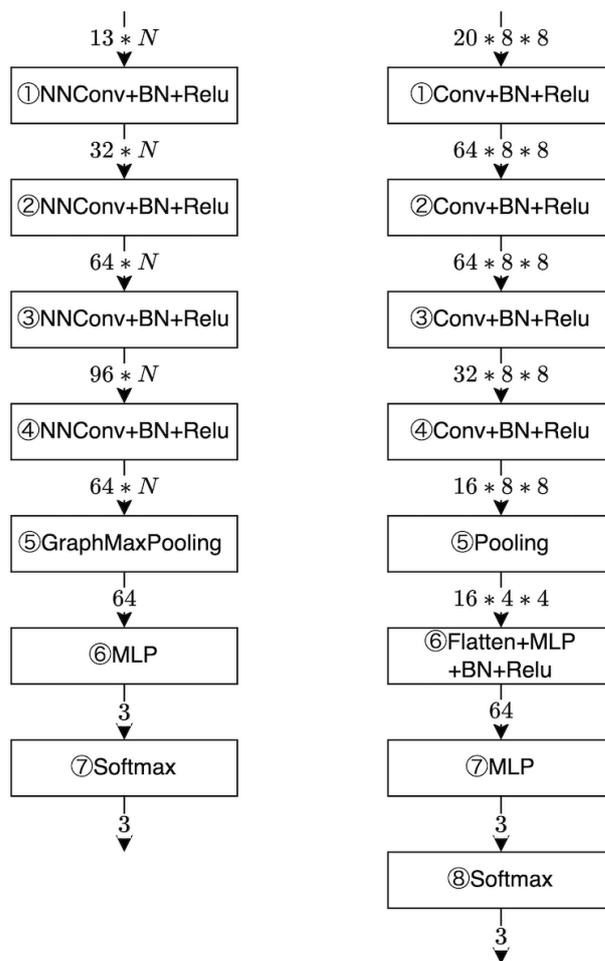


図 2 (左)GCN のモデル図 (右)CNN のモデル図

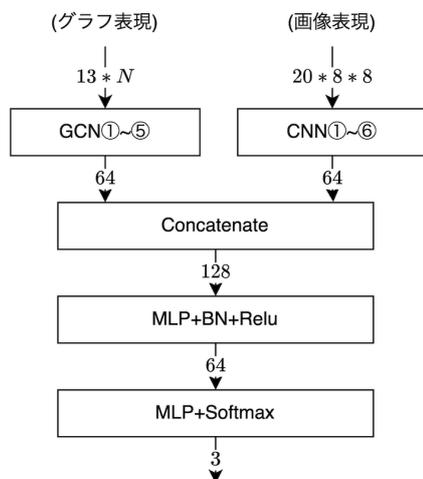


図 3 統合モデル A のモデル図

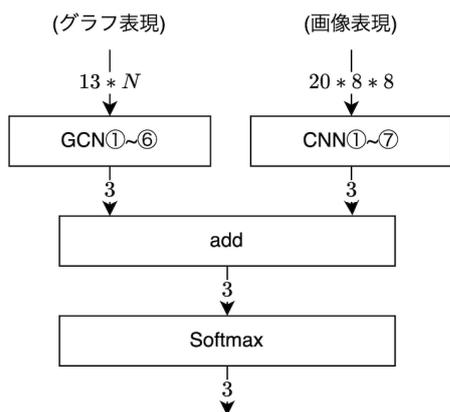


図 4 統合モデル B のモデル図

#### 4.4.3 統合モデル C

統合モデル C は、CNN が画像から抽出した特徴量を GCN のノード特徴量として利用するモデルである。モデル図を図 5 に示す。まず画像表現が CNN(図 2) の④まで処理され  $16 \times 8 \times 8$  次元の特徴量を得る。この  $8 \times 8$  次元は盤面の大きさ  $8 \times 8$  に対応するものであった。GCN の元のノード特徴量 13 次元に、対応する位置の CNN 特徴量 16 次元を連結することで、ノード特徴量を 29 次元に拡張する。その後、29 次元の特徴量が GCN(図 2) の①から⑧まで処理されて 3 次元の特徴量を得る。

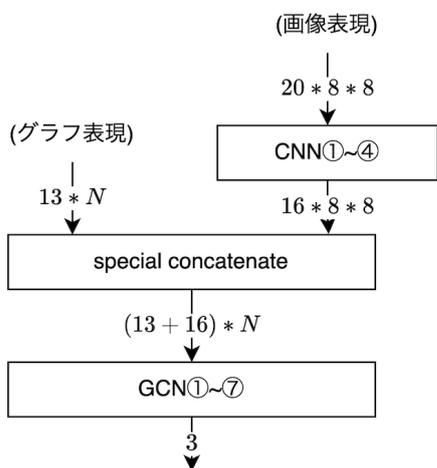


図 5 統合モデル C のモデル図

## 5. 実験

まず、対戦マップを複数生成しモンテカルロ法の AI を用いて自己対戦を行うことで棋譜を収集する。次に、得られた棋譜に登場する各盤面を学習データとして、最終的な勝敗を予測するように各モデルを学習し性能を比較する。

### 5.1 対戦マップ

学習データに用いるマップは、2種類の生成方法で 5000 マップずつ生成した。

一つ目は単純マップと名付けたマップである。マップの例を図 6 に示す。各チームごとに 3 体の歩兵と 1 体の戦車を持っていて、ランダムに配置される。歩兵の HP は 6 ~ 10 の範囲で、戦車の HP は 3 ~ 5 の範囲でランダムに初期化される。地形は各マスごとに 30% の確率で平原となり 70% の確率で林になる。16 ターンで決着がつかない場合は判定に移り、残りユニットの合計 HP の多寡で勝敗を決める。合計 HP が一致した場合のみを引き分けとする。

二つ目は複雑マップと名付けたマップである。マップの例を図 7 に示す。各チームごとに 2 体の歩兵と 1 体の戦車、1 体の自走砲を持っていて、ランダムに配置される。歩兵の HP は 6 ~ 10 の範囲で、戦車と自走砲の HP は 3 ~ 5 の範囲でランダムに初期化される。地形は 40% の確率で平原、30% の確率で林、30% の確率で山になる。複雑マップでは自走砲と山が存在するため、平原マップよりも地形を考慮することが重要となる。

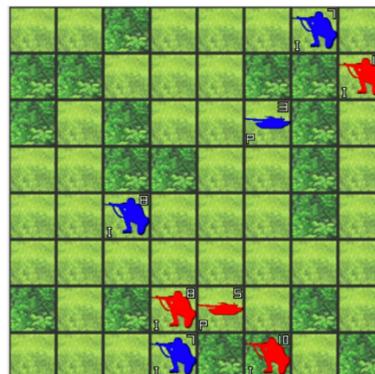


図 6 単純マップ

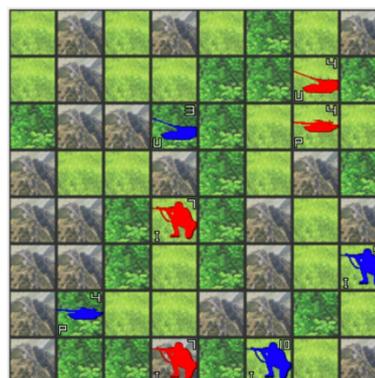


図 7 複雑マップ

## 5.2 棋譜の生成

2016年3月のUEC-GAT大会[5]の優勝プログラムであるM-UCT[6]を用いて、上記5000×2マップにおいて一度ずつ自己対戦を行い、対戦に現れたすべての盤面を学習データとして使用した。M-UCTはモンテカルロ木探索を用いたAIであり、現在TUBSTAPに標準搭載されているAIの一つである。最大ターンを16として、決着がつかなくなった場合は残ユニットのHP合計の多い方を勝利とし、一致する場合は引き分けとした。

標準マップの自己対戦からは117456個の盤面を得ることができた。そのうち、現在手番のプレイヤーが最終的に勝つ盤面は69121個、負ける盤面は45766個、引き分けとなる盤面は2569個であった。

複雑マップの自己対戦からは175533個の盤面を得ることができた。そのうち、現在手番プレイヤーが最終的に勝つ盤面は102831個、負ける盤面は66743個、引き分けとなる盤面は5959個であった。

## 5.3 学習

学習データのうち8割を訓練データとして、2割をテストデータとした。最適化アルゴリズムはAdam[7]を用いて、学習率は0.001とした。誤差関数には交差エントロピー誤差を用いた。バッチサイズを2000、エポックを50として十分収束するまで学習を行った。

## 5.4 結果

学習の結果として、各モデルの訓練正解率とテスト正解率を表1に示す。学習時に最もテスト正解率の高かったエポックのデータを記載している。

モデル	標準(訓練)	複雑(訓練)	標準(テスト)	複雑(テスト)
GCN	94.2%	91.7%	93.4%	90.0%
CNN	95.6%	96.8%	85.7%	82.2%
統合 A	95.8%	94.3%	92.4%	88.7%
統合 B	96.5%	94.9%	92.6%	89.1%
統合 C	95.5%	91.7%	87.9%	85.5%

表 1 学習結果

標準マップではGCNが最もテスト正解率が高く、統合モデルがB, A, Cの順に続き、CNNが最も低いという結果になった。複雑マップでは全体的に正解率が少し下がったものの順位は変わらなかった。

## 6. まとめ

本研究ではTUBSTAPにおいて地形の要素を考慮できる新たなグラフ表現を提案し、そのグラフ表現を用いるGCNのモデルと、画像表現を用いるCNNのモデルのそれぞれで盤面評価の学習を行った。その結果、TUBSTAPの盤面評価において地形の要素がある場合でもGCNがCNN

より有効な手法であることが示された。これはエッジ特徴量に攻撃ダメージと最小反撃ダメージを用いる提案手法が有効に働いたためと考えられる。

また、本研究ではユニット同士の関係を表すことに長けたグラフ表現に加えて、位置の情報を表すことに長けた画像表現も組みあわせることで、より性能の高いモデルになることを期待していた。GCNとCNNの組み合わせ方を変えた3つの統合モデルを提案し、それぞれのモデルで盤面評価の学習を行った。その結果、提案した統合モデルの中では、GCNとCNNそれぞれが出力した勝敗予想を単純に足し合わせ、最終的な勝敗予想とするモデルが最も良い性能であることが示された。しかしながら、最も性能の良い統合モデルでもGCNの性能を超えることはできなかった。統合モデルの訓練正解率はどれもGCNより高かったことから、リッチな情報を持つメリットよりもCNN部分が過学習の原因となるデメリットの方が大きかったためだと考えられる。

今後は、地形のあるTUBSTAPにおけるGCNの有効性をさらに調べるため、GCNとMini-Max法を用いた対戦プログラムを作成し、モンテカルロ法などを用いたプログラムと対戦し性能を比較したい。また、グラフ表現と画像表現の併用についても、過学習を抑えることでGCN単体の性能を超えられる手法を検討していきたい。

## 参考文献

- [1] 村山公志朗, 藤木翼, and 池田心. "学術研究用プラットフォームとしての大戦略系ゲームのルール提案." ゲームプログラミングワークショップ2013 論文集 (2013): 146-153.
- [2] Kipf, Thomas N., and Max Welling. "Semi-supervised classification with graph convolutional networks." arXiv preprint arXiv:1609.02907 (2016).
- [3] LI, Wanxiang, et al. Graph Convolutional Networks for Turn-Based Strategy Games. 2022.
- [4] Gilmer, Justin, et al. "Neural message passing for quantum chemistry." International conference on machine learning. PMLR, 2017.
- [5] GAT2016 TUBSTAP 大会 (<http://www.jaist.ac.jp/is/labs/ikedalab/tbs/competition.htm>)
- [6] 武藤孝輔, 西野順二. ターン制戦略ゲームにおけるUCTとファジィ評価の適用. In: 日本知能情報ファジィ学会ファジィシステムシンポジウム 講演論文集 第31回ファジィシステムシンポジウム. 日本知能情報ファジィ学会, 2015. p. 226-229.
- [7] Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2014).