

【背景・目的】 Geometric Deep Learning¹ とは、非ユークリッド構造を持つデータから学習可能な Neural Network の構築を目的とする研究分野である。とりわけ、グラフ・ネットワーク構造を扱う GNN(Graph Neural Network) モデルは、分子特性予測やソーシャルネットワークのコミュニティ予測において有効である。これらのモデルの多くは、ノード属性の集約により局所構造を学習する。このため GNN は、大域的な構造情報が重要な複雑ネットワークに対しては、適用限界がある。そこで本研究では、ネットワークの隣接行列より抽出される大域的構造情報のみに焦点を当てたネットワーク分類・変換手法の開発を目指した。

【方法】 隣接行列の画像化により、ネットワークデータが画像データを扱う既存の深層学習モデルで利用可能となる。更に、画像化に際して、隣接行列の次数による並べ替えを行う。以上の画像データ化により、ネットワークの分類・変換タスクが画像の分類・変換として実行される。得られた画像データを CNN・GAN を用いて学習した。ネットワーク分類では、合成データセットとベンチマークデータセットによる評価を行った。合成データセットは複雑ネットワークの代表モデルの Barabási-Albert モデル(BA), Erdős-Rényi モデルとその中間モデルを用いて作成し、ベンチマークデータセットは化学、生物、ソーシャルネットワークの3分野から4種用いた。またネットワーク変換では、合成データセットによる学習と現実ネットワークの変換による評価を行った。

【結果・議論】 表1にネットワーク分類のベンチマークデータセットに対する提案手法と GNN ベースの手法2種の精度を示した。CNNの添字(shuffle/sort)は隣接行列に対する並べ替えの有無を示している。提案手法は GNN ベースの手法と同等あるいはやや劣る精度である。また、次数による並べ替えはソーシャルネットワークのデータセットにおいて精度向上の効果が高い。図1は現実のネットワークを BA モデルに変換した際のネットワークを示している。点線のエッジは、変換により追加されたエッジであり、高次数ノードに対するエッジ追加傾向が高い。表2に変換前後のネットワーク特徴量を示した。変換によりハブノードの成長、最短経路長の低下を確認した。以上の結果は、本手法が現実ネットワークの成長に伴う大域的構造変換事象に有効であることを示唆している。

表1：ベンチマークデータに対する各モデルの精度

	MUTAG (Chemical)	D&D (Bio)	COLLAB (Social)	REDDIT-B (Social)
CNN(shuffle)	81.3 ± 9.2	73.8 ± 4.6	68.9 ± 1.8	76.9 ± 2.7
CNN(sort)	86.9 ± 5.3	72.8 ± 3.6	75.6 ± 1.4	89.5 ± 2.4
DGCNN	85.8 ± 1.7	79.4 ± 0.9	73.8 ± 0.5	87.8 ± 2.5
GIN	89.4 ± 5.6	75.3 ± 2.9	80.2 ± 1.9	92.4 ± 2.5

表2：変換前後の特徴量

	変換前	変換後
エッジ数	108	132
最大次数	8	19
平均次数	2.2	2.6
次数分散	3.5	9.2
平均クラスター係数	0.06	0.09
平均最短経路長	6.1	3.9
エッジ保存率	98.1 %	

図1：変換後のネットワーク

文献

1) Bronstein, Michael M., et al. *IEEE Signal Processing Magazine* 34.4 (2017): 18-42.