Brain Age Prediction using Topological Graph Augmentation and Graph Neural Networks

Ufuk Cem Birbiri[1,2,3]

[1] Université Côte d'Azur
[2] INRIA, Sophia Antipolis, France
[3] University of Amsterdam
ufuk-cem.birbiri@etu.univ-cotedazur.fr

Abstract. Many common neurological and neurodegenerative disorders, such as Alzheimer's disease, Parkinson's disease, or autism have been linked with anomalous patterns of biological aging of the brain. Discovering which brain regions are related to specific neurological disorders or cognitive stimuli became an essential field of neuroimaging research. Diffusion Tensor Images (DTI) can be used to reveal the biological underlying patterns behind the aging process, holds an individual's signs for various brain disorders, and propose new personalized-treatment approaches. In this work, we used the brain network graph representations of 998 patients. The nodes represent the brain regions and the edges are described as interactions between the regions. We trained a Graph Neural Network to predict four brain age classes [22-25, 26-30, 31-35, 36+]. To increase the accuracy of the model we propose a topological graph data augmentation method. We used the Mapper algorithm to convert brain graphs into the topological complex where each feature of the complex represents brain activations between regions. Based on our quantitative results, the GNN gained a 0.64 test accuracy using only raw graphs, and 0.776 test accuracy using both raw graphs and topological augmented graphs. Using topological graph representations of the brain helped to reveal the functional connectivity of the brain and increased the model performance.

Keywords: Brain age · Graph Neural Network · Topological complex.

1   Introduction and Related Work

The relation between the neurological and neurodegenerative disorders and deviational brain aging pathways has inspired a great number of research topics in brain age prediction using neuroimaging data [1–3]. The estimations of brain age gaps, described as the difference between the estimated brain age and the true biological age, have been related to brain diseases such as dementia and autism, Alzheimer's disease, Parkinson's, or Epilepsy [4, 5], and therefore could be beneficial in monitoring and treatment of disease.

Various studies used machine learning in brain age prediction, mostly using genetic information and magnetic resonance imaging(MRI) [3]. These studies mostly do not include other brain imaging modalities such as time-series data of functional MRI(fMRI), clinical test results, or the expertise of psychiatrists, or neuroscientists. Still, different modalities show improvement in the results [1, 6, 7].

To estimate brain age in a clinically applied fashion, we used Graph Neural networks(GNN). GNN provides a deeper way of learning of brain mapped as a graph structure and it is rapidly becoming the state-of-the-art approach in various network neuroscience problems. The brain is an

exceptionally complex system and understanding its functional organization is the goal of understanding various diseases. The motivation behind brain age estimation is to have a GNN network to predict human age by using the brain regions as nodes whose features will reveal brain activation in such patients.

Here we used Diffusion Tensor Images (DTI) of patients represented as a connectivity matrix of the brain. The connectivity matrix has been converted to a graph where the nodes in the graph are characterized as brain regions and edges as the functional connectivity between brain regions, computed as the pairwise interactions of DTI. We analyzed the effectiveness of GNN for brain age prediction by training Graph Convolutional Network [8] between four age groups. We additionally investigate adding topological graph representation of the brain networks as a data augmentation method to increase prediction accuracy.

2   Methods

In this section, we describe the dataset and the task description as well as the topological data augmentation method applied to brain networks.

2.1   Dataset

The data includes 998 patients belonging to Human Connectome Project (HCP) [9]. Each patient has a DTI scan represented as brain network activation during a period. The brain networks consist of brain regions and their interactions with each other. Each patient has a brain network scan represented as a connectivity matrix in which rows and columns are described as brain regions, the values represent the interactions between two regions. Connectivity matrices measure the number of fibers from one region to another region of the brain.

In addition, there are other clinical variables for the 998 patients in the dataset such as age, gender, diseases, past injuries, disease history, brain regions' volume, surface area, thickness, grey matter volume as well as cognitive test scores, motor skills, sensory, and emotion. The dataset includes very different clinical information for each patient. If we understand their brain network functioning, we could be able to understand specific cases for example an Alzheimer patient's brain functioning. Age prediction is important since age is the most effective and relatable parameter in such diseases [1]. Building a Graph network that can predict a patient's age correctly will contribute to disease analysis in neurological research.

2.2   Task Description

In this work, we represented the brain networks as connectivity matrices shape (116, 116) meaning that rows and columns are 116 brain regions. There is one matrix for each patient and each represents a weighted undirected graph. The elements of the matrix are the number of fibers in between the regions after normalization. The connectivity matrices are converted to undirect graphs using the Networkx library [10] where nodes represent the 116 regions, and edges represent interactions. Figure 1 shows an example of a connectivity matrices of a patient with its graph visualization. The different brain regions are shown in the matrice such as $Precentral\_L$ as the left Precentral gyrus, $Frontal\_Sup\_L$ as the left Superior frontal gyrus-dorsolateral, and $Frontal\_Sup\_Orb\_L$ as left Superior frontal gyrus-orbital.
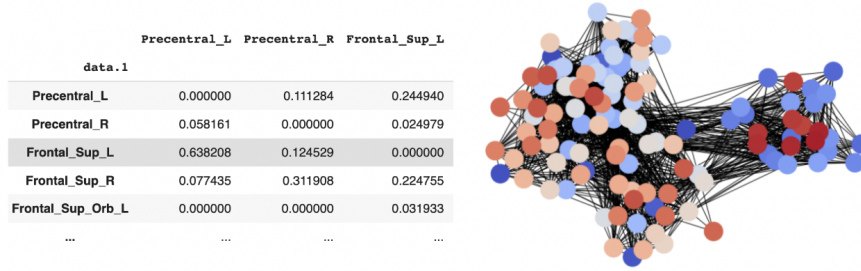
data.1	Precentral_L	Precentral_R	Frontal_Sup_L
Precentral_L	0.000000	0.111284	0.244940
Precentral_R	0.058161	0.000000	0.024979
Frontal_Sup_L	0.638208	0.124529	0.000000
Frontal_Sup_R	0.077435	0.311908	0.224755
Frontal_Sup_Orb_L	0.000000	0.000000	0.031933
...	...	...	...

Fig. 1: *((Left) An example of a brain connectivity matrix of size 116×116 that belongs to a male patient with age $36+$. The rows and columns represent brain regions. (Right) The graph visualization of the same patient's connectivity matrix where nodes represent brain regions.*

The patient's age differs from four categories [22-25, 26-30, 31-35, 36+] and the class ratio of the ages in the dataset are 17%, 35%, 25%, and 23%, respectively. The problem, therefore, turns into a multi-class graph classification problem with four different classes.

2.3 Dataset Preparetion

GNNs for graph classification have been used in network neuroscience [11], drug discovery to classify molecules as toxic or non-toxic [12], or in text classification [13]. To use our dataset in the GNN network we need to use the *torch.geometric Data* type. In our task, each brain network is represented as a graph with node and edge features. The edge attributes are assigned as interaction values and the node features are assigned as the eigenvector centrality of the nodes.

Eigenvector Centrality
The eigenvector centrality for node i is the ith entry of the vector characterized by the equation

$$Ax = \lambda x,$$

where A is the adjacency matrix of the graph G with eigenvalue λ. By the Perron–Frobenius theorem [2], one can conclude that there exists a unique solution x. If λ is the largest eigenvalue of the adjacency matrix A, we can say that the unique solution x has positive values in all of the entries [14]. As a result, the eigenvector centrality of each of the nodes is calculated G and assigned as the node attribute.

2.4 Graph Convolutional Network

The GNN model was adopted from the 'Semi-supervised Classification with Graph Convolutional Networks' [15] that uses the graph convolutional operator. In this network, three graph convolutional layers (GCNConv [16]) were used followed by a linear layer. The Relu activation function is added as an activation function. To overcome the over-fitting, dropout is used before the last linear layer. In addition, Pytorch geometric provides a pooling layer called *global_mean_pool* located before the dropout which returns batch-wise and graph-level outputs by taking the average of node features

across the node dimensions [17]. The GNN model is shown in the Appendix section.
To increase the number of graphs in the dataset and represent the brain networks in a better way, we propose the topological graph augmentation method.
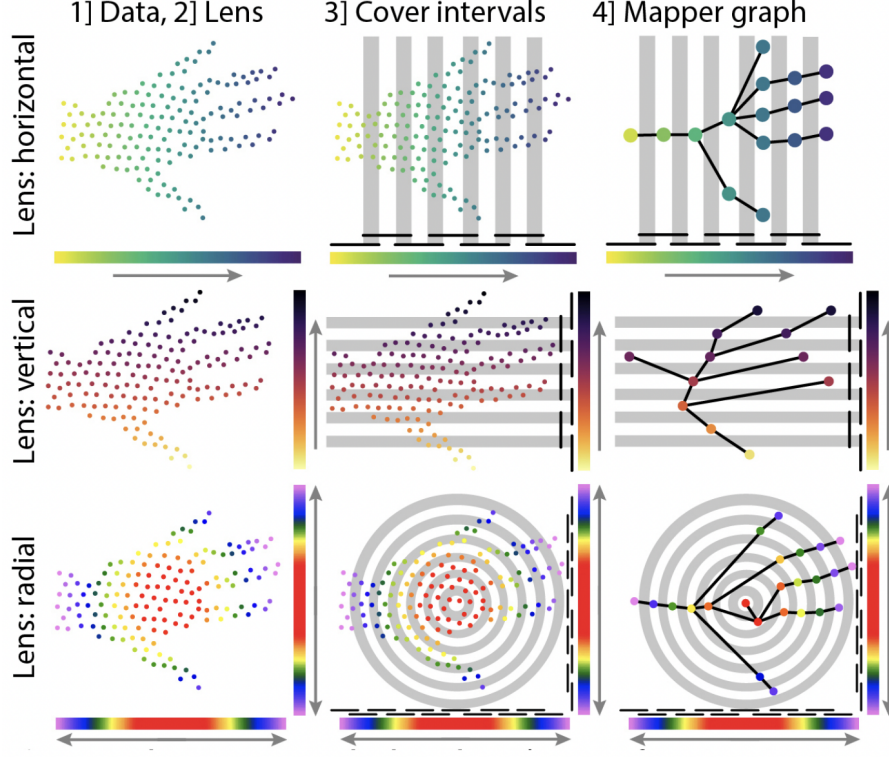

Fig. 2: *Three different lenses(filters) on the same point cloud data. The rows show the horizontal, vertical, and radial filter selection results, respectively. The data, the cover intervals(gain and resolution parameters), and the Mapper result are shown in the left, middle, and right columns respectively. The image is taken from [18].*

2.5   Data Augmentation Using Topological Complex

Understanding the *shape of data* brings us to topology, which is a branch of mathematics. Over the last 20 years, topological techniques have been used in various applied problems, one of these fields is called Topological Data Analysis (TDA). TDA provides detailed information about the characteristics of the data by providing some topological, statistical, and geometrical methods to infer complex topological structures such as connected components, holes, branches, or loops. This analysis method is very robust to outliers and noise where the data is usually represented as point clouds or distance matrices in a Euclidean metric space.

Each brain network graph is converted to a topological complex called the Mapper complex using the GUDHI Python library [19]. Mapper is a computational method for complex datasets to extract the simplest information in the form of $simplicial-complexes$ that preserve the underlying topological structures from the original data. It uses metrics and lenses to convert data into a topological network where nodes represent similar data points and edges represent topological relations. The key idea of the Mapper algorithm is to map data points into a metric space using some filters defined on data points while capturing topological and geometric information at a specified resolution and gain. The filter function (also called a lens) is the metric that allows us to map our data points. The choice of the filter is an important factor for achieving a good result in the Mapper. The results of different filter functions are shown in Figure 2. In the figure, three different filters (horizontal, vertical, and radial) are applied to the same data. It can be seen that different filter functions have different results. Therefore, selecting the filter functions for the Mapper algorithm plays a critical role.

The most commonly used filters are Singular Value Decomposition (SVD) and L-Infinity centrality based on their quantitative performances in topological data analysis [20–22]. The two filter functions are also selected as L-Infinity and SVD as the first and second filters, respectively. L-infinity centrality assigns a value to each data point which is the furthest distance from itself to another data point. Basically, for each data point y, it amounts to finding the maximum distance from y to any other point in the dataset. Large values of this filter put points that are far away from the center of the dataset. On the other hand, SVD is a linear dimension reduction technique that does not center the data points before the actual computation, meaning that it efficiently works with sparse data. We used the first component of the SVD in the Scikit-learn library [23]

The other parameters in the Mapper are the resolution and the gain. Resolution is the number of intervals required to cover each filtered image. If the resolution is increased, then the number of communities in the resulting graph will also increase. The communities refer to the nodes clustered together in the Mapper. This parameter is tested in the range [2-50] and selected as 20. The gain parameter is the percentage of overlaps between the intervals covering each filtered image. It should be between 0 and 1 and selected as 0.25. In addition, the clustering algorithm is selected as $DBSCAN$. Parameter selection was done based on a literature search and we choose the parameters that give the highest accuracy in the GNN model.
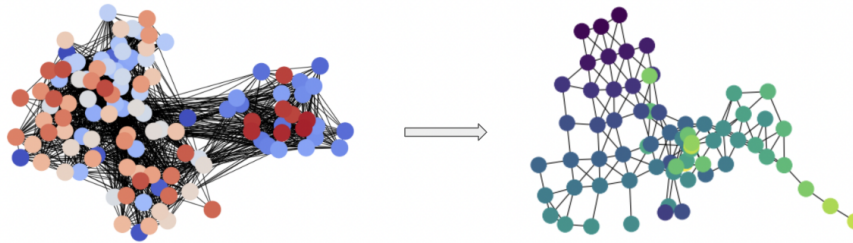

Fig. 3: *(Left) The brain network is displayed as a Networkx graph. (Right) The result of the Mapper algorithm.*

The Mapper algorithm is applied for each brain network graph and achieved a Mapper complex representation of the graphs. The Mapper complex is another Networkx graph with fewer nodes as edges since similar brain activations clustered together during the complex computation. The result of the same patient after the Mapper is shown in Figure 3. The nodes in the Mapper complex represent clustered brain regions and edges represent high-order interactions of the brain regions. The colors in both graphs do not mean anything and are used only for visualization. More Mapper results can be found in the Appendix section.

2.6   Training

Two different datasets were created for the GNN model. The first one consists of only 998 raw brain network graphs. The second dataset contains both raw brain network graphs(998) and topological graphs(998) resulting in 1996 graphs in total. The dataset was split into training(70%) and testing(30%). To obtain objective results, the test set included only raw graphs, and the testing data was only used in the testing phase. To achieve a better validation of the experiments, 4-fold cross-validation was performed. The models were trained with 150 epochs with a learning rate = 0.0005 using the cross-entropy loss function.

3   Results

The task is the multi-class graph classification with four different age groups in terms of [22-25, 26-30, 31-35, 36+]. It is a so-called balanced dataset that represents mostly the younger and mature phases of humans. The performance of the Graph Convolutional Network will be tested on both the *raw* graph dataset and the *raw + topological* graph dataset. The accuracy results are shown in Table 1.

Table 1: The training and testing accuracy results of multi-class classification of the network.

Training datasets	Train accuracy	Test accuracy
Raw graphs	0.634 ± 0.15	**0.640 $\pm$ 0.32**
Raw + Topological graphs	0.782 ± 0.12	**0.776 $\pm$ 0.20**

When we use only the raw brain graphs we achieve 0.634 training and 0.640 test accuracy. If we add the topological complex representations of the brain networks the age prediction accuracy achieved 0.782 in the training set and 0.776 in the testing set. The topological data augmentation not only increased the number of graphs in the dataset but also gave a chance to GNN to learn better from topological brain networks. The Mapper complex is the topological structure of the brain networks and its topological features represent brain activation. A topological feature can be a connected component, a 1D hole/loop, a 2D cavity, or more generally a d-dimensional "void" which preserves the interactions between different brain regions.

The proposed model could be used for the prediction of age in humans. The proposed graph representations capture informative features of the brain by linking the interaction among different brain regions.

4   Conclusion

The human brain is an enormously complex system and discovering its functional connectivity is the goal of neuroscience research. Many diseases can be related to brain age such as Epilepsy, stroke, Amyotrophic Lateral Sclerosis (ALS), or trauma. Understanding aging and its effects on various disorders can reveal much information about brain functioning. Diffusion Tensor Imaging is a novel imaging method that can reveal unique information about white matter within the brain regions. In this work, we presented brain networks of 998 patients as graphs created from DTI connectivity matrices. We used a Graph Convolutional Network to predict the brain age of patients in four different age categories. Also, we propose topological graph augmentation methods to analyze brain activation using the Mapper complex. The proposed method increased the GNN performance.

5   Perspectives

For the future of this study, we can different directions to understand complex brain structure and functioning. Each sample in the dataset was a DTI scan of a patient and it is represented by a graph. Significantly, the nodes are described as brain regions and edges are the functional connectivity between them. The Graph Convolutional layers can be analyzed with some feature engineering to understand which brain regions(nodes) were more effective for age prediction. Can we relate these regions with the patient's specific disease(Alzheimer, autism, etc)? Do these regions have anomalies in brain thickness, surface area, or (less) gray matter volume? We also used Topological data augmentation where we convert the brain network graphs into a topological complex called the Mapper complex. In the Mapper complex, the nodes were mapped by two filter functions and followed by clustering. Which brain regions are clustered and why? How to improve the classification accuracy by using different filters in Mapper. The subgroup analysis on brain regions to understand the interaction between brain regions [24, 25].

6   Acknowledgements

References

1. Tobias Kaufmann, Dennis van der Meer, Nhat Trung Doan, Emanuel Schwarz, Martina J Lund, Ingrid Agartz, Dag Alnæs, Deanna M Barch, Ramona Baur-Streubel, Alessandro Bertolino, et al. Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nature neuroscience*, 22(10):1617–1623, 2019.
2. Phillip Bonacich. Power and centrality: A family of measures. *American journal of sociology*, 92(5):1170–1182, 1987.
3. Katja Franke and Christian Gaser. Ten years of brainage as a neuroimaging biomarker of brain aging: what insights have we gained? *Frontiers in neurology*, page 789, 2019.
4. Michael A Nalls, Raquel Duran, Grisel Lopez, Marzena Kurzawa-Akanbi, Ian G McKeith, Patrick F Chinnery, Christopher M Morris, Jessie Theuns, David Crosiers, Patrick Cras, et al. A multicenter study of glucocerebrosidase mutations in dementia with lewy bodies. *JAMA neurology*, 70(6):727–735, 2013.

5. Abdulkadir Tunç. Early predictors of functional disability in guillain–barré syndrome. *Acta Neurologica Belgica*, 119(4):555–559, 2019.
6. Bianca Besteher, Christian Gaser, and Igor Nenadić. Machine-learning based brain age estimation in major depression showing no evidence of accelerated aging. *Psychiatry Research: Neuroimaging*, 290:1–4, 2019.
7. Lea Baecker, Rafael Garcia-Dias, Sandra Vieira, Cristina Scarpazza, and Andrea Mechelli. Machine learning for brain age prediction: Introduction to methods and clinical applications. *EBioMedicine*, 72:103600, 2021.
8. Rianne van den Berg, Thomas N Kipf, and Max Welling. Graph convolutional matrix completion. *arXiv preprint arXiv:1706.02263*, 2017.
9. Human Connectome data, `file:///Users/ufukcembirbiri/Downloads/Human`.
10. Networkx library Python, `https://networkx.org/`.
11. Alaa Bessadok, Mohamed Ali Mahjoub, and Islem Rekik. Graph neural networks in network neuroscience. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.
12. Oliver Wieder, Stefan Kohlbacher, Mélaine Kuenemann, Arthur Garon, Pierre Ducrot, Thomas Seidel, and Thierry Langer. A compact review of molecular property prediction with graph neural networks. *Drug Discovery Today: Technologies*, 37:1–12, 2020.
13. Lianzhe Huang, Dehong Ma, Sujian Li, Xiaodong Zhang, and Houfeng Wang. Text level graph neural network for text classification. *arXiv preprint arXiv:1910.02356*, 2019.
14. Markus Brede. Networks—an introduction. mark ej newman.(2010, oxford university press.) 65.38,£ 35.96 (hardcover), 772 pages. isbn-978-0-19-920665-0., 2012.
15. Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.
16. Graph convolutional networks. https://pytorch-geometric.readthedocs.io/en/latest/modules/nn.html. torch_geometric.nn.conv.GCNConv.
17. Graph convolutional networks. https://pytorch-geometric.readthedocs.io/en/latest/modules/nn.html#pooling-layers.
18. Frédéric Chazal and Bertrand Michel. An introduction to topological data analysis: fundamental and practical aspects for data scientists. *Frontiers in artificial intelligence*, 4:667963, 2021.
19. GUDHI, `https://gudhi.inria.fr/`.
20. Jessica L Nielson, Jesse Paquette, Aiwen W Liu, Cristian F Guandique, C Amy Tovar, Tomoo Inoue, Karen-Amanda Irvine, John C Gensel, Jennifer Kloke, Tanya C Petrossian, et al. Topological data analysis for discovery in preclinical spinal cord injury and traumatic brain injury. *Nature communications*, 6(1):8581, 2015.
21. Pek Y Lum, Gurjeet Singh, Alan Lehman, Tigran Ishkanov, Mikael Vejdemo-Johansson, Muthu Alagappan, John Carlsson, and Gunnar Carlsson. Extracting insights from the shape of complex data using topology. *Scientific reports*, 3(1):1–8, 2013.
22. Sagarika Banerjee, Tian Tian, Zhi Wei, Kristen N Peck, Natalie Shih, Ara A Chalian, Bert W O'Malley, Gregory S Weinstein, Michael D Feldman, James Alwine, et al. Microbial signatures associated with oropharyngeal and oral squamous cell carcinomas. *Scientific reports*, 7(1):1–20, 2017.
23. Svd - sklearn. https://scikit-learn.org/stable/modules/generated/sklearn.decomposition. TruncatedSVD.html.
24. Juan Vallarta. Topological data analysis to identify subgroups of type-2 diabetes mellitus patients.
25. Monica Nicolau, Arnold J Levine, and Gunnar Carlsson. Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival. *Proceedings of the National Academy of Sciences*, 108(17):7265–7270, 2011.

Appendix

In this section, we will demonstrate some visualizations of the Mapper complex. In Figure 4, four different patients' brain maps before and after the Mapper complex computation on the left and right columns, respectively. The age of the patients belongs to the following ranges respectively: [22-25, 26-30, 31-35, 36+].
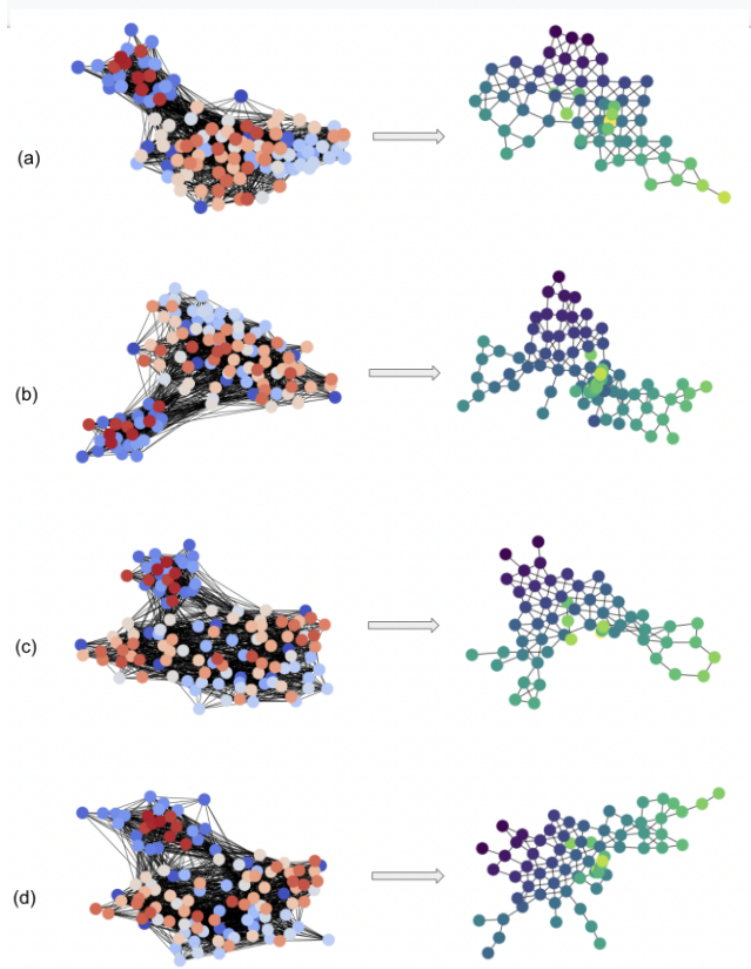

Fig. 4:

In Figure 5, the Graph Convolutional model that is used in this work is shown. There are three convolutional layers followed by a dropout and a linear layer.

```python
class GCN(torch.nn.Module):
    def __init__(self, hidden_channels):
        super(GCN, self).__init__()
        torch.manual_seed(12345)
        self.conv1 = GCNConv(1, hidden_channels)
        self.conv2 = GCNConv(hidden_channels, hidden_channels)
        self.conv3 = GCNConv(hidden_channels, hidden_channels)
        self.lin = Linear(hidden_channels, 4)

    def forward(self, x, edge_index, batch):
        # 1. Obtain node embeddings
        x = self.conv1(x, edge_index)
        x = x.relu()
        x = self.conv2(x, edge_index)
        x = x.relu()
        x = self.conv3(x, edge_index)
        # 2. Readout layer
        x = global_mean_pool(x, batch)   # [batch_size, hidden_channels]
        # 3. Apply a final classifier
        x = F.dropout(x, p=0.5, training=self.training)
        x = self.lin(x)
        return x

model = GCN(hidden_channels=8)
```

Fig. 5: *The GCN model for the graph classification task.*