



Budapest University of Technology and Economics

Faculty of Electrical Engineering and Informatics

Department of Automation and Applied Informatics

Medical data processing with graph neural networks

MASTER'S THESIS

Author

Anna Gergály

Advisor

dr. Luca Szegletes

May 30, 2025

Contents

Kivonat	i
Abstract	ii
1 Introduction	1
1.1 Structure	3
2 Background	5
2.1 Graph Neural Networks	5
2.1.1 Neural Networks	6
2.1.2 Working on Graph Data	7
2.1.2.1 Common tasks	9
2.1.2.2 Graph convolutions	11
2.1.3 Graph Neural Network Types	13
2.1.3.1 Graph Convolutional Networks	15
2.1.3.2 Graph Attention Networks	15
2.1.3.3 GraphSAGE	16
2.2 Medical data	17
2.2.1 Legal and ethical concerns	17
2.2.2 Types of medical data	18
2.2.3 Medical imaging	19
2.2.4 MRI and fMRI imaging	19
2.2.4.1 Clinical significance	21
2.2.4.2 Technical background	21

2.2.4.3	Processing	22
2.2.4.4	ROIs and connectomes	23
3	Technical background	24
3.1	Used Frameworks and Packages	24
3.1.1	Main Deep Learning Framework	25
3.1.2	Graph Neural Network Frameworks	25
3.1.3	Other Machine Learning Tools	25
3.1.4	Supporting packages	26
3.2	Using fMRI data	26
3.2.1	Data format	27
3.2.2	Extracting ROI Values and Connectomes	27
4	Methods and Implementation	28
4.1	Graph Neural Networks on a Group FC matrix	28
4.1.1	Dataset	28
4.1.2	Methodology	29
4.1.3	Results	30
4.2	Connectome-based diffusion	30
4.2.1	Dataset	32
4.2.2	Methodology	33
4.2.2.1	Data Preprocessing	33
4.2.2.2	Proposed Architecture	35
4.2.2.3	Graph U-net	36
4.2.2.4	Experiments	38
4.2.3	Results	41
4.2.4	Conclusions	42
4.3	Exploring device-to-device differences	43
4.3.1	Dataset	44
4.3.2	Methodology	45
4.3.3	Results	49

4.3.3.1	Full dataset results	49
4.3.3.2	Leave-one-out results	50
4.3.4	Conclusions	50
5	Summary	52
5.1	Summary	52
5.2	Future work	53
Acknowledgements		54
Bibliography		55

HALLGATÓI NYILATKOZAT

Alulírott *Gergály Anna*, szigorló hallgató kijelentem, hogy ezt a diplomatervet meg nem engedett segítség nélkül, saját magam készítettem, csak a megadott forrásokat (szakirodalom, eszközök stb.) használtam fel. minden olyan részt, melyet szó szerint, vagy azonos értelemben, de átfogalmazva más forrásból átvettettem, egyértelműen, a forrás megadásával megjelöltem.

Hozzájárulok, hogy a jelen munkám alapadatait (szerző(k), cím, angol és magyar nyelvű tartalmi kivonat, készítés éve, konzulens(ek) neve) a BME VIK nyilvánosan hozzáférhető elektronikus formában, a munka teljes szövegét pedig az egyetem belső hálózatán keresztül (vagy autentikált felhasználók számára) közzétegye. Kijelentem, hogy a benyújtott munka és annak elektronikus verziója megegyezik. Dékáni engedéllyel titkosított diplomatervek esetén a dolgozat szövege csak 3 év eltelte után válik hozzáférhetővé.

Budapest, 2025. május 30.

Gergály Anna
hallgató

Kivonat

A mélytanulás az elmúlt időszakban nagyon hasznos eszköznek bizonyult különböző tudományterületeken és a népszerűsége robbanásszerűen megnőtt. Míg az átlag emberek sokszor leginkább az LLM-re vagy chatbotokra gondolnak a mesterséges intelligencia halhatán, addig tudományos alkalmazási körökben gyakran specifikus feladatokra betanított modellek vezetnek nagyon jó eredményekre. Többek között az orvosi területen is fontos szerepet tölt be: orvosi képelemzésen keresztül megkönyíti a diagnózisok felállítását és gyógyszerkutatások elősegítésén keresztül hozzájárul a betegek jobb kezeléséhez.

Az MRI az egyik legnépszerűbb és legismertebb orvosi képalkotó eszköz. Egyik altípusa, az fMRI az agyi aktivitás rögzítésére alkalmas, érzékeli az oxigéndús vér jelenlétét az agyszövetben. Az így készült felvételek elemzésére több módszer is alkalmas, az egyik ilyen módszer az agy funkcionálisan elkülönülő területekre, ROI-kra (regions of interest) szegmentálása és egyfajta agyi gráf felépítése a régiók együttműködése alapján.

Az orvosi adatokkal való munka azonban nem minden egyszerű: mivel ezeket drága műszerekkel, specializált környezetben gyűjtik, sokkal nehezebben hozzáférhetők, mint például az egyszerű szöveges adatok vagy az általános képek. A másik korlátozó tényező az ilyen adatok érzékeny jellege. Az adatvédelmi irányelveket és előírásokat be kell tartani, és a legtöbb adathalmazt anonimizálják, mielőtt hozzáférhetővé tennék.

Munkám során az fMRI-felvételekből létrehozott agyi gráfokat gráf neurális hálózatokkal elemzem a neurodivergencia és kognitív feladatok felismeréséhez. Ezek a modellek lehetőséget kínálnak a mélytanulás gráfokon történő alkalmazására, és bizonyítottan képesek kiemelkedő eredményeket elérni komplex belső szerkezetű adatokon. Megvizsgálom, hogy az adatok szűkössége és a különböző vizsgálati helyekről összevont adathalmazok hogyan hatnak az fMRI-kutatásra, és lehetséges megoldásokat keresek.

Abstract

Deep learning has been a very useful tool in various sciences in the last few years and as a result it exploded in popularity. While a lot of people often mostly associate LLMs and chatbots with artificial intelligence, tailor-made machine learning solutions applicable to only their narrowly defined task have been quietly advancing science in the background. One of the domains embracing deep learning has been the medical domain: from streamlining diagnosis through medical image analysis to pioneering new treatments through enhanced drug discovery methods, machine learning has contributed to improving patient outcomes.

One of the most popular and well known medical imaging devices is the MRI. One of its subtypes, the fMRI is capable of capturing activity in the brain by detecting the flow of oxygenated blood through brain tissue. These scans can be analyzed in many ways, but one of them involves segmenting the brain into functionally distinct areas, ROIs (regions of interest) and building a sort of brain graph based on the interactions of these regions.

Working with medical data is not always easy however: since it is collected in highly specialized environments with expensive equipment, it is much more scarce than, for example, simple textual data or general images. The other limiting factor is the sensitive nature of such data. Data protection guidelines and regulations need to be respected and most datasets are therefore anonymised before becoming accessible.

In my work I employ Graph Neural Networks to analyze brain graphs created from fMRI scans for both neurodivergency detection and task prediction. These models offer the possibility of using deep learning on graphs and have demonstrated their ability to yield outstanding results on data with a complex inner structure. I explore the way data scarcity and datasets pooled from different sites impact fMRI research and look for potential solutions.

Chapter 1

Introduction

Graph Neural Networks (GNNs) are an area of machine learning that is fairly new, but gained a lot of popularity in the last few years. The reason for their increasing popularity can be found in the increased need for artificial intelligence to handle graph datasets: many phenomena can be captured well as graphs, things with complex structures, or cases where connections between items are important. However, common machine learning techniques are not well-suited for the internal complexity and variability of graphs.

These models have proven to be able capture complex connectivity patterns within graphs and have been found useful in many fields, from recommender systems to the medical domain. A lot of these architectures were created mirroring Convolutional Neural Networks (CNNs), which work on strict grid formats, and extrapolate the idea of convolutions to a non-Euclidian form.

There are earlier examples of graph methods as well, from classical mathematical algorithms to random-walk based models. Graph Convolutional Networks (GCNs) are capable of providing a computationally simple way of using convolutional layers on graph data, allowing for architectures that are capable of processing large amounts of data and slotting into backpropagation-based deep learning architectures. GCNs and most modern architectures fit into a message passing principle: information is aggregated in nodes as their neighbours pass messages in.

The collection of medical data has been incredibly important in the work of physicians for a very long time, but in the era of modern medicine its importance has become ever greater. Modern sensors and imaging devices, as well as statistical methods have made it so that the collection of medical data is faster and more accurate than ever, and this data is also more valuable than ever.

Of course this data collection is not entirely without drawbacks: it is a very hard task to keep the incredibly personal and sensitive data of patients safe and to obtain and use such data completely ethically and in line with the pertinent regulations. Such concerns are

doubly interesting in machine learning research applications: these methods are known to be data-hungry, requiring at least hundreds of samples, which are often not feasible to collect in controlled studies and have to be obtained in other ways.

This need for data is ubiquitous in the deep learning world, but in cases where the data is naturally occurring (such as general text and images) it is generally not as big of a problem. In case of medical imaging data however, due to the aforementioned privacy issues, the proportional rarity of data and the associated costs mean that the data hunger is much harder to satiate. This scarcity makes the use of data augmentation techniques that much more needed in this domain, but this poses its own set of challenges: traditional image augmentation techniques are often not as successful in this domain, due to the unique properties of medical images [25].

This means that solutions tailor-made to this type of task are necessary to achieve the best results. One of the methods where this can be done is synthetic data generation, where the already existing samples are used to inform a model of the properties and distribution of the dataset and this model is tasked with generating entirely novel samples based on the base set. I examined one such solution for brain fMRI scans together with Levente Sipos in our joint Scientific Students' Association Report [27]. In this thesis I will be detailing my work pertaining to the project, while also giving background information, explaining the goal and overarching approach of the project (See Section 4.2).

MRI (Magnetic Resonance Imaging) is a tool commonly used by physicians to see inside the human body in a non-invasive way, due to its ability to capture even soft tissues at a high-resolution and the lack of ionising radiation, which makes its repeated use in patients possible. One of the organs commonly examined via MRI is the brain: this can be done via an fMRI (functional Magnetic Resonance Imaging) for example. fMRIs show insight into the workings of the human brain, by providing detailed images of the oxygenated blood flow in the brain, which correlates with the neuronal activity. fMRIs are commonly taken either for a longer time with a lack of external stimulus (rs-fMRI, resting state fMRI) or specifically to examine the reaction to stimulus or the completion of a task (task based fMRI).

Machine learning provides valuable tools for the analysis of functional neuroimaging data, being effective in predicting neurological conditions, psychiatric disorders or cognitive patterns. There are limitations to using CNNs for this task however. It is very computationally intensive to process the 3 or 4 (accounting for time) dimensional data and brain activity is not well localized for this within the scan: most of it occurs in the cortical and subcortical structures and regions that need to be considered together are often not the ones close in Euclidian space.

Looking to better understand fMRI data, scans can be segmented into brain regions to reduce spatial complexity and noise. The segmentation occurs along brain atlases: these

outline certain Regions of Interest (ROIs) that have a similar function or are likely to be activated together. There are many atlases available that have been constructed with different methodologies and lead to different results from the same scan. Typically the activation of these regions are averaged and this can be used to create a time series of average activation for each ROI for the duration of the scan. These average activation values are often referred to as ROI values or BOLD (Blood-Oxygen Level Dependent) signals.

Altogether, these time series represent the average activation of a ROI area in the brain at every timestep during the fMRI. Using correlations of these signals can help us determine which regions work together frequently: effectively meaning ROIs with a high correlation are often collaborating/communicating, meaning they are "close" when we consider the functionality of the brain. When assessing the functional qualities of the brain this can be more useful than physical closeness, the attribute a CNN would consider. These correlations can be used to construct a connectome or FC (functional connectivity) matrix that is a sort of brain graph: the vertices are the ROIs and the correlation values correspond to the edge weights, stronger if the regions work together.

The use of graph neural networks on this connectome may seem straightforward at first glance, but there are major roadblocks. Despite their successes in data interpretation and predictive modeling, the application of these techniques to the neuroimaging domain has been challenging partly due to the expansive parameter search space in the dataset construction. With various preprocessing pipelines, denoising strategies, brain atlases available and options for node features and graph construction parameters, it is severely difficult to create datasets that GNNs perform well on, considering that even the network's hyperparameters need to be optimized.

In my work I explored various use-cases for the use of graph neural networks on connectomes extracted from fMRI scans: predictive and generative tasks, different preprocessing pipelines and examined the effect of different MRI machines in the dataset.

1.1 Structure

In the second chapter, I will be introducing machine learning and neural network concepts important to my work, with a special focus on graph neural networks, introducing their key components and grouping them by structure and function. In the second half of the chapter I introduce medical data, legal and ethical questions surrounding it, the types of medical imaging data and give a bit of background on brain fMRI data.

In the third chapter I will introduce the technical background of my work: the programming language, deep learning frameworks and libraries. In the second half of the chapter

I go into the technical parts of working with fMRI data. This includes the used data formats, preprocessing measures and the libraries necessary for these tasks.

In the fourth chapter, I detail the various use-cases of graph networks in fMRI processing that I worked on. For each section I will introduce the corresponding dataset and detail the approach and methodology used in an attempt to solve the problem. I present the results of this approach and the conclusions that can be drawn from the experiment.

In the last chapter, I will be detailing possible ways to enhance the models' capabilities and directions this research could take in the future and giving a summary of my work.

Chapter 2

Background

2.1 Graph Neural Networks

Many real life problems can be best modelled with graphs and these representations can code a lot of information if assessed correctly. By developing techniques that work to extract this information and extrapolate from it effectively, we can build powerful systems that can predict, classify and advise based on graph data.

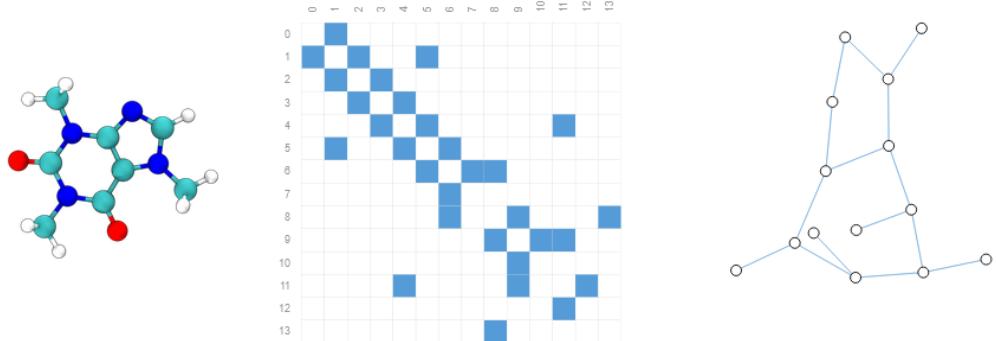


Figure 2.1: From left to right: 3D representation of a caffeine molecule, its adjacency matrix, its graph representation. [59]

Graph algorithms have a long history in mathematics and there is a rich variety of graph related problems that are best solved using these classical discrete mathematical methods: pathfinding algorithms, breadth or depth-first search algorithms are used prominently both in everyday IT applications and infrastructure, and also in research. But in certain cases these methods may not be well-equipped to handle the task at hand. With the large amount of data collected every day and the specialized tasks that need to be fulfilled there is reason to bring machine learning into graphs, and the discipline has exploded in popularity in the last five years or so [26].

2.1.1 Neural Networks

Neural networks are an especially interesting and useful area of machine learning and data analysis: as referenced in their name, they were created in resemblance of the human brain's structure, mimicking the interconnected neurons. In recent years, they became the flagship of AI research because of their ability to work on incredibly large datasets effectively and produce results never seen before.

The basis of a neural network are neurons which sum up incoming "signals" multiplied by weights and apply an activation function to their output. The network learns by updating these weights and biases based on the training data, to match the desired output to each piece of input. The activation function adds non-linearity to the model, making it capable of learning more complicated relationships in the data.

Neurons are typically organized into layers in a neural network and the models usually benefit from having a large number of layers: modern models for more complex tasks can get very 'deep' which lead to the popularization of the term deep learning when talking about such models. But since having more layers makes a model more computationally intensive, leading to higher costs and slower inference, researchers are often looking to prune models or find architectures that deliver similar results while having a lower parameter count [46].

Parameters in a model are updated through a process called backpropagation [57]. Here, for labeled training data, the error of the prediction of the model is calculated through the use of a loss function. This should be representative of how far off the model was from the ground truth and also easy to work with derivation-wise and numerically. This is then "propagated back" through the model by deriving what each of the model's parameters contributed to this loss (see Equation 2.1). This is what's known as the gradient and it is used to perform gradient descent: based on the resulting derivative we have a direction to move in to lower the loss.

$$\frac{\partial E}{\partial w_{i,j}} = \frac{\partial E}{\partial o_j} \frac{\partial o_j}{\partial net_j} \frac{\partial net_j}{\partial w_{i,j}}, \quad (2.1)$$

In case of larger datasets it is impossible to calculate the gradient for the entire dataset in one go, so a process called stochastic gradient descent [3] is used, where randomly selected subsets (batches) of the dataset are used to calculate it during training. Having a gradient, the other part needed is the learning rate: how much to change weight in the given direction. This can be considered a hyperparameter of the model, but modern optimizers (such as ADAM [40]) use adaptive learning rate optimization techniques to adjust learning rate during training to ensure a smooth and fast convergence.

From the most fundamental concepts of the original perceptron and the breakthrough idea of using backpropagation, neural networks became a widespread and varied phenomenon: there are architectures for different types of data optimized for different kinds of tasks; these models have proved to be applicable in almost any area [61].

A very active and widely used branch of neural networks are convolutional neural networks (CNNs) [45]. These are most commonly used on image data and work by convolving learnable kernels with the input, to extract features. This is useful for capturing spacial relationships allowing for better pattern detection while also greatly decreasing the number of trainable parameters compared to a fully connected feed-forward network. Attention-based networks [72] are another important model class that became very successful in recent times. Transformer architectures are used in many disciplines and are capable of state-of-the-art results. Both of these concepts turn up in graph neural networks as well.

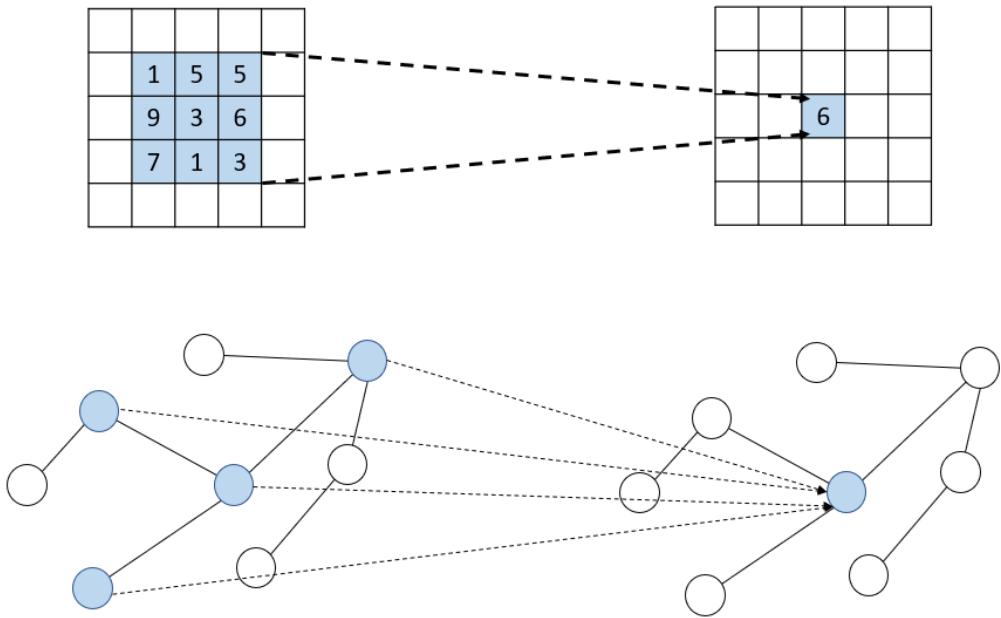


Figure 2.2: Convolution on images versus graphs.

2.1.2 Working on Graph Data

Graphs are a very versatile mathematical concept because of the way they lend themselves very neatly as a generalization. They have great expressive power when it comes to describing relations, groups and things with a rich inner structure. We can find graph-based datasets in a variety of domains: molecules in chemistry, interaction graphs in social sciences, knowledge graphs and computer networks [82]. The level of abstraction

they provide makes them suitable for use in many fields, as connections between items, molecules or people are vital in understanding complex natural systems.

One area that is particularly rich with such examples is biology and medicine [81]. Interactions between drugs, relationships between species and contact tracing in epidemiology are all important facets that can benefit from the ability to analyse graph datasets. We can also find interesting networks to analyse inside of organisms: the focus of this thesis is the analysis of network formed by brain regions, and the interactions that happen in the human brain.

Graphs are mathematically defined as a set of nodes (or vertices) and edges (or links) which run between two nodes. They can be categorized by their edge structure: directed graphs specifically code the information of a starting and end node, while undirected graphs do not. In terms of graph neural networks this is a very important distinction as it fundamentally alters the network's meaning. Certain types of networks are designed with a specific type of graph in mind and might need special normalisation.

Certain types of graphs with special properties can be especially interesting in certain areas. Trees (which contain no cycles) can be useful for describing containment hierarchies or sentence structures, for example. Based on semantic meaning we can also talk about heterogeneous graphs where nodes may represent entirely different things, for example purchase or interaction graphs in a recommender system where a vertex could be a user or an item to be purchased. In this case it is important to make this information available to the network.

A graph dataset often codes much more information than just the mathematical structure itself. Ideally some information is available for each node; a feature descriptor vector for each node supplies more information for the network and also helps identify the node it is attached to: since graphs are order agnostic, it may help to have identifying information about individual nodes.

The contents of such a descriptor vector must be domain specific. For example in case of a social network it would code information about a given person: age, gender, height. In case of a molecule prediction scenario it could be atomic weight or covalent radius. For the brain fMRI analysis case most relevant to this thesis, it could be information about a given brain region, average activation or even an activation time series [71]. In case there is no suitable information available for the node feature vectors it is common practice to use rows of the identity matrix: in absence of additional data this can be useful to serve as an identification tool. In some cases information might be available in other ways, for example as edge weights or edge feature vectors. These could be descriptive of the users' interactions in a social network example.

Node feature vectors can be concatenated together to form a matrix and together with the adjacency matrix (or edge weight matrix) of the graph they form a very neat representation of the graph. This makes the highly variable structure of the graph containable within a constant, well-structured manner that is crucial for machine learning applications. But since graphs can be used to represent complex structures and hierarchies, where this inner construction is more pronounced and relevant than in other cases, they require specialized methods when it comes to machine learning and neural networks. A lot of these methods take after image recognition or object detection solutions [80].

The reason for this is that while images do not have explicitly structured data in the same way as graphs do, their contents can be in complex spatial relations with each other. A computer vision model needs to take these into account when performing complex tasks such as segmentation or creating a full description of what is happening in a photo. Even for simple object detection, more complex objects have hierarchical structures that the model needs to 'understand' in some way.

2.1.2.1 Common tasks

The motivation behind working with graph datasets is of course to perform some sort of inference using the patterns extracted from the training data. Since these datasets are diverse both in semantics and in structure this could mean a lot of different types of tasks. In this section the most common of these will be described with examples and commonly used methods for solving them.

Most graph architectures work by assigning a vector to each node on their output. This is often called a node embedding: an n -dimensional vector is attached to the node that attempts to convey all available information about the node and its role in the graph structure, effectively embedding it in an n -dimensional space. Node embeddings are not strictly only generated in deep learning models, there are for example random-walk based methods for this purpose, such as node2vec [31].

In case of deep learning models the creation of node embedding is an iterative process as there is an updated vector attached to the node after every layer. In this process the original feature vectors can be thought of as a sort of 0th iteration of the embedding, purely representing information about the node itself. With more and more layers aggregating information on a larger and larger neighbourhood of a node, this slowly transforms into containing information of the structure of the network as well.

Node prediction. In this type of task the goal is to infer some information about individual nodes of the graph. This could be sorting into already existing categories (classification) or estimating a value related to the node (regression). Both types of tasks can work with datasets that are either graphs made up of labeled nodes that have a certain

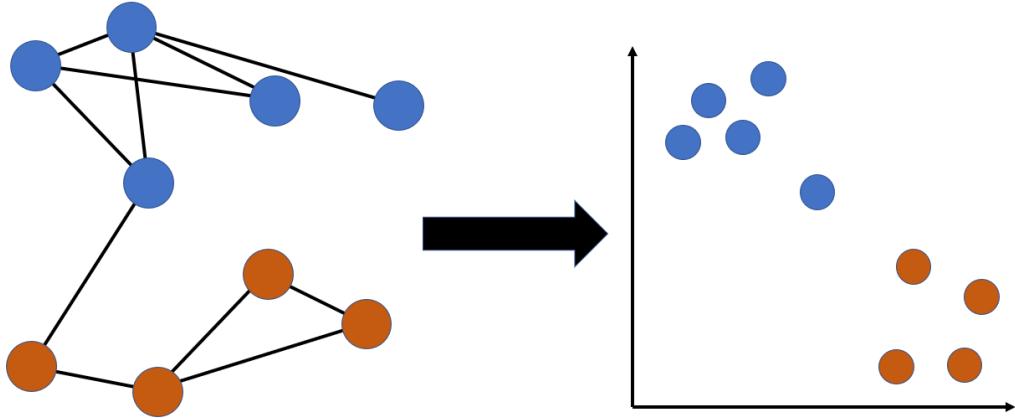


Figure 2.3: Visual representation of node embeddings in 2 dimensional space.

number of missing values or with training graphs that have fully labeled nodes, performing inference on graphs the model has not seen before. The second type of task is much more difficult, meaning models often performs significantly worse on such tasks, but oftentimes this reflects the real world scenarios and use-cases for models better.

Since the network output is representative of a node and its role in the structure in the first place, this is the most straightforward type of task to solve using a deep learning model. In case of node classification a commonly used technique is to match the dimensions of the final embedding to the number of possible classes and have each of the values represent probability of class membership: this can work for both multi-label and simple classification using different activation functions. Another technique is to use a small fully connected MLP on the embedding for prediction. This is also useful in case of regression and gives a bit more control over the final output of the model.

Node classification tasks are common in social datasets where they are used to make predictions about the characteristics of individuals. It also shows up in the biomedical domain, for example in protein feature prediction based on protein interactions [79]. Another common use-case is citation networks, where features of scientific publications can be predicted [47].

Graph prediction. In case of this task the goal is to predict some sort of information about the entire graph. Node embeddings can be aggregated in a lot of ways to predict graph classes or labels. These aggregation methods are commonly referred to as a ‘readout function’: the values attached to nodes are ‘read’ together to reveal information about the entire graph. More common, basic readout methods are similar to the simpler pooling

methods used in CNNs: node features can be averaged, summed or there can be a minimum or maximum choosing operation.

This is a common task in cases where a single thing is represented by one whole graph: molecules' property prediction based on structure [76] or, more relevantly here, predictions about a person based on a graph of their brain activity [74, 44].

Edge prediction. There are two main types of edge prediction: either, similarly to nodes, there is a specific feature or classification of edges we want to predict or the goal is to find edges in the graph that are not present in the structure right now but are most likely to exist; the second type is a very common task in recommender systems.

There are multiple commonly used techniques for edge prediction using node embeddings generated by a GNN that are capable of providing either continuous or binary results. These methods usually rely on using the information about a pair of nodes to determine their 'compatibility' or other feature of their shared edge [43].

This type of task is commonly solved in recommender systems for e-commerce or online content platforms [78], but friend recommendations in social networks are also very common [23].

Clustering, influence prediction. Clustering in this domain usually refers to clustering the nodes of the graph to find groups more 'similar' or 'closely connected'. There are a lot of classical solutions for this problem as well such as the Girvan-Newman algorithm [50] which optimises an edge-betweenness metric and the Louvain method [8] which uses modularity. In case a GNN is used in this task, it is done in an unsupervised or semi-supervised manner, where it is used to generate embeddings and then the embeddings are clustered using a common clustering ML solution such as K-means or DBSCAN [20]. Hybrid solutions are also possible: a neural network could be used to create edge weights that a classical algorithm could incorporate.

The most common application of this is community detection [9]. Influence prediction is also an unsupervised task that seeks to find a node or multiple nodes that have the most influential role in the graph, finding hotspots of activity [10].

2.1.2.2 Graph convolutions

The Laplacian of a graph is a square matrix $n \times n$ in size (where n is the number of nodes in the graph). It can be calculated from the adjacency matrix and diagonal node degree matrix:

$$L = D - A,$$

where L is the Laplacian, D is the diagonal node degree matrix and A is the adjacency matrix. The diagonal node degree matrix is similar to the identity matrix but in each row instead of the one we have the degree of the node corresponding to said row. The Laplacian can be used to build polynomials that can be used on node features [15].

$$p_w = w_0 I_n + w_1 L + w_2 L^2 + \cdots + w_d L^d = \sum_{i=0}^d w_i L$$

These polynomials can be thought of as analogues of 'filters' in CNNs, and the coefficients w_i as the weights of the 'filters'. There exists a type of network that is built directly on the concept of Laplacian polynomials: ChebNet [16] used Chebyshev polynomials and normalized Laplacians to be more numerically stable and 'stackable'.

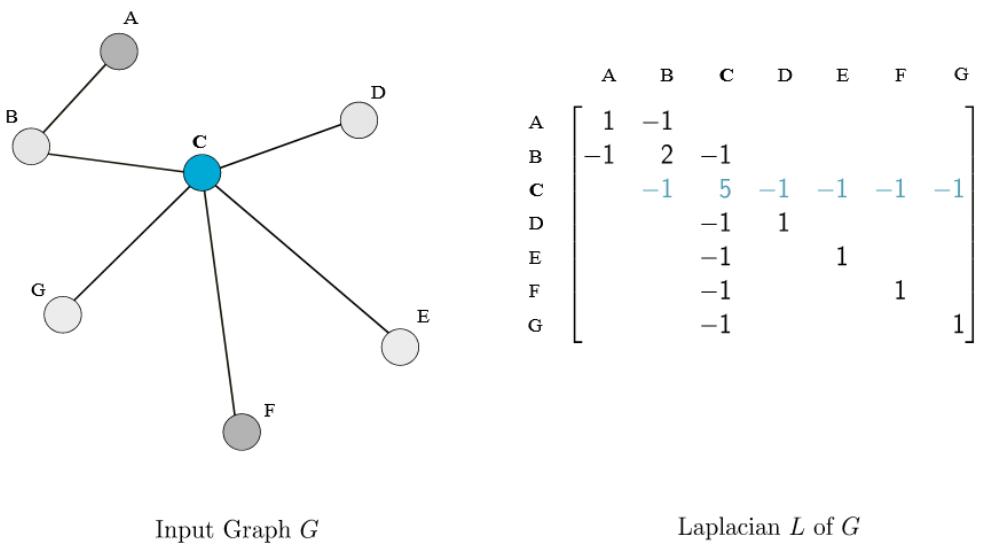


Figure 2.4: The Laplacian L for an undirected graph G , with the row corresponding to node C highlighted. Zeros in L are not displayed above. The Laplacian L depends only on the structure of the graph G , not on any node features. [15]

ChebNet was a big step in ushering research into graph networks, as it motivated many to think about graph convolutions from a different perspective. Current graph neural networks utilize graph convolutions in ways that make use of 'computational shortcuts': bypassing costly operations such as eigenvalue calculations can make a model much more scalable, both in terms of data throughput and model size.

These shortcuts often mean calculating local approximations of the graph convolutions. Local graph operations are often the so-called 'message passing' operations [28]: every node in the graph sends a 'message' to its neighbours based on its own data and then each node aggregates the received messages. In practice this means each node vector is updated in a layer based on its neighbours.

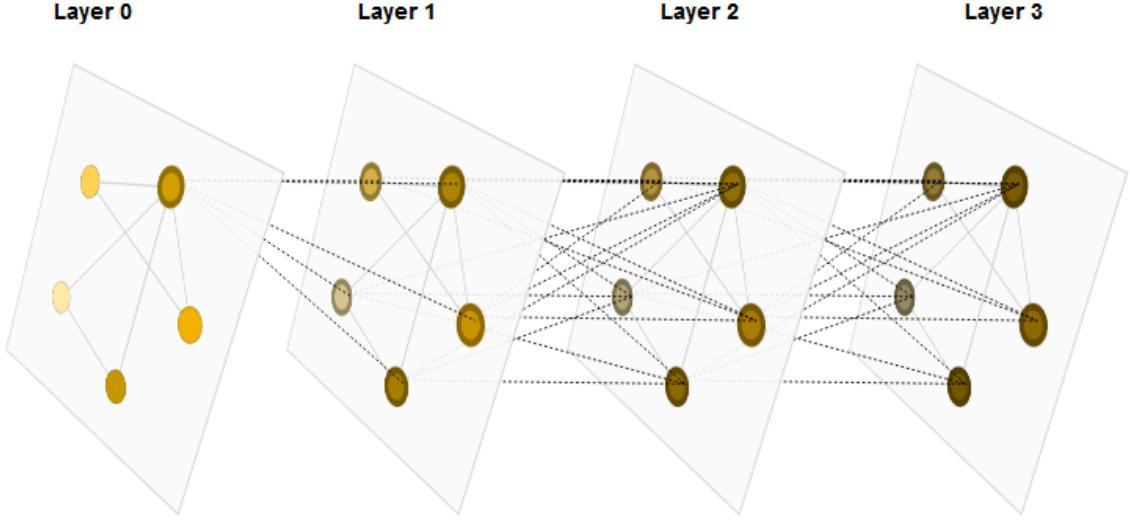


Figure 2.5: Visualization of a GNN accumulating information via message passing. [59]

2.1.3 Graph Neural Network Types

There are many architectures that fall under the umbrella of graph neural networks and they can be grouped by many different criteria with interesting similarities between them. In [82] the authors break down the most common parts of graph architectures: a way of **propagating** information between nodes, a way of **sampling** relevant nodes when graphs and neighbourhoods get large, and, in case it is needed for the task at hand, a way of **pooling** information to represent subgraphs or graphs.

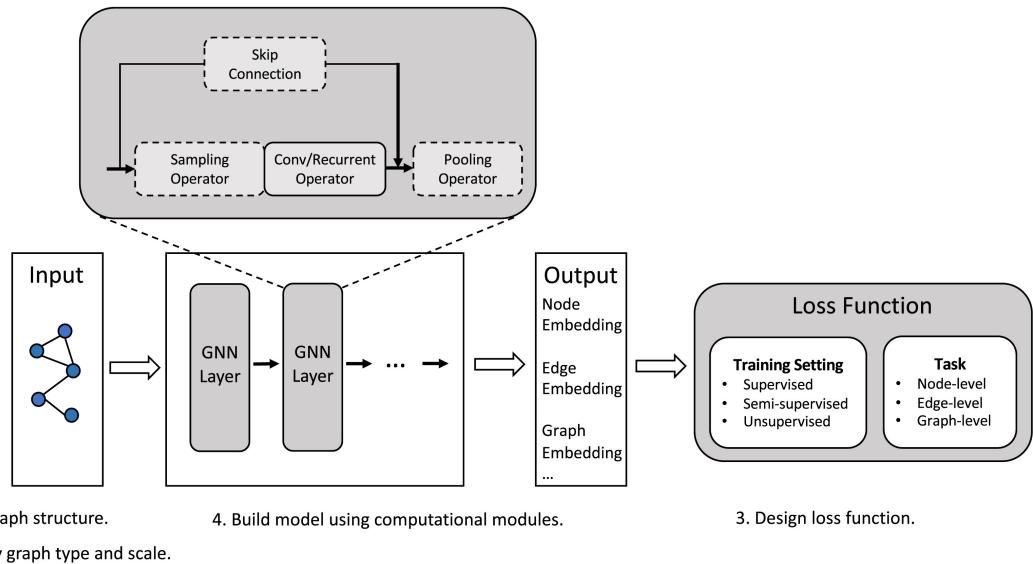


Figure 2.6: A generalized schematic of GNN architectures [82]

We can also differentiate architectures based on the type of graphs they operate on, the previously discussed task of the network (see Section 2.1.2.1) or the used learning type: supervised (using labeled data), semi-supervised (using a small amount of labeled data to guide training) and unsupervised (using unlabeled data to find patterns).

The most commonly used network types/layer types include GCNs [42] (Graph Convolutional Network), GATs [73] (Graph Attention Network), GraphSAGE [32] (Graph Sample and AGggregateE). These are all convolution-based solutions with different approaches to the implementation. GCNs try to stay true to spectral convolutions following ChebNet's footsteps, GATs leverage the very powerful concept of attention for convolutions and GraphSAGE focuses on building a good way to sample large neighbourhoods in large graphs.

In the realm of convolution based models MPNNs[28] (Message Passing Neural Network) can be seen as a generalization of the concept, extracting common features from other methods, serving as a framework that can instantiate any of those models. Other than convolutions another type of architecture that has been used in graph networks with success has been recurrent networks.

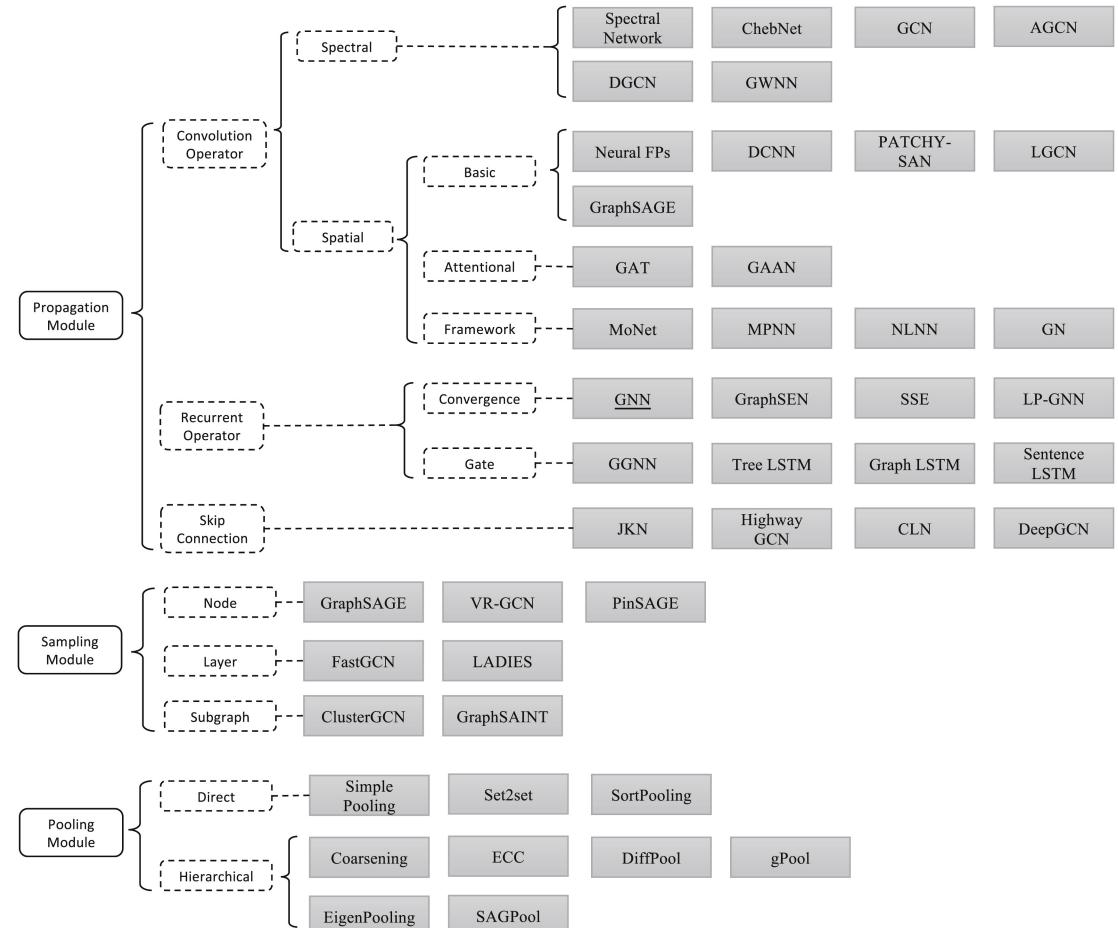


Figure 2.7: GNN architectures grouped by characteristics [82].

2.1.3.1 Graph Convolutional Networks

Graph Convolutional Networks [42] (GCNs) expand on the idea of CNNs, which are commonly used in image processing. Observing from a graph perspective, an image can be considered a very special case of a graph, where pixels are nodes and their neighbours can be determined from the grid. This type of network generalizes the concept of local convolutions from the image domain to general graphs.

$$H^{(l+1)} = f(H^{(l)}, A) = g\left(\hat{D}^{-\frac{1}{2}} \hat{A} \hat{D}^{-\frac{1}{2}} H^{(l)} W^{(l)}\right) \quad (2.2)$$

Equation 2.2 shows how the next layer's node feature vectors are calculated from the previous ones: A is the adjacency matrix, \hat{A} is the adjacency matrix with self-loops added in ($\hat{A} = A + I$), \hat{D} is the diagonal node degree matrix with the purpose of normalisation, W is the learnable weight matrix and g is some type of non-linear activation function, most commonly ReLU or leaky ReLU in practice.

This computation model is a nice analogue for convolutions on images, and approximates spectral convolutions while remaining computationally stable and non-resource intensive. Layers can also be stacked nicely to allow information to travel further than the immediate neighbourhood of a node.

2.1.3.2 Graph Attention Networks

Graph Attention Networks (GATs) leverage masked self-attention layers to address problems present in earlier architectures of graph convolutional networks [73]. By using attention to implicitly assign different weights to nodes in a neighbourhood, the model can gain a deeper understanding of the graph structure without the use of costly matrix operations.

In a layer of a GAT architecture input node features ($h \in \mathbb{R}^{N \times F}$, where N is the number of nodes and F is the number of features) are turned into output node features by first applying a linear transformation via a weight matrix ($W \in \mathbb{R}^{F' \times F}$) and then performing self-attention (a) on the nodes. The full equation is as follows:

$$h'_i = \sigma \left(\sum_{j \in \mathcal{N}_i} \text{softmax}(a(Wh_i, Wh_j)) Wh_j \right).$$

Where σ is the non-linear activation function and h_i and h_j refer to the node features belonging to node i and j . Self-attention coefficients $e_{ij} = a(Wh_i, Wh_j)$ indicate the importance of node j 's features to node i . If the model would allow every node to attend every node, that would mean dropping all structural information. To avoid this, attention

is masked: e_{ij} is only computed for nodes j that are in the neighbourhood of node i ($j \in \mathcal{N}_i$). This is generally the first-order neighbourhood, but it can be parametrized.

The attention mechanism is a feedforward neural network, parametrized by the concatenated transformed node features of the two nodes. The calculated coefficients are then normalized across the choices of j using the softmax function. Once obtained, the normalized coefficients are used to compute the linear combination of the features corresponding to them, resulting in the final output of the layer (after applying an activation function). Multi-headed attention is commonly used to stabilize the learning process.

2.1.3.3 GraphSAGE

GraphSAGE (graph SAmpLE and Aggregate) was created as a general inductive framework to leverage node feature information and efficiently generate node embeddings for previously unseen data [32]. Each node of the graph is represented as some aggregation of its neighbours, therefore, even if a new unseen node is encountered, it can be represented as some aggregation of its neighbourhood.

GraphSAGE aggregates information from the neighbours of nodes similarly to other architectures, but also takes a step to sample this neighbourhood: a fixed-size set gets uniformly sampled from the available neighbourhood. This is resampled at each iteration when calculating embeddings. Weight matrices are learned using gradient descent based on random walks: the graph-based loss function encourages nearby nodes to have similar representations, while enforcing that the representations of disparate nodes are highly distinct.

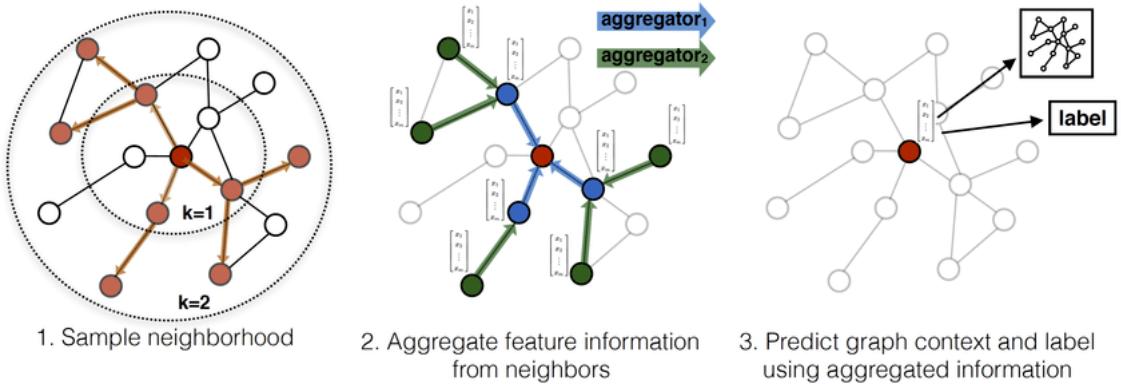


Figure 2.8: Visual illustration of the GraphSAGE sample and aggregate approach [32].

The aggregation method from the neighbourhood is an important feature in a GraphSAGE model, a symmetrical function is most desired, since node neighbours do not have a set ordering. Based on the used aggregation function, GraphSAGE can be very similar to GCNs: the mean aggregation method results in the convolution propagation rule of GCNs.

2.2 Medical data

We can call anything medical data that relates to human healthcare in one way or another. This most commonly means data collected by healthcare institutions about patients using sensor equipment, but it could also be information about a patient's habits or general environment, data from drug test trials or even location data in case of epidemiological contact tracing.

2.2.1 Legal and ethical concerns

This kind of data is often very personal and regarded as sensitive data: people are entitled to equal treatment regardless of medical status and as such information about the health status of an individual is protected. Many patients do not want employers or other third parties who could use such information against them to obtain details about their conditions and treatment.

For this reason doctors are required to not give out patient information unless it is strictly necessary and/or they have the informed consent of the person. This puts medical researchers in an interesting position; medical researchers using machine learning even more so. Data collected specifically for experiments where subjects can consent to their data being used is a straightforward situation, but there is data collected every day, worldwide in hospitals that could be very useful for furthering medicine, but doing so poses data privacy concerns. This is especially important in case of machine learning and neural network research as these disciplines require a very large amount of data that is very hard to collect using only organized experiments.

Developments in IT have lead to concerns about the effectiveness of data protection laws and in the European Union, for the sake of a more consistent and comprehensive protection of personal data, a General Data Protection Regulation (GDPR) has been enacted [21]. There are fears that the combination of strict requirements and very limited research exemptions will stifle and severely restrict medical research, and that the current consent or anonymise approach is not suitable for data-intensive research [49].

On the consent side of the equation, there is great difficulty in obtaining consent for personal data to be available for linkage, re-use and analysis for largely undetermined future research purposes [4]. It is a question whether meaningful or freely given and informed consent can be given at the time of data collection as it may not be possible to foresee or comprehend the possible consequences of consenting: a person consenting to a DNA sample has no way of knowing what could be determined from their sample in 10 years time.

Since personal data is defined as data that can directly or indirectly identify the individual [11], the act of anonymisation seeks to remove identifiers that link datapoints to specific individuals to make the data suitable for use in research without infringing any of the patients' rights. A simple deletion of the name and address of the subject is usually not enough, as other characteristics can be enough to identify an individual. A commonly used technique is pseudonymisation where identifiers are removed and replaced with a unique keycode, but this is not truly anonymous as someone with the requisite key can link it back to the individual [56].

Anonymisation requires extensive stripping of datasets and makes data linkage and update impossible in almost all scenarios, which makes datasets much less useful for research networks and projects. In addition to this, even in anonymised datasets there can be information that could negatively affect groups and lead to discrimination and stigmatisation [48].

There are currently new EU regulations underway for the construction of a European Health Data Space Regulation (EHDS), which aims to establish a common framework for the use and exchange of health data across the EU [22]. This seeks to empower individuals to be in charge of their medical data and to enable secure and trustworthy reuse for research and innovation.

2.2.2 Types of medical data

We can organize most medical/healthcare data into 7 categories that have differing use-cases and are of different interest to researchers [1]. Electronic Health Records (EHR) encompass the digital version of a patient's chart in a doctor's office, containing medical and treatment histories, but take the concept a step further offering a comprehensive view of patient health. Administrative data is similarly collected at clinical offices, detailing admissions, home care and prescriptions.

Claims data refers to the data that is transferred to the insurance provider (public or private) for coverage and compensation. This contains information about procedures and tests as well as costs. Patient/disease registries help collect secondary data associated specifically with patients who share a particular diagnosis. These are particularly important for chronic conditions such as cancer or diabetes.

Health surveys are conducted in the general population to assess the health of the population and estimate the prevalence of certain diseases. Clinical trial data refers to data collected in experiments and studies specifically for the purpose of research. Genomic data involves the molecular sequence in an organism's genes, the role of each gene, the regulatory elements managing gene expression, and the connections among various genes.

All of these types of data have their place in healthcare research and they occasionally overlap in their subjects, but they generally fall under different providers and legal protections. The data types most commonly used in machine learning and big data applications are EHRs, aggregated clinical trial data, administrative data, genomic data [49] and medical imaging data [68].

2.2.3 Medical imaging

Non-invasive medical imaging techniques, such as MRI, X-ray, and CT scans, are essential tools in medicine, allowing physicians to understand the internal structures and functions of the body without the need for surgery. These techniques are invaluable for their ability to produce high-resolution images of different types of tissues, making it ideal for assessing the state and function of internal organs [39].

These techniques vary by advantages, disadvantages, risks, image quality, safety, costs and applications. Choosing the right technique for the examination of a particular patient or for data collection in an experiment is a delicate balancing act that requires a physician well-versed in the particular niche.

The basic concept of a medical imaging system consists of a source of energy that can penetrate the human body and as the energy passes through the body it is absorbed or deflected at varying levels according to the density and structure of tissues and organs. The energy is then detected by special detectors compatible with the energy source and the signals are then converted into images using the correct algorithm matching the system. Techniques are often categorized based on the type of energy passing through the body.

The earliest discovered medical imaging tool was the X-ray, discovered by Röntgen in 1895, followed by the CT imaging technique in 1963 [39]. Research into the MRI started in the 1970s and the first prototypes were tested in 1980 [39].

2.2.4 MRI and fMRI imaging

Magnetic Resonance Imaging (MRI) is a diagnostic technology that uses magnetic and radio frequency fields to image tissues [5]. It produces high-fidelity images, while using no ionizing radiation, making its repeated use in patients possible.

The procedure can be performed without contrast (no possible allergy problems) and is painless, but due to the long scan time and the design of machines it can induce claustrophobia and young children might require sedation to remain still. Comparatively to other imaging techniques, MRI is also rather expensive, which is an important factor in deep learning applications where a large volume of data is needed.

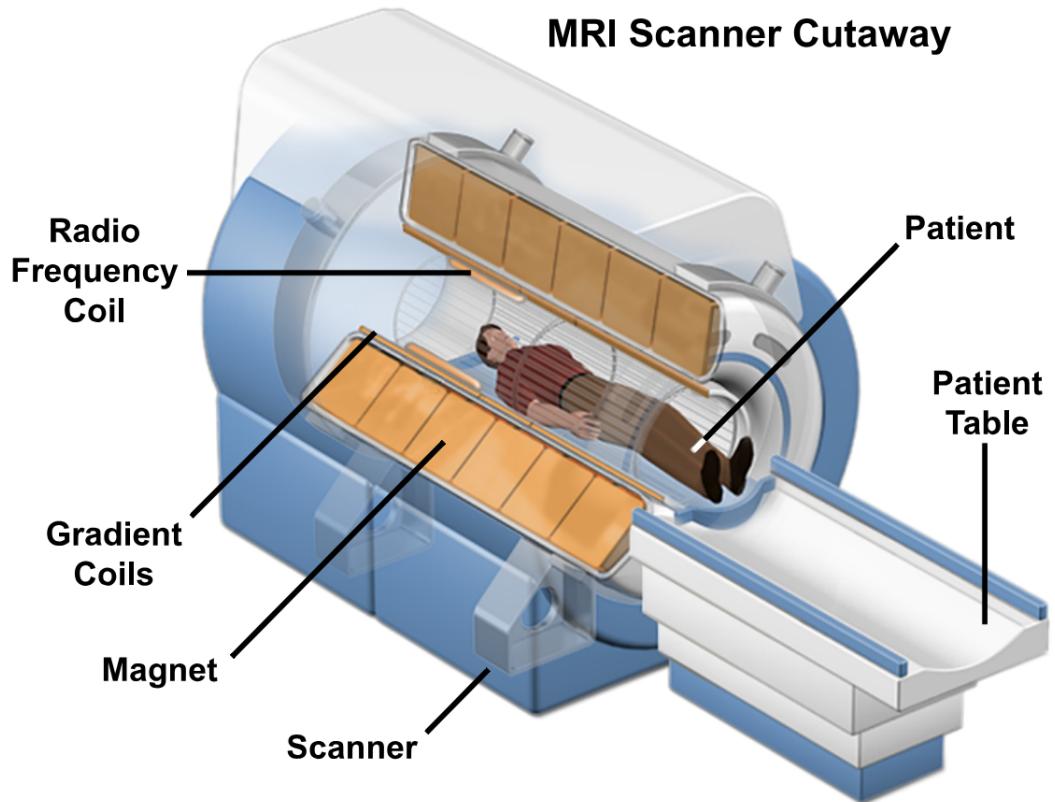


Figure 2.9: Schematic of an MRI machine [12].

MRI is well suited for examining abdominal organs, such as the liver, joints and finding tumours, cysts and other abnormalities and unhealthy tissue in various parts of the body. It is also good for providing a view of the cranium and examining the brain and the spinal cord.

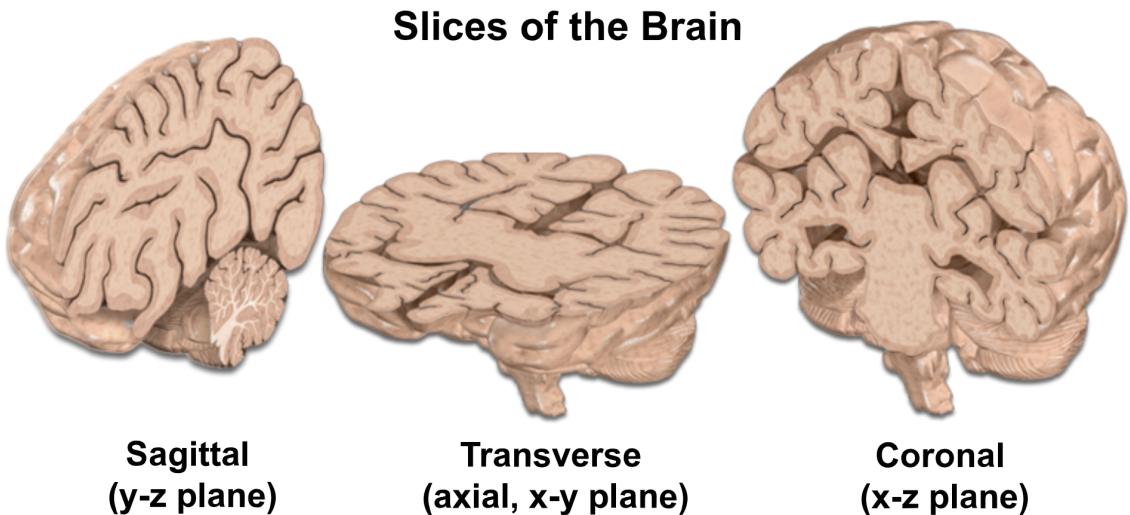


Figure 2.10: Types of slices in brain MRIs [12].

Functional MRI (fMRI) extends the capabilities of MRI by measuring brain activity, via a method called Blood Oxygen Level Dependant Constant [30]. This technique works

because of the haemodynamic response: when a part of the brain is activated, blood flows to the region (after an initial dip). The surge of oxygenated blood increases the ratio of oxygenated haemoglobin.

Oxygenated haemoglobin has different magnetic properties than deoxygenated haemoglobin, which causes measurable changes in the magnetic field, thus changing the readings of the MRI in the activated region, thus allowing detection of neural activity. The ability to capture neural activity in a non-invasive way is incredibly valuable in both clinical and research settings.

There are two main types of fMRI scans: resting state and task-based fMRI. Resting state is self explanatory: patients are asked to remain still and relax while the scan is taken. For task-based scans, patients are given some sort of stimulus: either a task to complete, such as imagining moving a part of their body (for examining motor function activity) or holding a conversation, or simply being shown an image or text.

2.2.4.1 Clinical significance

MRI findings are commonly used to both confirm initial diagnoses and to investigate further, when physicians cannot find a diagnosis in initial examinations. According to an observational study [36], where clinicians filled out questionnaires before the MRI stating initial diagnoses, diagnostic confidence and treatment plans and a similar questionnaire after imaging, in the light of MRI findings. In seven of nine diagnostic groups MRI findings were associated with a diagnostic impact. Diagnoses were revised or discarded following normal MRI findings and diagnostic confidence was increased by confirmative MRI findings. In five of nine diagnostic groups MRI findings had a clear impact on treatment plans.

2.2.4.2 Technical background

MRI technology builds on the physical phenomenon known as Nuclear Magnetic Resonance. Certain atomic nuclei, when exposed to a strong magnetic field, align themselves with this field. The hydrogen atom also shares this property and is abundant in all parts of the human body.

The aligned nuclei can then be perturbed by a radio frequency signal, which causes them to absorb the energy of the signal and go out of alignment with the original field. The nuclei then start to precess, during which they emit the absorbed energy, which is unique in different types of tissues in the body [53].

MRI machines provide a strong, constant magnetic field, then send short, localized pulses of radio frequency signals to cause the misalignment. As these atoms realign themselves

the machine uses the signals emitted to construct a three dimensional image of the area, comprised of individual cuboid elements, called voxels, based on the measurements.

The number and size of voxels reflects the quality of the scan and high-resolution images permit the detailed analysis of brain responses. fMRIs recorded this way reflect neural activity: the presented task increases local neuron activity, which in turn increases metabolic rate, which also means increased blood flow, which translates into an increased MRI signal. The time it takes for this process to occur also means that the changes in the fMRI are delayed compared to the stimuli of the patient by 1-2 seconds [18].

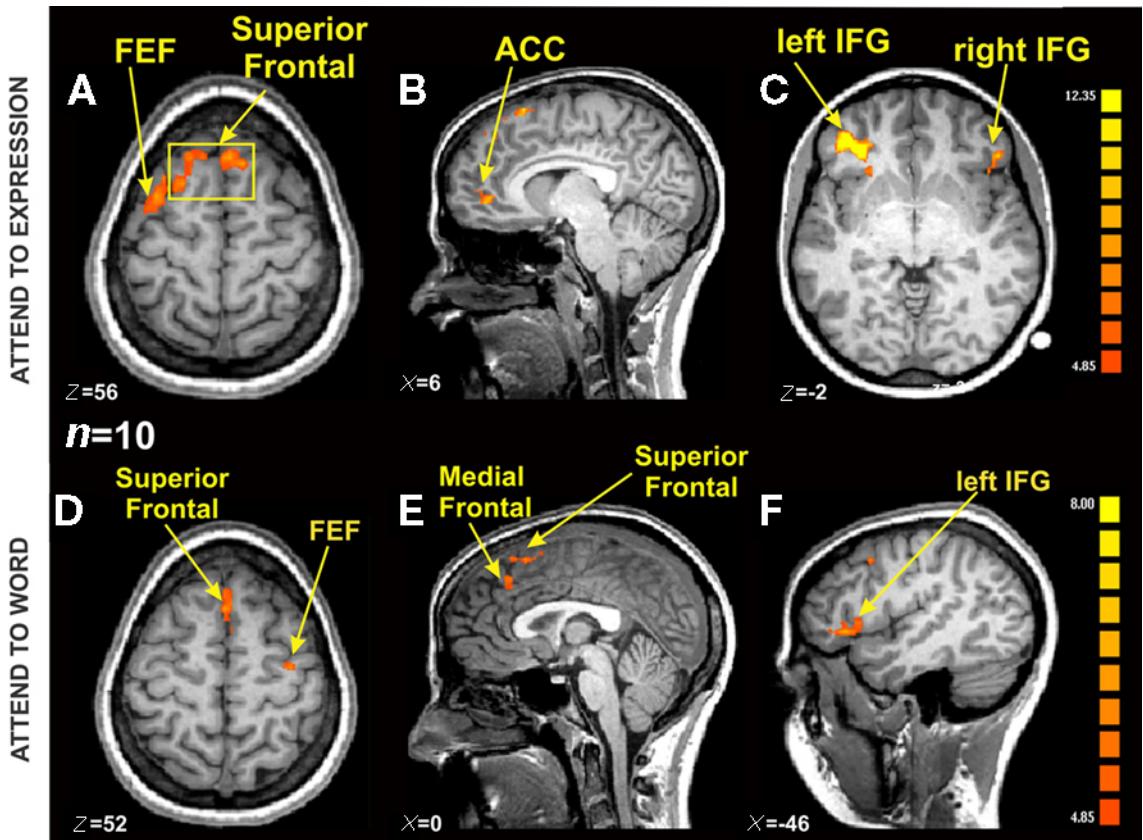


Figure 2.11: Example of an fMRI scan showing different activations on expressions and speech [52].

2.2.4.3 Processing

An fMRI dataset from a single session can either be thought of as t volumes, one taken every few seconds, or as v voxels, each with an associated time series of t time points [63]. This raw four-dimensional data needs to be thoroughly preprocessed as a first step, before any meaningful conclusions can be drawn from it. The beginning of this procedure is very similar to other image or medical imaging preprocessing methods. There is however no consensus currently on the best methods for preprocessing fMRI data, so picking a

strategy and parameters often works on a case-by-case basis, there are a few well-known pipelines that are available and commonly experimented with [66].

Despite their differences, most pipelines accomplish most of the same few tasks. Slice-timing correction makes sure that when viewing a volume, all of the voxels have been acquired at the same time, motion correction is applied to make sure each volume is aligned with every other one. Smoothing methods are applied to reduce noise and artifacts. Some methods also try to reduce variance in the scans by aligning them to a particular template.

2.2.4.4 ROIs and connectomes

For better analysis of the functional brain scans, it is important to have a certain understanding of the functional areas of the brain to help understand the significance of certain regions firing. Brain atlases are sophisticated templates created by medical researchers as the average of many brains, giving a common coordinate system and also important information about the region at every voxel. With these atlases brain scans can be segmented into different functional centers which are commonly referred to as Regions of Interest (ROIs).

If the segmentation is completed for an entire scan, the average activation of these regions can be computed and used to draw conclusions about how this brain region works normally or how it participates in a given task. Even more information can be extracted if we are able to analyse how the regions interact with each other [54].

The correlations between the fluctuations of activation in regions can reveal information about how these segments of the brain interact. These correlations change depending on the subject or the performed task, showing that regions interact differently under different conditions. Several methods have been proposed for assessing connectivity, most commonly using the time series of ROI activations to calculate covariance or correlation [71].

The goal is to yield a complete matrix of connectivity values between the regions, which can be thought of as a sort of weighted brain graph describing how connected these regions are, often referred to as connectomes. A typical brain connectome is comprised of nodes for each brain ROI and edge weights that denote the connectivity strength between two ROIs [7].

Chapter 3

Technical background

In my work I used the Python programming language both for implementing machine learning solutions and for creating pipelines; downloading, sorting and preparing data, as well as various scripting needs. For running deep learning tasks I have employed the Google Colab¹ environment, an online compute platform that lets users run Jupyter notebooks with limitations, one of my own laptops, equipped with an Nvidia GTX 1050Ti portable, and for the particularly resource intensive diffusion task (See Section 4.2) the project received help from a corporate compute cluster.

In this chapter I detail the frameworks and packages used in my work, as well as the technical aspects of using and preprocessing fMRI data.

3.1 Used Frameworks and Packages

For machine learning tasks I have employed various libraries commonly used for such tasks to avoid reimplementing standard solutions, both for creating the different models for comparison and for data exploration and visualization purposes.

For the purpose of organizing the used Python packages and ensuring a consistent environment I have used a various virtual environment management tools in my work: venv, uv and conda. In these types of projects package management is a common painpoint. Since machine learning is a very active field with new discoveries and optimizations packages are updated quite often. Backwards compatibility is not guaranteed usually and issues of certain packages needing another package in a certain version, but not being labeled accurately come up commonly.

Even when all needed package versions are documented correctly in a project, there can still be problems for people trying to recreate the results. Older package versions of

¹<https://colab.research.google.com/>, accessed: 2025-05-29

ten become unsupported and unavailable, and package managers might disallow certain combinations of installations.

3.1.1 Main Deep Learning Framework

The main base deep learning framework I have used in all of my work is Pytorch². Pytorch is one of the most popular foundational machine learning libraries (alongside TensorFlow³) used for a wide variety of applications, such as computer vision or NLP.

At it's core is the provided tensor computing capabilities and an automatic differentiation system, which enables backpropagation. Moreover, Pytorch also has a neural networks module with a lot of widely used lower level concepts: abstract model class, basic layers (e. g.: fully connected, convolutional), optimizers.

3.1.2 Graph Neural Network Frameworks

It is completely possible to implement graph neural network layers and complete graph networks only using the Pytorch library and I have also done so, but there are also frameworks built to enable users to create these networks more efficiently. These libraries have implementations of commonly used GNN layers (such as GCN or GAT) and usually have their own format for graph storage. This can speed up model creation greatly but working purely from Pytorch offers the greatest flexibility.

I have tried two specific libraries in my work: DGL⁴ and PyG⁵. DGL (Deep Graph Library) is a framework agnostic and scalable solution. It represents graphs with a DGLGraph object which stores all graphs as directed using an edge list. It is also possible to add both edge and node features.

PyG (Pytorch Geometric) was specifically created to enhance Pytorch with better GNN capabilities. It aims to follow the design principles of vanilla Pytorch and provide a good interface to both researchers and first-time users.

3.1.3 Other Machine Learning Tools

Sklearn⁶ (Scikit Learn) is a comprehensive library of machine learning solutions, offering a wide variety of models for classification, regression, clustering and dimensionality reduction, as well as model selection and preprocessing tools.

²<https://pytorch.org/>, accessed: 2025-05-29

³<https://www.tensorflow.org/>, accessed: 2025-05-29

⁴<https://www.dgl.ai/>, accessed: 2025-05-29

⁵<https://pytorch-geometric.readthedocs.io/en/latest/>, accessed: 2025-05-29

⁶<https://scikit-learn.org/stable/>, accessed: 2025-05-29

XGBoost⁷ (eXtreme Gradient Boosting) is a gradient boosting library that provides a well optimized and very performant gradient boosting machine implementation.

Denoising-diffusion-pytorch⁸ is a package commonly used for experimenting and training diffusion models. It provides an easy interface to train diffusion models with custom networks and enables easy DDIM sampling.

Weights & Biases⁹ (WandB) is a library designed to track machine learning work. By integrating WandB into a project the results of a run can be tracked and hyperparameter sweeps can be ran. Experiments can be tracked through a dashboard with a wide array of functionalities and sweeps can be configured completely freely to automatically run experiments and find optimal hyperparameter tunings.

3.1.4 Supporting packages

Numpy¹⁰ is a commonly used computing framework with great features regarding N-dimensional arrays. Most machine learning use it to some degree and are compatible with its formats.

Matplotlib¹¹ is a powerful plotting and visualisation library for Python. It helps in creating easy to understand and colourful graphs from data.

Nibabel¹² and nilearn¹³ are specifically neuroimaging libraries: the first one for handling the associated file formats and providing access to neuroimages and the latter for enabling analysis of brain volumes.

Tsfresh¹⁴ is a tool designed to calculate time series characteristics. It also contains methods to evaluate the importance of these features for regression and classification tasks.

3.2 Using fMRI data

fMRIs produce a very special four dimensional data mass, that is not easily readable by default in most cases. Fortunately there have been specific data formats and libraries created to enable researchers to transport, view and extract information from these scans.

⁷https://xgboost.readthedocs.io/en/release_3.0.0/, accessed: 2025-05-29

⁸<https://github.com/lucidrains/denoising-diffusion-pytorch/tree/main>, accessed: 2025-05-29

⁹<https://wandb.ai/>, accessed: 2029-05-29

¹⁰<https://numpy.org/>, accessed: 2025-05-29

¹¹<https://matplotlib.org/>, accessed: 2025-05-29

¹²<https://nipy.org/nibabel/>, accessed: 2025-05-29

¹³<https://nilearn.github.io/stable/index.html>, accessed: 2025-05-29

¹⁴<https://tsfresh.readthedocs.io/en/latest/>, accessed: 20025-05-29

3.2.1 Data format

NIFTI (Neuroimaging Informatics Technology Initiative) is a file format created to store brain imaging data [77]. It was agreed upon as a replacement to the previous ANALYZE format, which was widespread but had problems: the main issue, that NIFTI aimed to solve was that there was not adequate information about orientation.

Instead of separate files for meta-information and the actual image (like in ANALYZE), this format allows storage as a single .nii file. Since there are very often large areas of a single colour or masked sections, these files can be compressed with great results. Most available datasets are downloadable as nii.gz files.

The first three dimensions are reserved for spatial dimensions (x , y and z) and the fourth one is for time, t . The remaining dimensions from fifth to seventh can be used to store other information. The fifth dimension, however, can still have some predefined uses, such as to store voxel-specific distributional parameters or to hold vector-based data.

The first 348 bytes (or for the later/current version NIFTI-2 500 bytes) of the file are reserved for the metadata header which is full of information on how to interpret the actual data in the file. Great care has been taken to make ANALYZE, NIFTI-1 and NIFTI-2 as compatible with each other as possible, while advancing the user experience and representation capabilities of the file format. Nibabel is compatible with all of these formats, providing an easy way to access this data from Python.

3.2.2 Extracting ROI Values and Connectomes

Using the nilearn package it is possible to load brain atlases and use them to segment the brain scan loaded from a NIFTI file using nibabel. It is also possible to create masks and use only certain parts of the brain/atlas. This allows for the creation of average ROI activation time series for each region. This means we end up with a tensor that is number of patients \times timesteps \times number of ROIs in size.

Nilearn also has tools to perform the approximation of connectomes from these time series. The ConnectivityMeasure class is capable of calculating various types of functional connectivity matrices on multiple subjects. It can be parameterized with a covariance estimator object for more control over the resulting connectomes.

Chapter 4

Methods and Implementation

4.1 Graph Neural Networks on a Group FC matrix

4.1.1 Dataset

The ABIDE (Autism Brain Imaging Data Exchange) dataset [19] focuses on the furthering research into ASD (Autism Spectrum Disorder) through neuro-imaging and neuroscience. The dataset contains in total 1112 resting state fMRIs of subjects both with ASD (539 individuals) and typically developing controls (573 individuals).

The imaging data has been collected by 16 institutions, who collaborated to create a publicly accessible anonymised dataset to allow the broader scientific community to take part in ASD related research while preserving the privacy of the participants. As such, the dataset does not contain any protected health information.

Since fMRI preprocessing techniques are very complex image processing operations and heavily resource intensive, an initiative also formed to provide a preprocessed version of the dataset. At this point I was not comfortable with taking on the connectome extraction process and so following [74], their already preprocessed group FC matrix was utilized.

This matrix was extracted by first using the Craddock200 atlas to receive the ROI time series and using their average correlation to construct the matrix. After this a k -NN graph was constructed, keeping only the top k edges for any given node, resulting in a sparser graph and creating a hyperparameter that can tune graph density.

In their work they used this same graph for all subjects and all time frames, created only using training data. For the node features they used the activation levels corresponding to the ROIs and built a complex model that works with the temporally changing nature of the fMRI.

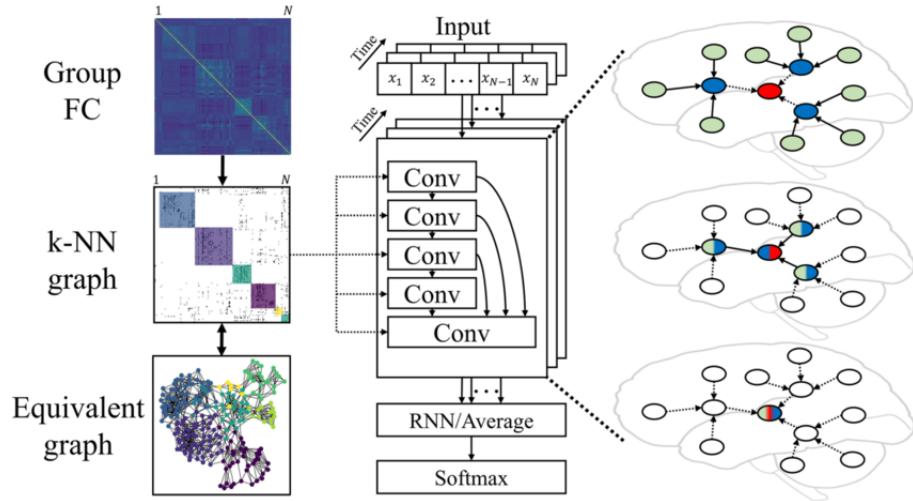


Figure 4.1: The used model architecture in [74].

The very close ratio of affected vs control subjects (539 to 573) means the dataset is well-balanced in terms of a classification task on the diagnosis of subjects.

4.1.2 Methodology

In the paper [74] by Wang et al., researchers worked with the ABIDE dataset to create a graph convolutional network based architecture for predicting the diagnosis of a patient based on the resting state fMRI. They also used the same architecture to perform a different task on a different dataset: identification of individuals based on the rs-fMRI.

In their model, convolutional layers were used to extract information from the neighbouring nodes with skip connections. Outputs from these convolutional layers were followed by an RNN (Recurrent Neural Network) layer to combine spatial representations from all frames.

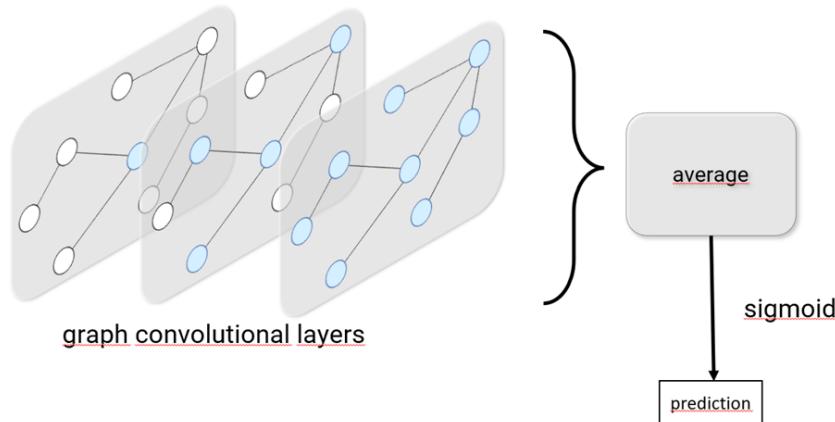


Figure 4.2: My own model for this task.

In my model, as seen on Figure 4.2, I strived to examine how a much simpler model would fare on this dataset compared to the research paper's. I simply used 3 GCN layers with ReLU activation with an average pooling readout function and sigmoid activation at the end for the binary classification task. To be able to compare the results to the ones presented in the paper I also performed a 10-fold crossvalidation.

4.1.3 Results

As it is visible on Figure 4.3, my model was less performant than the one presented in the paper. This was to be expected due to the simpler construction, but this model comes closer than expected, despite being much simpler and less computationally taxing and not considering timesteps. This is somewhat in line with findings in [58], where it was found that dynamic datasets (although in that case separate graphs were constructed for time windows) do not offer better performance.

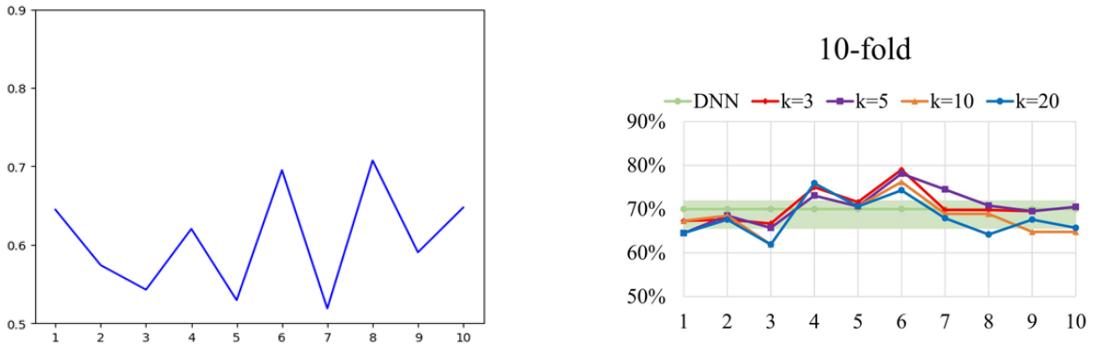


Figure 4.3: Tenfold crossvalidation results from my model (left) and the cGCN model from [74] (right).

4.2 Connectome-based diffusion

Deep learning methods are generally known for needing a large number of samples to produce the outstanding results they are capable of [2]. When working with naturally occurring data such as text and general images this is not quite as large of an issue, however medical imaging data, such MRI and CT scans, poses a problem, as it is much harder to access.

In case of fMRI scans this is partly due to the cost associated with the machines and taking scans, and the cost of the many hours of work by experts which is required for creating high-quality datasets. Furthermore, accessing these datasets is difficult due to privacy concerns and the sensitive nature of said data (See Section 2.2.1).

Data augmentation techniques can be used to counteract these hurdles. An important one of these is data synthesis; creating entirely new samples based previously collected data.

The goal of this project was to generate synthetic three dimensional fMRI brain scans. By employing diffusion generative models it is possible to produce high-fidelity images based on previous scans.

Generative models are capable of generating novel samples that they have not seen previously, based on the distribution learned from the training samples. The two types of generative models used in this architecture are Variational Autoencoders (VAEs) and diffusion models.

Variational Autoencoders combine deep neural networks with probabilistic modelling for image generation [41]. Instead of learning deterministic encoding, like standard autoencoders [6], VAEs learn a probabilistic representation, where the latent space follows a distribution which allows for controlled sampling and generation.

A VAE is made up of two parts, an encoder and a decoder. The encoder usually consists of a series of convolutional and pooling layers to downsample the data to create a bottleneck. From this latent space, the decoder works with convolutional and unpooling (upsampling) layers to try and reconstruct the input.

Once trained, the model is capable of generating new samples by sampling from the prior distribution of the latent variable $p(z) = \mathcal{N}(0, I)$ and passing in the sample at the bottleneck to the decoder part of the model.

Denoising Diffusion Probabilistic Models take inspiration from thermodynamics (particularly Langevin dynamics [75]) and instead of generating a sample in one step, they use a process of iterative noising and denoising [35]. These models have become state of the art for various image synthesis tasks, producing high-fidelity results. This does come at a cost though: these models are very computationally expensive, even with further improvements, such as DDIMs (Denoising Diffusion Implicit Models [64]).

During the forward noising process noise is sampled from a Gaussian distribution and added to the input image according to a noise scheduler. This process is repeated usually hundreds of times as the image shifts to a pure noise version (see Figure 4.4 for examples) as the time t runs from 0 to T .

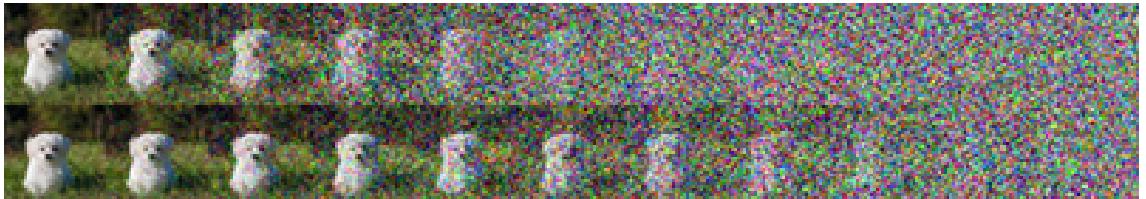


Figure 4.4: Linear (top) and cosine (bottom) noise schedulers being applied to an image of a small white dog [51].

The backwards process is the task of iteratively removing noise from the destroyed image, starting from random noise. This does not have an exact solution due to information loss,

so it is approximated by a neural network. The network receives the current timestep in the noising process t and the image, aiming to calculate the noise, so it can be subtracted.

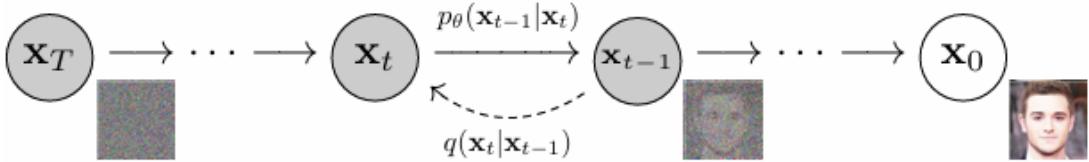


Figure 4.5: The forward and backward process of diffusion [35].

Since generating in the original fMRI domain is incredibly resource intensive as the data is always at least 3 dimensional, only very small resolution samples could be generated within reasonable computation demands. The idea was to take the generative part of the model to a different space, by using a more compact representation of the brain scan: the average activation values of brain regions (ROIs) and the generated connectomes seemed like a good fit for this purpose.

4.2.1 Dataset

The HCP dataset (Human Connectome Project) was collected by WU-Minn HCP consortium. The goal of the human connectome project is to facilitate research into mapping out the human brain; for this reason the project collected high-quality neuroimaging data (MRI, PET and MEG scans) in various studies targeting a variety of problems: Alzheimer's disease, normal brain development or ageing. In doing so, they have not only supplied high quality data for neurological researchers, but also established a protocol for such data collection [70].

One of the collected datasets is HCP Young Adult which has scans of over 1100 healthy young adults for the free use of researchers. The dataset encompasses both resting state and task-based MRIs and was collected by 10 institutions in different parts of the world.

This work focuses on the task-based portion of the dataset, where subjects were instructed to imagine moving various parts of their body while the fMRI took place, creating 5 classes in the data based on the body part - tongue, left and right hand, left and right foot respectively. The data can be downloaded from their ConnectomeDB¹ in an already image-wise preprocessed state as well.

Creating an easy to use dataset, where the end user would not need to have all the necessary medical knowledge and computation power to do the initial, required preprocessing steps was also a focus of the project. For this purpose common automated preprocessing pipelines were created to accomplish low-level tasks, such as surface generation, artifact removal and alignment to standard space [29].

¹<https://db.humanconnectome.org/app/template/Login.vm>

4.2.2 Methodology

The ROI+connectome representation first has to be extracted from each sample of the dataset and then they are used as the input for the diffusion model. Since the connectome at this point is not an image but rather a sort of brain graph, a decision has been made to create a model different from most DDPMs, by creating a specialized graph U-net, that can serve as the neural network inside the diffusion model that predicts the noise. To do this, I adapted the original graph U-net [24] to work in this setting, by adding time and class conditioning and edge prediction techniques.

The resulting generated connectomes and ROI values are then used as conditioning in the bottleneck of a VAE, which has been trained to generate MRI scans with the original images of the same dataset, using the original ROI and connectome values as conditioning during training.

Here, the goal was to make sure the VAE learns how to generate scans that correspond to the information fed at the bottleneck; this way it is possible to generate ROI and connectome samples, which can be turned into high-fidelity MRI samples with significantly less resource usage than if we were generating the MRIs in the first place.

4.2.2.1 Data Preprocessing

The architecture requires three separate types of data: the fMRI scan itself, the ROI values and connectomes corresponding to the scan. Of these, the HCP dataset only contains the preprocessed fMRI scan, the other two needed to be computed in a preprocessing pipeline (See Section 3.2.2). The ROI values were extracted with the help of an atlas, which describes the regions of the brain corresponding to the ROI values. Since the focus of the project was on sensory motor tasks, the atlas was filtered for brain regions activated during these tasks (in practice this means 21 ROIs for each side of the brain, so 42 altogether).

With the filtered atlas, a mask can be constructed to block out unwanted parts of the brain during the extraction. Understanding the format and meaning of the data used in the preprocessing (e.g.: what ROIs and connectomes are and how to calculate them) is difficult as it requires knowledge in the medical domain.

The correct preprocessing was possible thanks to communication from doctors though my advisors, the documentation of the nilearn package and high-quality HCP dataset. Due to the preprocessing of the dataset, it works best with the Glasser360 brain atlas [65], which was constructed using HCP data and identifies 180+180 (in the left and right hemisphere) regions of interest according to changes in cortical architecture, function, connectivity, and/or topography.

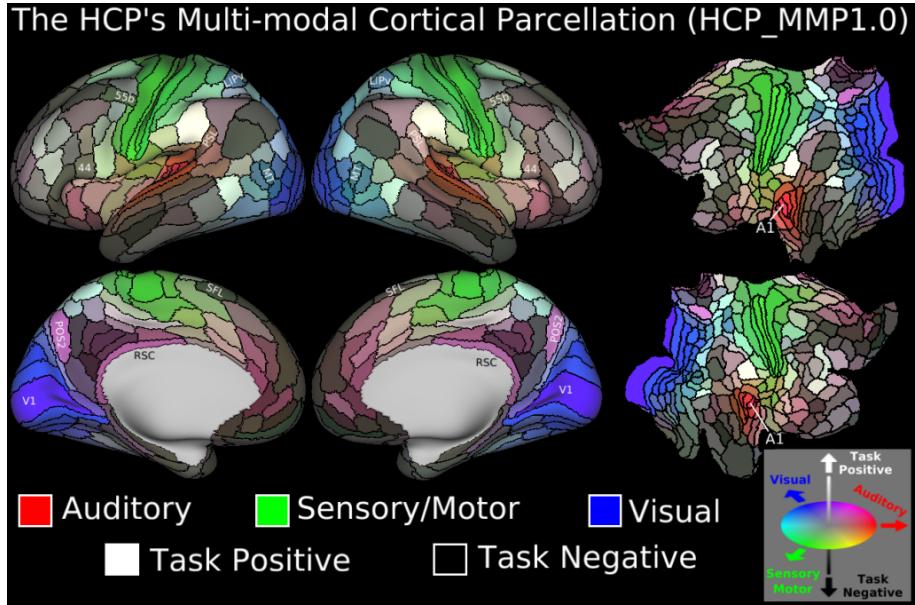


Figure 4.6: The Glasser360 brain atlas created based on the HCP healthy adults dataset [65].

After the extraction is complete we get a time series of ROI values for each patient. This means a tensor of size $284 \times 16 \times 42$, (time \times repeatedMeasurements \times ROIs). To validate the correctness of the preprocessing a Support Vector Machine was used to classify the ROIs based on the performed task. The ROI values need to be standardized per subject, and separated by task. For this a Python script from previous works was used that generates a csv file with the times slices and the task performed at that time interval.

This file is essential in separating the relevant parts of the scans from the noise where no task is performed. The scans contain measurements for 5 tasks performed twice, so in total 10 for each patient, with 16 repeated measurements taken by the MRI machine (As a further preprocessing step, before the data was used in models the mean value of these 16 steps was taken). This means 124 of the 284 long scan is noise.

Once the preprocessing was correct the SVM achieved 90-95% accuracy on the classification task based on performed tasks. For calculating the connectomes from the ROI values the nilearn package's ConnectivityMeasure class was used.

The subject-wise calculated ROI values were separated by tasks and connectomes were calculated for each task separately. Since measurements across different tasks differ significantly, calculating the connectomes with the tasks mixed would not yield meaningful results.

4.2.2.2 Proposed Architecture

The model consists of a diffusion block and an autoencoder. The diffusion block is trained on the ROI values and connectomes to also generate ROI values and connectomes, which are used to condition the autoencoder at the bottleneck.

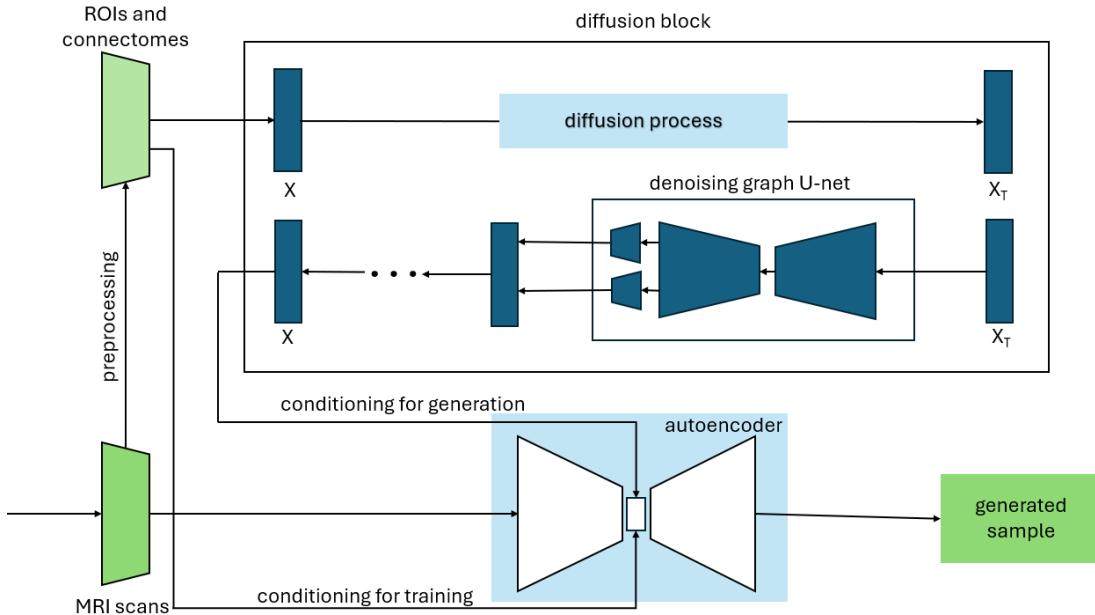


Figure 4.7: Visualization of the model architecture, focusing on the flow of data.

A big advantage of this model is that the two parts of the model can be trained separately. The autoencoder is trained to be able to reproduce scans with the corresponding connectome. At sampling time the two models are used in conjunction: the diffusion model generates a connectome and only the decoder part of the autoencoder is used to generate novel samples.

For the DDPM and DDIM diffusion blocks the implementation in the denoising-diffusion-pytorch package was used to enable quick experimentation. The package provides an easy interface to train diffusion models with custom networks, though it required some unused parameters to be added to model, that were unnecessary in this use-case.

Another advantage of this package is that it enables easy DDIM sampling. This could be used with 200 steps, instead of the 1000 used for training the DDPM. This package sadly lacks the capability for training a network with class conditioning, so for experiments involving class conditioning it was extended manually.

For a VAE implementation, since it is not a focus of this work, a choice was made to use an implementation provided by the latent-diffusion Github repository². The model this library provides has the option to train a VAE that learns representations on 3D data.

The class was modified in minor ways to fit in with the rest of the codebase and two major changes had to be made to it. Firstly, the loss function was recalibrated to enable the model to train on the given dataset. Secondly, it was extended with the capability to condition the bottleneck with the ROI values and connectomes.

For the conditioning a small CNN was implemented with a fully connected layer at the end, that takes the $2 \times 42 \times 42$ size ROI and connectome tensor and creates an embedding with the same size as the bottleneck. The conditioning is then implemented by performing Adaptive Instance Normalization [38] on the embedding and the values in the bottleneck.

4.2.2.3 Graph U-net

U-nets are a very successful subtype of CNN architectures, originally invented with the purpose of biomedical image segmentation [55]. It is widely used in segmentation tasks and other cases where the output needs to have the same dimensions as the input, due to its symmetrical nature. This is the same reason why they're commonly used in diffusion models as the noise prediction neural network [35].

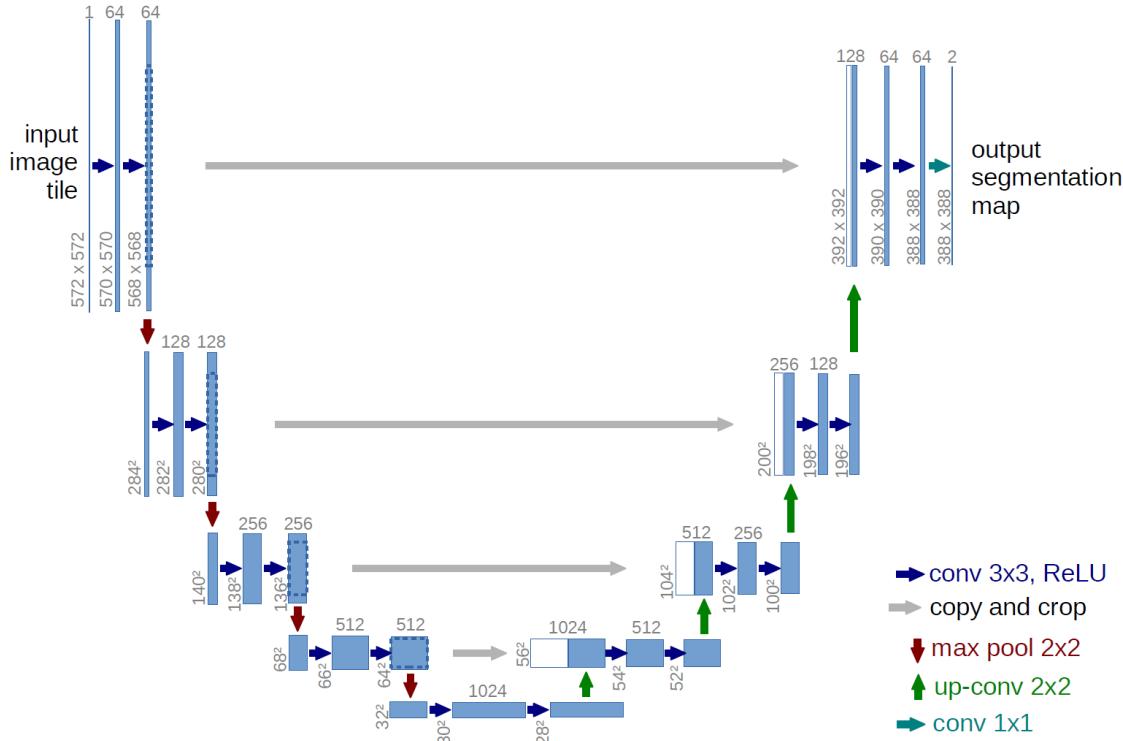


Figure 4.8: Visualisation of the U-net architecture [55].

²<https://github.com/CompVis/latent-diffusion>

U-nets are named due to their distinctive architecture: the network consists of a contracting and expansive path, which together form a u-shape (Figure 4.8). The contracting path is a typical CNN with convolutions and max pooling. Conversely, the expansive path has upsampling and convolutions, while skip connections are also utilized where dimensions match.

Since U-nets have been such a successful model type in convolutional networks, there has been an effort to find an analogue in graph convolutional networks as well. Graph U-nets [24] follow the original principles of the U-net architecture: the encoder-decoder structure, the skip connections and the general architecture, as seen on Figure 4.9.

The convolutional layers are replaced by GCN layers and for standard pooling and unpooling layers which do not have an obvious graph equivalent, since they rely on locality in images that is not present in general graphs, novel layers have been developed for graph pooling and unpooling.

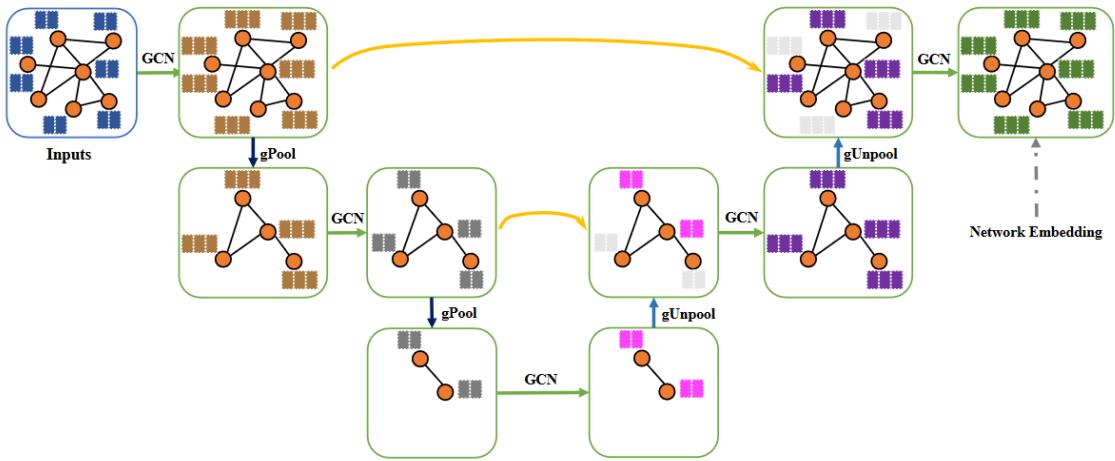


Figure 4.9: Visualization of the graph U-nets architecture [24].

Graph pooling layers enable downsampling on graphs by adaptively selecting nodes to build a smaller graph from existing nodes, while preserving edges. The selection method is similar to a traditional k -max pooling: the highest k values are preserved after the operation. The top k nodes are selected based on a learned 1D projection, to a vector \mathbf{p} :

$$y_i = x_i \mathbf{p} / \|\mathbf{p}\|,$$

the nodes are kept which retain most of their information when projected to \mathbf{p} . After the selection has been made, the adjacency matrix is updated to only include these nodes and the corresponding rows are extracted from the feature matrix via a gate operation, using matrix multiplication and a sigmoid operation, which makes \mathbf{p} trainable by backpropagation.

The unpooling layer, unlike image upsampling layers, can only work in connection with a graph pooling layer for which it performs an opposite action: the adjacency matrix is restored to its state prior to the the pooling layer and the node feature matrix is distributed into a shape corresponding to this larger number of nodes, having zero values for newly reappearing nodes.

The graph U-net used in the model is a modified version of the architecture proposed in [24]. The connectome is used as the adjacency matrix in the U-net, the connectivity values between ROIs are the weights of edges in the graph.

Since we have singular ROI values available for each of the nodes, these were distributed along an identity matrix to create a better node feature input for the network, and since the U-net requires symmetry and we will be looking to generate both connectomes and ROI values, the full input is the concatenation of the ROI filled identity matrix and the connectome.

In this case, since the connectomes are 42×42 , the input of this part of the network is batch size $\times 42 \times 84$.

The model has the same number of blocks in both directions, comprised of one GCN layer and one graph pooling layer in the down direction, and one graph unpooling layer and one GCN layer in the up direction, with a connecting bottom GCN layer.

Corresponding blocks of the same size are connected via skip connections. Conditioning (with the main concern being time conditioning) is applied at every block by calculating scale and shift vectors via a small MLP.

The output of the last block is thus the same size as the input of the first one and it gets separated into node features that are used for ROI values calculation and node features that are used for connectome calculation. The ones used for ROIs are sent through a final GCN layer where they are reduced to a single value per node and then distributed along an identity matrix to preserve the symmetry of the data.

For the connectome however, as each value represents an edge between two regions, the other half of the output gets pairwise concatenated and passed through an MLP to determine one value per edge for the connectome.

4.2.2.4 Experiments

To find out which choices work best for the model, there were multiple experiments with various architecture and parameter ideas. This was mainly focused on the graph U-net, as it is the central part of the architecture. The VAE remained largely the same, with only minor changes to accommodate the conditioned training and sampling.

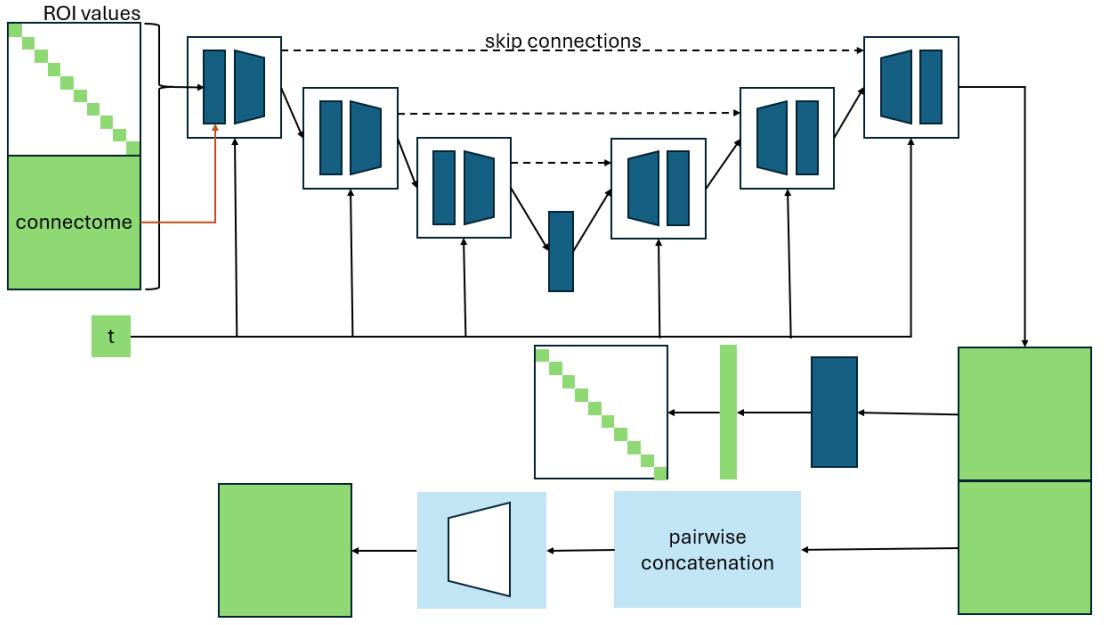


Figure 4.10: Visualization of the graph U-net architecture, focusing on the flow of data

Graph U-net Since this kind of graph U-net and diffusion architecture does not show up in literature often, experiments were done to try and figure out which setups would lead to better generation results in the end. There were wide variety of options to try and the lack of guidance on which ones could be successful in which combination meant trying out new ideas in a fail fast method would be best: the results of the model were inspected after a set number of epochs and if the connectome visually worsened that method could be avoided in the future.

The usual output of a graph U-net is just node features, and there are various methods to convert them into edge predictions [43], therefore it was important try a few to see what would work well with the model. In the implementation of the graph U-net I have included a switch that can take None, 'scalar' and 'mlp' values and makes experimenting with this easier.

The None value keeps the original vector of the output, without any transformations applied. In this setting the network tries to learn the noise as if they were features of the nodes, both for ROI values and edge weights. This is the easiest method to implement but this way it is not taken into consideration how each edge represents the connection between two nodes.

To better represent the way a connection is formed between two nodes, the 'scalar' method takes each resulting node embedding and computes the scalar product of each node pair to build the new adjacency/edge weight matrix. In case of the 'mlp' option the node embeddings get pairwise concatenated to represent their respective edges, and then are

passed through a small MLP to calculate an edge strength for the pair of nodes. Here the concatenation makes sure that both nodes are considered together in the calculation.

The most usable and best performant solution for calculating the connectome ended up being the MLP, as it yielded better results than the base method and the scalar product approach seemed to lead to generated samples looking almost identical.

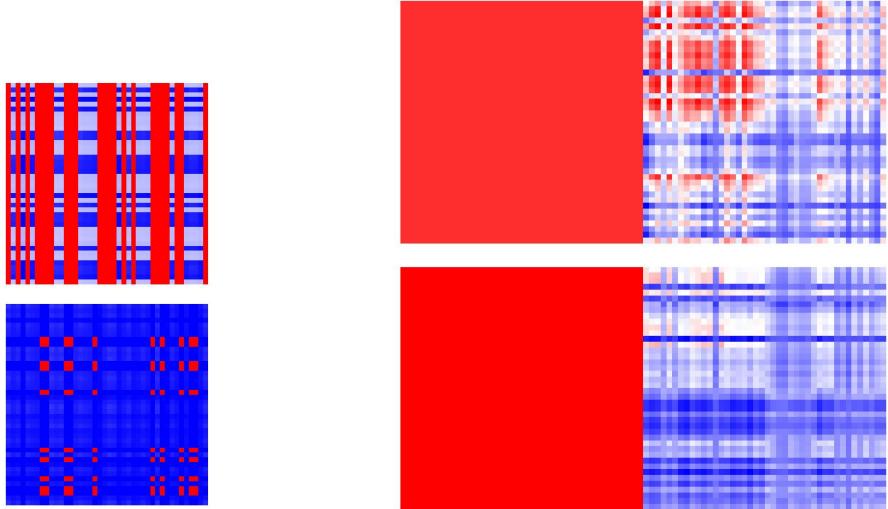


Figure 4.11: Bad generated ROI activation values (left) and connectomes (right).

Class conditioning Conditional diffusion models enable generating images with specific features. In case of class conditioning the model is tasked with generating a sample from a category from a list that is already known ahead of time. This can be achieved by passing an embedding of the class labels to the noise predicting network. This approach guides the model to learn the different characteristics of each class, which usually leads to better overall image quality. It also grants the ability to generate images for specific classes, which is especially useful in the medical domain.

Since the used dataset has labels in form of the body part corresponding to the performed task, it was possible to add this technique into the model. A very simple implementation of this was constructed, by passing the class label to the U-net, where it was turned into a one-hot vector and concatenated to the time conditioning to perform both at once.

Unfortunately this implementation did not result in improved performance, the resulting generated samples tended to have prominent bands of a single value instead of resembling the original samples. It could still be a good idea to further experiment with class conditioning, perhaps with a different implementation method.

Since the connectome is supposed to be a symmetric matrix, the graph U-net's loss function could be changed in some way to reflect this in some form. A version of loss calculation was tried out, where only the lower triangle matrix of both the result and target were

considered during calculation: keeping the generation of the upper half as a sanity check, but concentrating on the lower half. This did not noticeably impact the resulting generated samples, so it was removed, choosing to keep the model simpler if possible.

4.2.3 Results

To evaluate image generation tasks, it is best to use multiple different approaches. On one hand, quantitative metrics like Fréchet Inception Distance [34] (FID), Peak Signal to Noise Ratio [37] (PSNR), and Structural Similarity Index Measure [37] (SSIM) are important to obtaining an objective view of the quality of the generated images. On the other hand, especially in image synthesis tasks, qualitative evaluation is important to assess the visual quality of the samples, as numerical measurements don't always align with visual representations.

FID uses an InceptionV3 [69] network to calculate the distance between the extracted features of real and generated samples, PSNR measures the ratio of the signal (i.e., the important parts of the image) and noise in the reconstructed image, and SSIM, unlike PSNR, measures not the pixel values, but structural aspects of the samples.

As a baseline, we can compare these metrics to those measured on a pixel-based 3D DDPM architecture [62]:

Method	FID ↓	PSNR ↑	SSIM ↑
Proposed	400.8233	13.4702	0.0535
Pixel-based	20.4152	27.5278	0.8442

Table 4.1: Comparison of results on quantitative metrics between proposed method and a pixel-based DDPM method.

From these numerical metrics it is obvious that the results of the architecture are not comparable to the results achieved by using the diffusion to generate the fMRI sample directly.

As it can be seen on Figure 4.12, the trained VAE was capable of generating reasonably good samples in validation epochs when conditioned with one of the original connectomes. It has some artifacting and small chunks outside of the brain, but it looks very similar to the input images. The left sample however, the one conditioned by synthetic data is very noisy and does not resemble samples from the training dataset. This comparison points to our generated ROIs having the opposite of the desired effect, as they completely destroy the image.

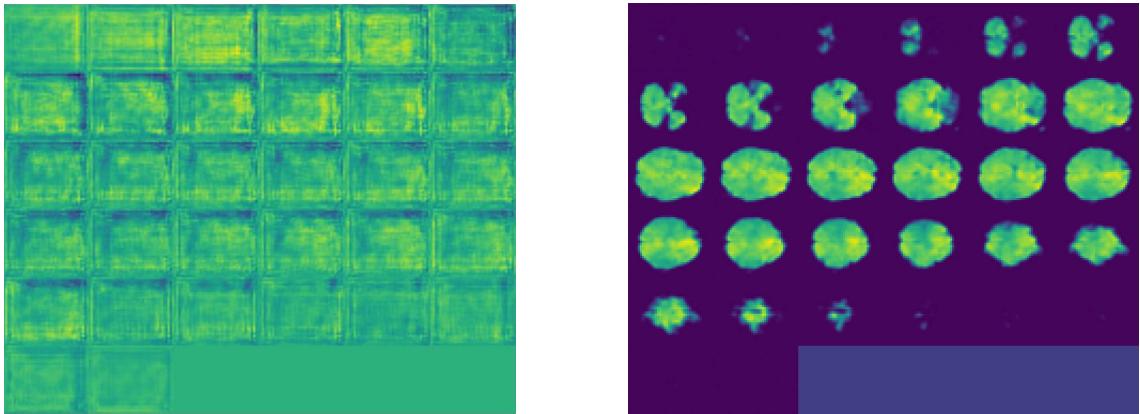


Figure 4.12: Samples from the VAE, on the left a sample generated by the decoder conditioned by generated BOLD signals, on the right a sample generated during a validation epoch.

4.2.4 Conclusions

To understand the poor performance of image generation, we have to examine the results of the connectome and ROI value generation. The training of the model was very unstable, often devolving into a mass of low and high values in various patterns (as seen on Figure 4.11). This has become worse in later epochs, as the model could produce promising results after ~ 20 epochs of training, as seen on Figure 4.13.

Overall, this seems to suggest that there is potential for good outcomes in the model architecture but there is some sort of numerical or gradient instability that causes it not to converge. This is backed up by how the loss behaves during training. It can be seen on Figure 4.14 that while the loss initially drops many orders of magnitude at the start of training, after zooming in it is obvious that the loss still jumps around and does not actually converge.

While the network is not capable of generating consistently good, novel samples in its current state, there is reason to believe that with the right adjustments this architecture could be capable of learning meaningful representations of the data. Other than that, this idea is a great step in resource intensiveness. It has about 80% less parameters than a pixel-based diffusion model using 3D-Unet would. It also decouples the diffusion process from the slower, more costly pixel-based VAE, which enables rapid experimentation: once trained, a VAE can be used to test out many GNN models.

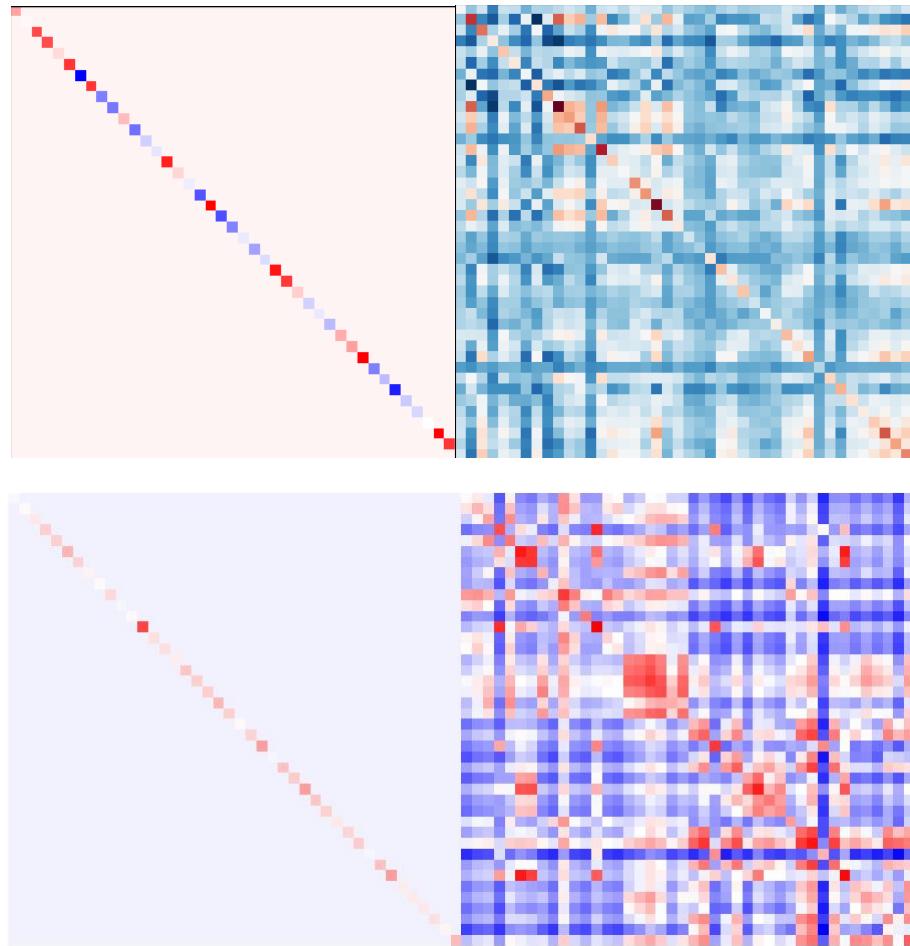


Figure 4.13: On the bottom is one of the better samples generated by the model, and on the top a sample from the training dataset. ROIs are on the left and connectomes are on the right.

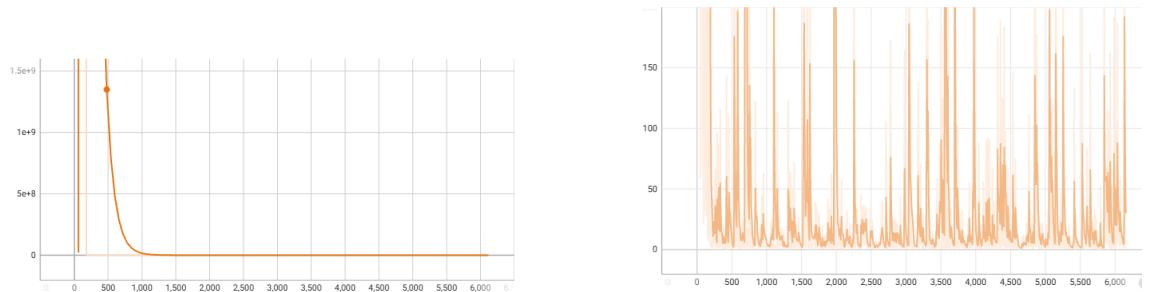


Figure 4.14: Loss during training.

4.3 Exploring device-to-device differences

Most publicly available larger fMRI datasets are a result of large-scale collaborations of various healthcare and academic institutions (like the ones described in Section 4.1.1 and 4.2.1), since collecting this many scans is both very resource intensive and takes a long

time. These collaborations are incredibly valuable data sources, but this also results in a crucial characteristic: the fMRI scans are from different machines.

These machines are very precise and well-engineered devices and the scan settings are coordinated to be the same across all study participants, but there can be variances in the different machines and environmental factors that lead to small differences in scans [67] that can lead to different performance in machine learning applications. If this is the case it could be helpful to introduce a preprocessing step that neutralizes this issue.

4.3.1 Dataset

My experiments were conducted on the ABIDE dataset [19] (same as in Section 4.1.1). More specifically to avoid costly preprocessing steps, but to be able to be keep in control of the connectomes and other facets of the resulting data, I turned to the Preprocessed Connectomes Project (PCP) initiative’s ABIDE preprocessed dataset [13].

Since there is not a widely accepted best practice method for preprocessing fMRI images, the participating five teams have all used their preferred methods to clean the data and various derivatives of this process are downloadable from the provided Amazon S3 bucket. From my perspective, the most important of these is the ROI activation time series: for each preprocess method, the resulting preprocessed scans were segmented using 7 different brain atlases.

I built the used dataset from these activation time series using methods outlined in a paper about the NeuroGraph dataset [58]. In this paper researchers argue that the reason applying GNNs in the neuroimaging domain has been so challenging is the expansive number of different preprocessing pipelines and the large parameter search space. To counteract this problem, they benchmarked various representations with the most commonly used GNN architectures to offer guidance for dataset construction.

For their experiments they used the HCP dataset (See Section 4.2.1) and constructed five different tasks, each with corresponding datasets with different parameters. Through their benchmarking efforts they have found that a larger number of regions of interests, creating sparser graphs and using the correlation matrix as node features leads the best results across tasks and network types. This is invaluable information for further research and it would have been very beneficial, had I known about it in previous endeavours.

For the construction of this dataset, I have followed these suggestions. First the time series needed to be normalized to a zero mean and unit variance for each subject individually. With the help of nilearn, the correlation was used to create a graph structure and the node features. For the graph structure, only edges running between nodes with a positive correlation were kept and a hyperparameter was introduced to keep only the top $n\%$ of

edges with the highest correlation leading to sparser graphs. For edge features the original correlation matrix was kept.

In [58], researchers tested three different sizes of the Schaefer atlases [60], with 100, 400 and 1000 ROIs and the results indicated better model performance on the larger graphs created using the 1000 regions. Since the preprocessed dataset did not carry time series created with this type of atlas, I opted to try two of the available atlases, the Harvard-Oxford atlas [17] with 111 ROIs and the Craddock 400 atlas [14] with 392 ROIs. Just based on these numbers, I had a general expectation to see the second one perform better, but the atlases were created with significantly differing methodologies and often it is hard to predict what works better with a given dataset.

I experimented with creating a baseline, a not deep learning-based model to compare the results to using tsfresh for feature extraction from the time series and XGBoost and various sklearn models to make the predictions. These were not successful however: this dataset and objective are not an easy classification problem, even complex GNN-based solutions can only perform at about a 70% accuracy at most, but these alternate solutions could not consistently perform better than random chance, so I scrapped that part of the experiment.

4.3.2 Methodology

In the ABIDE preprocessed dataset samples are also categorized by the performing institution. To better understand the impact of the specific machine being used on the data, in this experiment I have used the most numerous sites to create a dataset that focuses not only predicting diagnoses, but also on this aspect. As it is visible on Figure 4.15, there are only a few sites that have a considerable amount of samples on their own. The decision was to use the ones with over 100 samples: NYU (NYU Langone Medical Center), UM_1 (University of Michigan: Sample 1) and USM (University of Utah School of Medicine).

In creating a model for this task, the guidance from [58] was followed similarly as to before. This study found that residual- (or skip-) connections play a pivotal role in enhancing the performance of brain networks. My model features 3 GCN layers with ReLU activation and residual connections after each layer, which are then all subjected to average pooling and then concatenated. I have also employed dropout between layers and used Kaiming-He uniform weight initialization [33]. The output from this is then fed through two fully connected layers in a similar fashion and finally a sigmoid activation is applied for the binary classification task.

For loss, the matching binary cross entropy was utilized and the ADAM optimizer [40]. I have used a set seed in PyTorch to make sure the experiments are repeatable. I added the

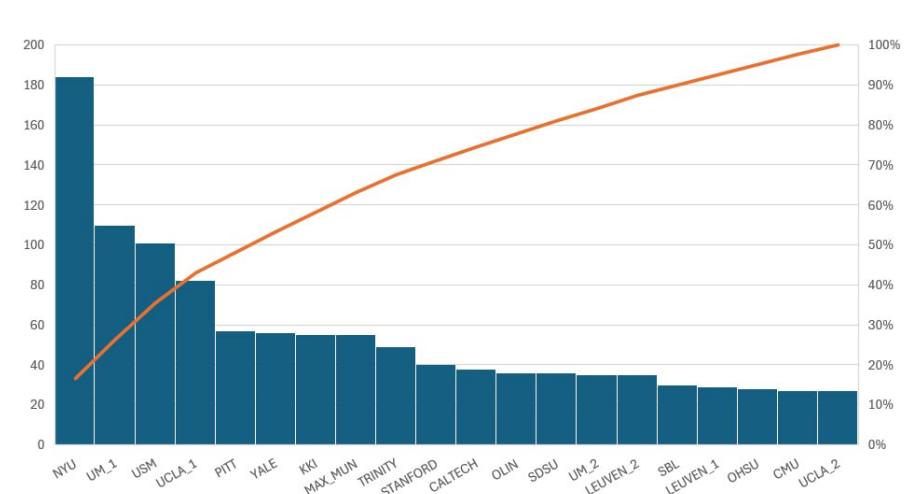


Figure 4.15: The distribution of capturing site within the ABIDE dataset, according to the summary spreadsheet.

WandB library to the project and used it both for experimenting with hyperparameters to narrow the search space and find better setups and to conveniently run my experiments.

The tests were performed using the three institutions' samples: on one hand, with each of the three sets separately and together with the test set randomly separated from the train set, on the other hand with a leave-one-out method. In the latter, one of the sites is left out of the training set and kept exclusively for testing: this can be a good way to measure how effectively the model can learn patterns that are also applicable in data from previously unseen sites, how good it is at generalizing when taking into account device-to-device differences.

As mentioned before, my original plan was to use the three most populous sites from the dataset to examine the individual, combined and leave-one-out results and compare them, to get a sense for how a different measuring device impacts the impact the classification performance. I was unable to realize the plan in this capacity however, as the model would not converge and produced wildly varying results on different seeds and settings.

I ran test runs using WandB, where I experimented with a few different values for parameters in the individual and combined dataset scenarios. As it can be seen on Figure 4.17a, the accuracy on the test set was in a lot of cases not really higher than random chance and showed very wide variance, showing significant instability in the training process.

I tested the model with the NeuroGraphDataset [58], to see if the problem was with the model or with dataset. The model performed well on that set, leading me to believe that the problem was that there is simply not enough samples for the model to train effectively on those three sites combined, let alone on a single one of them. The NeuroGraphDataset has 1,078 graphs, while these have a bit more than 100 each. Tests on this dataset have

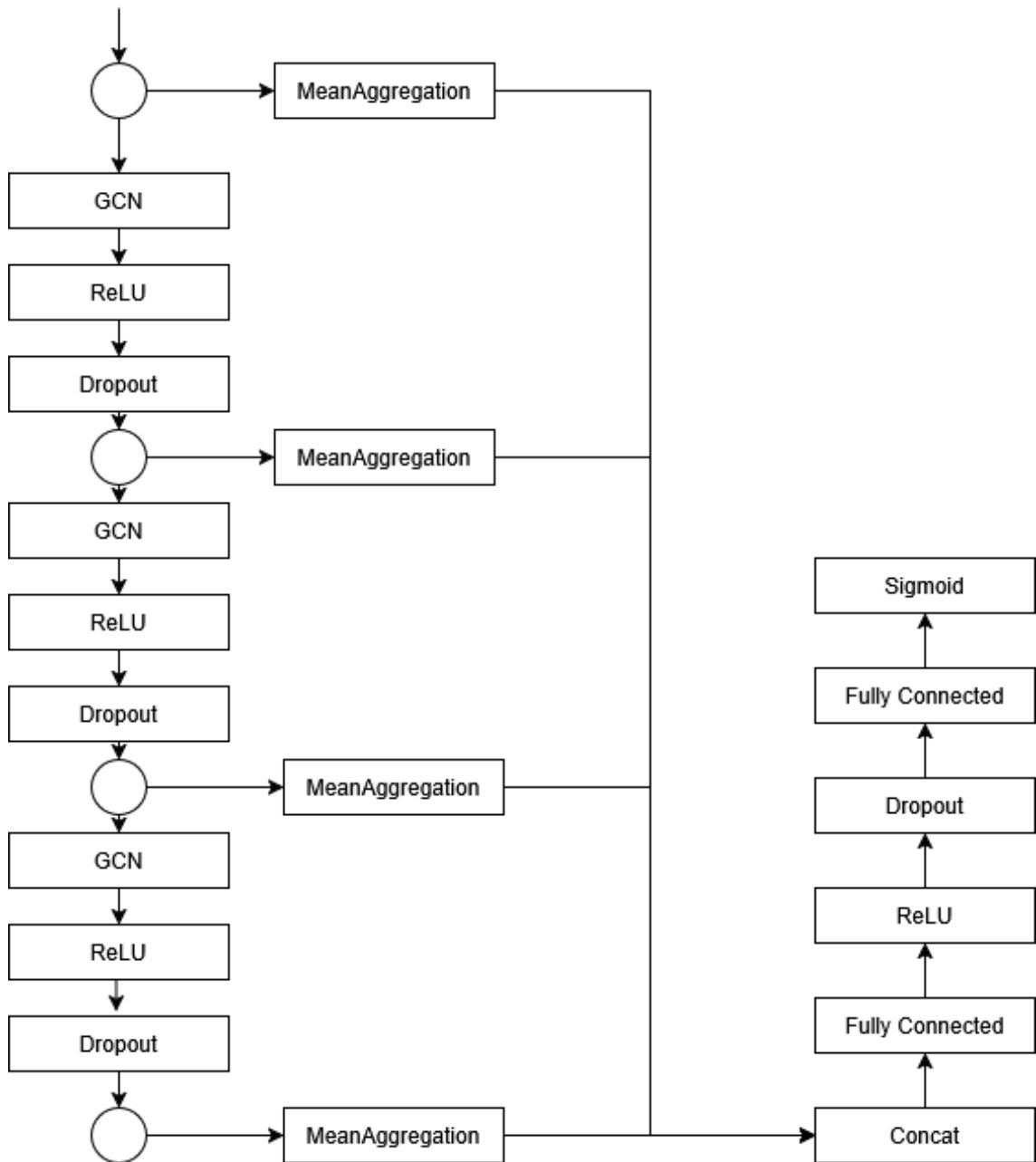
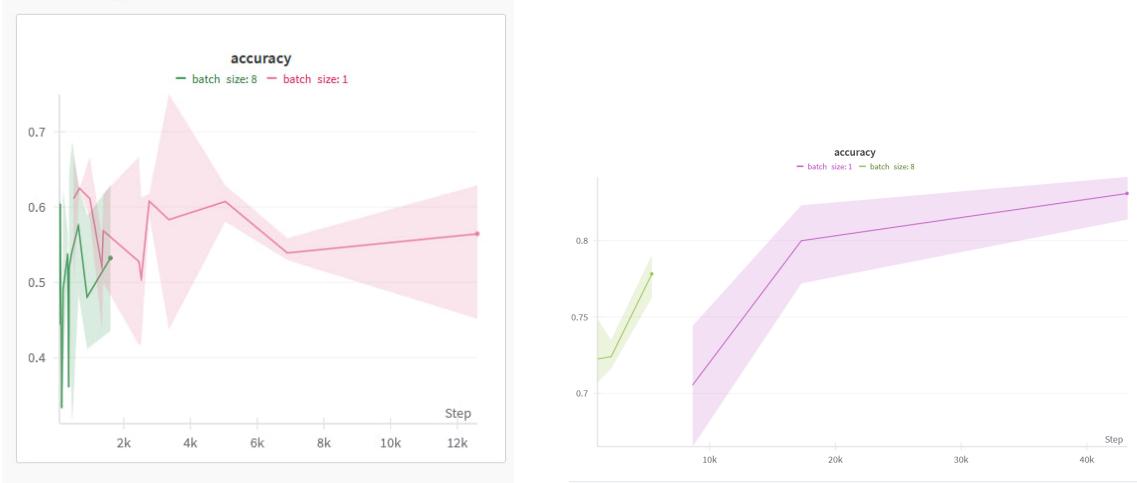


Figure 4.16: Visualization of my used model architecture.

also shown that the model does significantly better with a batch size of 1, so going forward test were run with this parameter.

Initial tests of using the entire ABIDE dataset seemed much more promising, so I decided on running the experiment by including all sites and verifying the model in a leave-one-out way to still be able to find answers to the original question. I separated the data from the rest of the sites into a training and validation set and kept remaining set as the test set. These results would be compared to each other and to the results on a site-wise mixed and randomly into training, validation and test separated run.



(a) Visualization of the accuracy on the test set using the first iteration of the ABIDE dataset, separated for batch size of 1 and 8.

(b) Visualization of the accuracy on the test set using the NeuroGraphDataset, separated for batch size of 1 and 8.

Figure 4.17: Visualisation of the results on two datasets.

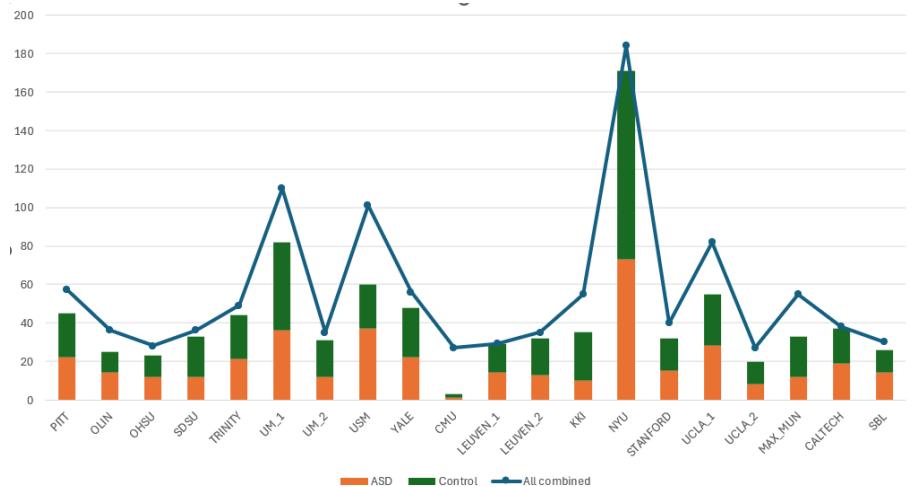


Figure 4.18: The distribution of capturing site within the ABIDE dataset, stacked columns representing the used ASD and control samples, line representing the supposed total number of samples according to the spreadsheet.

The final distribution of samples among sites was in the end not exactly the same as the spreadsheet results have suggested. There were some missing files, it was suggested not to use samples where the functional image's mean framewise displacement is above 0.2 and some of the time series had been deleted due to uneven lengths. The final distribution compared to the original information can be seen on Figure 4.18. In line with this, to keep the results more representative, I decided to exclude sites from the leave-one-out scheme where there were an outlying number of samples: CMU only had 3, NYU had 171 and UM_1 82.

4.3.3 Results

In this section I want to share the accuracy results of the runs performed both on the whole dataset and in the leave-one-out method. The results of multiple runs are represented as box plots.

4.3.3.1 Full dataset results

Figure 4.19 has the results of running the model on the full dataset in a 0.7-0.15-0.15 train-validation-test split. Runs had either 30 or 50 epochs, either the Craddock 400 or the Harvard-Oxford brain atlas with either 5 or 10% graph density.

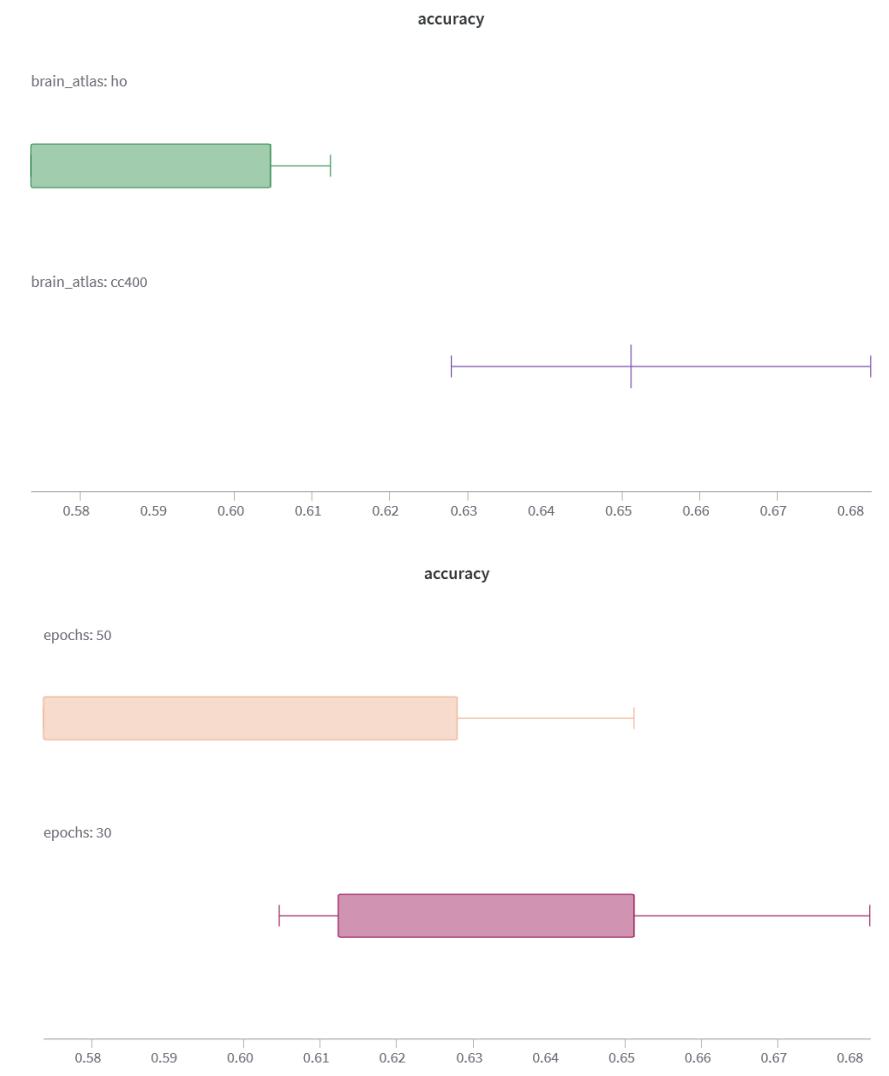


Figure 4.19: Accuracy in the randomly separated full dataset. On the right grouped by used brain atlas, on the left by number of training epochs

4.3.3.2 Leave-one-out results

Figure 4.20 shows the results of the leave-on-out experiments. Each box-plot is labelled with the site that was used as the testing section. Each of these were run with similar settings to the full dataset tests.



Figure 4.20: Accuracy in the leave-one-out scenario

4.3.4 Conclusions

Overall the accuracy of the model was not particularly outstanding, but these results still show that the capturing site of fMRI scans does impact how the model performs. The leave-one-out experiment results show a much larger variance than the random test set and it is also visible from the plots that certain sites perform significantly better or worse than others.

KKI is a particularly egregious example of this: accuracy hovers around the expected accuracy for random chance, meaning whatever the model had learned from all of the other sites' data is not very applicable in this scenario. No other site dataset performed this bad, but there are very stark differences among these as well.

This is generally bad news for both similar experiments and the prospect of creating a usable system for diagnostic purposes. It would be unacceptable for such a system to perform at wildly varying accuracies in different institutions. It seems currently that models cannot work well with this type of variance in the input data. To alleviate this problem, either a new preprocessing step would need to be introduced to the already complicated pipeline, or the model would need to be enhanced in a way that counteracts this problem.

Chapter 5

Summary

In this chapter I will give a short overview of my work and how these projects fit into the broader research space and provide some ideas of future enhancements.

5.1 Summary

Using statistical and machine learning methods in the medical domain has been around for quite some time now, but the latest artificial intelligence boom has really sped up its adoption. While a lot of solutions are still in a research phase, since diagnostic solutions can only be safely deployed once their accuracy is on par with humans, we are much more likely to see machine learning solutions in healthcare now, than a few years ago.

Analysing medical imaging scans is a very active part of this process. This is often done by computer vision solutions, such as CNNs, but in case of fMRI footage, there are challenges that may be better overcome by methods involving graph convolutional networks. fMRIs can be represented as a brain graph, by first segmenting them into ROIs using a brain atlas and then finding correlations in their activation. Graph neural networks are a class model that use the concept of message passing (nodes aggregating features from their neighbours) to enable deep learning on graphs.

In my work I have used GNNs on multiple dataset created from fMRI scans. These datasets are large from a medical standpoint and were created by the international joint effort of multiple medical and research institutions. Still, these datasets are quite small from a deep learning standpoint, containing a few thousand samples each.

I have contributed to a project on generating novel fMRI samples by using the connectome and ROI representations in a diffusion model and using these generated samples to condition a VAE to create fMRI samples. I created a new graph U-net architecture to be used inside the diffusion model.

Using the ABIDE dataset I created a preprocessing pipeline to go from ROI time series to data usable in models and multiple models that predict whether a patient has ASD. I examined the effect of different collection sites on the outcomes of a model.

5.2 Future work

As continuation of this work I would like to revisit the connectome based diffusion project and apply the things I have learnt while working on the device-to-device differences project. Since the paper about benchmarking optimal preprocessing of fMRI for use with graph neural networks [58] became available too late to be used in this project, many of the preprocessing steps have been suboptimal: using the mean ROI values for the node features instead of correlation and not taking care to make the graph more sparse by thresholding correlations.

By preparing the dataset in a more favourable way and setting up a similar WandB sweeping system to try out both model and preprocessing parameters it may be possible to stabilize the training loop of the graph U-net and generate good novel samples. Since the graph part of the net can be trained separately, this possibly would not take a lot of computational resources.

Following the device-to-device differences project, it would be beneficial to use the information gathered about how this impacts models to devise a method to counteract these differences. If all of the samples could be translated to a uniform space, it could be easier for a network to effectively work on them and produce better classification results. For more pointers on how to transform data to achieve this, it might be beneficial to analyse and visualise the final layer of activations on case of different origin sites.

Acknowledgements

Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

I would like to thank Levente Sipos and his advisor Szabolcs Torma for our collaboration on the "Connectome-based diffusion using graph neural networks for generating high fidelity synthetic fMRI data" Scientific Students' Association Report. Their insight and work has been invaluable. I would like to thank Karz Gergely and Artillence Kft. for providing us access to their hardware and allowing us to do our experiments for the Scientific Students' Association Report.

I would like to extend my immense gratitude to my advisor, dr. Luca Szegletes for the constant support and for trusting me even when I did not trust myself.

Bibliography

- [1] Types of healthcare data: A comprehensive overview. URL <https://kms-healthcare.com/blog/types-of-healthcare-data/>. Accessed: 2025-04-07.
- [2] Laith Alzubaidi, Jinshuai Bai, Aiman Al-Sabaawi, Jose Santamaría, Ahmed Shihab Albahri, Bashar Sami Nayyef Al-Dabbagh, Mohammed A Fadhel, Mohamed Manoufali, Jinglan Zhang, Ali H Al-Timemy, et al. A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications. *Journal of Big Data*, 10(1):46, 2023.
- [3] Shun-ichi Amari. Backpropagation and stochastic gradient descent method. *Neurocomputing*, 5(4-5):185–196, 1993.
- [4] RJ Anderson. The collection, linking and use of data in biomedical research and health care: ethical issues. 2015.
- [5] M Stella Atkins and Blair T Mackiewich. Fully automatic segmentation of the brain in mri. *IEEE transactions on medical imaging*, 17(1):98–107, 2002.
- [6] Dor Bank, Noam Koenigstein, and Raja Giryes. Autoencoders. *Machine learning for data science handbook: data mining and knowledge discovery handbook*, pages 353–374, 2023.
- [7] Danielle S Bassett and Olaf Sporns. Network neuroscience. *Nature neuroscience*, 20(3):353–364, 2017.
- [8] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.
- [9] Joan Bruna and Xiang Li. Community detection with graph neural networks. *stat*, 1050:27, 2017.
- [10] Qi Cao, Huawei Shen, Jinhua Gao, Bingzheng Wei, and Xueqi Cheng. Popularity prediction on social platforms with coupled graph neural networks. In *Proceedings of the 13th international conference on web search and data mining*, pages 70–78, 2020.

- [11] European Commission. Opinion 4/2007 on the concept of personal data., 2007. URL https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf. Accessed: 2025-04-07.
- [12] Kristen Coyne. Mri: A guided tour. URL <https://nationalmaglab.org/magnet-academy/read-science-stories/science-simplified/mri-a-guided-tour/>. Accessed: 2025-04-11.
- [13] Cameron Craddock, Yassine Benhajali, Carlton Chu, Francois Chouinard, Alan Evans, András Jakab, Budhachandra Singh Khundrakpam, John David Lewis, Qingyang Li, Michael Milham, et al. The neuro bureau preprocessing initiative: open sharing of preprocessed neuroimaging data and derivatives. *Frontiers in Neuroinformatics*, 7(27):5, 2013.
- [14] R Cameron Craddock, G Andrew James, Paul E Holtzheimer III, Xiaoping P Hu, and Helen S Mayberg. A whole brain fmri atlas generated via spatially constrained spectral clustering. *Human brain mapping*, 33(8):1914–1928, 2012.
- [15] Ameya Daigavane, Balaraman Ravindran, and Gaurav Aggarwal. Understanding convolutions on graphs. *Distill*, 2021. DOI: [10.23915/distill.00032](https://doi.org/10.23915/distill.00032). <https://distill.pub/2021/understanding-gnns>.
- [16] Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. Convolutional neural networks on graphs with fast localized spectral filtering. *Advances in neural information processing systems*, 29, 2016.
- [17] Rahul S Desikan, Florent Ségonne, Bruce Fischl, Brian T Quinn, Bradford C Dickerson, Deborah Blacker, Randy L Buckner, Anders M Dale, R Paul Maguire, Bradley T Hyman, et al. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *Neuroimage*, 31(3):968–980, 2006.
- [18] Edgar A DeYoe, Peter Bandettini, Jay Neitz, David Miller, and Paula Winans. Functional magnetic resonance imaging (fmri) of the human brain. *Journal of neuroscience methods*, 54(2):171–187, 1994.
- [19] Adriana Di Martino, Chao-Gan Yan, Qingyang Li, Erin Denio, Francisco X Castellanos, Kaat Alaerts, Jeffrey S Anderson, Michal Assaf, Susan Y Bookheimer, Mirella Dapretto, et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry*, 19(6):659–667, 2014.

- [20] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, pages 226–231, 1996.
- [21] European Parliament and Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council. URL <https://data.europa.eu/eli/reg/2016/679/oj>. Accessed: 2025-04-13.
- [22] European Parliament and Council of the European Union. Regulation (eu) 2025/327 of the european parliament and of the council of 11 february 2025 on the european health data space and amending directive 2011/24/eu and regulation (eu) 2024/2847 (text with eea relevance), 2025. URL https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202500327. Accessed: 2025-04-07.
- [23] Wenqi Fan, Yao Ma, Qing Li, Yuan He, Eric Zhao, Jiliang Tang, and Dawei Yin. Graph neural networks for social recommendation. In *The world wide web conference*, pages 417–426, 2019.
- [24] Hongyang Gao and Shuiwang Ji. Graph u-nets. In *international conference on machine learning*, pages 2083–2092. PMLR, 2019.
- [25] Fabio Garcea, Alessio Serra, Fabrizio Lamberti, and Lia Morra. Data augmentation for medical imaging: A systematic literature review. *Computers in Biology and Medicine*, 152:106391, 2023.
- [26] Thomas Gaudelet, Ben Day, Arian R Jamasb, Jyothish Soman, Cristian Regep, Gertrude Liu, Jeremy BR Hayter, Richard Vickers, Charles Roberts, Jian Tang, et al. Utilizing graph machine learning within drug discovery and development. *Briefings in bioinformatics*, 22(6):bbab159, 2021.
- [27] Anna Gergály and Sipos Levente. Connectome-based diffusion using graph neural networks for generating high fidelity synthetic fmri data, 2024. URL <https://tdk.bme.hu/conference/VIK/2024/sessions/orvosi/paper/Konnektomalapu-graf-neuralis-haloval>.
- [28] Justin Gilmer, Samuel S Schoenholz, Patrick F Riley, Oriol Vinyals, and George E Dahl. Message passing neural networks. In *Machine learning meets quantum physics*, pages 199–214. Springer, 2020.
- [29] Matthew F Glasser, Stamati N Sotiroopoulos, J Anthony Wilson, Timothy S Coalson, Bruce Fischl, Jesper L Andersson, Junqian Xu, Saad Jbabdi, Matthew Webster, Jonathan R Polimeni, et al. The minimal preprocessing pipelines for the human connectome project. *Neuroimage*, 80:105–124, 2013.

- [30] Gary H Glover. Overview of functional magnetic resonance imaging. *Neurosurgery Clinics of North America*, 22(2):133, 2011.
- [31] Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 855–864, 2016.
- [32] William L. Hamilton, Rex Ying, and Jure Leskovec. Inductive representation learning on large graphs, 2018. URL <https://arxiv.org/abs/1706.02216>.
- [33] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, pages 1026–1034, 2015.
- [34] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30, 2017.
- [35] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.
- [36] William Hollingworth, Christopher J Todd, Matthew I Bell, Qais Arafat, Simon Girling, Kanti R Karia, and Adrian K Dixon. The diagnostic and therapeutic impact of mri: an observational multi-centre study. *Clinical radiology*, 55(11):825–831, 2000.
- [37] Alain Hore and Djemel Ziou. Image quality metrics: Psnr vs. ssim. In *2010 20th international conference on pattern recognition*, pages 2366–2369. IEEE, 2010.
- [38] Xun Huang and Serge Belongie. Arbitrary style transfer in real-time with adaptive instance normalization. In *Proceedings of the IEEE international conference on computer vision*, pages 1501–1510, 2017.
- [39] Hany Kasban, MAM El-Bendary, DH Salama, et al. A comparative study of medical imaging techniques. *International Journal of Information Science and Intelligent System*, 4(2):37–58, 2015.
- [40] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [41] Diederik P Kingma, Max Welling, et al. Auto-encoding variational bayes, 2013.
- [42] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.
- [43] Juanhui Li, Harry Shomer, Haitao Mao, Shenglai Zeng, Yao Ma, Neil Shah, Jiliang Tang, and Dawei Yin. Evaluating graph neural networks for link prediction: Current

- pitfalls and new benchmarking. *Advances in Neural Information Processing Systems*, 36:3853–3866, 2023.
- [44] Xiaoxiao Li, Yuan Zhou, Nicha Dvornek, Muhan Zhang, Siyuan Gao, Juntang Zhuang, Dustin Scheinost, Lawrence H Staib, Pamela Ventola, and James S Duncan. Braingnn: Interpretable brain graph neural network for fmri analysis. *Medical Image Analysis*, 74:102233, 2021.
 - [45] Zewen Li, Fan Liu, Wenjie Yang, Shouheng Peng, and Jun Zhou. A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE transactions on neural networks and learning systems*, 33(12):6999–7019, 2021.
 - [46] Tailin Liang, John Glossner, Lei Wang, Shaobo Shi, and Xiaotong Zhang. Pruning and quantization for deep neural network acceleration: A survey. *Neurocomputing*, 461:370–403, 2021.
 - [47] Jiaying Liu, Feng Xia, Xu Feng, Jing Ren, and Huan Liu. Deep graph learning for anomalous citation detection. *IEEE Transactions on Neural Networks and Learning Systems*, 33(6):2543–2557, 2022.
 - [48] Brent Daniel Mittelstadt and Luciano Floridi. The ethics of big data: current and foreseeable issues in biomedical contexts. *The ethics of biomedical big data*, pages 445–480, 2016.
 - [49] Menno Mostert, Annelien L Bredenoord, Monique CIH Biesaart, and Johannes JM Van Delden. Big data in medical research and eu data protection law: challenges to the consent or anonymise approach. *European Journal of Human Genetics*, 24(7):956–960, 2016.
 - [50] Mark EJ Newman and Michelle Girvan. Finding and evaluating community structure in networks. *Physical review E*, 69(2):026113, 2004.
 - [51] Alexander Quinn Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models. In *International conference on machine learning*, pages 8162–8171. PMLR, 2021.
 - [52] Shima Ovaysikia, Khalid A Tahir, Jason L Chan, and Joseph FX DeSouza. Word wins over face: emotional stroop effect activates the frontal cortical network. *Frontiers in human neuroscience*, 4:234, 2011.
 - [53] Donald B Plewes and Walter Kucharczyk. Physics of mri: a primer. *Journal of magnetic resonance imaging*, 35(5):1038–1054, 2012.
 - [54] Baxter P Rogers, Victoria L Morgan, Allen T Newton, and John C Gore. Assessing functional connectivity in the human brain by fmri. *Magnetic resonance imaging*, 25(10):1347–1357, 2007.

- [55] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, proceedings, part III* 18, pages 234–241. Springer, 2015.
- [56] John Mark Michael Rumbold and Barbara Pierscionek. The effect of the general data protection regulation on medical research. *J Med Internet Res*, 19(2):e47, Feb 2017. ISSN 1438-8871. URL <http://www.jmir.org/2017/2/e47/>.
- [57] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986.
- [58] Anwar Said, Roza Bayrak, Tyler Derr, Mudassir Shabbir, Daniel Moyer, Catie Chang, and Xenofon Koutsoukos. Neurograph: Benchmarks for graph machine learning in brain connectomics. *Advances in Neural Information Processing Systems*, 36:6509–6531, 2023.
- [59] Benjamin Sanchez-Lengeling, Emily Reif, Adam Pearce, and Alexander B. Wiltschko. A gentle introduction to graph neural networks. *Distill*, 2021. DOI: [10.23915/distill.00033](https://doi.org/10.23915/distill.00033). <https://distill.pub/2021/gnn-intro>.
- [60] Alexander Schaefer, Ru Kong, Evan M Gordon, Timothy O Laumann, Xi-Nian Zuo, Avram J Holmes, Simon B Eickhoff, and BT Thomas Yeo. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri. *Cerebral cortex*, 28(9):3095–3114, 2018.
- [61] Ajay Shrestha and Ausif Mahmood. Review of deep learning algorithms and architectures. *IEEE access*, 7:53040–53065, 2019.
- [62] Levente Sipos. Orvosi adatgenerálás mély neurális generatív modellekkel, 2024. URL <https://diplomaterv.vik.bme.hu/hu/Theeses/Orvosi-adatgeneralas-mely-neuralis-generativ>.
- [63] Stephen M Smith. Overview of fmri analysis. *The British Journal of Radiology*, 77(suppl_2):S167–S175, 2004.
- [64] Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. *arXiv preprint arXiv:2010.02502*, 2020.
- [65] Olaf Sporns, Giulio Tononi, and Rolf Kötter. The human connectome: a structural description of the human brain. *PLoS computational biology*, 1(4):e42, 2005.
- [66] Stephen C Strother. Evaluating fmri preprocessing pipelines. *IEEE Engineering in Medicine and Biology Magazine*, 25(2):27–41, 2006.

- [67] Bradley P Sutton, Joshua Goh, Andrew Hebrank, Robert C Welsh, Michael WL Chee, and Denise C Park. Investigation and validation of intersite fmri studies using the same imaging hardware. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 28(1):21–28, 2008.
- [68] Kenji Suzuki. Overview of deep learning in medical imaging. *Radiological physics and technology*, 10(3):257–273, 2017.
- [69] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.
- [70] David C Van Essen, Kamil Ugurbil, Edward Auerbach, Deanna Barch, Timothy EJ Behrens, Richard Bucholz, Acer Chang, Liyong Chen, Maurizio Corbetta, Sandra W Curtiss, et al. The human connectome project: a data acquisition perspective. *NeuroImage*, 62(4):2222–2231, 2012.
- [71] Gaël Varoquaux and R. Cameron Craddock. Learning and comparing functional connectomes across subjects. *NeuroImage*, 80:405–415, 2013. ISSN 1053-8119. URL <https://www.sciencedirect.com/science/article/pii/S1053811913003340>. Mapping the Connectome.
- [72] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- [73] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.
- [74] Lebo Wang, Kaiming Li, and Xiaoping P Hu. Graph convolutional network for fmri analysis based on connectivity neighborhood. *Network Neuroscience*, 5(1):83–95, 2021.
- [75] Max Welling and Yee W Teh. Bayesian learning via stochastic gradient langevin dynamics. In *Proceedings of the 28th international conference on machine learning (ICML-11)*, pages 681–688. Citeseer, 2011.
- [76] 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.
- [77] Anderson Winkler. The nifti file format, 2012. URL <https://brainder.org/2012/09/23/the-nifti-file-format/>. Accessed: 2025-04-26.

- [78] Shiwen Wu, Fei Sun, Wentao Zhang, Xu Xie, and Bin Cui. Graph neural networks in recommender systems: a survey. *ACM Computing Surveys*, 55(5):1–37, 2022.
- [79] Mingda Xu, Peisheng Qian, Ziyuan Zhao, Zeng Zeng, Jianguo Chen, Weide Liu, and Xulei Yang. Graph neural networks for protein-protein interactions—a short survey. *arXiv preprint arXiv:2404.10450*, 2024.
- [80] Si Zhang, Hanghang Tong, Jiejun Xu, and Ross Maciejewski. Graph convolutional networks: a comprehensive review. *Computational Social Networks*, 6(1):1–23, 2019.
- [81] Xiao-Meng Zhang, Li Liang, Lin Liu, and Ming-Jing Tang. Graph neural networks and their current applications in bioinformatics. *Frontiers in genetics*, 12:690049, 2021.
- [82] Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. Graph neural networks: A review of methods and applications. *AI Open*, 1:57–81, 2020. ISSN 2666-6510. URL <https://www.sciencedirect.com/science/article/pii/S2666651021000012>.