Supplementary Material: Whole Brain Vessel Graphs: A Dataset and Benchmark for Graph Learning and Neuroscience (VesselGraph)

Johannes C. Paetzold†‡, Julian McGinnis†, Suprosanna Shit†, Ivan Ezhov†, Paul Büschl†, Chinmay Prabhakar*, Mihail I. Todorov‡, Anjany Sekuboyina*, Georgios Kaissis†, Ali Ertürk‡, Stephan Günnemann†, Bjoern H. Menze*
†Technical University of Munich, ‡Helmholtz Zentrum München, *University of Zürich
johannes.paetzold@tum.de

Abstract

Biological neural networks define the brain function and intelligence of humans and other mammals, and form ultra-large, spatial, structured graphs. Their neuronal organization is closely interconnected with the spatial organization of the brain's microvasculature, which supplies oxygen to the neurons and builds a complementary spatial graph. This vasculature (or the vessel structure) plays an important role in neuroscience; for example, the organization of (and changes to) vessel structure can represent early signs of various pathologies, e.g. Alzheimer's disease or stroke. Recently, advances in tissue clearing have enabled whole brain imaging and segmentation of the entirety of the mouse brain's vasculature. Building on these advances in imaging, we are presenting an extendable dataset of whole-brain vessel graphs based on specific imaging protocols. Specifically, we extract vascular graphs using a refined graph extraction scheme leveraging the volume rendering engine *Voreen* and provide them in an accessible and adaptable form through the *OGB* and *PyTorch Geometric* dataloaders. Moreover, we benchmark numerous state-of-the-art graph learning algorithms on the biologically relevant tasks of *vessel prediction* and *vessel classification* using the introduced vessel graph dataset. Our work paves a path towards advancing graph learning research into the field of neuroscience. Complementarily, the presented dataset raises challenging graph learning research questions for the machine learning community, in terms of incorporating biological priors into learning algorithms, or in scaling these algorithms to handle sparse,spatial graphs with millions of nodes and edges.[1]

1 Introduction

Human intelligence and brain function are defined by the cerebral biological neuronal network, the so-called *connectome*. The entirety of all single neurons forms an ultra-large, spatial, hierarchical and structured graph. Imaging and reconstructing these whole-brain graphs on a single-neuron level is one of the key problems in neuroscience. Neuronal organization is closely linked to the vascular network, as vessels supply the neurons with nutrients (e.g. oxygen). Specifically, the vessel topology determines the maximum metabolic load and determines neural growth patterns [1]. Vascular organisation, particularly in regards to vessel sizes and numbers of capillary links, varies substantially between brain regions, see Supplementary Figure 5 and 6. Moreover, its organization and changes to its structure are early signs for the development of specific diseases, e.g. Alzheimer's disease [2, 3]

[1] All datasets and code are available for download at `https://github.com/jocpae/VesselGraph`.
Ali Ertük, Stephan Günnemann and Bjoern H. Menze share last authorship.

35th Conference on Neural Information Processing Systems (NeurIPS 2021) Track on Datasets and Benchmarks.

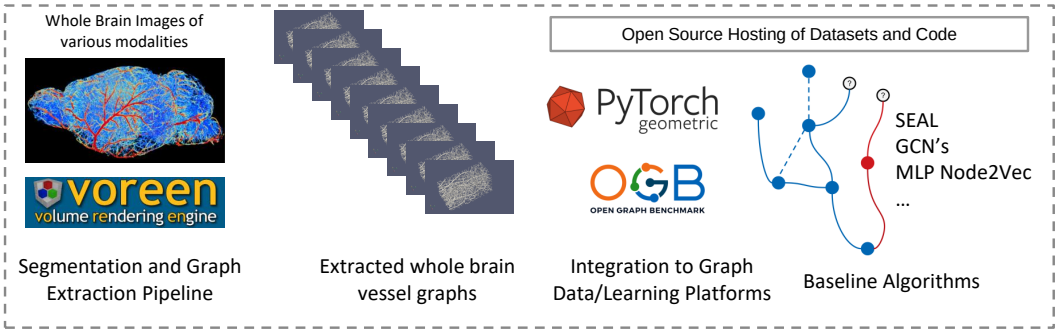

Figure 1: Graphical Abstract of *VesselGraph*.

or even COVID-19 encephalopathy [4]. As an initial step towards understanding the neuronal and vascular connectome (also known as the *angiome* [5]), reliable imaging and segmentation methods are required. To this day, whole-brain imaging and segmentation of all neurons in the brain remains elusive. On the other hand, advances in tissue clearing and deep learning have enabled imaging and segmentation of the whole murine brain vasculature down to the microcapillary level [6, 7].

Nonetheless, a binary segmentation of the vasculature is insufficient for a full, abstract description of the vascular connectome. To enable a comprehensive hierarchical description of the spatial vessel structure and anatomy, a graph representation of the brain with detailed features is required. This work provides the first large-scale, reproducible graph dataset thereof.

We believe that such a graph representation can facilitate research and understanding in many fields. The correction of imperfect vascular imaging and segmentation based on such an enhanced vascular graph, could one day enable the simulation of blood-flow (hemodynamic modeling), the study of vessel anatomy, connectivity, collateralization/anastomosis and structural abnormalities. Future studies using enhanced datasets could find our approach useful to study pathologies associated with neurovascular disorders, such as stroke and dementia, given that obstacles such as plaques would be accounted for.

Evidently, the study of such spatial graphs with millions of nodes requires its own set of methods; we believe that the recent rise of advanced machine learning methods for graphs will provide suitable approaches to efficiently and accurately permit drawing deep insight from vascular graphs. This, in turn, will foster the development of methods capable of dealing with massive, but sparsely connected circular graphs, for inference on these graphs, and inference under structural and functional prior constraints that are present in such spatial physical 3D networks.

In this work we benchmark two exemplary and biologically relevant tasks using both traditional approaches and advanced graph learning. First, in order to improve the structure and anatomical fidelity of the extracted graphs, we benchmark vessel (link) prediction. As a second task, we benchmark vessel (node) classification into the three main classes (capillaries, arterioles/venules, and arteries/veins), which represent biologically meaningful classifications by vessel size, and whose relevance for hemodynamics has been demonstrated in stroke and oxygenation modeling [8].

1.1 Whole brain vascular imaging and segmentation

Novel imaging methods, e.g. tissue-clearing-based methods [9–12], VesSAP [6], Tubemap [7] and the work by diGiovanna et al. [13] have enabled the imaging of the full vascular structure on a whole-brain scale [1].

The segmentation of the resulting ultra-large and unbalanced images with thousands of pixels in each dimension (e.g. $3096 \times 4719 \times 1867$ pixels [6]) is a challenging computer vision task which is strongly affected by technical imaging imperfections. The best-performing segmentation approaches rely on deep learning, e.g. using the U-Net architecture, and are only trained on selected, manually annotated sub-volumes of the whole brain images [6, 7, 1], leading to further imperfections in the segmentation masks.

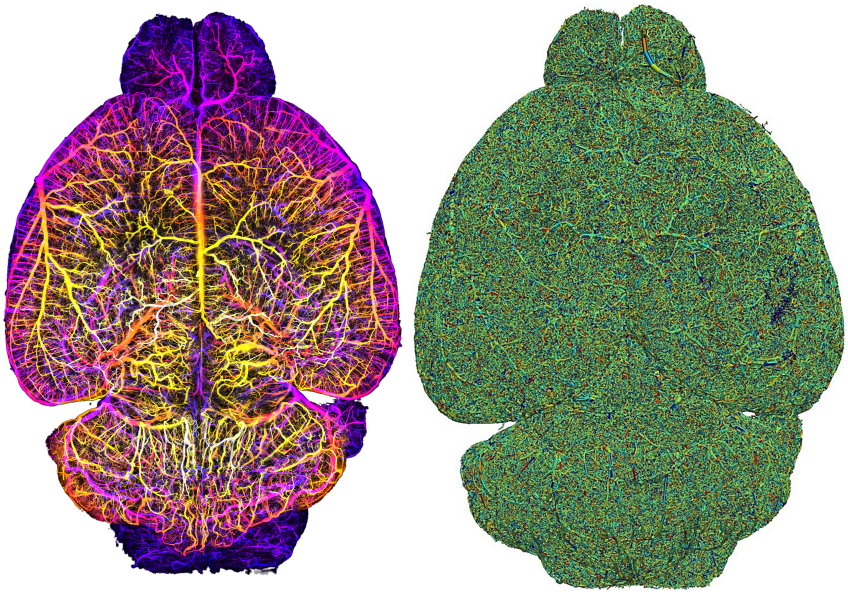

Figure 2: Left: 3D imaging of the whole mouse brain vasculature [6] and right; the corresponding rendering of our whole brain spatial vessel graphs; the edges (vessels) are rendered with the average radius feature.

The process presented in our work commences with segmentations of whole-brain vessel images, for which we use publicly available data from lightsheet microscopy (VesSAP), two-photon microscopy and a synthetic blood vessel dataset. For details refer to Appendix B.4. In the future, we will continuously increase the dataset with whole-brain images and segmentation as they become publicly available.

1.2 Graph learning

Machine learning on graphs is a highly relevant research field which aims to develop efficient machine learning algorithms exploiting the unique properties of graphs, such as structure neighborhoods and the sparse representation of complex systems. Our work concerns a particularly challenging domain - spatial, structured and ultra large biological graphs. In this paper we utilize and benchmark two fundamental graph learning tasks: node classification and link prediction to study the biological properties of the vascular connectome.

A widely recognized concept for node classification is the adaption of deep learning techniques to graphs via graph convolutional networks (GCN) [14], a concept which was adapted and extended for many of the algorithms that we implemented, such as such as GNNs, GCNs, and GAEs [15–23]. A key approach for link prediction is a so-called *labeling trick* [24], which is a concept to generate sensible training data. The SEAL labeling trick used in our work constructs a subgraph for two candidate nodes (enclosing subgraph) and aims to learn a functional mapping on the subgraph to predict link existence [25].

1.3 Our contribution

Our main contributions are:

1. We extract a set of standardized whole-brain vessel graphs based on whole, segmented murine brain images.
2. We publicly release said dataset in an easily accessible and adaptable format for use in graph learning benchmarking by implementing the *open graph benchmark* (OGB) [26] and *PyTorch Geometric* data loaders [27].
3. In addition to our standard vessel graph, in which bifurcation points are nodes and vessels are edges, we propose an alternative representation of the vascular connectome as a line

graph (where vessels become nodes), enabling the use of a multitude of advanced *node classification algorithms* for vessel property prediction.

4. We extensively benchmark graph algorithms for the biologically important tasks of **link prediction** and **node classification**, which can serve as baselines for further research efforts in graph learning and neuroscience.

The rest of the paper is organized as follows: In Section 2, we describe our refined graph generation process and provide implementation details for the used *voreen* framework and compare to other graph generation methods. We introduce the structure of our 3D brain vessel graph and provide statistics on the different extracted graphs from different brains in Section 3.1. We describe how we generated an alternative line graph representation in Section 3.2. In Section 4, we benchmark the link prediction task and in Section 5, we benchmark the node classification task on a multitude of baseline algorithms. We conclude with a focused discussion of our contribution and outline future perspectives and topics related to dataset maintenance.

2 Graph extraction from segmentations

Our graph extraction protocol begins with a given segmented whole-brain vascular dataset. Independent of segmentation method used (deep learning or filter-based), we tested the following state-of-the-art graph extraction algorithms: 1) the TubeMap method [7] which uses pruning on a 27-neighborhood skeletonization after a deep learning based tube-filling algorithm, based on a modified DeepVesselNet architecture [28]; 2) the metric graph reconstruction algorithm by Aanjaneya et al. [29] which reduces linear connections of a skeleton to form a more compact and topologically correct graph and 3) the *Voreen* vessel graph extraction method [30, 31]. We tested the graph extraction algorithms on different imaging modalities, varying brain areas, and the synthetically generated vascular trees [32].

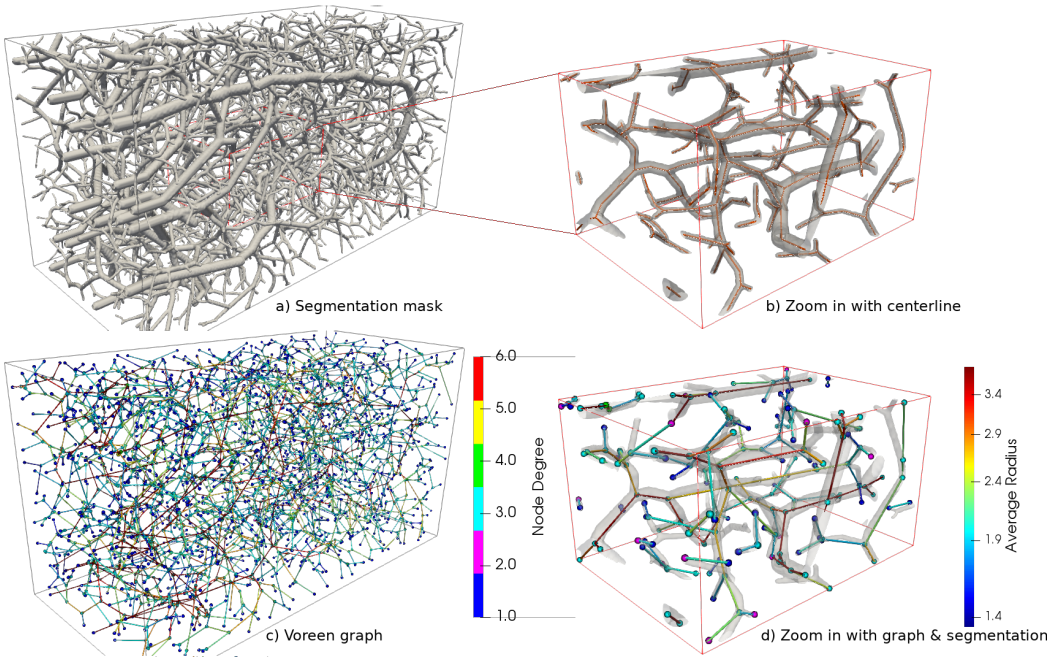

Figure 3: Extracted spatial vessel graph on a synthetic vessel volume [32]; the graph is extracted using the *Voreen* software [30]; a) the original vascular segmentation rendered in rendered in grey; b) depiction of the centerlines in red for a zoomed-in section; c) the nodes with a discrete colorbar encoding their degree; d) depiction of the segmentation with the edges and a continuous colorbar encoding the radius.

After expert-level evaluation of the extracted graphs in terms of feature quality, graph robustness and pipeline parameters, and of the algorithms in terms of scalability, runtime and resource constraints, we

selected *Voreen* [31] for our graph generation. For details and comparisons we refer to Supplementary section C.1.

Voreen (Volume Rendering Engine) is a general framework for multi-modal volumetric dataset visualization and analysis purposes. One key advantage of *Voreen* compared to other graph generation algorithms, is that its graph extraction process is deterministic, robust and scalable. It has successfully been applied to cosmological visualization [33], visualization of large volumetric multi-channel microscopy data [34], 3D visualization of the lymphatic vasculature [35], 3D histopathology of lymphatic malformations [36] and velocity mapping of the aortic flow in mice [37].

Our graph extraction follows a four-stage protocol:

1. Skeletonization: The binary segmentation volume is reduced to a skeleton based representation by applying a standard topological thinning algorithm by Lee et al. [38].
2. Topology Extraction: memory efficient algorithms extract the vessel centerlines [39]. *Voreen* allows to store this intermediate representation in a combination with the graph.
3. Voxel-Branch Assignment: Computing of mapping between the so-called protograph (i.e. the initial graph) and the voxels of the binary segmentation.
4. Feature Extraction: On basis of the protograph and the mapping, several features can be computed from the foreground segmentation.

Multiple iterations of the four-stage protocol refine and improve the graph quality and prune small, spurious branches. The key optimization parameter for the graph structure in terms of node representation, and node statistics is the *bulge size*. Expert neuroscientists determined the parameter (bulge size = 3, a parameter choice in line with previous work [31]) by statistically comparing the resulting graphs, and visually interpreting the vascular connections in varying brain regions (compare Supplementary Figure 6). Still, known limitations of topological thinning-based methods for graph extraction exist [31], motivating our first baseline task, presented in Section 4.

3 3D vessel graph dataset

Complete Datasets			
Name	Number of Nodes	Num of Edges	Node Degree
BALBc1 [6]	3,538,495	5,345,897	3.02
BALBc2	3,451,306	5,193,775	3.01
BALBc3	2,850,347	4,097,953	2.88
C57BL/6-1	3,820,133	5,614,677	2.94
C57BL/6-2	3,439,962	5,070,439	2.95
C57BL/6-3	3,318,212	4,819,208	2.90
CD1-E-1	3,645,963	5,791,309	3.18
CD1-E-2	1,664,811	2,150,326	2.58
CD1-E-3	2,295,360	3,130,650	2.73
C57BL/6-K18 [1]	4,284,051	6,525,881	3.05
C57BL/6-K19	3,948,612	5,999,958	3.04
C57BL/6-K20	4,165,085	6,317,179	3.03
Synth. Graph 1 [32]	3159	3234	2.05
Synth. Graph 2	3349	3421	2.04
Synth. Graph 3	3227	3310	2.05
Synth. Graph 4	3178	3251	2.05
Synth. Graph 5	3294	3376	2.05

Table 1: Total number of edges, nodes and average node degree for the different whole brain graphs.

Our 3D vessel dataset features 17 graphs from 2 different imaging modalities as well as 5 sets of synthetic vascular graphs. We found the smaller synthetic graphs useful for prototyping since they are smaller in size and cover all three classes of vessels (arteries, arterioles and capillaries). For all real vessel graphs, the full 3D images and binary segmentations are also publicly available. An overview of the notation used throughout the following sections alongside typical values can be found in Table 2.

3.1 Vessel graph G

The output of the *Voreen* graph extraction pipeline represents our primary unweighted and undirected graph or "intuitive" vessel graph. Let this graph be denoted as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of nodes and $\mathcal{E}$ is the set of all the edges of the graph.

Nodes: From a biological perspective, each node $n \in \mathcal{V}$ in our graph either represents end points of the vessel branches or the bifurcation of vessel branches, (see Figure 4). Bifurcation points are the points where a larger vessel branches into two or more smaller vessels (in case of an artery) or smaller vessels merge into a large vessel (in case of a vein). The number of vessels branching from a bifurcation point defines the degree of that particular node. Bifurcation points have node degree of 3 or higher. In some cases, our graphs also have vessel endpoints, which are encoded as nodes of degree 1. Further, degree 2 nodes are generated by the graph extraction in cases when vessels exhibit a large curvature. These nodes are important to preserve the vessel curvature in its graph representation. For a statistical evaluation of the node degree please see Supplementary Figure 9.

Feature Overview			
Name	Feature Type	Value	Description
x_n	node feature	$[178, 3096]$ *	x-coordinate
y_n	node feature	$[808, 4719]$ *	y-coordinate
z_n	node feature	$[0, 1866]$ *	z-coordinate
a_n	node feature	$\{0, 1\}^{71}$	Allen mouse brain atlas region
μ_{ij}^{r}	edge feature	$[0.5, 38.65]$	mean of minimum radii
σ_{ij}^{r}	edge feature	$[0.0, 12.49]$	std. of minimum radii
$\mu_{ij}^{\bar{r}}$	edge feature	$[0.79, 38.65]$	mean of average radii
$\sigma_{ij}^{\bar{r}}$	edge feature	$[0.0, 11.99]$	std. of minimum radii
μ_{ij}^{R}	edge feature	$[0.91, 44.12]$	mean of maximum radii
σ_{ij}^{R}	edge feature	$[0.0, 23.64]$	std. of minimum radii
μ_{ij}^{o}	edge feature	$[0.04, 1.99]$	mean of roundness
σ_{ij}^{o}	edge feature	$[0.0, 1.0]$	std. of roundness
l_{ij}	edge feature	$[2, 322.81]$	vessel length
d_{ij}	edge feature	$[1.77, 300.36]$	shortest distance
ρ_{ij}	edge feature	$[0.18, 27.43]$	curvature
α_{ij}	edge feature	$[0.29, 1587.49]$	mean crosssection area
v_{ij}	edge feature	$[1.0, 119459]$	Volume of vessel
nv_{ij}	edge feature	$[0.0, 256] \cap \mathbb{N}$	no. of voxel in vessel
ν_{ij}^{1}	edge feature	$[1, 14] \cap \mathbb{N}$	degree of n_i of edge e_{ij}
ν_{ij}^{2}	edge feature	$[1, 14] \cap \mathbb{N}$	degree of n_j of edge e_{ij}

Table 2: Systematic overview of the notation of the existing node and edge features in our spatial vessel graphs. All features besides the Allen brain atlas region and the node degree are spatial and extracted using Voreen, discrete ranges are given for the *Balbc1* brain (* subject to imaging resolution).

Node features: We extract two important features for the nodes of graph $\mathcal{G}$. For each node, the key features are the physical location in the coordinate space and the anatomical location in reference to the Allen brain atlas [40]. For the physical location feature, we denote real valued coordinates $[x_n, y_n, z_n] \in \mathbb{R}^3 \; \forall n \in \mathcal{V}$ where $[x_n, y_n, z_n]$ is the location of node n in 3D space. Further, multiple prior works have shown that regional differences in vessel geometry can be observed in different brain regions [1, 41, 6]. This motivates us to include anatomical location features for the nodes. Hence, we register the whole segmentation volume to the Allen brain atlas. Our reference Atlas uses the ontology the Allen mouse brain atlas (CCFv3 201710). We use the average template. After appropriate downsampling of the Allen brain atlas and the images, we apply a two-step-rigid and deformable registration using *elastix*. Our protocol is thus identical to the Vessap paper[6]. Subsequently, we assign the brain region where a particular node is located in the brain atlas as anatomical node location feature, see Supplementary Figure 7. Formally, the anatomical location feature $a_n = c \; \forall n \in \mathcal{V}$ if $[x_n, y_n, z_n] \in \; A_c$, where A_c is the cth region of the brain atlas. The atlas includes 71 brain regions which are hierarchically clustered from > 2000 subregions. The anatomical location feature is embedded as a one-hot encoded vector.

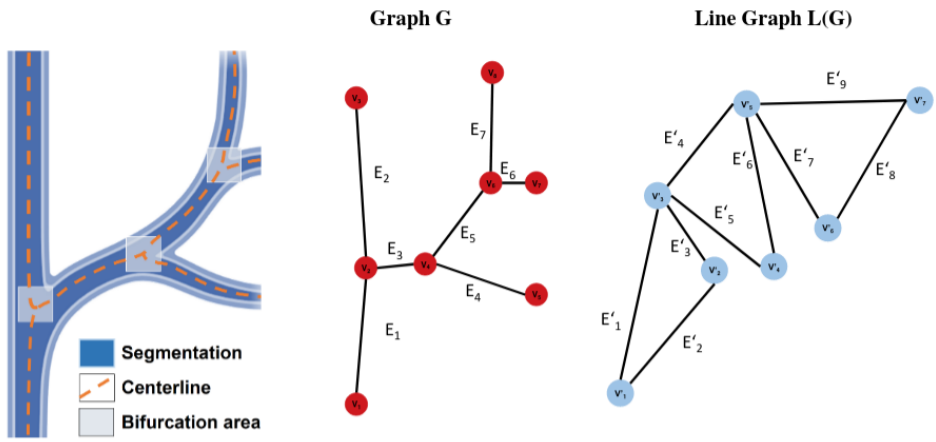

Figure 4: Depiction of an exemplary vessel tree with the the spatial vessel Graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ with nodes $(\mathcal{V})$ and edges $(\mathcal{E})$; additionally, a line graph $L(\mathcal{G})$ of the spatial vessel graph $\mathcal{G}$; where each node (bifurcation point) becomes an edge; two nodes of $L(\mathcal{G})$ are adjacent if and only if their edges are incident in $\mathcal{G}$.

Edges: Each edge $e_{ij} \in \mathcal{E}$ in our graph represents vessels or vessel segments which connect two nodes $\mathcal{V}$, see Figure 4. These edges (vessels) determine the structure of the whole brain network and represent the core aspect of our research questions. The edges exhibit the following rich set of features, which are extracted based on the shape and topology of the given segmented images.

Edge features: We extract geometric properties for each of the edges. For that, we determine the maximum diameter inscribed circle, least square reference circle, and minimum circumscribed circle on the discretized cross-section of a vessel branch and compute their radius as $\{r_{ij}^k\}, \{\bar{r}_{ij}^k\}$ and $\{R_{ij}^k\}$ where $k = 1 : K$ for K number of cross section of the edge e_{ij}, respectively. From this, we compute the mean and standard deviation of the minimum, average and maximum radius for each edge e_{ij} as follows. Specifically, $\mu_{ij}^r, \sigma_{ij}^r$ denotes the mean and standard deviation of minimum radius of edge e_{ij}. We extend the same notation for mean and standard deviation for $\{\bar{r}_{ij}^k\}$ and $\{R_{ij}^k\}$ as $\mu_{ij}^{\bar{r}}, \sigma_{ij}^{\bar{r}}, \mu_{ij}^R, \mu_{ij}^R$ respectively. We compute the roundness of each cross section as $o_{ij}^k = \frac{r_{ij}^k}{R_{ij}^k}$. We denote the mean and standard deviation of roundness as μ_{ij}^o and σ_{ij}^o, respectively. Further, we extract the vessel length l_{ij}, shortest distance between two nodes of an edge d_{ij}, curvature $\rho_{ij} = \frac{l_{ij}}{d_{ij}}$, mean cross section are α_{ij}. Moreover, we use the degree of the nodes n_i and n_j for an edge e_{ij} as ν_i and ν_j, respectively. The complete set of edge features can be found in Table 2.

3.2 Line vessel graph L(G)

As an alternative representation of whole brain vessel graphs, we convert our vessel graphs $\mathcal{G}$ to a corresponding line-graph representation, $L(\mathcal{G})$[42]. A line graph (depicted in Figure 4) is a graph where the edges of the base graph $\mathcal{G}$ become nodes and an edge between the new nodes is created if and only if their edges are incident in $\mathcal{E}$. Edges are the most important aspects in our graph $\mathcal{E}$ because of their one-to-one correspondence to the vessels. Therefore, we wish to apply another set of graph-learning algorithms, namely node classification algorithms, to study their biological properties based on the rich set of vessel features. Hence, we construct an alternative representation with the help of line graph $L(\mathcal{G})$. We formally define $L(\mathcal{G}) := (\mathcal{V}', \mathcal{E}')$ where $\mathcal{V}' = \mathcal{E}$ and $\mathcal{E}' = \{\{e_{ij}, e_{ik}\}$ if $\exists\, (e_{ij}, e_{ik}) \in \mathcal{E}\}$.

Nodes: Now, the nodes in the line graph $\mathcal{V}'$ represent vessels or vessel segments, see Figure 4.

Node features: Thus, all edge features of $\mathcal{G}$ can now be used as node features for $L(\mathcal{G})$, see Table 2. One of the key advantages of constructing the line graph is that we can now leverage a large number

of prior techniques presented in node classification literature such as the use of vessel features in message passing.

Edges: Edges are defined as pairwise adjacencies of two nodes (vessels) *if and only if* the corresponding edges in $\mathcal{G}$ are connected to a node $\mathcal{V}$. In practice, this means that nodes in $\mathcal{G}$ which are of degree 1 disappear in $L(\mathcal{G})$ and that each node in $\mathcal{G}$ with a degree ≥ 2 will create multiple edges in $L(\mathcal{G})$.

Edge features: The spatial location given as node features in $\mathcal{G}$ can now be added as an edge feature.

4 Benchmarking link prediction

The formal goal of link prediction is to train a classifier $\mathcal{F}$ which predicts links in $\mathcal{E}_{pred}$ as positive and negative labels, it can be formalized as follows $\mathcal{F} : \mathcal{E}_{pred} \rightarrow \{0, 1\}$.

From a biological perspective this task is relevant to correct missing and imperfect vessel graph connections, because the extracted graph may be over- or under-connected, due to artifacts and shortcoming of the segmentation and network extraction.

In order to provide initial baselines for vessel (link) prediction, we implemented 10 models. The following graph learning baselines were trained without edge features: the GCN by Kipf et al. [14], a GNN using the GraphSAGE operator [17] and the SEAL GNN, a network aiming to learn general graph structure features from the local subgraph [25]. Furthermore, we trained a multilayer perceptron (MLP) on full batches based on Node2Vec features [43]. Apart from these, more traditional, heuristic-based methods were implemented for the task of link prediction, which include the Katz index [44], Common Neighbour, Page Rank and Adamic Adar [45], a measure which computes the closeness of nodes. These traditional methods make predictions based on the graph structure itself.

4.1 Dataset curation - SEAL

Link sampling strategy: The curation of a balanced training dataset requires the introduction of two types of edges. Similar to the SEAL paper [25], we use the notion of *positive edges* and *negative edges*. Generally, positive edges are random samples of existing links and negative edges are samples of non-existent links between randomly chosen nodes of the dataset (which are included in the adjacency matrix). For positive edges, we utilize random samples of the existing edges of each graph. However, since our dataset includes 3D coordinates as the node features, their spatial nature makes selecting negative samples more challenging. A trivial random selection, which has been used in other state-of-the-art methods such as SEAL, would lead to biologically implausible edges, e.g. an edge between two nodes in different brain hemispheres. These can be easily distinguished based on the coordinates and thus would not provide useful information to the model. As such, models trained with trivial random sampling struggle with the link prediction task. To address this issue, we restrict negative edge sampling to a coordinate space which spatially surrounds the source node, and choose the target node by randomly selecting nodes that are located within the following cubic space around the source node: $\delta = \overline{l_{i,j}} + 2\sigma$, where $\overline{l_{i,j}}$ denotes the average vessel length in $\mathcal{G}$. We note that this link sampling strategy is a first baseline and could be improved upon in future work.

Experiment: For our GCN based architectures we did an extensive grid-search of hyper-parameter combinations on a subset of the whole brain graph. We subsequently trained on the whole brain graphs. This intermediate step was necessary because exploring thousands of hyper-parameter combinations on the whole brain dataset is computationally infeasible. Implementation details and details on the hyperparameter search are indicated in supplementary Table 5.

For the main experiment we sample all edges from one whole brain graph as positive edges $\mathcal{G}(\mathcal{V}, \mathcal{E})$ (BALBc-1, Vessap, see Table 1)and randomly assign these to the training, validation and test set (80/10/10 split). Moreover, we sample an identical number of negative edges, i.e. non-existent but theoretically probable links according to the curation criterion described above. Next, we randomly shuffle all negative edges. Thus, we mitigate any bias in the negative train, validation and test splits and ensure a region-independent distribution. Subsequently, we randomly assign the negative edges

to the train, validation and test set (80/10/10 split). This provides us with a balanced datast in regards to positive and negative edges.

We choose to only use the spatial node features for our experiment: x_n, y_n, z_n. This task is very hard because the algorithm essentially has to learn the vascular graph hierarchy purely on undirected relational and spatial information.

Table 3: Results for the link prediction baselines.

Algorithm	ROC AUC	
	validation	test
Adamic Adar	48.49	48.49
Common Neighbors	48.50	48.49
Resource Allocation	48.49	48.50
Matrix Factorization	50.07	50.08
MLP	57.98	58.02
GCN GCN	50.69	50.72
GCN GCN + embeddings	51.32	51.13
GCN SAGE + embeddings	52.81	52.88
GCN SAGE	59.37	59.23
SEAL	**91.01**	**90.96**

Generally, traditional methods and simple GCN models performed poorly. Among the traditional methods tested, the MLP performed best. On the other hand, the SEAL implementation reached a superior performance and a strong inductive bias (ROC AUC $> 90\%$). This improvement is in line with recent literature [24], which found a considerable performance improvement as a result of the employed labeling trick. This highlights that complex, dedicated graph-learning concepts need to be developed to address biologically inspired spatial graph challenges. A detailed experimental description and interpretation can be found in the Supplementary material, section D.1.

5 Benchmarking vessel attribute classification

Our formal goal of node classification is to train a classifier $\mathcal{F}$ which predicts a class label $\mathcal{Y}$ out of a set of possible classes $\mathcal{N}_n$ of a node $\mathcal{V}$, it can be formalized as follows $\mathcal{F} : \mathcal{V} \rightarrow \mathcal{Y} \in \mathcal{N}_n$.

Biologically, this task is relevant because the vessel radius is one of the most important parameters for blood flow; any task associated with flow modelling (such as stroke diagnosis and treatment) is heavily dependent on the diameter of the affected vessel. For example in stroke, a different treatment option is chosen based on the size of the vessel in the context of its local network topology. Therefore, reliably classifying vessel segments into categories such as arteries/veins, arterioles/venules and capillaries is relevant.

For the secondary task of vessel radius (node) classification we implemented 7 graph and non-graph learning baselines discussed in the OGB paper [26]. Among them node classification using an MLP initialized on N2Vec [43], a simple GCN [14], a GNN using the GraphSAGE operator [17], the GraphSAINT algorithm which includes a mini-batch GCN[16], the Scalable Inception Graph Neural Networks (SIGN)[18] and the Cluster-GCN algorithm[20]. Furthermore we implemented SpecMLP-W + C&S and SpecMLP-W + C&S + N2Vec, which use shallow models ignoring graph structure and standard label propagation techniques from semi-supervised learning methods [22].

Experiment: We split our three classes according to the minimum radius feature μ_{ij}^r into classes of $\mu_{ij}^r < 15\mu m; 15-40\mu m$ and $> 40\mu m$. Defined by the anatomy and properties of oxygen distribution these three classes are highly imbalanced. E.g. for the *Vessap* datasets the distribution is roughly 95%, 4% and 1%. Similarly to the link prediction task we carried out a grid search for optimal hyper-parameters, see Supplementary Table 6. We randomly split the nodes into train, validation and test sets of (80/10/10) of one whole mouse brain (BALBc-1, Vessap, see Table 1). We choose to use the following node features for our experiment: l_{ij} , d_{ij} and ρ_{ij}.

For node classification, we find acceptable to high performance in our baselines by all the methods we tested. More complex graph models such as GraphSAGE and Cluster-GCN outperform simple GCNs on average over all metrics. According to the metrics which account for class imbalance i.e.

Table 4: Results of the implemented node classification baselines. The performance scores are the weighted F1 score, one versus rest ROC AUC, class balanced accuracy and total accuracy (ACC).

	F1 Score		ROC AUC		Balanced ACC		ACC	
	valid	test	valid	test	valid	test	valid	test
GCN	75.74	75.75	67.23	66.46	58.38	56.83	62.94	62.92
GraphSAGE	81.98	81.98	**77.35**	**77.18**	**71.82**	**71.33**	72.02	71.98
GraphSAINT	77.46	77.40	71.38	70.71	63.74	62.51	64.88	64.84
SIGN	74.46	74.49	67.26	66.04	57.90	55.88	61.25	61.27
Cluster-GCN	**86.10**	**86.06**	**77.91**	**77.43**	**72.23**	**71.87**	**77.47**	**77.41**
MLP	76.11	76.11	58.08	57.79	42.36	41.72	63.65	63.61
SpecMLP-W + C&S	84.48	84.55	58.12	58.54	42.20	42.93	75.84	75.91
SpecMLP-W + + N2Vec	80.53	80.63	66.69	66.20	59.04	57.90	69.99	70.10

ROC AUC and balanced ACC, graph neural networks outperform non-graph learning methods, for a detailed interpretation see Supplementary D.2.

6 Discussion

In this work, we introduce and make publicly available a large dataset of vessel graphs representing the most comprehensive and highest resolution representation of the whole vascular connectome to-date. We provide this set of graphs as a new "baseline dataset" for machine learning on graphs and make it re-usable and easily accessible by leveraging widely employed open standards, such as the *OGB* and *PyTorch Geometric* dataloaders.

To provide an example for the utilization of our dataset and to promote graph machine learning research in neuroscience, we provide two benchmarks: First, we benchmark vessel (link) prediction to improve the vascular connectome; second, we implement vessel (node) classification into three main anatomical categories on the line graph. We thus show that graph learning-based methods outperform traditional methods for vessel (node) classification. Moreover, we demonstrate that link prediction based solely on the spatial organization is a difficult task for most algorithms. However, we provide evidence that the combination of an appropriately chosen, complex GNN model (SEAL) with a labeling trick can achieve high accuracy on this task, paving the way for dedicated machine learning research on spatial (biological) graphs as a key to unlocking biological insight.

Dataset bias: While the dataset and the evaluation we provide are thorough, we note the following bias in our work: Our vascular graphs are constrained by the technical bias and limitations inherent to experimental imaging, such as artifacts in the clearing protocol and physical limitations concerning the resolution and isotropy of the microscopy. All specimen imaged in this study are males. Moreover, even state-of-the-art deep learning methods for segmentation presented in literature are only trained on incomplete sets of labeled data, leading to a model bias in segmentation. Further problems can occur from the known limitations of topological thinning-based methods for graph extraction [31].

Limitations: The sum of these effects and bias can impair the usefulness of our dataset for certain, highly specialised tasks, such as flow simulations using the Navier-Stokes equations, which are strongly dependent on accurate radius measurements.

Moreover, benchmarking all available features, data and concepts was beyond the scope of our work. For instance, an extension to heterogeneous graph representations [46, 47], the utilization of more features, the inclusion of more than one graph or of weighted graphs, where e.g. all edges (vessels) are weighted depending on an embedding of their radius, may facilitate an improved interpretation. In summary, we are convinced that both the machine learning concepts and the biological insight arising from our work can be translated to other tasks, such as graph extraction and refinement on different vascular or neuronal imaging techniques, artery and vein classification, and even vessel classification in inherently different medical imaging protocols such as angiography for stroke diagnosis. We are thus hopeful that our provision of high-quality data and strong baselines will stimulate future research in this area.

Acknowledgments and Disclosure of Funding

We thank Dominik Drees for his advice and help with setting up the *Voreen* pipeline, and Muhan Zhang for his advice on enhancing SEAL. Moreover, we would like to thank Mattias Fey, Weihua Hu and the OGB Team for providing PyTorch geometric and baseline implementations.

References

[1] Xiang Ji, Tiago Ferreira, Beth Friedman, Rui Liu, Hannah Liechty, Erhan Bas, Jayaram Chandrashekar, and David Kleinfeld. Brain microvasculature has a common topology with local differences in geometry that match metabolic load. *Neuron*, 109(7):1168–1187, 2021.

[2] Eszter Farkas, Gineke I De Jong, Rob AI de Vos, ENH Jansen Steur, and Paul GM Luiten. Pathological features of cerebral cortical capillaries are doubled in alzheimer's disease and parkinson's disease. *Acta neuropathologica*, 100(4):395–402, 2000.

[3] Rachel E Bennett, Ashley B Robbins, Miwei Hu, Xinrui Cao, Rebecca A Betensky, Tim Clark, Sudeshna Das, and Bradley T Hyman. Tau induces blood vessel abnormalities and angiogenesis-related gene expression in p301l transgenic mice and human alzheimer's disease. *Proceedings of the National Academy of Sciences*, 115(6):E1289–E1298, 2018.

[4] Marjolaine Uginet, Gautier Breville, Jérémy Hofmeister, Paolo Machi, Patrice H Lalive, Andrea Rosi, Aikaterini Fitsiori, Maria Isabel Vargas, Frederic Assal, Gilles Allali, et al. Cerebrovascular complications and vessel wall imaging in covid-19 encephalopathy—a pilot study. *Clinical neuroradiology*, pages 1–7, 2021.

[5] Pablo Blinder, Philbert S Tsai, John P Kaufhold, Per M Knutsen, Harry Suhl, and David Kleinfeld. The cortical angiome: an interconnected vascular network with noncolumnar patterns of blood flow. *Nature neuroscience*, 16(7):889–897, 2013.

[6] Mihail Ivilinov Todorov, Johannes Christian Paetzold, Oliver Schoppe, Giles Tetteh, Suprosanna Shit, Velizar Efremov, Katalin Todorov-Völgyi, Marco Düring, Martin Dichgans, Marie Piraud, et al. Machine learning analysis of whole mouse brain vasculature. *Nature methods*, 17(4):442–449, 2020.

[7] Christoph Kirst, Sophie Skriabine, Alba Vieites-Prado, Thomas Topilko, Paul Bertin, Gaspard Gerschenfeld, Florine Verny, Piotr Topilko, Nicolas Michalski, Marc Tessier-Lavigne, et al. Mapping the fine-scale organization and plasticity of the brain vasculature. *Cell*, 180(4):780–795, 2020.

[8] Franca Schmid, Giulia Conti, Patrick Jenny, and Bruno Weber. The severity of microstrokes depends on local vascular topology and baseline perfusion. *Elife*, 10:e60208, 2021.

[9] Hiroki R Ueda, Ali Ertürk, Kwanghun Chung, Viviana Gradinaru, Alain Chédotal, Pavel Tomancak, and Philipp J Keller. Tissue clearing and its applications in neuroscience. *Nature Reviews Neuroscience*, 21(2):61–79, 2020.

[10] Ali Ertürk et al. Three-dimensional imaging of solvent-cleared organs using 3DISCO. *Nature Protocols*, 7(11):1983, 2012.

[11] Kwanghun Chung and Karl Deisseroth. Clarity for mapping the nervous system. *Nature methods*, 10(6):508–513, 2013.

[12] Nicolas Renier, Zhuhao Wu, David J Simon, Jing Yang, Pablo Ariel, and Marc Tessier-Lavigne. idisco: a simple, rapid method to immunolabel large tissue samples for volume imaging. *Cell*, 159(4):896–910, 2014.

[13] Antonino Paolo Di Giovanna et al. Whole-brain vasculature reconstruction at the single capillary level. *Scientific reports*, 8(1):12573, 2018.

[14] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

[15] Thomas N. Kipf and Max Welling. Variational graph auto-encoders, 2016.

[16] Hanqing Zeng, Hongkuan Zhou, Ajitesh Srivastava, Rajgopal Kannan, and Viktor Prasanna. Graphsaint: Graph sampling based inductive learning method. *arXiv preprint arXiv:1907.04931*, 2019.

[17] William L Hamilton, Rex Ying, and Jure Leskovec. Inductive representation learning on large graphs. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pages 1025–1035, 2017.

[18] Fabrizio Frasca, Emanuele Rossi, Davide Eynard, Ben Chamberlain, Michael Bronstein, and Federico Monti. Sign: Scalable inception graph neural networks. *arXiv preprint arXiv:2004.11198*, 2020.

[19] Johannes Klicpera, Aleksandar Bojchevski, and Stephan Günnemann. Predict then propagate: Graph neural networks meet personalized pagerank. *arXiv preprint arXiv:1810.05997*, 2018.

[20] Wei-Lin Chiang, Xuanqing Liu, Si Si, Yang Li, Samy Bengio, and Cho-Jui Hsieh. Cluster-gcn: An efficient algorithm for training deep and large graph convolutional networks. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 257–266, 2019.

[21] Kezhi Kong, Guohao Li, Mucong Ding, Zuxuan Wu, Chen Zhu, Bernard Ghanem, Gavin Taylor, and Tom Goldstein. Flag: Adversarial data augmentation for graph neural networks, 2020.

[22] Qian Huang, Horace He, Abhay Singh, Ser-Nam Lim, and Austin R Benson. Combining label propagation and simple models out-performs graph neural networks. *arXiv preprint arXiv:2010.13993*, 2020.

[23] Johannes Klicpera, Janek Groß, and Stephan Günnemann. Directional message passing for molecular graphs. In *International Conference on Learning Representations*, 2019.

[24] Muhan Zhang, Pan Li, Yinglong Xia, Kai Wang, and Long Jin. Revisiting graph neural networks for link prediction, 2021.

[25] Muhan Zhang and Yixin Chen. Link prediction based on graph neural networks, 2018.

[26] Weihua Hu, Matthias Fey, Marinka Zitnik, Yuxiao Dong, Hongyu Ren, Bowen Liu, Michele Catasta, and Jure Leskovec. Open graph benchmark: Datasets for machine learning on graphs. *arXiv preprint arXiv:2005.00687*, 2020.

[27] Matthias Fey and Jan Eric Lenssen. Fast graph representation learning with pytorch geometric. *arXiv preprint arXiv:1903.02428*, 2019.

[28] Giles Tetteh et al. Deepvesselnet: Vessel segmentation, centerline prediction, and bifurcation detection in 3-d angiographic volumes. *arXiv preprint arXiv:1803.09340*, 2018.

[29] Mridul Aanjaneya, Frederic Chazal, Daniel Chen, Marc Glisse, Leonidas J Guibas, and Dmitriy Morozov. Metric graph reconstruction from noisy data. In *Proceedings of the twenty-seventh annual symposium on Computational geometry*, pages 37–46, 2011.

[30] Jennis Meyer-Spradow, Timo Ropinski, Jörg Mensmann, and Klaus Hinrichs. Voreen: A rapid-prototyping environment for ray-casting-based volume visualizations. *IEEE Computer Graphics and Applications*, 29(6):6–13, 2009.

[31] Dominik Drees, Aaron Scherzinger, René Hägerling, Friedemann Kiefer, and Xiaoyi Jiang. Scalable robust graph and feature extraction for arbitrary vessel networks in large volumetric datasets. *arXiv preprint arXiv:2102.03444*, 2021.

[32] Matthias Schneider et al. Tissue metabolism driven arterial tree generation. *Med Image Anal.*, 16(7):1397–1414, 2012.

[33] Aaron Scherzinger, Tobias Brix, Dominik Drees, Andreas Völker, Kiril Radkov, Niko Santalidis, Alexander Fieguth, and Klaus H Hinrichs. Interactive exploration of cosmological dark-matter simulation data. *IEEE computer graphics and applications*, 37(2):80–89, 2017.

[34] Tobias Brix, Jörg-Stefan Praßni, and Klaus Hinrichs. Visualization of large volumetric multi-channel microscopy data streams on standard pcs. *arXiv preprint arXiv:1407.2074*, 2014.

[35] Cathrin Dierkes, Aaron Scherzinger, and Friedemann Kiefer. Three-dimensional visualization of the lymphatic vasculature. In *Lymphangiogenesis*, pages 1–18. Springer, 2018.

[36] René Hägerling, Dominik Drees, Aaron Scherzinger, Cathrin Dierkes, Silvia Martin-Almedina, Stefan Butz, Kristiana Gordon, Michael Schäfers, Klaus Hinrichs, Pia Ostergaard, et al. Vipar, a quantitative approach to 3d histopathology applied to lymphatic malformations. *JCI insight*, 2(16), 2017.

[37] Philipp Rene Bovenkamp, Tobias Brix, Florian Lindemann, Richard Holtmeier, Desiree Abdurrachim, Michael T Kuhlmann, Gustav J Strijkers, Jörg Stypmann, Klaus H Hinrichs, and Verena Hoerr. Velocity mapping of the aortic flow at 9.4 t in healthy mice and mice with induced heart failure using time-resolved three-dimensional phase-contrast mri (4d pc mri). *Magnetic Resonance Materials in Physics, Biology and Medicine*, 28(4):315–327, 2015.

[38] Ta-Chih Lee et al. Building skeleton models via 3-D medial surface axis thinning algorithms. *CVGIP: Graphical Models and Image Processing*, 56(6):462–478, 1994.

[39] Martin Isenburg and Jonathan Shewchuk. Streaming connected component computation for trillion voxel images. In *Workshop on Massive Data Algorithmics*, volume 2, 2009.

[40] Susan M Sunkin, Lydia Ng, Chris Lau, Tim Dolbeare, Terri L Gilbert, Carol L Thompson, Michael Hawrylycz, and Chinh Dang. Allen brain atlas: an integrated spatio-temporal portal for exploring the central nervous system. *Nucleic acids research*, 41(D1):D996–D1008, 2012.

[41] Artur Hahn, Julia Bode, Allen Alexander, Kianush Karimian-Jazi, Katharina Schregel, Daniel Schwarz, Alexander C Sommerkamp, Thomas Krüwel, Amir Abdollahi, Wolfgang Wick, et al. Large-scale characterization of the microvascular geometry in development and disease by tissue clearing and quantitative ultramicroscopy. *Journal of Cerebral Blood Flow & Metabolism*, 41(7):1536–1546, 2021.

[42] Jonathan L Gross, Jay Yellen, and Mark Anderson. *Graph theory and its applications*. Chapman and Hall/CRC, 2018.

[43] 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.

[44] Leo Katz. A new status index derived from sociometric analysis. *Psychometrika*, 18(1):39–43, 1953.

[45] Lada A Adamic and Eytan Adar. Friends and neighbors on the web. *Social networks*, 25(3):211–230, 2003.

[46] Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne Van Den Berg, Ivan Titov, and Max Welling. Modeling relational data with graph convolutional networks. In *European semantic web conference*, pages 593–607. Springer, 2018.

[47] Shikhar Vashishth, Soumya Sanyal, Vikram Nitin, and Partha Talukdar. Composition-based multi-relational graph convolutional networks. *arXiv preprint arXiv:1911.03082*, 2019.

[48] Eva L Dyer, William Gray Roncal, Judy A Prasad, Hugo L Fernandes, Doga Gürsoy, Vincent De Andrade, Kamel Fezzaa, Xianghui Xiao, Joshua T Vogelstein, Chris Jacobsen, et al. Quantifying mesoscale neuroanatomy using x-ray microtomography. *Eneuro*, 4(5), 2017.

[49] Jingpeng Wu, Yong He, Zhongqin Yang, Congdi Guo, Qingming Luo, Wei Zhou, Shangbin Chen, Anan Li, Benyi Xiong, Tao Jiang, et al. 3d braincv: simultaneous visualization and analysis of cells and capillaries in a whole mouse brain with one-micron voxel resolution. *Neuroimage*, 87:199–208, 2014.

[50] Evan Calabrese, Alexandra Badea, Gary Cofer, Yi Qi, and G Allan Johnson. A diffusion mri tractography connectome of the mouse brain and comparison with neuronal tracer data. *Cerebral cortex*, 25(11):4628–4637, 2015.

[51] Tianqi Li, Chao J Liu, and Taner Akkin. Contrast-enhanced serial optical coherence scanner with deep learning network reveals vasculature and white matter organization of mouse brain. *Neurophotonics*, 6(3):035004, 2019.

[52] Arttu Miettinen, Antonio Zippo, Alessandra Patera, Anne Bonnin, Sarah Shahmoradian, Gabriele Biella, and Marco Stampanoni. Micrometer-resolution reconstruction and analysis of whole mouse brain vasculature by synchrotron-based phase-contrast tomographic microscopy. *bioRxiv*, 2021.

[53] Solomon Kullback and Richard A Leibler. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86, 1951.

[54] Rex Ying, Dylan Bourgeois, Jiaxuan You, Marinka Zitnik, and Jure Leskovec. Gnnexplainer: Generating explanations for graph neural networks. *Advances in neural information processing systems*, 32:9240, 2019.