TOPOLOGICAL DATA ANALYSIS FOR SELECTION OF MACHINE LEARNING MODELS IN CEREBRAL STROKE DETECTION WITH LIMITED RESOURCES

Aleksandra Vatian, Natalia Gusarova, Ivan Tomilov, Pavel Brunko
and Alexey Zubanenko
ITMO University, Sankt-Petersburg 197101, Russia

ABSTRACT

Rapid and reliable diagnosis of cerebral stroke is a vital necessity, and among them, ischemic stroke is the most difficult to recognize on MRI images. Increasing the efficiency of stroke diagnosis is associated with the transition to increasingly "heavy" AI tools working in 3D mode, processing MRI images not pixel-by-pixel, but voxel-by-voxel and using complex multifactor information processing algorithms. The implementation of such products requires large computing resources, which are often unavailable outside large medical centers. In addition, existing explainable AI tools identify the affected area very roughly and generally, which reduces the doctor's confidence in the diagnostic result. On the other hand, one can use AI models operating in 2D mode - weaker, but faster and less demanding on computing resources. The article uses a Siamese neural network as a base model. To improve classification efficiency, a model pretrain on "light" synthetic data based on Perlin noise is proposed. To objectify the choice of mode, the article uses the apparatus of topological data analysis, namely, changes in persistent entropy and Renyi entropy of data embeddings on fully connected layers of a neural network. It is experimentally confirmed that using a 2D model, when trained on slices with maximum lesion visibility, produces ROC-AUC values no worse than using a full-scale 3D SOTA models, while allowing the clinician to selectively evaluate the individual slices he or she selects. It is experimentally confirmed that simple model that contains only locally useful features can support neural network training to a level comparable to much more complex and resource-intensive generative model.

KEYWORDS

Cerebral Stroke, Topological Data Analysis, Persistent Entropy, Pre-Training, Perlin Noise

1. INTRODUCTION

The World Health Organization describes cerebral stroke as a clinical syndrome with rapidly increasing clinical signs of local or global impairment of brain function without an obvious cause other than vascular origin (Mosisa et al., 2023). The annual mortality rate from stroke is almost six million people worldwide. Stroke is the second most common cause of death among adults over 60 years of age, and the fifth most common among people aged 15 to 59 years, accounting for 11.8% of all deaths (National Collaborating Centre for Chronic Conditions, 2008; Writing Group Members et al., 2016).

The predominant diagnostic tools for early detection of stroke in modern medicine is magnetic resonance imaging (MRI) (Valliani et al., 2019). As a result, a set of sections of brain tissue is formed, which the doctor can view individually (in 2D mode), as well as in the form of a pseudo-volume image (in 3D mode) and make his diagnostic conclusion. Although stroke is a disease caused by a disruption in the blood supply to brain tissue, there are two types of strokes - hemorrhagic and ischemic, which manifest differently on MRI images. In the case of a hemorrhagic stroke, a separate vessel ruptures, and the outflowing blood immediately forms a hematoma, which is visible on MRI images as a neoplasm with structure and characteristics differing from the surrounding tissues. In ischemic stroke, on the contrary, a spasm of the vessel occurs, and, as a result, the necrosis of the tissues being supplied before. MRI manifestations of this process increase gradually and may be different in structure and characteristics depending on the location and level of the lesion. Therefore, ischemic stroke is more difficult to recognize on MRI images, and at the same time, it is vitally important to carry out the correct diagnosis at the earliest possible stages of the disease.

To objectify the diagnostic process, as well as to help the radiologist, a constantly expanding arsenal of artificial intelligence (AI) tools is being developed. The main trend in research is associated with the transition to increasingly "heavy" machine learning models. As a rule, such models work in 3D mode, i.e., they process MRI images not pixel-by-pixel, but voxel-by-voxel, and implement complex multifactor information processing algorithms, which allows for increased automation of the diagnostic process, but comes at the cost of the expansion of information and computing resources used, especially when processing high resolution images. However, in real medical practice there are several factors that impede such a transition. First, as a rule, the deployment of such models requires computer equipment with sufficient software and hardware resources that allow you to quickly calculate a large amount of data typical for 3D models. Outside a specialized medical center, such equipment is most often not available.

Second, practitioners who are directly responsible for the patient's life place high demands on the explainability of diagnostic AI tools, and for this a wide range of explainable AI tools (Speith, 2022; Saranya and Subhashini, 2023) have been proposed. However, as an analysis of the literature shows, XAI methods in the diagnosis of ischemic stroke make it possible to identify the affected area very roughly and generally, which is not enough to make a diagnostic decision. Therefore, to increase confidence in the diagnostic result, a combined intelligence procedure can be proposed: the doctor sequentially views slices of the MRI image and, relying on his expert knowledge and experience, identifies the most "suspicious" slices; on them, a selection of the affected area is carried out using AI and compared with an expert assessment. In this case, one can use AI models operating in 2D mode - weaker, but faster and less demanding on computing resources. Besides, the diagnostic process remains completely transparent to the doctor.

Obviously, the fundamental difference between 2D and 3D models lies in the amount of information simultaneously processed by the network, so it is reasonable to use an information approach to select a specific model. In this work, for this purpose, topological data analysis (Chazal and Michel, 2021) is used, specifically, the apparatus of persistent homologies.

The main results of our work are as follows:

- We have shown that in conditions of small datasets, which is typical for MRI images of ischemic stroke, the use of pre-training of difference models on "light" synthetic data allows to leverage efficiency of 2D-mode processing up to 3D-mode.

- We have shown that changes in topological characteristics (persistent entropy and Renyi entropy) of data embeddings on fully connected layers of a neural network can serve as an objective assessment of the effectiveness of the network architecture and, therefore, be used for its selection in conditions of limited resources.

- We have shown that using a 2D model, when trained on slices with maximum lesion visibility, produces ROC-AUC values no worse than using a full-scale 3D model, while allowing the clinician to selectively evaluate the individual slices he or she selects.

- We have experimentally confirmed that simple model that contains only locally useful features can support neural network training to a level comparable to much more complex and resource-intensive generative model.

The rest of the article is organized as follows. The section Background and related works presents existing work on automating stroke diagnostics, briefly describes the use of topological data analysis and persistent homologies for it, and justifies the choice of the SOTA model. The Method and materials section describes the methods and datasets used, the Results and discussion section describes the experimental results obtained and compared with the existing SOTA model. The Conclusion section completes the work...

2. BACKGROUND AND RELATED WORKS

A plethora of AI solutions, primarily based on machine learning (ML), are used for automated recognition of stroke and its localization (Ruksakulpiwat et al., 2023). All researchers note that the most difficult problem is recognizing ischemic stroke. The most effective solutions include the use of deep convolutional neural networks (Zhang et al., 2021; Stib et al., 2020; Shinohara et al., 2020), simultaneous use of images from different modalities (Nishio et al., 2020; Lu et al., 2022), combining disparate algorithms into a single pipeline (Zhang et al., 2021), full-scale 3D image processing (Oman et al., 2019; Wang et al., 2021), etc. The efficiency of the best models is estimated as ROC-AUC = 0.93 (Oman et al., 2019) and ROC-AUC = 0.90±0.04 (Wang et al., 2021). To overcome the problem of small training samples, which is characteristic for the recognition of ischemic stroke, using generative models (Zhu et al., 2020) is proposed. In general, all above-mentioned models allow you to take into account many influencing factors, but they are "heavy" and do not fit well into the equipment available to practicing physicians.

An elegant way to "lighten" the model is based on physical features of medical artifacts, namely the symmetry of brain hemispheres. Symmetrical neural networks (Cui et al., 2021), first Siamese neural networks (Chen et al., 2020), are widely employed, and their combination with the few-shot learning (Barman et al., 2019; Shovkoplias et al., 2023) allows to work on small datasets. In determining of stroke, the best results are ROC-AUC = 0.914 with 0.88–0.95

confidence interval in three-dimensional mode (Barman et al., 2019) and ROC-AUC = 0.794–0.921 depending on the type of lesion in two-dimensional mode (Shovkoplias et al., 2023). As one can see, symmetric ML models are potentially not inferior in efficiency to heavier ones, but are more critical to the model settings, primarily to the choice of the analysis area. Therefore, for them the problem of choosing and setting model parameters is acute.

The dominant approach here is the heuristic one (see, for example, Thomas et al., 2020), which relies on the intuition and experience of developers and does not generalize well. More formalized methods use data statistics, like nested cross-validation (Larracy, Phinyomark and Scheme, 2021) and Cohen's kappa metric (Castelo, 2022), but they are data-driven and single-criterion, which does not meet the real conditions of stroke detection. The choice of an ML model can also be considered as a decision-making process, which allows the use of widely known mathematical methods, such as linear and non-linear programming, genetic algorithms, AutoML, etc. (Enns et al., 2015; Romero et al., 2022). However, the effectiveness of these approaches is limited by the completeness of considering all parameters of the problem, which is difficult to achieve. In real practice, generalized parameters is more in demand, which would allow, on a strictly formal basis, to obtain comparative assessments of the prospects of specific architectural and parametric solutions. For example, in the work (Shovkoplias et al., 2023) the theory of Probably Approximately Correct learning (Valiant, 2013) is used as a methodological basis. Accordingly, the generalized parameter here is the Vapnik-Chervonenkis dimension (Vapnik and Chervonenkis, 2015), for which it is difficult to offer a meaningful interpretation that is meaningful for a practitioner.

It seems more natural to use topological data analysis (Chazal and Michel, 2021; Wasserman, 2018), which allows us to directly consider the topological features of medical artifacts involved in the construction of Siamese neural networks. In recent years, TAD has established itself as one of the most advanced means of studying the internal structure of the feature space of a neural network (Carlsson and Gabrielsson, 2020; Hensel, Moor and Rieck, 2021; Moroni and Pascali, 2021; Zia et al., 2023). The main tools in TDA are simplicial complexes, forming a kind of multidimensional clusters in input data structure or feature space of ML model. To visualize it, persistent homologies (PH) (Zomorodian and Carlsson, 2005) are used - a set of diagrams in the form of segments (barcodes), which describe the relative position of points in feature space through the distance parameter between points (filtering) ε. At the initial moment ($birth_t$) all points are separated from each other, which corresponds to the birth of individual homologies. As ε increases, some points unite into clusters, which is considered as the moment of disappearance of homology ($death_t$).

Formation and evaluation of PHs are the most widely used applications of the TDA apparatus to the analysis of neural networks. In work (Goibert, Ricatte and Dohmatob, 2022), the topology of the induced graph describing the propagation of the computational process from the input to the output of the NN is studied, in works (Chen et al., 2018; Ramamurthy, Varshney and Mody, 2018) - the topology of the interface of classifiers based on NN, in works (Gabrielsson and Carlsson, 2019; Rieck et al., 2019) - the topology of NN inner space. In the last case, the subject of consideration is the spatial-topological configuration of embeddings formed by the NN on its intermediate layers.

Most well-known works study the actual topological changes performed by NNs. The connection between the topological characteristics of a neural network and its typical metrics is considered only in a few works. Namely, in Gabrielsson and Carlsson, 2019, the connection between accuracy on test data (accuracy) and the lifetime of the most persistent one-dimensional

homology of the first layer of the neural network is shown. Rieck et al., 2019, estimates of the minimum achievable error in NN training were obtained depending on the topological characteristics of the input data and the first layers of fully connected NNs. The use of TDA for comparative assessment of various NN architectures is not presented in the literature available to the authors.

3. METHOD AND MATERIALS

3.1 Neural Network Configuration

For the study, we used a Siamese neural network (SNN) of typical architecture (Barman et al., 2019) (Fig. 1), which is designed to compare the symmetry of MRI images of 2D and 3D dimensions. SNN has two parallel branches, which include four inception modules with identical weights, as well as a common branch containing convolutional and pooling modules. An MRI image of the brain is supplied to the network input in the form of a set of slices, potentially containing the affected area. Sections of the affected area falling within the slices are highlighted in red.

During preprocessing, the input image is divided into two symmetrical halves (left and right hemispheres), each of which is fed into its own channel. Thus, two branches of the network process the left and right images independently, and in the common branch the difference between the feature representations of the input data is calculated. The architectures of both branches are identical, they include four InceptionModule (IM) layers, each of which extracts features from the input data x_1, x_2 through parallel paths $f_1(x_1)$ and $f_2(x_2)$. Both branches have the same sets of weights $W * f_1(x_1)$ and $W * f_2(x_2)$ and identical convolution and pooling operations. The common branch includes multiple IM layers, as well as pooling and fully connected layers for the final output. When solving a classification problem, the network generates an output of 1 if a pathology is detected, and 0 otherwise.

The network was trained in two modes: with and without preliminary training on a simulating dataset.
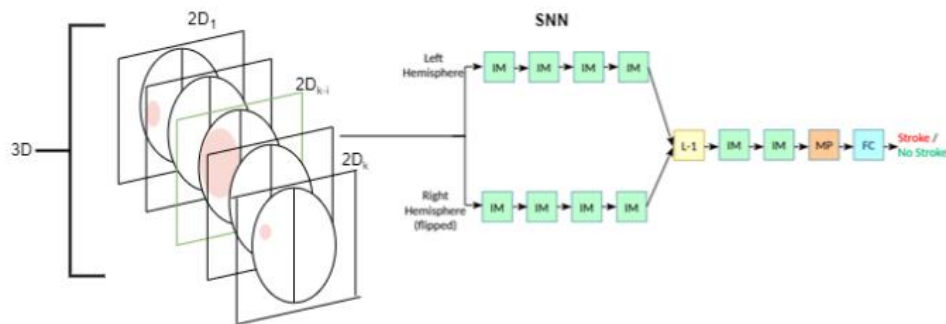

Figure 1. Basic model architecture: IM – Inception Module, L-1 – Merge Layer with L-1 difference, MP –Max Pooling, FC – Fully Connected Layer. Basic inputs: 3D – all slices per patient; 2D – one slice per patient

3.2 Metrics Used

To assess the classification properties of SNN, we used typical metric, namely ROC-AUC score:

$$roc_{auc} = \int_0^1 TPR(FPR^{-1}(t))dt \tag{1}$$

where TPR – true positive rate, FPR – false positive rate.

To assess the topological properties of SNN, we built zero PHs for embeddings generated by SNN on each of the last four layers, and then we calculated the following metrics for each layer:

–   Persistent entropy:

$$entropy = -\Sigma_{i=1}^n p_i \log p_i, p_i = \frac{death_i - birth_i}{\Sigma_{i=1}^n[death_i - birth_i]} \tag{2}$$

–   Rényi Entropy for α=0.5, 2, 3:

$$renyi\_entropy_\alpha = \frac{1}{1-\alpha} log\Sigma_{i=1}^n p_i^\alpha, p_i = \frac{death_i - birth_i}{\Sigma_{i=1}^n[death_i - birth_i]} \tag{3}$$

–   The difference in persistent entropies between adjacent layers

$$delta_{entr} = |entropy_l - entropy_{l-1}|, l - layer \tag{4}$$

–   The difference in Rényi entropies between adjacent layers

$$delta_{renyi_\alpha} = |renyi\_entropy_{\alpha_l} - renyi\_entropy_{\alpha_{l-1}}|, l - layer \tag{5}$$

To quantify similarity of datasets, Fréchet inception distance is used:

$$FID = ||\mu_1 - \mu_2||^2 + Tr\left(\Sigma_1 + \Sigma_2 - 2*(\Sigma_1\Sigma_2)^{\frac{1}{2}}\right) \tag{6}$$

where $\mu_1, \Sigma_1, \mu_2, \Sigma_2$ are respectively mean and covariance matrix calculated based on embeddings of 1 and 2 groups of data in resnet18.

3.3 Experimental data and scenarios

Three datasets were used in the experiments. The native dataset consisted of MRI images of the brain tissue in FLAIR format; the images were desculled, normalized and divided into left and right hemispheres; the total size of the dataset is 200 images, of which 100 are normal and 100 with pathologies; image size was 64×92×26 for 3D variant and 64×92 for 2D variant. The dataset generated by diffusion model was taken from (Tasci, 2023).

Figure 2 shows slices of a real MRI image of the brain, including areas of ischemic stroke (highlighted in ovals). It is easy to see that the image contains other hyperintense zones of a similar configuration, which, however, are not affected. This makes it difficult to train a neural network and reduces its potential efficiency, especially when working on individual slices, i.e., in 2D mode. To overcome this difficulty, we used model pretraining. To do this, we generated a synthetic dataset based on Perlin noise (Etherington, 2022). Its advantages over generative models for simulating brain tissue are its low computational and memory requirements.

Moreover, using Perlin noise allows to simulate healthy and affected fragments of brain tissue separately and to combine them in a controlled manner, i.e., create more realistic conditions for preliminary training of the model.

Examples of images from the synthetic dataset are shown in Figure 3. Simulation parameters were selected with the participation of a radiologist to reproduce the typical problems encountered in distinguishing between hyperintense areas, normal and stroke-affected. In Figure 3a shows a simulation of healthy brain tissue, Fig. 3b – tissues with normal hyperintense areas, in Figure 3c – tissues with areas of ischemic stroke. The dataset for pretraining was composed of randomly selected images of the second and third groups (Figure 3b and Figure 3c), taken in equal proportions, the image sizes were normalized identically to the native dataset, the total volume of the dataset was 200 images.

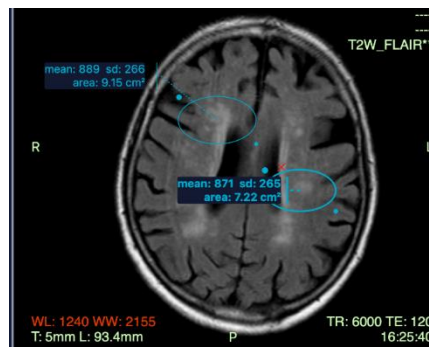Example of image from dataset generated by diffusion model is shown in Figure 4.

Figure 2. Slice of MRI image of a brain with ischemic stroke. Damage zones are marked with ovals; the remaining hyperintense zones are not stroke zones
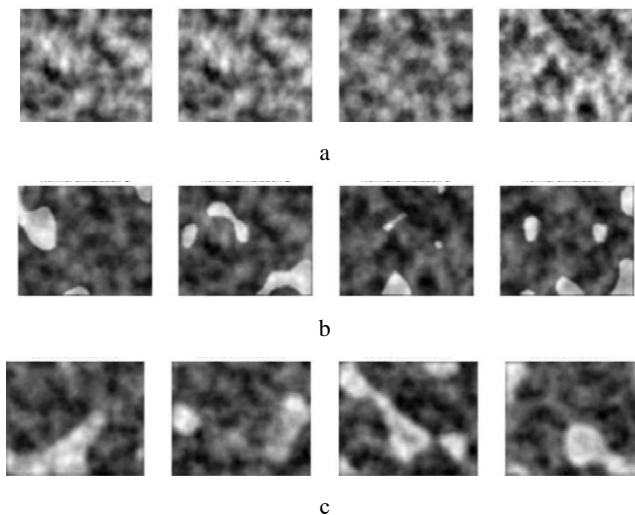

Figure 3. Examples of simulated images from a synthetic dataset: a - healthy brain tissue, b – tissue with normal hyperintense areas, c – tissue with areas of ischemic stroke
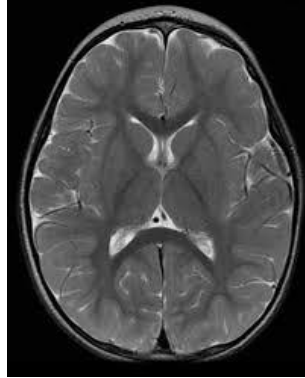
Figure 4. Example of brain MRI slice generated by diffusion model

The experiment was carried out according to three scenarios:

I. SNN was trained without pretraining on a native dataset and was used to classify stroke in 2D and in 3D modes;

II. SNN was pre-trained on a synthetic dataset and trained on a native dataset afterwards; then the wholly trained SNN was used to classify stroke in 2D mode, and the central frame from the MRI package was used as the analyzed frame;

III. SNN was pre-trained on a synthetic dataset and trained on a native dataset afterwards; then the wholly trained SNN was used to classify stroke in 2D mode, and the analyzed frame was a frame from the MRI package, selected by the radiologist as the most informative.

Then, for each scenario, the differences between persistent entropies (4) and Renyi entropies (5) were calculated for four adjacent layers of the common branch of the network (Figure 1). In addition, for each scenario the classification efficiency was assessed using the ROC-AUC score metric (1).

4. RESULTS AND DISCUSSION

An example of the ratio of PH lifetimes for 2D and 3D SNN operating modes is shown in Figure 5. It is easy to see that the lifetime of homologies in the first case is significantly greater than in the second, and, as our experiments have shown, this ratio persists for all SNN layers, regardless of the preprocessing of the dataset. This confirms that, all other things being equal, 3D SNN operating mode provides greater data separability than 2D one. A more detailed comparison of topological characteristics, carried out using metrics (4) and (5), is graphically presented in Fig. 6.a.-6.c., namely, on Fig 6.a there are results of PH calculations obtained from SNN without pretraining, on Fig 6.b there are results of PH calculations obtained from SNN pretrained on Perlin noise dataset and on Fig 6.c there are results of PH calculations obtained from SNN pretrained on dataset generated by diffusion model (Tasci, 2023).
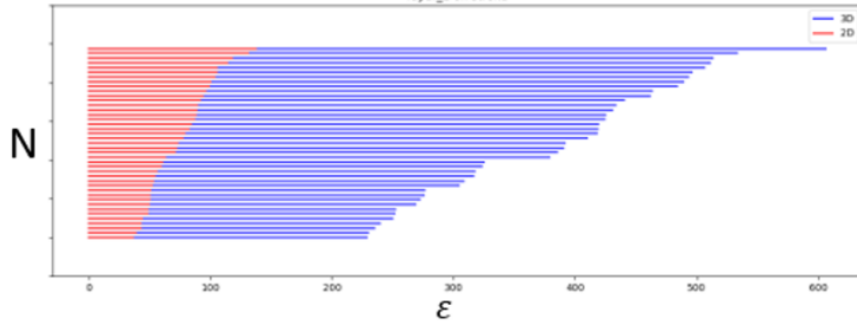
Figure 5. An example of the ratio of zero GHGs on one layer for 2D (blue lines) and 3D (red lines)
SNN operating modes: ε – filtration parameter, N – homology number in birth order

A comparison of the graphs shows that with pretraining of 3D SNN on Perlin noise data, the persistent entropy is practically unchanged across layers, while in 2D SNN such a change takes place. It is worth noting that pretraining on data generated by diffusion model only smooths out changes in persistent entropy but saves trend.
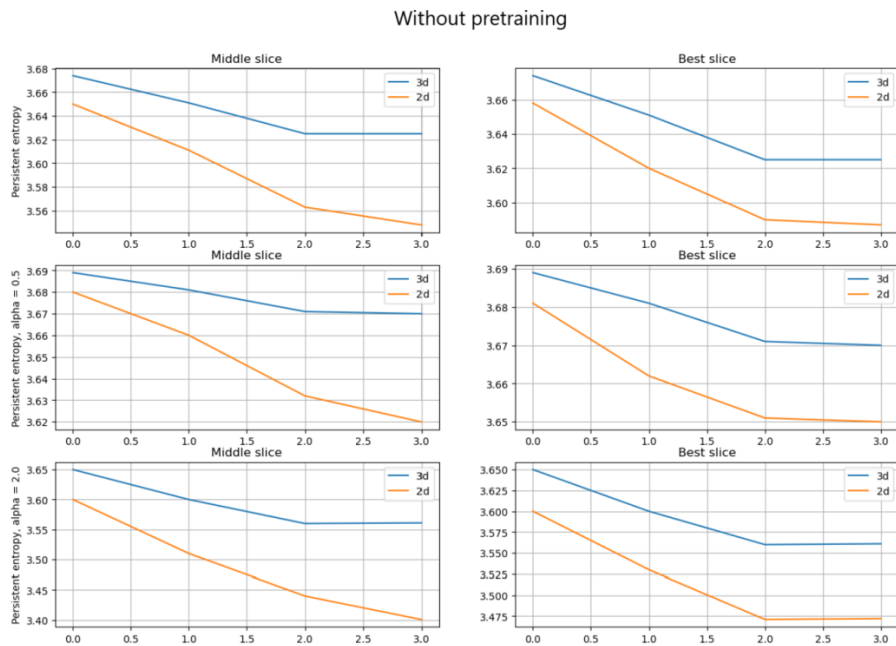

Figure 6a. Comparison of persistent entropies between adjacent layers for
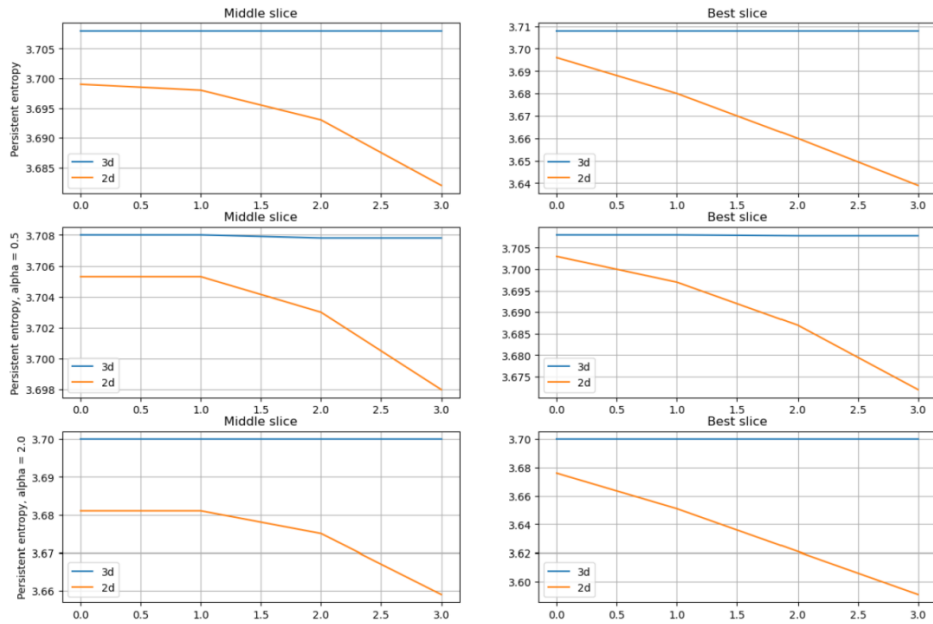different experimental scenarios

Figure 6b. Comparison of persistent entropies between adjacent layers for different experimental
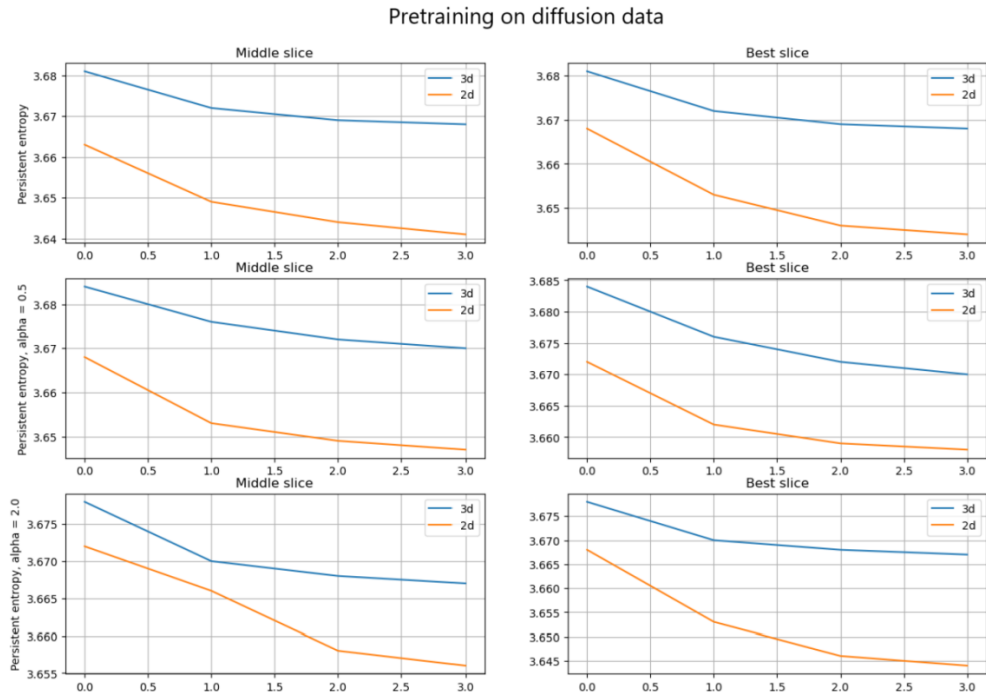scenarios for model pretrained on data obtained from Perlin noise

Figure 6c. Comparison of persistent entropies between adjacent layers for different experimental scenarios for model pretrained on data generated by diffusion model

Moreover, in 2D SNN without pretrain, the use of the best slices provides a smaller drop in entropy than in the middle ones, and in 2D SNN with pretrain the opposite effect is observed.

Results shows that pretraining on synthetic data generated from Perlin noise allows the model to extract more information from the data during additional training. In other words, with the same architectural solutions, the 3D model extracts redundant information, which can reduce the efficiency of its configuration, while a simpler 2D model works with less information, but uses it more efficiently. When we use synthetic data generated by diffusion model, we see similar effect, but it affects trend in 3D case to a lesser extent. It means that globally consistent data created by diffusion model fills 3D SNN with fewer redundant detail compared to Perlin noise data.

Figure 7 shows results of pairwise Frechet Inception distance calculation between two synthetic and one real dataset are presented in graphical form. We see that data generated by diffusion model is closer to real data than data obtained from Perlin noise, but both synthetic datasets are sufficiently closer to each other. This indicates that, while diffusion model captures visual appearance of data better (see Fig 2. vs Fig 4.), data generated based on Perlin noise still contain crucial features of real dataset needed for SNN training.
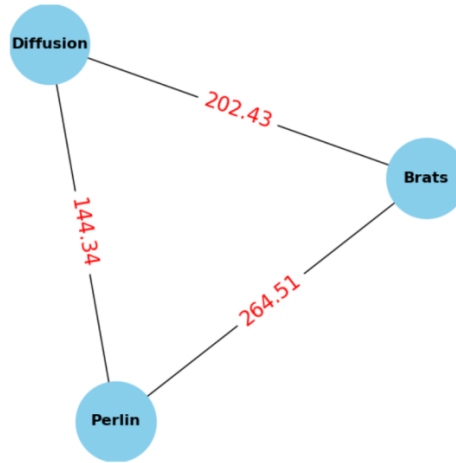
Figure 7. Comparison of Frechet distances between datasets

The conclusions obtained are confirmed by comparing ROC-AUC metrics for 2D and 3D models presented in Table 1. Rows represent model type (either 3D SNN or 2D SNN), and columns represent experiments configurations respectively: the first column contains results obtained by SOTA models, the second one contains results obtained by SNN model without pretraining, the third one contains results obtained by SNN model with pretraining on data generated based on Perlin noise, the last one contains results obtained by SNN model with pretraining on data generated by diffusion model.

Table 1. Experimental assessment of the ROC-AUC score for the classification of ischemic stroke during operation of the SNN in different scenarios

Model type	SOTA	SNN, not pretrained		SNN, Perlin pre-trained		SNN, Diffusion pretrained	
		middle slice	best slice	middle slice	best slice	middle slice	best slice
2D	0, 794	0,851	0,969	0,890	0.984	0,899	0,989
3D	0,899	0,973		0.995		0.996	

We see that pretraining on synthetic data improves 2D SNN quality when we evaluate it both on middle slices and on best slices of MRI. Diffusion model data improves ROC-AUC more significantly than Perlin noise data: on middle slices we get 5.6% vs 5% and on best slices we get 2% vs 1.5% respectively. However, percent gains in ROC-AUC obtained by both methods are comparable (difference between them in both cases is approximately 0.5%).

As we can see, all the results from three different experiments (Figures 5a – 5c, Figure 6., Table 1) agree and support the following conclusion from different perspectives: a simple model that contains only locally useful features can support neural network training to a level comparable to much more complex and resource-intensive generative models. This is especially important in the context of small sizes of available datasets and limitations in computing resources, which is typical for real medical practice.

5. CONCLUSION

In this study, the use of Siamese neural networks (SNNs) was shown to be effective for ischemic stroke classification in resource-limited settings.

The paper proposes a method for assessing the degree of informativeness of SNNs using topological data analysis, which allows us to objectify the differences between 2D and 3D modalities and thereby optimize the choice of neural network architecture when classifying medical images depending on specific conditions.

It is experimentally confirmed that the use of pre-training of difference models on "light" synthetic data allows leveraging the efficiency of 2D-mode processing up to 3D-mode.

It is experimentally (from three different perspectives) confirmed that simple model that contains only locally useful features can support neural network training to a level comparable to much more complex and resource-intensive generative model.

In addition, allowing the clinician to selectively use slices of greatest interest while performing automated classification based on them increases the explainability of AI systems in medicine and staff confidence in their work.

ACKNOWLEDGEMENT

REFERENCES

Barman A. et al., 2019. Determining ischemic stroke from CT-angiography imaging using symmetry-sensitive convolutional networks. *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, Venice, Italy, pp. 1873-1877.

Carlsson, G. and Gabrielsson, R. B., 2020. Topological Approaches to Deep Learning. In N. Baas, G. Carlsson, G. Quick, M. Szymik and M. Thaule (eds.) *Topological Data Analysis. Abel Symposia* (Vol. 15). Springer, Cham. https://doi.org/10.1007/978-3-030-43408-3_5

Castelo, G. P., 2022. *Minimum Sample Size Estimation in Machine Learning* (Master's thesis, Universitat Oberta de Catalunya). Available at: https://openaccess.uoc.edu/bitstream/10609/146007/8/gprolcasteloTFM0622report.pdf

Chazal, F. and Michel, B., 2021. An Introduction to Topological Data Analysis: Fundamental and Practical Aspects for Data Scientists. *Frontiers in Artificial Intelligence*, Vol. 4, 667963. https://doi.org/10.3389/frai.2021.667963

Chen, C. et al., 2018. A Topological Regularizer for Classifiers via Persistent Homology. *arXiv:1806.10714v3*

Chen, H. et al., 2020. Anatomy-Aware Siamese Network: Exploiting Semantic Asymmetry for Accurate Pelvic Fracture Detection in X-ray Images. In A. Vedaldi, H. Bischof, T. Brox and J. M. Frahm (eds.) *Computer Vision – ECCV 2020. Lecture Notes in Computer Science* (Vol. 12368). Springer, Cham. https://doi.org/10.1007/978-3-030-58592-1_15

Cui, L. et al., 2021. Deep symmetric three-dimensional convolutional neural networks for identifying acute ischemic stroke via diffusion-weighted images. *Journal of X-Ray Science and Technology,* Vol. 29, No. 4, pp. 551-566. doi: 10.3233/XST-210861

Enns, E. A. et al., 2015. Identifying best-fitting inputs in health-economic model calibration: a Pareto frontier approach. *Medical Decision Making*, Vol. 35, No. 2, pp. 170-182.

Etherington, T. R., 2022. Perlin noise as a hierarchical neutral landscape model. *Web Ecology*, Vol. 22, No. 1. https://doi.org/10.5194/we-22-1-2022

Gabrielsson R. B. and Carlsson G., 2019. Exposition and Interpretation of the Topology of Neural Networks. *arXiv:1810.03234v3*

Goibert, M., Ricatte, T. and Dohmatob, E., 2022. An Adversarial Robustness Perspective on the Topology of Neural Networks. *arXiv.org/abs/2211.02675v1*

Hensel, F., Moor M. and Rieck B., 2021. A Survey of Topological Machine Learning Methods. *Frontiers in Artificial Intelligence*, Vol. 4. https://doi.org/10.3389/frai.2021.681108

Larracy, R., Phinyomark, A. and Scheme, E., 2021. Machine Learning Model Validation for Early Stage Studies with Small Sample Sizes. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Vol. 2021), pp. 2314-2319. doi: 10.1109/EMBC46164.2021.9629697

Lu, J. et al., 2022. Identification of early invisible acute ischemic stroke in non-contrast computed tomography using two-stage deep-learning model. *Theranostics*, Vol. 12, No. 12, pp. 5564-5573. doi: 10.7150/thno.74125

Moroni, D. and Pascali M. A., 2021. Learning topology: bridging computational topology and machine learning. *Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications*, Vol. 31, pp. 443-453. https://doi.org/10.1134/S1054661821030184

Mosisa, W. et al., 2023. Survival status and predictors of mortality among adult Stroke patients admitted to Jimma University Medical Center, South west Ethiopia: A retrospective Cohort study. *Vascular Health and Risk Management*, Vol. 19, pp. 527-541.

National Collaborating Centre for Chronic Conditions (UK), 2008. *Stroke: National Clinical Guideline for Diagnosis and Initial Management of Acute Stroke and Transient Ischaemic Attack (TIA)*. Royal College of Physicians (UK), London, UK.

Nishio, M. et al., 2020. Automatic detection of acute ischemic stroke using non-contrast computed tomography and two-stage deep learning model. *Computer Methods and Programs in Biomedecine*, Vol. 196, 105711. doi: 10.1016/j.cmpb.2020.105711

Oman, O. et al., 2019. 3D convolutional neural networks applied to CT angiography in the detection of acute ischemic stroke. *European Radiology Experimental*, Vol. 3, 8. doi: 10.1186/s41747-019-0085-6

Ramamurthy, K. N., Varshney, K. R. and Mody, K., 2018. Topological Data Analysis of Decision Boundaries with Application to Model Selection. *arXiv:1805.09949v1*

Rieck, B. et al., 2019. Neural persistence: A complexity measure for deep neural networks using algebraic topology. *International Conference on Learning Representations (ICLR 2019)*. *https://arxiv.org/abs/1812.09764*

Romero, R. A. et al., 2022. Benchmarking AutoML frameworks for disease prediction using medical claims. *BioData Mining,* Vol. 15, 15. https://doi.org/10.1186/s13040-022-00300-2

Ruksakulpiwat, S. et al., 2023. Using Neural Networks Algorithm in Ischemic Stroke Diagnosis: A Systematic Review. *Journal of Multidisciplinary Healthcare,* Vol. 16, pp. 2593-2602. https://doi.org/10.2147/JMDH.S421280

Saranya, A. and Subhashini R., 2023. A Systematic Review of Explainable Artificial Intelligence Models and Applications: Recent Developments and Future Trends. *Decision Analytics Journal,* Vol. 7, 100230.

Shinohara, Y. et al., 2020. Development of a deep learning model to identify hyperdense MCA sign in patients with acute ischemic stroke. *Japanese Jounal of Radiology*, Vol. 38, No. 2, pp. 112-117. doi: 10.1007/s11604-019-0089

Shovkoplias, G. et al., 2023. Proactive selection of machine learning models for small sample sizes in cerebral stroke detection based on PAC-learning theory. *Procedia Computer Science*, Vol. 229, pp. 80-90.

Speith, T., 2022. A Review of Taxonomies of Explainable Artificial Intelligence (XAI) Methods. *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*, Seoul, Republic of Korea. https://doi.org/10.1145/3531146.3534639

Stib, M. T. et al., 2020. Detecting large vessel occlusion at multiphase CT angiography by using a deep convolutional neural network. *Radiology,* Vol. 297, No. 3, pp. 640-649. doi: 10.1148/radiol.2020200334

Tasci, B., 2023. Automated ischemic acute infarction detection using pre-trained CNN models' deep features. *Biomedical Signal Processing and Control*, Vol. 82, 104603.

Thomas, R. M. et al., 2020. Chapter 14 - Dealing with missing data, small sample sizes, and heterogeneity in machine learning studies of brain disorders. *Machine Learning: Methods and Applications to Brain Disorders*, pp. 249-266.

Valiant, L., 2013. *Probably Approximately Correct: Nature's Algorithms for Learning and Prospering in a Complex World*. Basic Books.

Valliani, A. A., Ranti, D. and Oermann, E. K., 2019. Deep Learning and Neurology: A Systematic Review. *Neurology and Therapy*, Vol. 8, No. 2, pp. 351-365.

Vapnik, V. N. and Chervonenkis, A. Y., 2015. On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities. In V. Vovk, H. Papadopoulos and A. Gammerman (eds) *Measures of Complexity*. Springer, Cham. doi:10.1007/978-3-319-21852-6_3

Wang, C. et al., 2021. Deep learning-based identification of acute ischemic core and deficit from non-contrast CT and CTA. *Journal of Cerebral Blood Flow and Metabolism*, Vol. 41, No. 11, pp. 3028-3038. doi: 10.1177/0271678x211023660

Wasserman, L., 2018. Topological data analysis. *Annual Review of Statistics and Its Application,* Vol. 5, pp. 501-532.

Writing Group Members et al., 2016. Heart Disease and Stroke Statistics – 2016 Update: A Report from the American Heart Association. *Circulation*, Vol. 133, No. 4, pp. 38-360.

Zhang, S. J. et al., 2021. Stroke lesion detection and analysis in MRI images based on deep learning. *Journal of Healthcare Engineering*, Vol. 2021, 5524769. https://doi.org/10.1155/2021/5524769

Zhu, L. et al., 2020. Make-A-Volume: Leveraging Latent Diffusion Models for Cross-Modality 3D Brain MRI Synthesis. *arXiv: 07.10094v1*

Zia, A. et al., 2023. Topological Deep Learning: A Review of an Emerging Paradigm. *Artificial Intelligence Review*, Vol. 57, No. 77. https://doi.org/10.1007/s10462-024-10710-9

Zomorodian, A. and Carlsson G., 2005. Computing persistent homology. *Discrete & Computational Geometry,* Vol. 33, No. 2, pp. 249-274. https://doi.org/10.1007/s00454-004-1146-y