Федеральное государственное автономное образовательное учреждение
высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»
Физтех-школа Прикладной Математики и
Информатики Кафедра интеллектуальных систем

Направление подготовки / специальность: 03.04.01 Прикладные математика и физика
Направленность (профиль) подготовки: Математическая физика, компьютерные технологии и математическое моделирование в экономике

АВТОМАТИЧЕСКАЯ ДЕТЕКЦИЯ ФОКАЛЬНОЙ КОРТИКАЛЬНОЙ ДИСПЛАЗИИ В СЛУЧАЕ РАЗРЕЖЕННОГО ПРЕДСТАВЛЕНИЯ ДАННЫХ

(магистерская диссертация)

Студент:
Гребенькова Ольга Сергеевна

_____
(подпись студента)

Научный руководитель:
Бурнаев Евгений Владимирович,
доктор физ.-мат. наук, доц.

_____
(подпись научного руководителя)

Консультант *(при наличии)*:
_____
_____
(подпись консультанта)

Москва 2023

Abstract

Фокальная кортикальная дисплазия (ФКД) является одним из наиболее распространенных поражений головного мозга, вызывающих эпилепсию. Обычно она диагностируются рентгенологом при МРТ-сканировании головного мозга пациента. Её хирургическое удаление может привести к снижению потребности в противоэпилептических препаратах, снижению частоты или к полному отсутствию приступов. Цель проекта — создать удобную и надежную систему для выявления врачом структурных изменений в коре головного мозга на основании данных МРТ. Представленное исследование посвящено использованию глубоких нейронных сетей для этой задачи. Предлагается использовать фреймворк Minkowski Engine для МРТ данных, представленных в виде облака точек. Полученный метод был применен к набору из 183 МРТ изображений пациентов с ФКД, собранных в сотрудничестве с медицинскими партнерами. Также в работе представлено сравнение с современными подходами к обнаружению ФКД.

Ключевые слова: Фокальная кортикальная дисплазия, эпилепсия, МРТ, облако точек, сети Минковского, нейронные сети

Abstract

Focal cortical dysplasia (FCD) is one of the most common brain lesions that cause epilepsy, and it is usually diagnosed by a radiologist inspecting an MRI scan of the patient's brain. Surgical resection can result in reduced need for anti-epileptic medication, reduced frequency or most commonly complete absence of seizures. The purpose of the project is to create a reliable system for the doctor to detect structural changes in the cerebral cortex based on MRI data. This work focuses on deep learning part of the pipeline. I propose to use Minkowski Engine for the point cloud representation of brains. The resulted method is applied to the dataset of 183 labeled FCD patients collected in a collaboration with medical partners and compared with the state-of-the-art approaches for FCD detection.

Keywords: Focal Cortical Dysplasia, Epilepsy, MRI, Minkowski Engine, Point Cloud, Neural Network

Contents

Chapter 1

Introduction

1.1 Background

Epilepsy is a neurological disorder in which brain activity becomes abnormal, causing seizures or periods of unusual behavior, sensations and sometimes loss of awareness. Anyone can develop this disease as it affects both males and females of all races, ethnic backgrounds and ages, with an incidence of approximately 50 new cases per year per 100,000 population [19]. Treatment with medications or sometimes surgery can control seizures for the majority of people with epilepsy. Some people require lifelong treatment to control seizures, but for others, symptoms eventually go away. Some children with epilepsy may outgrow the condition with age.

Focal cortical dysplasia (FCD) is one of the most common epileptogenic lesions associated with cortical development malformations. The exact cause of FCD is not fully understood. However, it is believed to be the result of abnormal neuronal migration during fetal development. This leads to the formation of abnormal clusters of neurons, which can disrupt the normal functioning of the brain. The severity of symptoms associated with FCD can vary widely depending on the location and extent of the malformation.

Although medication is the primary treatment for epilepsy caused by focal cortical dysplasia (FCD), it is effective for only a small number of patients [20]. Therefore, the treatment of FCDs represents significant clinical challenges due to their medication-resistant nature. When one or more antiepileptic drugs fail to control seizures, surgical resection of the affected area may be considered [15]. This procedure can reduce the need for medication, minimize seizure frequency, and in many cases, eliminate seizures altogether. However, identifying the precise location of the affected tissue remains a challenge in many cases.

Diagnosis of FCD typically involves a combination of imaging studies, EEG, and neurological evaluation. MRI scans are the most commonly used imaging studies to detect FCD. However, between 50 and 80% of FCDs are too subtle to detect by conventional radiological analysis of MRI scans. When compared to other forms of lesions that may be identified through MRI analysis, the rate of missed detection is significant. To illustrate, less than 4% of neck and head cancer remains uncaught on first MRI scanning [12]. Furthermore, the visual interpretation of MRI is very subjective and correlated with radiologist's skills [10].

Therefore in recent years, there has been increased interest in developing automated tools for detecting FCD. These tools could improve the accuracy and consistency of presurgical evaluation, leading to better surgical outcomes for individuals with FCD. Additionally, ongoing research into the underlying causes of FCD may lead to new treatments and therapies for this challenging condition.

1.2 Aim of the research

The development of a reliable system for detecting structural changes in the cerebral cortex based on MRI data is an important step towards improving the diagnosis and treatment of focal cortical dysplasia (FCD). This project aims to use deep learning techniques to provide precise maps of possible FCD areas, which can aid in presurgical evaluation and improve surgical outcomes for individuals with this condition.

One of the challenges in MRI segmentation is the size of the whole MR image, which can be computationally demanding. To address this issue, this paper proposes to represent MRI images in the form of Point Cloud. This approach can significantly reduce the computational cost while maintaining the accuracy of the model.

The use of Minkowski Engine Networks for potentially epileptogenic foci localization built on multimodal MRI data is a promising direction for this research. Minkowski Engine is a PyTorch extension that provides an extensive set of neural network layers for sparse data, including Point Clouds. This technology can help improve the accuracy and consistency of predictions.

1.3 Novelty of the project

To the best of our knowledge it is the first work, where for the detection of Focal Cortical Dysplasia on MR data the sparse representation of the whole brain is used. Furthermore, additional morphometry and contrast features were added to MRI analysis. The whole pipeline was validated on the one of the biggest labeled dataset of people with FCD, that was collected during the project and compared with state-of-the-art method. Our pipeline showed the comparable performance, making it a promising tool for clinical applications.

Chapter 2

Current state of the art and difficulties

Many studies have addressed the challenges of identifying FCD through manual examination of MRI scans. FCD patterns can involve thickening of the cortex, loss of differentiation between gray and white matter, focal atrophy, and changes in MR signal of gray or white matter [8, 21]. There are three popular approaches to detect FCD on MRI: voxel-based morphometry (VBM), surface-based classification (SBC) and deep learning algorithms.

2.1 Voxel-based morphometry approaches

VBM approach [25] is based on a voxel-wise statistical analysis. VBM aims to distinguish variations in brain tissue composition at a local level, while disregarding differences in overall anatomy. This is accomplished through the normalization of all structural images to standard stereotaxic space [32], segmentation of the normalized images into gray and white matter, smoothing of the gray and white matter images, and finally, statistical analysis to localize significant differences between individuals with FCD and healthy controls.

Images can also be divided into regions of interest (ROIs) that correspond to specific areas of the brain affected by FCD. These ROIs are then smoothed and used to create statistical feature maps. By selecting an optimal threshold for these maps, FCD regions can be identified. In [24] four voxel-based morphometric methods were compared, including three T1 image-based approaches (gray matter volume, gray matter concentration, and white-gray matter junction map) and one FLAIR-based method. The study evaluates their effectiveness in detecting FCD lesions using positive and negative MRI.

But VBM has important caveats. One of the limitations of VBM is that it relies on accurate segmentation of brain tissue into gray matter and white matter. Errors in segmentation can lead to inaccurate results and false positives [7]. Additionally, VBM is a relatively complex and computationally intensive technique that requires specialized software and expertise to perform and interpret. Finally, VBM is limited by the resolution of the imaging modality used, which may not be able to detect subtle structural changes in the brain. In addition, statistical maps depend on the control population used for analysis [22]

2.2 Surface-based classification

Surface-based classification is a method of identifying FCD through the analysis of the cortical surface of the brain. The main part of SBC techniques is the cortical reconstruction to obtain inner- and outer-cortical surfaces. After this at each vertex of surface authors extract features classify lesion vertices using the machine learning methods, such as support vector machine (SVM). [17] proposed an automatic FCD lesion detection algorithm using surface-based and intensity-based features in T1, FLAIR, and FLAIR/T1 images. Another paper [14] compared three widely used classifiers (decision tree, support vector machine, and artificial neural network) for classification method selection and show the importance of surface-based MRI morphometry in combination with machine learning techniques.

SBC is computational expensive, but more stable then voxel-based morphometry approaches when applied to MRI data with anatomical artifacts. One potential problem with surface-based classification methods is that they rely on accurate and precise imaging data. If the imaging data is of poor quality or resolution, it may be difficult to accurately detect abnormalities in the cortical surface. Additionally, these methods may not be able to detect subtle or small abnormalities that may be missed by traditional manual examination methods. Finally, there may be variations in the cortical surface between individuals that could lead to false positives or false negatives in the classification process.

2.3 Deep learning methods

The ability to learn the appropriate features/representations from the training data without any human interventions make deep-learning-based methods suitable for FCD detection. In [6] authors use U-Net-like architecture for building the FCD segmentation model. They train the network using sagittal slices and beat previous statistical approaches.

But, most of the state-of-the-art FCD segmentation approaches use two-dimensional (2D) convolutional neural networks(CNN) models despite the availability of 3D Magnetic resonance imaging (MRI) volumes, and hence fail to leverage the inter-slice information present in the MRI volumes. In [26] the model uses a 3D version of U-Net with residual blocks that works on shallow depth 3D sub-volumes generated from MRI volumes. In [16] the cascade of 3D patch-based CNN is proposed. One of the networks is responsible for removing false positives (FPs) while maintaining optimal sensitivity. This gives the superior performance over the state-of-the-art FCD segmentation methods. However, this approach still has low quality in terms of numbers of false positive regions. So it is hard to visually find the exact FCD location. Therefore a new more precise methods must be developed.

The major challenge of working with MRI is the size of one data object. It is always impossible to apply network even to one brain due to GPU memory limits. For this reason [26, 16] apply models to brain patches. In this paper we propose to represent MR images in the from of a point cloud.

Point cloud segmentation. 3D point cloud segmentation is the process of classifying point clouds into multiple homogeneous regions, the points in the same region will have the same properties. First works in this field [34, 11] used a dense representation where the empty space is represented as 0. This representation is intuitive and supported by all popular deep learning libraries. However, it is computational demanding. In [27] the Octree structure for representing 3D space and faster convolution on it was proposed. Another approach in this field is PointNet-based methods [28, 29]. PointNets use a set of input coordinates as features for a multi-layer perceptron. But they can be applied only to limited number of points at once, so these approaches have rather limited receptive field size. In [4] authors addresses an issue of inefficient point cloud data processing in convolutional neural networks. They showed that usage of classic architectures leads to dilation of sparse data. So authors tried to construct a CNN that can keep the same sparsity pattern throughout the layers. The new technique was called valid sparse convolution. It discards any computations on the empty regions: all the atomic operations are calculated with indices generation algorithms. In this paper we propose to use a 4D spatio-temporal convolutional neural network called MinkowskiNet [9] for 3D video perception. This model uses generalized sparse convolution to effectively process high-dimensional data. Also, authors defined other common CNN layers such as pooling and transposed convolutions for sparse data.

Chapter 3

Methods

3.1 Main goal

This thesis is a part of the EpiDetect project in Skoltech. The main aim of the project is to provide doctors with reliable visualisation system to detect FCD. In order to do this we need to collect training data, preprocess it, make research on different models and statistical ideas to provide in future heat maps of FCD probabilities and finally make good visualiser convenient for doctors. The goal of current thesis is to build the deep learning pipeline to produce heat maps for doctors with the probabilities of FCD.

3.2 Data

3.2.1 Dataset

The major limitation for using deep learning approaches in the task of detecting focal cortical dysplasia is small datasets. It is hard to find suitable patients and professionals, who can precisely label MR images. Furthermore, there is no public-available data because it contain information that could compromise the privacy of research participants. In order to collect required data we partnered with two leading medical centers in Russia:

- The Scientific Center of Obstetrics, Gynecology and Perinatology named after academician V. I. Kulakov
- The National Medical and Surgical Center named after academician N. I. Pirogov

The dataset included patients who met the following criterias:

1. they were diagnosed with pharmacoresistant epilepsy;
2. the study sequences were obtained using an epilepsy-specific protocol;
3. there are MR-positive radiological signs of FCD;

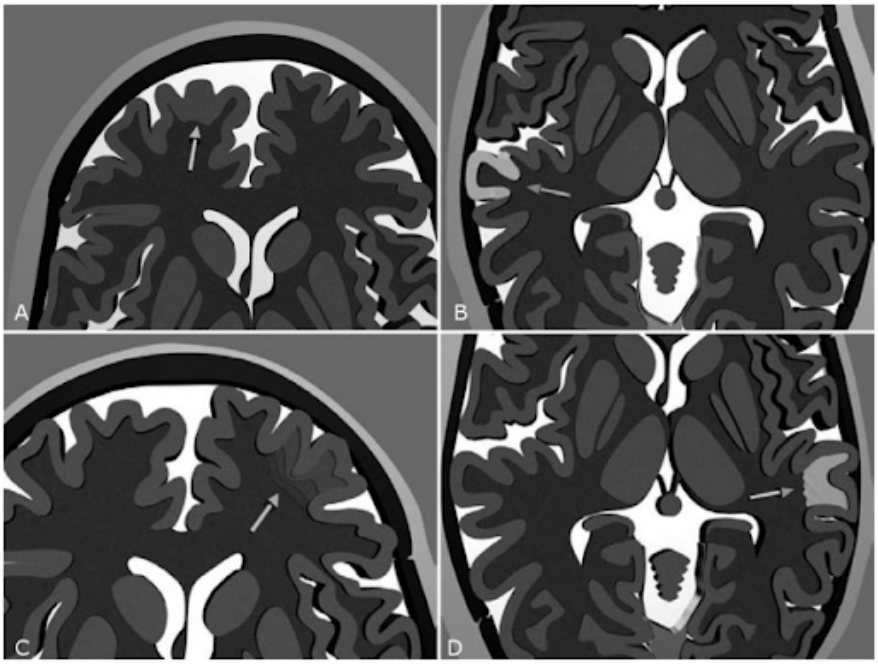

Figure 3.1: Features of FCD: A – cortical thickening; B – alteration of gray matter signal; C – loss of differentiation between gray and white matter; D – alteration of white matter signal.

4. the location of FCD was correlated with the findings of the electroencephalography (EEG) study.

Possible radiological signs of FCD are presented on Fig. 3.1.

With the help of our partners we have collected 183 subjects diagnosed with pharmacoresistant epilepsy including T1, T2 and FLAIR modality for each. Two professional radiologists labelled FCD area in one projection slide-by-slide, therefore we get voxel-wise label in 3D space. They independently performed manual segmentations on T1w images, incorporating information from T2w and T2-FLAIR MR images. Therefore, the following data was acquired for every subject:

- 3D T1-weighted
  - TR = 2300 ms for sedated and non-sedated studies
  - TE = $2, 38$ ms for sedated studies and $2, 33$ ms for non-sedated studies
  - Flip angle = $8°$ for sedated and non-sedated studies
  - FOV = 255 mm2 for sedated studies and 250 mm2 for non-sedated studies
  - Voxel size = $0.8 \times 0.8 \times 0.8$ mm for sedated and $0.8 \times 0.8 \times 0.9$ mm for non-sedated studies
  - matrix = 256×256 for sedated studies and 320×320 for non-sedated studies
- 3D T2-weighed
  - TR = 8000 ms for sedated studies and 7000 ms non-sedated studies
  - TE = 463 ms for sedated studies and 384 ms for non-sedated studies

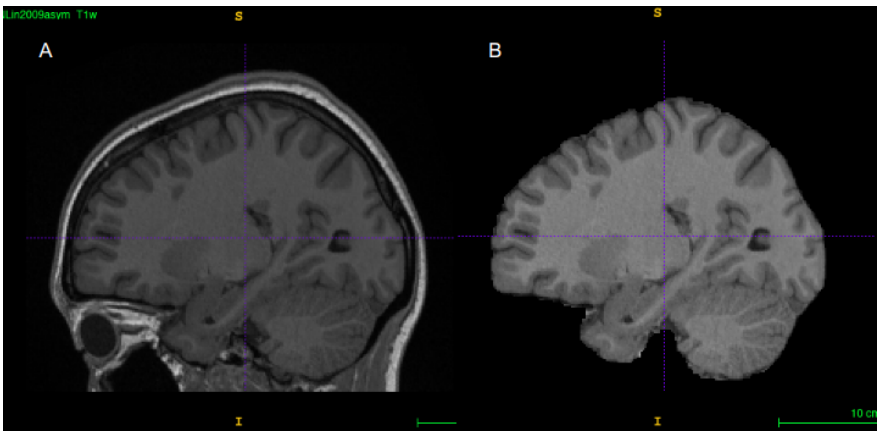

Figure 3.2: Example of preprocessing: A – T1w sequence after preliminary preprocess; B – the same final T1w sequence.

  - Flip angle = -0.28° for sedated studies and -0.12° non-sedated studies
  - FOV = 230 mm2 for sedated and non-sedated studies
  - Voxel size = $0.4 \times 0.4 \times 1$ mm for sedated and $0.9 \times 0.9 \times 1$ mm for non-sedated studies
  - matrix = 256×256 for sedated studies and 230×230 for non-sedated studies
- 3D fluid-attenuated inversion recovery FLAIR.

To the best of our knowledge we have the biggest dataset from the limited number of clinics as total number of subjects in [16] is 148, in [14] is 58, in [2] is 31 and in [3] is 15.

3.2.2 Data preprocessing

In medical image analysis, image preprocessing is a crucial step that must be performed prior to the use of deep neural networks. For each patient, all 3D MRI sequences (T1w, T2w and FLAIR) underwent registration, normalization, morphometry feature extraction and contrast feature extraction. The raw DICOM [5] files were first anonymized and converted to the NIfTI (Neuroimaging Informatics Technology Initiative) format. In order to extract morphometric features we used Freesurfer v7.0.2 [13]. Then obtained surface thickness, curvature, and sulcation maps were translated to volumetric maps with further normalization. For contrast feature extraction, preliminary image processing was performed by registering the T1w NIfTI images to a symmetric template (MNI152) and translating the T2w and FLAIR images to MNI space using evaluated linear mapping on T1w. After this the non-brain tissues were striped, bias corrected and all brains were aligned to standard atlas.

As a result we have following features for deep learning pipeline:

1. T1w intensity;
2. T2w intensity;

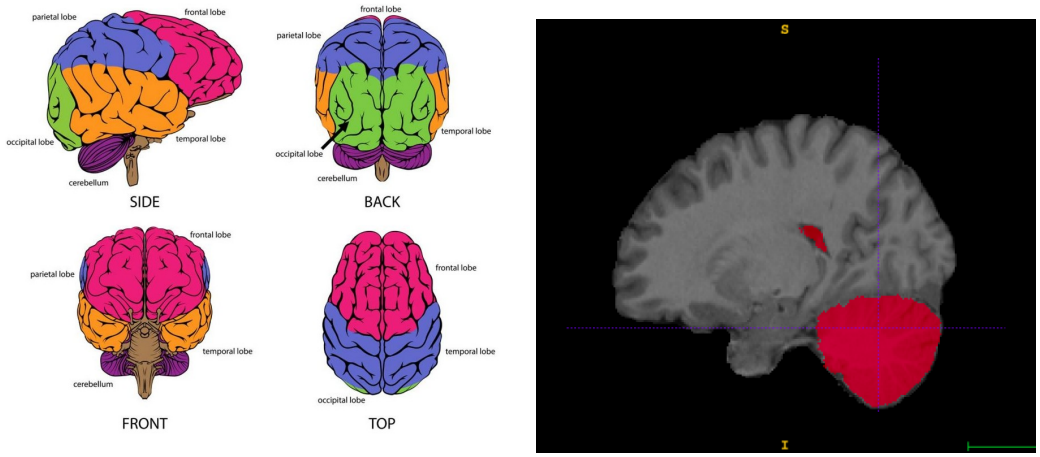

Figure 3.3: *Left.* Brain diagram. Possible options of FCD location are temporal, occipital, frontal and parietal lobes. *Right.* Example of cerebellum mask on of the subjects.

3. FLAIR intensity;
4. cortical thickness;
5. sulcal depth;
6. curvature;
7. T1w-Blurring;
8. T2w-Blurring;
9. T2Flair-Blurring;
10. T2w-CR;
11. T2Flair-CR;
12. Entropy;
13. Variance.

3.2.3 Experimental Setup

Apart from the 3D MR data provided by the radiologists, our dataset also includes valuable meta information about each subject, such as the location of FCD lesions (Fig. 3.3 *Left*) and their specific features, observed during MRI analysis. With this information, we create cerebellum masks for every brain (Fig. 3.3 *Right*), as FCDs can not appear in this region. In certain experiments, we have removed the brain points located under these masks, which we refer to as masked experiments.

To ensure that our deep learning models do not overfit the data, we propose using cross-fold validation. We divide the subjects from each cite into 9 folds based on the available meta information (Fig. 3.4). For each experiment, we train the model on 8 folds and test it on the remaining fold. This approach allows us to evaluate the performance of our models on different subsets of the data and obtain more reliable results.

3.3 Methodology

As it was mentioned above one of the main problem of 3D MR data is the size. In order to analyze the whole MR scan the system should work on GPU clusters, which is not possible in the majority of situations. Previous papers suppose to make a training sequence from crops of single brain. However, this approaches can lead to misdetecting of small leisures on the boarders of such crops. Furthermore, models can not use global information. This thesis focuses on the solution of this problems. We suppose to use sparse representation of the whole brain – point clouds. My main goal is to train a neural network that takes brain MRI in the form of a point cloud as an input and outputs probabilities for each point to be a part of the FCD. One of the most popular approach to work with sparse tensors for segmentation is Minkowski Engine. It is a special framework that provides a faster computations with sparse data.

3.3.1 Point cloud

Point clouds are a collection of points in 3D space that are used to represent the shape and structure of an object or environment. They can be processed and analyzed to extract features such as edges, surfaces, and volumes, which can then be used for further analysis or modeling. Point clouds can also be used as a sparse representation of 3D objects. In this context, ”sparse” refers to the fact that not every point in the space is represented, but rather a subset of points that capture the essential features of the object or environment. Sparse point clouds can be generated using different sampling techniques.

From mathematical point of view point cloud is a tensor of size $N \times k$, where N is the number of points in a point cloud and k stands for 3 space coordinates of a point in an original representation and the last $(k-3)$ are additional features like intensity in T1, T2 modalities in our task (Fig. 3.5).

$$\mathbf{P} = \{(x_i, y_i, z_i, \mathbf{f}_i) | i \in [0, N]\}$$

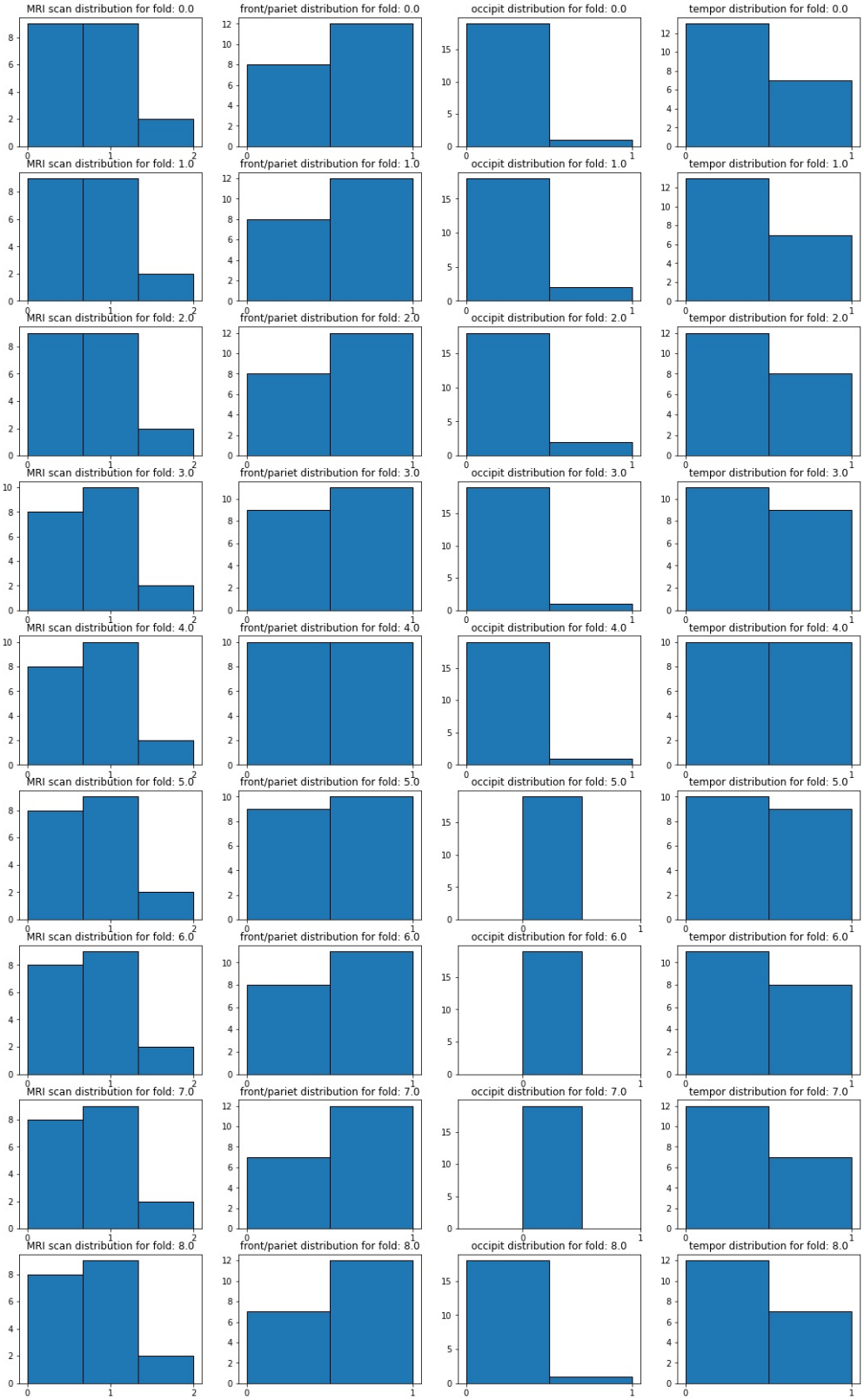

Figure 3.4: Distributions of subjects for every fold based on the site of MRI (first column) and location of FCD (other columns).

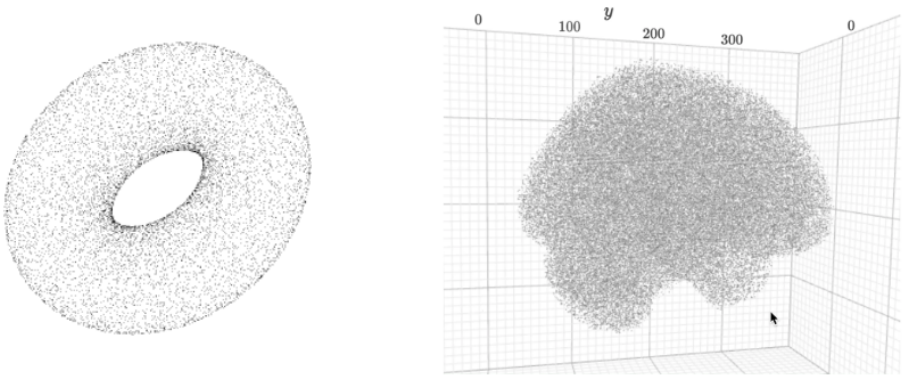

Figure 3.5: *Left.* Example of the Point Cloud. *Right.* Point Cloud of subject from our dataset.

3.3.2 Minkowski Engine

In order to apply neural networks for sparse representations we need to generalize basic operations from Convolutional Neural Networks (CNN). In this thesis we use the functions from Minkowski Engine [9].

Let $\mathbf{f}_{\mathbf{u}}^{\text{in}} \in \mathbb{R}^{K_{\text{in}}}$ – K_{in}-dimensional input feature-vector in a 3-dimensional space at $\mathbf{u} \in \mathbb{R}^3$ (a 3-dimensional coordinate). In this case convolution kernel weights can be represented as

$$\mathbf{W} \in \mathbb{R}^{K^D \times K_{\text{out}} \times K_{\text{in}}}.$$

In [9] authors proposed to break down the weights into spatial weights with K^D matrices of size $K_{\text{out}} \times K_{\text{in}}$. Then, the traditional dense convolution in three dimensions can be represented as:

$$\mathbf{f}_{\mathbf{u}}^{\text{out}} = \sum_{\mathbf{i} \in \mathcal{V}^3(K)} \mathbf{W}_{\mathbf{i}} \mathbf{f}_{\mathbf{u}+\mathbf{i}}^{\text{in}}; \quad \mathbf{u} \in \mathbb{R}^3,$$

where $\mathcal{V}^3(K)$ is the list of offsets in 3-dimensional hypercube centered at the origin. For example, $\mathcal{V}^1(3) = \{-1, 0, 1\}$.

Then the generalized convolution can be written as

$$\mathbf{f}_{\mathbf{u}}^{\text{out}} = \sum_{\mathbf{i} \in \mathcal{N}^D(\mathbf{u}, \mathcal{C}^{\text{in}})} \mathbf{W}_{\mathbf{i}} \mathbf{f}_{\mathbf{u}+\mathbf{i}}^{\text{in}}; \quad \mathbf{u} \in \mathcal{C}^{\text{out}},$$

where $\mathcal{N}^D(\mathbf{u}, \mathcal{C}^{\text{in}}) = \{\mathbf{i} | \mathbf{u} + \mathbf{i} \in \mathcal{C}^{\text{in}}, \mathbf{i} \in \mathcal{N}^D\}$ is a set of offsets that define the shape of a kernel, $\mathcal{C}^{\text{in}}$ and $\mathcal{C}^{\text{out}}$ are predefined input and output coordinates of sparse tensors. Notably, the input and output coordinates may differ.

This formula generalizes various special scenarios, including the dilated convolution and typical hypercubic kernels (Fig. 3.6). For our purposes we will use the version where the input and

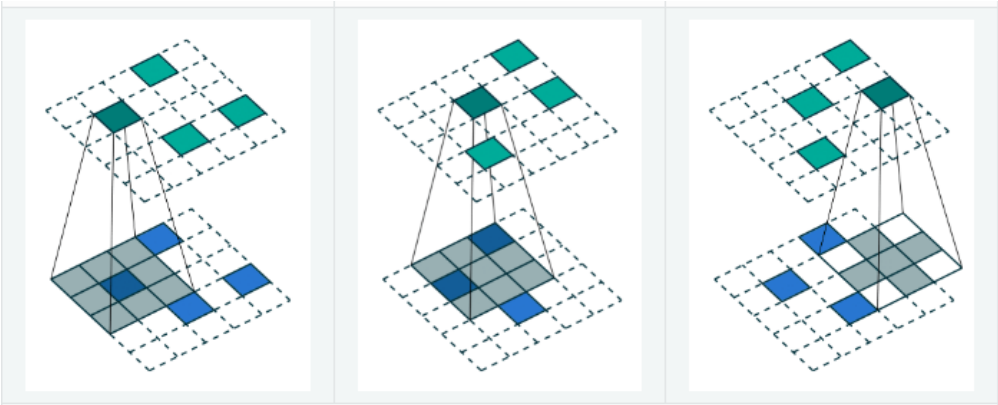

Figure 3.6: Examples of convolution. *Left.* Same in/out coordinates. *Middle.* Arbitrary in/out coordinates. *Right.* Generalized convolution. [9]

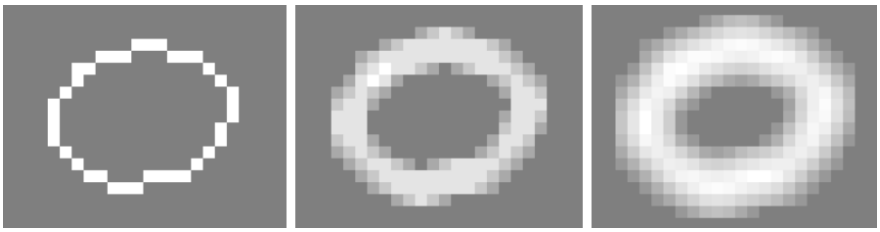

Figure 3.7: Result of applying classic convolution for sparse tensor two times. The sparsity of the grid rapidly disappears. [18]

output coordinates are the coordinates of non-zero elements on a sparse tensor.

If we do not make a restriction on output coordinates, regular convolutions will substantially reduce the sparsity of the features with each convolutional layer. The example of such process is presented on Fig. 3.7.

3.3.3 Model architecture

Many papers empirically proved that U-Net [30] is one of the most suitable for medical segmentation architecture in deep learning [23, 31]. For this reason it was decided to use it sparse modification for point clouds — MinkUNet34C [9]. Instead of simple CNN encoder, this implementation exploits ResNet18 neural network [1] with all layers replaced by sparse ones. So, a block of encoder consists of two subblocks connected by local skip-connection. Each subblock is presented by two sparse convolutions. Between each block on encoder a sparse convolution is used to decrease a spatial size of data. Decoder repeats the structure of the encoder, though here blocks are connected with transposed sparse convolutions. Each ResNet block of the encoder is connected with a skip-connection to block of decoder and a concatenation is applied to aggregate two blocks of features. For more details see Fig. 3.8.

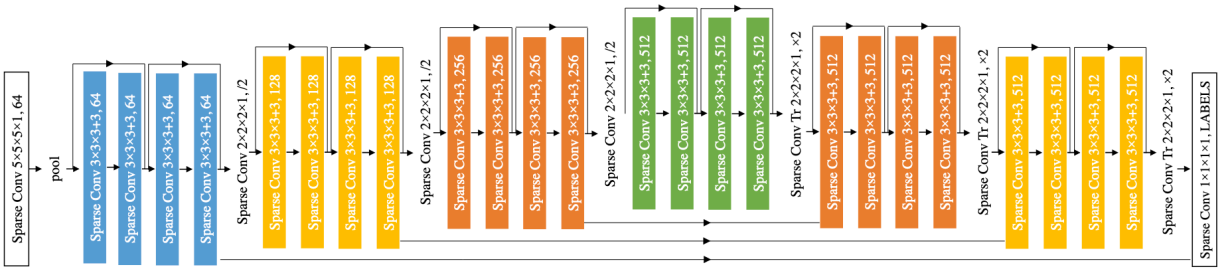

Figure 3.8: MinkUNet

3.3.4 Loss functions.

The main difficulty of the task is the presence of a strong imbalance of classes. So it is very important to choose the appropriate loss function. In this work we proposed to use functions that are described below.

First loss function is an application for imbalanced classification task of classical cross-entropy function - weighted cross-entropy (WCE) function. The formula of this function for 2 classes is presented below.

$$L(y, \hat{y}) = -\frac{1}{N}\sum_{i}^{N} w_1 y_i log(\hat{y}_i) + w_0(1 - y_i)log(1 - \hat{y}_i)$$

y - true binary vector of labels, $\hat{y}$ - predicted vector in which i component is probability that x_i belongs to the class 1 (FCD in our case). w_0 and w_1 are weights of classes. These weights help model to more concentrate on examples of rare class.

The second function I reviewed is Unified Focal loss [35]. This function is hybrid because consist of 2 parts. The first part is modified asymmetric focal loss. Original Focal loss applies a modulating term with focus parameter γ to the cross entropy loss in order to focus learning on hard misclassified examples. Formula of this symmetric loss is the following:

$$L_F(y, \hat{y}) = \frac{-1}{N}\sum_{i}^{N}(1 - \hat{y}_i)^{\gamma} y_i log(\hat{y}_i) + \hat{y}_i^{\gamma}(1 - y_i)log(1 - \hat{y}_i)$$

The modified asymmetric Focal loss removes the focal parameter for the component of the loss relating to the rare class r (FCD in our case). Such asymmetry enables selective enhancement or suppression using the focal parameter by assigning various losses to each class, and this overcomes the both the harmful suppression of the rare class and enhancement of the background class. The formula of the loss is as follows:

$$L_{amF}(y, \hat{y}) = -\frac{\delta}{N}\sum_{c=r} log(\hat{y}_{i,c}) + \frac{1 - \delta}{N}\sum_{c \neq r} \hat{y}_{i,r}^{\gamma} log(\hat{y}_{i,c})$$

$\hat{y}_{i,c}$ means the probability that x_i is belongs to the class c. Coefficient $\delta \in [0, 1]$ reweights components of the loss. Condition $c = r$ means that we sum only prediction for x_i that belongs to class r.

Second component of the Unified Focal loss is modified asymmetric Focal Tversky loss. Focal Tversky loss is a generalization of the Dice loss [33] that allows to reweight the importance of FP and FN metrics while training. Formula of this loss is presented below:

$$L_{FT}(y, \hat{y}) = (1 - TI_0(y, \hat{y}))^{1-\gamma_1} + (1 - TI_1(y, \hat{y}))^{1-\gamma_1}$$

$$TI_c(y, \hat{y}) = \frac{\sum_i^N y_i \hat{y}_{i,c} + \epsilon}{\sum_i^N (y_i \hat{y}_{i,c} + \alpha \hat{y}_{i,c}(1 - y_i) + \beta \hat{y}_{i,\bar{c}} y_i) + \epsilon}$$

α and β - weight that controls importance of FP and FN, their sum must be is equal to 1. $\bar{c}$ means other class (if $c = 0$, then $\bar{c} = 1$). In asymmetric case of the loss the focal parameter γ remains only for component which corresponds to the rare class r:

$$L_{amFT}(y, \hat{y}) = (1 - TI_0(y, \hat{y}) + (1 - TI_1(y, \hat{y}))^{1-\gamma_1}$$

Finally the Unified Focal loss constructed as a weighted sum of modified asymmetric Focal loss and modified asymmetric Focal Tversky loss:

$$L_{maUF} = \lambda L_{amF} + (1 - \lambda) L_{amFT}$$

For our specific dataset we empirically calculated hyperparameters of described losses to equalize the contribution of its components. For WCE the weights were as follows: $w_0 = 0.001$ and $w_1 = 0.999$. For Unified Focal loss we consider the following hyperparameters: $\lambda = 0.999$, $\gamma = 0.2$, $\gamma_1 = 0.5$, $\delta = 0.999$, $\alpha = 0.9$ and $\beta = 0.1$. The final value of the Unified Focal loss was multiplied by 100 to better stability of calculations.

3.3.5 Metrics

To evaluate our proposed models we consider several metrics: recall, precision, Dice, specificity and top-k metric.

Precision is a commonly used metric in medical image segmentation, as it provides a measure of the accuracy of the model in identifying true positives. A high precision value indicates that the model is able to accurately identify positive instances, while a low precision value indicates that the model is prone to false positives.

Dice is another popular metric in medical image segmentation, as it provides a measure of the overlap between the predicted and ground truth segmentations. A high Dice value indicates a

high degree of similarity between the predicted and ground truth segmentations, while a low Dice value indicates poor overlap. In our study, this metric can be relatively small as we solve two task at once. First of all it is classification and localization task. It means that it is more vital to show possible regions of FCD, while professional radiologists will make final decision. And the second task is precise segmentation of the region. Therefore, even dice value greater 0.1 can be interpreted as detected case of FCD.

Sensitivity and specificity are also important metrics in medical image segmentation. They provide a measure of the ability of the model to correctly identify positive and negative instances. Sensitivity measures the proportion of true positives that are correctly identified by the model, while specificity measures the proportion of true negatives that are correctly identified by the model.

$$\textit{Sensitivity} = \frac{TP}{TP + FN} \in [0, 1]$$

TP - the number of right predicted FCD points. FN - number of misclassified FCD points.

As classic metrics not appropriate for the medical task, team of the project made a new metric top-k in collaboration with radiologists.

The algorithm for calculating this metric involves the following steps:

- Dividing the heat-map into crops of a specific size that intersect with each other.
- Calculating the average confidence for each crop.
- Sorting the crops based on their average confidence.
- Identifying epileptogenic lesions as detected if at least one of the top k crops intersects with the ground truth FCD area by more than 20%.

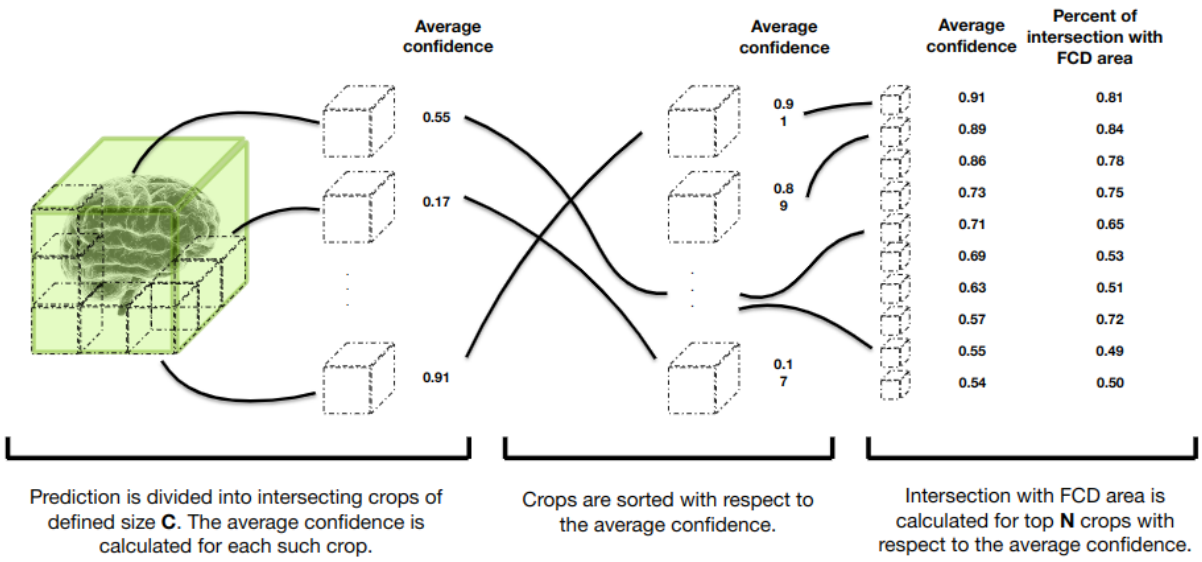

Figure 3.9: Scheme of top-k metric.

Chapter 4

Results and Discussion

4.1 Results

First of all, all approaches were tested on the stratified dataset from one hospital to reduce effects of radiologists' subjectivity and differences in medical equipment. This part consists of 90 brains from previously described dataset.

As shown in 4.1 the best results in terms of dice metric showed model that was trained with BCE loss, where we used cerebellum masks for data and give no information about statistical features. This set of hyperparameters were used as basis. Model from the state-of-the-art-approach [16] was fine-tuned and also evaluated on the collected dataset.

The results of our best model on the stratified sample from one cite have shown promising outcomes for the proposed approach in detecting epileptogenic lesions in MRI scans. On Fig. 4.1 we can see that the top-10 metric proves the effectiveness of the method in more than 70% of the brains, as the first 10 crops intersected with the ground truth label. This indicates that the proposed method can help doctors to analyze the MRI scans. Moreover, in 35% of the brains, all FCD areas were found to lie in the most probable crops, which further highlights the effectiveness of the proposed approach.

In addition, the results of top-10 metric is comparable to the performance of the state-of-the-art approach [16] (orange on Fig. 4.1).

Furthermore, the dice metric was found to be better for 63 out of 90 brains in comparsion with SOTA, which indicates that the proposed approach is able to provide more accurate segmentation results compared to the state-of-the-art approach. The recall metric was found to be better for almost all brains, which further highlights the effectiveness of the proposed approach.

Input	Loss	Masked	Number of detection by Dice
T1w, T2w, FLAIR + 10 features	Focal	Yes	29
T1w, T2w, FLAIR + 10 features	BCE	Yes	28
T1w, T2w, FLAIR	BCE	Yes	37
T1w, T2w, FLAIR	BCE	No	27

Table 4.1: Result of experiments with input combinations and different losses.

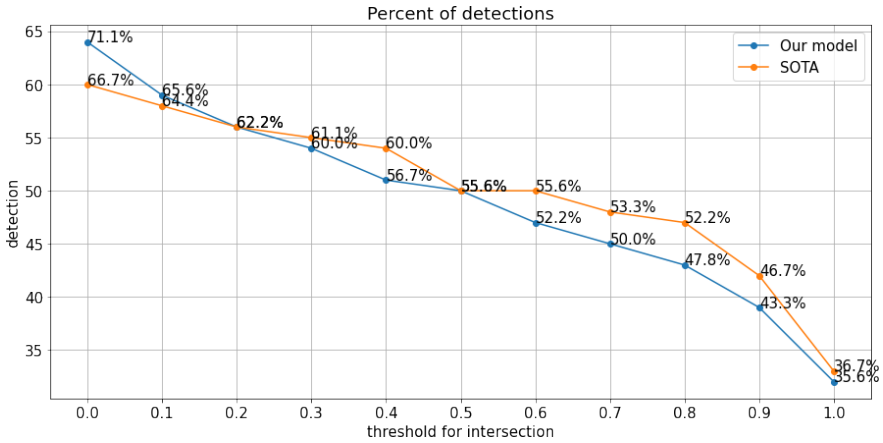

Figure 4.1: Results for the top-10 metric for the best set of hyperparameters for the first hospital.

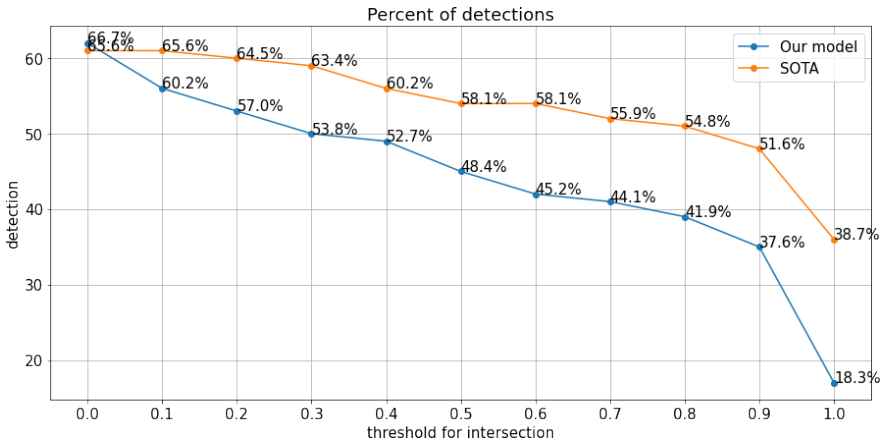

Figure 4.2: Results for the top-10 metric for the best set of hyperparameters for the second hospital.

The results for the second part of the dataset can be seen on Fig. 4.2. In terms of top-10 metric for more than 60% of the brains the model was effective, as the first 10 crops intersected with the ground truth label. However, for this sample the state-of-the-art approach showed better average performance.

Generalized results are provided on Fig. 4.3. We can say that our model is comparable to the state-of-the-art model in general.

4.2 Visual Comparison

On Fig. 4.4, 4.5, 4.6, 4.7, 4.8, 4.9 green region is ground truth, blue one is prediction from proposed pipeline and red is prediction of SOTA model [16]. On Fig. 4.4 and Fig. 4.5 our pipeline outperforms SOTA model. It make more precise predictions, while producing less false positives. These is really vital for support system as doctors will need less time to find FCD region.

On Fig. 4.6 we can see an example of subject with two FCD regions. Our model detect the bigger one and produce no other predictions. The SOTA model produces many false positives, but have predictions in both regions.

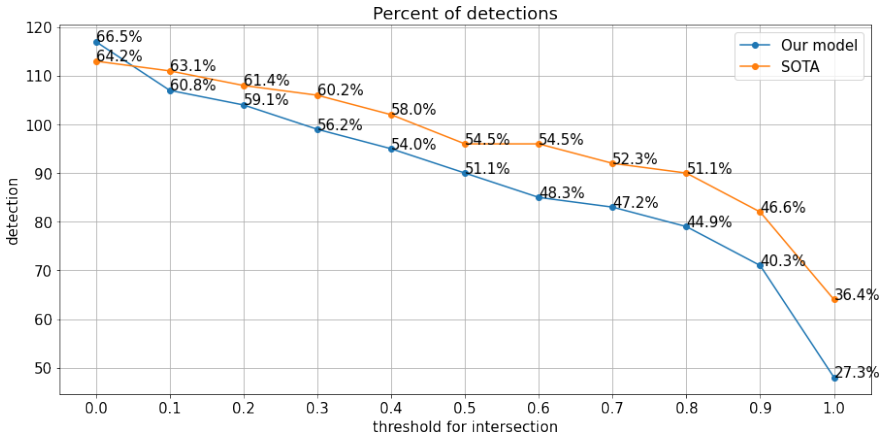

Figure 4.3: Results for the top-10 metric for the best set of hyperparameters for the both hospitals.

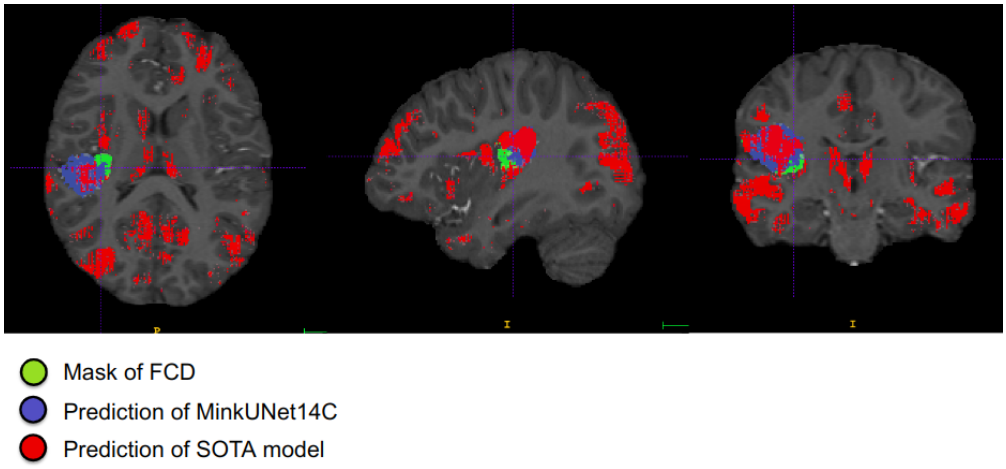

Figure 4.4: Example of prediction, where proposed approach beats SOTA model.

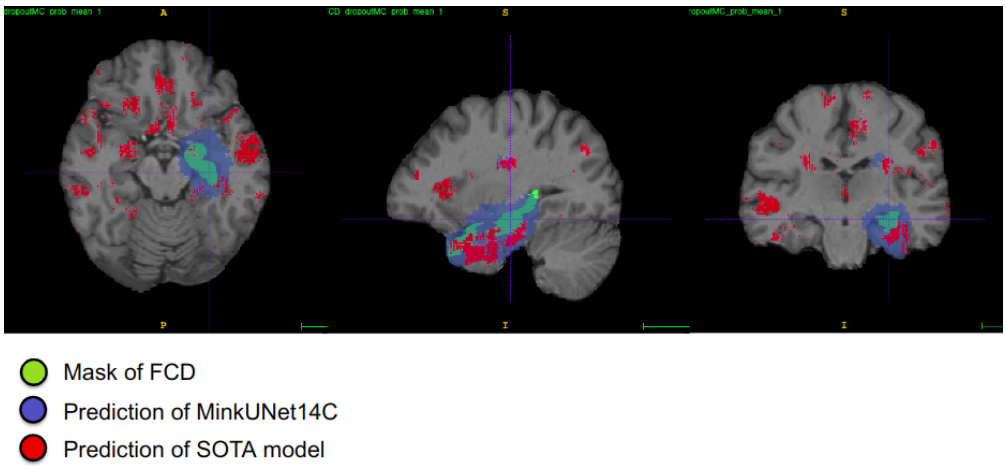

Figure 4.5: Example of prediction, where proposed approach beats SOTA model.

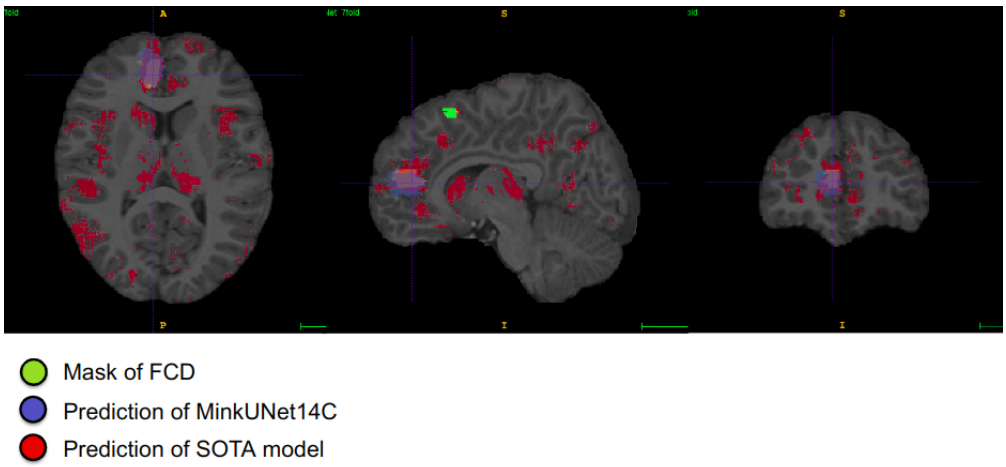

Figure 4.6: Example of prediction, where proposed approach works comparable to SOTA model.

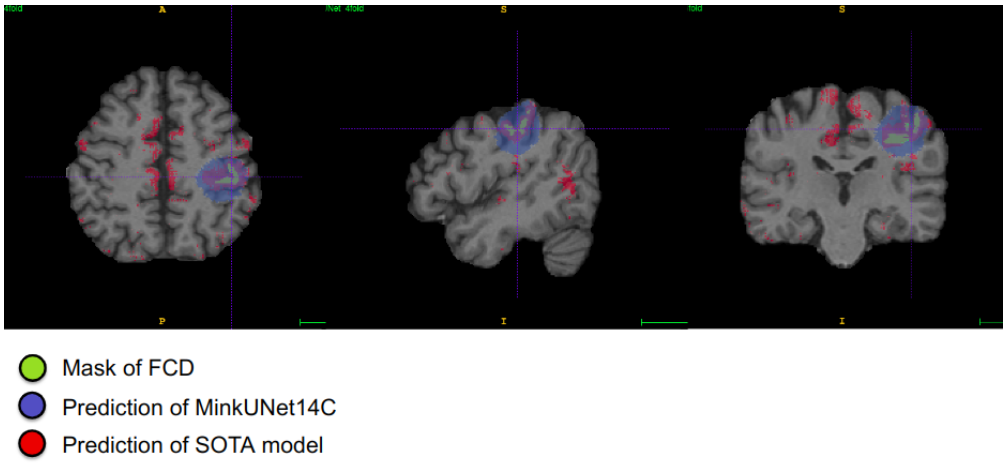

Figure 4.7: Example of the subject, where dice metric is not an appropriate.

On Fig. 4.7 there is an example of the subject, where our prediction covers the whole FCD region. Although the FCD is detected, but the dice score for this subject is relative small (less than 0.1). This illustrate the idea of using different metrics for our task.

On Fig. 4.8 there is an example of the subject, where both approaches can not detect FCD. The reason for this is relatively small volume of FCD region. However, our approach proposes less false positives. Furthermore, the confidence of the prediction is small, so if we use high threshold for prediction, there will be no prediction for our model at all.

On Fig. 4.9 there is an example of the subject, where our approach failed, while SOTA model works relatively good. That means that one of the possible use cases is ensemble of approaches, where predictions can be averaged or united to make more stable system.

4.3 Discussion

The approach showed results better than the SOTA model in many cases. At the same time sets of non-detected subjects differ between models. So it is possible to use them as ensemble and improve

Figure 4.8: Example of the subject, where both models can not detect small FCD leisure.

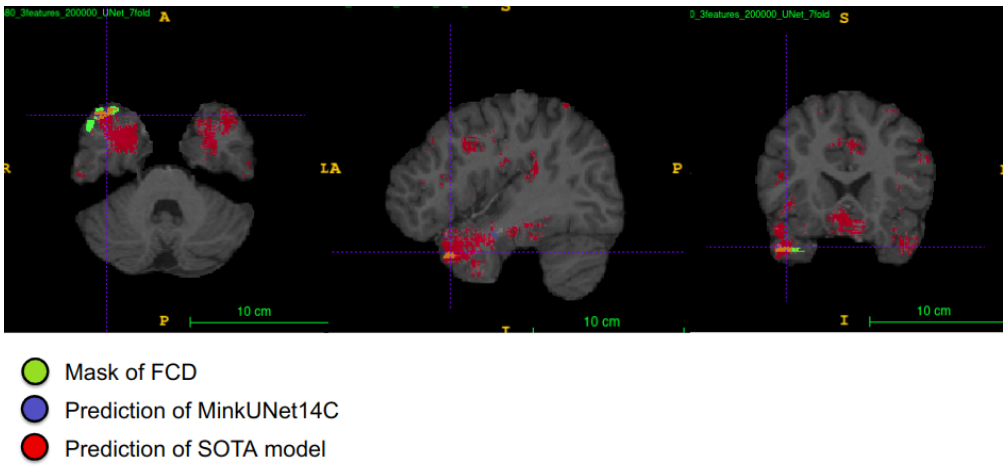

Figure 4.9: Example of the subject, where our model failed in comparison with SOTA one.

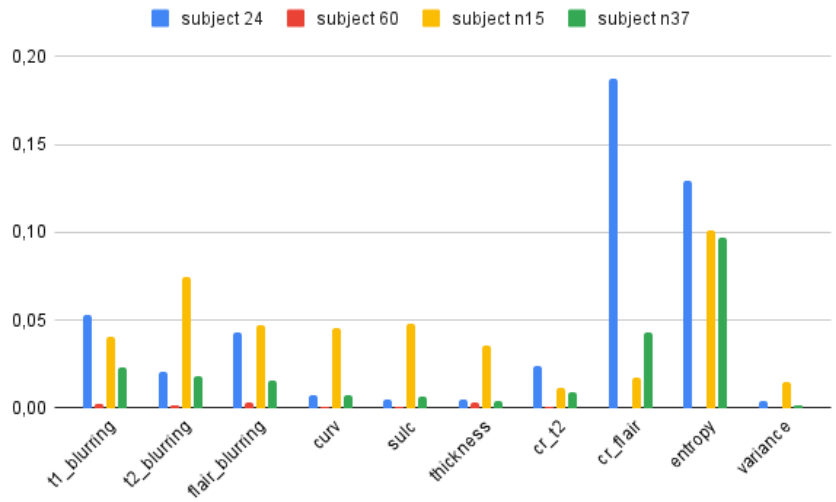

Figure 4.10: Distribution of recall metric for different features.

results.

One of the main problems of prediction is low dice with recall metric around 1. This means that our prediction covers FCD region, but propose lots of false positives. For our task it is better, than have no prediction at all, as our task is to build support system for doctors. So we propose radiologists several possible regions and then doctors make final decision about FCD. Furthermore, false positives can be reduced by smart post processing as deleting of too small clusters.

Another interesting question is correlation between statistical features, that can be count for every subject, and predictions of our model. We took heat maps of detected subjects and compared it with statistical feature maps. The importance of feature for subjects are presented in Fig. 4.10.

As we can see the recall metrics for blurrings and cortical thickness are always relatively high for all subjects in sample. While sulcal depth, curvature and other features can correlate or not with the predicted heat map. It is vital to notice, that for some subjects entropy is the feature with the highest recall, but it is not a rule. So predictions and entropy heat maps may have no intersections at all as for subject 60.

Our proposed approach demonstrated high levels of accuracy in identifying FCD in medical images, indicating its potential for clinical use. However, it is important to note that this model is less interpretable than statistical feature maps. Further research is needed to address these factors and improve the performance of the model.

Chapter 5

Conclusions

1. New method in the task of detecting Focal Cortical Displasya was proposed.
2. The effectiveness of this approach was demonstrated in the experiments on collected dataset.
3. FCD areas were detected in 108 out of 183 cases, which is comparable to the state-of-the-art approach.
4. The sets of non-detected subjects differed between proposed approach and the state-of-the-art one, suggesting that an ensemble of these models could be used to obtain more reliable predictions.
5. Overall, neural networks for point clouds segmentation have proven to be useful for MRI analysis.

Chapter 6

Author Contribution

1. Research on the new method of MRI representation;
2. Implementation of training pipeline for sparse data representation of MRI;
3. Checking the statistical properties of sampled points;
4. The experiments with different models, losses, features;
5. Making of cerebellum masks;
6. The evaluation of metrics, interpolation of prediction to MRI format and MRI visual comparison.

Chapter 7

Acknowledgements

I would like to thank my scientific colleges and advisors from the Skoltech's research group:

1. Professor Maxim Sharaev;
2. Vyacheslav Yarkin;
3. Professor Evgeny Burnaev;
4. Professor Alexander Bernstein.

and our project medical partners from scientific centers named after V. I. Kulakov and named after N. I. Pirogov:

1. Egor Syrkashev;
2. Aleksei Marinets.

for their consistent support and guidance during the project.

Bibliography

[1] Deep residual learning for image recognition. 12 2015.

[2] Bilal Ahmed, Carla E Brodley, Karen E Blackmon, Ruben Kuzniecky, Gilad Barash, Chad Carlson, Brian T Quinn, Werner Doyle, Jacqueline French, Orrin Devinsky, and Thomas Thesen. Cortical feature analysis and machine learning improves detection of "MRI-negative" focal cortical dysplasia. *Epilepsy Behav.*, 48:21–28, July 2015.

[3] Ruslan Aliev, Ekaterina Kondrateva, Maxim Sharaev, Oleg Bronov, Alexey Marinets, Sergey Subbotin, Alexander Bernstein, and Evgeny Burnaev. Convolutional neural networks for automatic detection of focal cortical dysplasia, 2020.

[4] Laurens van der Maaten Benjamin Graham. Submanifold sparse convolutional networks. 06 2017.

[5] W D Bidgood, Jr, S C Horii, F W Prior, and D E Van Syckle. Understanding and using DICOM, the data interchange standard for biomedical imaging. *J. Am. Med. Inform. Assoc.*, 4(3):199–212, May 1997.

[6] K.M. Bijay Dev, Pawan S. Jogi, S. Niyas, S. Vinayagamani, Chandrasekharan Kesavadas, and Jeny Rajan. Automatic detection and localization of focal cortical dysplasia lesions in mri using fully convolutional neural network. *Biomedical Signal Processing and Control*, 52:218–225, 2019.

[7] Fred L Bookstein. "voxel-based morphometry" should not be used with imperfectly registered images. *Neuroimage*, 14(6):1454–1462, 2001.

[8] Fernando Cendes, William H Theodore, Benjamin H Brinkmann, Vlastimil Sulc, and Gregory D Cascino. Neuroimaging of epilepsy. *Handb. Clin. Neurol.*, 136:985–1014, 2016.

[9] Chris Choy, JunYoung Gwak, and Silvio Savarese. 4d spatio-temporal convnets: Minkowski convolutional neural networks. pages 3070–3079, 06 2019.

[10] O Colliot, N Bernasconi, N Khalili, S B Antel, V Naessens, and A Bernasconi. Individual voxel-based analysis of gray matter in focal cortical dysplasia. *Neuroimage*, 29(1):162–171, January 2006.

[11] Angela Dai, Angel X. Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Nießner. Scannet: Richly-annotated 3d reconstructions of indoor scenes. In *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE*, 2017.

[12] Lu F and Lysack JT. Lessons learned from commonly missed head and neck cancers on cross-sectional imaging. *Canadian Association of Radiologists journal = Journal l'Association canadienne des radiologistes*, 08 2022.

[13] Bruce Fischl. Freesurfer. *NeuroImage*, 62(2):774–781, 2012. 20 YEARS OF fMRI.

[14] Zohreh Ganji, Mohsen Aghaee Hakak, Seyed Amir Zamanpour, and Hoda Zare. Automatic detection of focal cortical dysplasia type ii in mri: Is the application of surface-based morphometry and machine learning promising? *Frontiers in Human Neuroscience*, 15, 2021.

[15] Garrett L. Garner, Daniel R. Streetman, Joshua G. Fricker, Nicholas E. Bui, Chenyi Yang, Neal A. Patel, Nolan J. Brown, Shane Shahrestani, India C. Rangel, Rohin Singh, and Julian L. Gendreau. Focal cortical dysplasia as a cause of epilepsy: The current evidence of associated genes and future therapeutic treatments. *Interdisciplinary Neurosurgery*, 30:101635, 2022.

[16] Ravnoor Gill, Hyo-Min Lee, Benoit Caldairou, Seok□Jun Hong, Carmen Barba, Francesco Deleo, Ludovico D'Incerti, Vanessa Coelho, Matteo Lenge, Mira Semmelroch, Dewi Schrader, Fabrice Bartolomei, Maxime Guye, Andreas Schulze-Bonhage, Horst Urbach, Kyoo Cho, Fernando Cendes, Renzo Guerrini, Graeme Jackson, and Andrea Bernasconi. Multicenter validation of a deep learning detection algorithm for focal cortical dysplasia. *Neurology*, 97:10.1212/WNL.0000000000012698, 09 2021.

[17] Ravnoor S. Gill, Seok-Jun Hong, Fatemeh Fadaie, Benoit Caldairou, Boris Bernhardt, Neda Bernasconi, and Andrea Bernasconi. Automated detection of epileptogenic cortical malformations using multimodal mri. In M. Jorge Cardoso, Tal Arbel, Gustavo Carneiro, Tanveer Syeda-Mahmood, João Manuel R.S. Tavares, Mehdi Moradi, Andrew Bradley, Hayit Greenspan, João Paulo Papa, Anant Madabhushi, Jacinto C. Nascimento, Jaime S. Cardoso, Vasileios Belagiannis, and Zhi Lu, editors, *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 349–356, Cham, 2017. Springer International Publishing.

[18] Benjamin Graham, Martin Engelcke, and Laurens van der Maaten. 3d semantic segmentation with submanifold sparse convolutional networks. pages 9224–9232, 06 2018.

[19] W Allen Hauser. Epilepsy: frequency, causes and consequences. *Epilepsy foundation of America*, 1990.

[20] Philip H Iffland and Peter B Crino. Focal cortical dysplasia: gene mutations, cell signaling, and therapeutic implications. *Annual Review of Pathology: Mechanisms of Disease*, 12:547–571, 2017.

[21] Joanna Kabat and Przemysław Król. Focal cortical dysplasia - review. *Pol. J. Radiol.*, 77(2):35–43, April 2012.

[22] Patrick Kwan, Steven C Schachter, and Martin J Brodie. Drug-resistant epilepsy. *New England Journal of Medicine*, 365(10):919–926, 2011.

[23] Haoran Lu, Yifei She, Jun Tie, and Shengzhou Xu. Half-unet: A simplified u-net architecture for medical image segmentation. *Frontiers in Neuroinformatics*, 16, 2022.

[24] Pascal Martin, Gavin P. Winston, Philippa Bartlett, Jane de Tisi, John S. Duncan, and Niels K. Focke. Voxel-based magnetic resonance image postprocessing in epilepsy. *Epilepsia*, 58(9):1653–1664, 2017.

[25] A Mechelli, C J Price, K J Friston, and J Ashburner. Voxel-based morphometry of the human brain: Methods and applications. *CURRENT MEDICAL IMAGING REVIEWS*, 1(2):105–113, June 2005.

[26] S. Niyas, S. Chethana Vaisali, Iwrin Show, T.G. Chandrika, S. Vinayagamani, Chandrasekharan Kesavadas, and Jeny Rajan. Segmentation of focal cortical dysplasia lesions from magnetic resonance images using 3d convolutional neural networks. *Biomedical Signal Processing and Control*, 70:102951, 2021.

[27] Oscar Perdomo, Sebastian Otálora, Fabio A. González, Fabrice Meriaudeau, and Henning Müller. Oct-net: A convolutional network for automatic classification of normal and diabetic macular edema using sd-oct volumes. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 1423–1426, 2018.

[28] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. *arXiv preprint arXiv:1612.00593*, 2016.

[29] Charles R Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *arXiv preprint arXiv:1706.02413*, 2017.

[30] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[31] Nahian Siddique, Paheding Sidike, Colin Elkin, and Vijay Devabhaktuni. U-net and its variants for medical image segmentation: A review of theory and applications. *IEEE Access*, PP:1–1, 06 2021.

[32] Henrik Skibbe, Muhammad Febrian Rachmadi, Ken Nakae, Carlos Enrique Gutierrez, Junichi Hata, Hiromichi Tsukada, Charissa Poon, Kenji Doya, Piotr Majka, Marcello G. P. Rosa, Hideyuki Okano, Tetsuo Yamamori, Shin Ishii, Marco Reisert, and Akiya Watakabe. The brain/minds marmoset connectivity atlas: exploring bidirectional tracing and tractography in the same stereotaxic space. *bioRxiv*, 2022.

[33] Carole H Sudre, Wenqi Li, Tom Vercauteren, Sebastien Ourselin, and M Jorge Cardoso. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep learning in medical image analysis and multimodal learning for clinical decision support*, pages 240–248. Springer, 2017.

[34] Lyne P. Tchapmi, Christopher B. Choy, Iro Armeni, JunYoung Gwak, and Silvio Savarese. Segcloud: Semantic segmentation of 3d point clouds, 2017.

[35] Michael Yeung, Evis Sala, Carola-Bibiane Schönlieb, and Leonardo Rundo. Unified focal loss: Generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation. *Computerized Medical Imaging and Graphics*, 95:102026, 2022.