

**MEDICAL IMAGES CLASSIFICATION FOR SMALL DATASETS
EXISTING APPROACHES AND CHALLENGES**

Kayali Ismail (ITMO University)

Supervisor –Associate Professor Ph.D. – Gusarova Natalia (ITMO University)

Abstract. In the world of medical imaging, achieving accurate imaging diagnoses is critical to the success of patient treatment. Imaging diagnostics through medical research. With technological advancements, the use of images in healthcare has increased quickly, and medical imaging has become more sophisticated and accurate, enabling doctors to detect and diagnose diseases more accurately. The complexity and sparsity of the data still pose a challenge to the precise classification and analysis of these images. In this report we will discuss the medical image classification for small datasets, existing approaches, and challenges.

Introduction. Medical imaging presents several challenges when it comes to an accurate diagnosis. One of the biggest challenges is the scant amount of data that is currently available for image diagnostics. It is challenging to make an accurate diagnosis because medical imaging data sets are frequently small and constrained in size. Furthermore, the high levels of noise present in most medical images can make it difficult to diagnose diseases precisely.

The need for a medical image classification dataset is evident in the current state of healthcare. With ever-increasing amounts of data being generated, it is becoming increasingly difficult to manage and interpret that data.

A well-designed annotated medical image classification dataset can help address these issues by providing a way to automatically extract useful information from images. This would allow for better management of medical images and could lead to lower healthcare-related costs.

Main part. Existing Approaches to Medical Image Classification

Classifying medical images can be done in many different approaches, each with advantages and disadvantages of their own.

Typical techniques include:

1. Traditional machine learning: this approach uses a set of labeled training data to train the classifier, which can then be used to label the new data. This approach is effective but can be limited by the volume and quality of training data.
2. Deep learning: this approach uses a deep neural network to learn features from raw data. This can be efficient but computationally intensive and requires large amounts of training data.
3. Transfer learning: this approach uses a pre-trained model to learn features from new data. This can be effective when the available training data is limited.
4. Hybrid methods: this approach combines multiple methods to improve classification performance. This can be effective but requires careful design and tuning of various components.

Image Classification with Small Datasets:

To accurately diagnose diseases in medical imaging, it is necessary to use image classification techniques. Image classification is the process of categorizing images into different classes, such as healthy and diseased. Image classification with small datasets is a challenge, as it can be difficult to accurately classify images with limited data.

One approach to image classification with small datasets is to use transfer learning. Transfer learning is the process of applying knowledge gained from one task to another. In the context of image classification, transfer learning involves taking knowledge gained from a large dataset and applying it to a smaller dataset. This allows for the accurate classification of images with small datasets.

Another approach to image classification with small datasets is to use data augmentation. Data augmentation is the process of artificially increasing the amount of data available for training. This

can be done by rotating, flipping, and cropping images to create new images. This allows for the accurate classification of images with small datasets.

Challenges in Small Medical Image Classification Datasets:

There are several issues that make small medical image classification datasets difficult to work with. Firstly, the images in these datasets are usually of lower quality than those in larger datasets. This makes it harder to extract useful features from them and can lead to poorer results.

Secondly, the class labels in small medical image datasets are often imbalanced, meaning that there are more images of one class than another. This can make it difficult for a classifier to learn from the data and can again lead to poorer results.

Finally, small medical image datasets are often very heterogeneous, meaning that they contain a wide variety of different types of images. This can make it hard to train a classifier that is able to generalize well to new data.

Conclusions. Image diagnosis in medical imaging is an essential part of medical imaging, and it is becoming increasingly difficult to achieve accurate diagnosis with the limited amount of data available. In conclusion, the development of effective medical image classification techniques for small datasets is an important area of research. Existing approaches are limited by the availability and size of datasets, and existing challenges need to be addressed in order to reduce errors due to overfitting and improve generalization ability. Therefore, by understanding the challenges and exploring approaches to achieving accurate image diagnosis, medical professionals can ensure that patients receive the best possible treatment. With advancements in computer vision technologies, deep learning methods can be used to address these challenges and provide more accurate results for medical image classification on small datasets.

Kayali Ismail (Author)

Signature

Gusarova Natalia (Supervisor)

Signature