A comprehensive study of Semi-Supervised Learning in Medical Imaging

1. General Info

Project Title: A comprehensive study of Semi-Supervised Learning in Medical Imaging

Contact Person: Tariq Bdair

Contact Email: t.bdair@tum.de

2. Project Abstract

In the last few years, deep neural networks (DNNs) have shown impressive success in different real-life problems such as computer versions, speech recognition, and medical imaging. DNNs demand a huge amount of high-quality labeled data, which is very expensive and difficult to find. To overcome the above limitations, the semi-supervised learning (SSL) paradigm has been widely used in the literature. SSL methods utilize few labeled data in a combination with a large amount of the unlabeled data in the training process. SSL methods can be divided into different categories. Generative models, Graph-based models, Self-training, Entropy minimization, Consistency regularization. However, the current state-of-art (SOTA) methods are considered holistic approaches that combine consistency regularization, pseudo labeling, and data augmentation and regularization methods. While these methods have proven their success in different real-life situations, no previous work has conducted a fair and comprehensive analysis using unified baseline architecture and environments in medical imaging. Recently, semi-supervised federated learning has gin much intention due to the ability to leverage the scarcity of the label data through training on distributed data in a privacy-preserved manner. Thus, understanding and benchmarking these methods are of high importance for any researcher especially when it comes to the medical data.

3. Project goals

1. Conducting a fair and comprehensive analysis of the current SOTA SSL methods in unified architecture following the guidelines of [7].
2. Showing the applicability of these methods in medical image classification and segmentation.
3. Showing the applicability of these methods in a semi-supervised federated learning setting.

4. Technical Prerequisites

1. Solid background in Machine/Deep Learning.
2. Familiar with Unet, ResNet, CNN, ...etc.
3. Sufficient knowledge of Python programming language and libraries (Scikit-learn, NumPy, Tensorboard, tSNE...).

4. Experience with a mainstream deep learning framework such as PyTorch or Tensorflow.
5. Machine/Deep learning hands-on experience

6. Benefits:

1. Learning semi-supervised learning deep learning.
2. Learning federated semi-supervised learning.
3. Get on-hand experience and implementing SSL method for medical image classification and segmentation.
4. Get on-hand experience and implementing federated learning for medical images..
5. Possibility of writing a scientific paper.

7. Students' Tasks Description

Group 1 (Two students):

- Familiarize yourself with the literature on semi-supervised and federated learning paradigms.
- Implementing a unified/baseline model for medical image classification.
- Implementing a SOTA SSL method from each category in SSL.
- Conducting a realistic evaluation of the Implemented methods following the guidelines of [7] (not all experiments).
- Applying and implementing the above methods in the basic federated semi-supervised setting.
- Conducting a realistic evaluation of the Implemented methods following the guidelines of [7] in a federated semi-supervised setting. (Optional).
- Writing and summarizing the findings in the project report.

Group 2 (Two students):

- Familiarize yourself with the literature on semi-supervised and federated learning paradigms.
- Implementing a unified/baseline model for medical image segmentation.
- Implementing a SOTA SSL method from each category in SSL.
- Conducting a realistic evaluation of the Implemented methods following the guidelines of [7] (not all experiments).
- Applying and implementing the above methods in the basic federated semi-supervised setting.
- Conducting a realistic evaluation of the Implemented methods following the guidelines of [7] in a federated semi-supervised setting. (Optional).
- Writing and summarizing the findings in the project report.

8. Work-packages and Time-plan:

	Description	#Students	From	To
WP1	Familiarize yourself with the literature on semi-supervised and federated learning paradigms	4		
WP2	Implementing a unified/baseline model for medical image classification/segmentation	2/2		
WP3	Implementing a SOTA SSL method from each category in SSL	2/2		
WP4	Conducting a realistic evaluation of the Implemented methods following the guidelines of [7] (not all experiments)	2/2		
M1	Intermediate Presentation II	4		
WP5	Applying and implementing the above methods in the basic federated semi-supervised setting	2/2		
WP6	Conducting a realistic evaluation of the Implemented methods following the guidelines of [7] in a federated semi-supervised setting. (Optional)	2/2		
WP7	Writing and summarizing the findings in the project report	4		
M2	Final Presentation	4		

[1] Van Engelen JE, Hoos HH. A survey on semi-supervised learning. Machine Learning. 2020 Feb;109(2):373-440.

[2] Schmarje L, Santarossa M, Schröder SM, Koch R. A survey on semi-, self-and unsupervised learning for image classification. IEEE Access. 2021 May 27;9:82146-68.

[3] Yang X, Song Z, King I, Xu Z. A survey on deep semi-supervised learning. arXiv preprint arXiv:2103.00550. 2021 Feb 28.

[4] Zhu XJ. Semi-supervised learning literature survey.

[5] Song Z, Yang X, Xu Z, King I. Graph-based semi-supervised learning: A comprehensive review. IEEE Transactions on Neural Networks and Learning Systems. 2022 Mar 18.

[6] Li T, Sahu AK, Talwalkar A, Smith V. Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine. 2020 May 1;37(3):50-60.

[7] Oliver A, Odena A, Raffel CA, Cubuk ED, Goodfellow I. Realistic evaluation of deep semi-supervised learning algorithms. Advances in neural information processing systems. 2018;31.