HARAMAYA UNIVERSITY

Building the Basis for Development

HARAMAYA UNIVERSITY

SCHOOL OF POSTGRADUATE STUDIES

COLLEGE OF COMPUTING AND INFORMATICS

DEPARTMENT: COMPUTER SCIENCE

PROGRAM: MASTER IN COMPUTER SCIENCE

COURSE NAME: SEMINAR IN COMPUTER SCIENCE

COURSE CODE: COSC672

SEMINAR PAPER REPORT: "SEMI-SUPERVISED MEDICAL IMAGE SEGMENTATION VIA CROSS TEACHING BETWEEN CNN AND TRANSFORMER"

NAME:- Yetsedaw Worku Tadesse

ID: PGP/710/15

Advisor: Mr. Leweyehu Y**.**

SUBMITTED TO: COMPUTER SCIENCE DEPARTMENT

HARAMAYA, ETHIOPIA

Submission date: 05/10/2024

Acknowledgement

I would like to thank my sincere gratitude to all those who have contributed to the completion of this seminar report. First and foremost, thank God Almighty for giving strength, courage and blessings to complete this work.

I am grateful to Mr. Leweyehu Yirsawi(Seminar Advisor) for his support, and encouragement. His feedbacks and insights have been instrumental in shaping this work.

I also would like thank my senior friends for their encouragement and giving direction during the course of this report development. Their moral support has been a constant source of motivation.

Finally, I am grateful to all those who have contributed to the completion of this seminar report.

Your support has been invaluable in bringing this report to be successful.

List Of Definition Of Terms and Acronyms

Acronyms	Description
ACDC	Automated Cardiac Diagnosis challenge
CNN	Convolutional neural network
DSC	'Dice Similarity Coefficient
HD95	95% Hausdorff Distance
MM	Millimeter.
MIS	Medical Image Segmentation
MRI	Magnetic resonance imaging
SSSM	Selective State Space Model
SSL	semi-supervised Learning
SGD	Stochastic Gradient Descent
SOTA	State-of-the-art
Vmamba	Visual Mamba

Abstract

This is a seminar report on the paper "Semi-Supervised Medical Image Segmentation via cross teaching between CNN and Transformers.". The paper introduced a cross-teaching method that leverages the strengths of CNNs and transformers to enhance the performance of the model. The framework is evaluated on the public benchmark dataset ACDC with less labeled and more unlabeled data, and the results show that it outperforms eight existing semi-supervised methods. Deep learning-based CNNs and transformers have shown promising results, but they depend on fully supervised medical image segmentation. Attaining good performance for them with limited labeled data poses a challenge as well as labeling is resource intensive, which prompts the researcher to return to semi-supervised learning.

Again, this report highlights the strong points of the paper, including its transparency, innovative approach, simplicity, and effectiveness. The report provides a detailed explanation of the methods, gaps, and results, including the cross-teaching strategy and the evaluation metrics employed. However, challenges in the paper, such as the achievement of a 90.1% F-1 score less than the current state of the art, a lack of generalizability due to the limited diversity of datasets used for validation, and a lack of information on the computational cost of the proposed method, The gaps in the proposed method should be addressed, and I propose the Semi-Vmamba-CNN (UNet++) via pixel-level contrastive and pixel-level cross-supervised solutions to fill these gaps and improve the overall performance.

Keywords: Mamba, Segmentation, pixel, ACDC(Automated Cardiac Diagnosis challenge)

Table of Contents

1.Introduction

Image segmentation is a technique of dividing images or classifying pixels with respect to their local features or texture features [1-2] which are very important to understand the image content (what is or where in an image). The medical image segmentation method derived from image segmentation aims to create visual representations of the body of the patient for clinical analysis [3]. This is one of the most complicated tasks in computer vision, automated medical image segmentation, which aims to act like experienced physicians to identify types of tumors, cardiac image segmentation, and delineate different sub-regions of organs on medical images [4]. The segmented images obtained from the MIS process serve a broad range of clinical applications, including treatment planning, disease monitoring, and surgical navigation.

In the last few years, researchers have made great efforts in image segmentation [5] and achieved outstanding performance with a large amount of annotated data. Earlier approaches were built on traditional methods such as edge detection filters and mathematical methods [6]. Then, machine learning approaches to extracting handcrafted features have become a dominant technique for a long time. In the 2000s, owing to hardware improvements, deep learning approaches came into the picture and started to demonstrate their considerable capabilities in medical image segmentation tasks [7]. Entirely supervised based deep learning approaches, mainly CNN have become more advanced and prevalent in medical image analysis, attaining state-of-the-art (SOTA) performance in image segmentation tasks.

Despite these advancements, CNN based medical image segmentation networks have failed to capture long range dependencies due to the inherent locality of the convolution operation or spatial invariance [8].

As a result, contemporary research has shifted its focus towards self-attention-based approaches [9], exploring the potential of vision transformers to model long-range dependencies for medical image segmentations, while the above method and the reset supervised way of doing image segmentation mainly rely on large labeling of sufficient data sets and couldn't meet efficiency requirements in clinics with limited annotations for training. Their remarkable success is based on the substantial quantity of labeled data that is fed to the model during training. However, the labeling of medical image data is typically dependent on medical doctors or radiologists, and annotating a huge amount of data is tedious and time-consuming.

To overcome these challenges, researchers turn their attention to research on semi-supervised (SS) medical image segmentation, which can leverage a large amount of unlabeled data in conjunction with a small set of labeled data to meet requirements in the clinic.

Semi supervised learning is halfway between supervised learning and unsupervised learning [10], where the algorithm utilizes a limited annotated data set with a large amount of data.

"Semi-Supervised Medical Image Segmentation via Cross Teaching Between CNN and Transformer" (2022), published by Xiangde Luo, Minhao Hu, T. Song, G. Wang, and Shaoting Zhang is the subject of this seminar report that I will be talking about. In this research, a cross-teaching method between CNN (Unet) and the Transformer (Swin-net) model is used to address the challenge of achieving good performance observed in fully supervised medical image segmentation with limited annotations. Instead of just using CNN, the researchers in this paper combine CNN with a transformer to compensate for each other's weaknesses. It was evaluated by employing two evaluation metrics on the public ACDC benchmark: dice similarity coefficient (DSC) and 95% Hausdorff distance.

The result obtained was 90.1% dice (F1-Score), showcasing the potential for future advancement in this area.

2.Paper review

2.1. Description of the paper

This seminar report gives a detailed review of the paper titled "Semi-Supervised Medical Image Segmentation via Cross Teaching between CNN and Transformer," published by Xiangde Luo, Minhao Hu, Tao Song, Guotai Wang, and Shaoting Zhang. This work presents a cross teaching between CNN and transformer models for semi supervised medical image segmentation that were studied to overcome the supervised based deep learning model, which heavily rely on labeled data. This is primarily attributed to the expensive and time-consuming process of obtaining the required labeled data. The inefficiency of the CNN-based method in modeling global or long-range convolution operations is one of drawbacks that were mentioned by the researchers but they combine with transformer model due to it's capability of capturing long range dependencies. During model training, two types of inputs were fed to both CNN and transformer: labeled and unlabeled images, and they were introduced perturbations at both the learning paradigm and output levels of CNN and transformer. They used predictions of unlabeled images generated by CNN or transformer to update the parameters of the transformer or CNN, respectively. There were two predictions one from the CNN model and one from the transformer model but different learning paradigm. Implicit consistency regularization is the study's advantage over explicit consistency regularization in that it can yield pseudo labels that are more accurate and stable. All experiments were done on the public benchmark data set ACDC, which contains annotated cardiac cine MRI images. The evaluation metrics were used in the study these are dice similarity coefficient (DSC) and 95% Hausdorff distance. The training involved two networks, CNN and transformer.

All networks were trained using the SGD optimizer with a batch size of 16, where half of them are labeled in batch for semi-supervised learning, and the performance of the proposed method was compared with eight previous methods. The experiment show that the cross-teaching of CNN and transformer models can complement each other, leading to improved performance compared to using CNN alone.

2.2. Statement of the problem

Medical image segmentation can significantly assist in different stages of clinical practice. Therefore, the reliable automation of this process has widespread implications. Many researches conducted under deep learning model mostly in CNN and later transformer that has shown significant improvement and achieved state of the art performance in medical image segmentation tasks such as cardiac segmentation. But these methods require pixel or voxel-level expert labeling, which is expensive, tedious and time-consuming. Evidence show that achieving satisfactory performance with limited annotations for training deep learning model remain a significant challenge. This leads to the need for semi-supervised methods to train model efficiently with limited labeled and large unlabeled data set. The CNN and transformer model failed modeling of long range and local semantic information interaction respectively. The authors were motivated by the challenge of achieving good performance in MIS with few annotations, prompting them to explore semi-supervised learning methods as a solution to leverage both labeled and unlabeled data.

To overcome the above challenges, researchers proposed semi supervised based method by combining CNN and transformer. They have called their study " **Semi-Supervised Medical Image Segmentation via Cross Teaching between CNN and transformer"**. The aim is to introduce a transformer to the semi-supervised medical image segmentation via presenting cross teaching between CNN(Unet) and transformer(Swi-net) to use unlabeled data. Even though limitations are addressed in this study, achieving higher segmentation accuracy to be balanced or exceed the current SOTA of fully supervised deep learning, reducing computational demand, discriminate unnecessary categories when the background resembles the foreground categories in medical image segmentation and generalization issues need to be addressed. Transformer is based on self attention mechanism has quadratic complexity($\mathbf{L^2}$) resulting high computational demand[11].

To bridge these gap I propose semi-supervised medical image segmentation based on combination of Visual Mamba and CNN (Unet++) via with pixel-level contrastive and pixel-level cross-supervised.

2.3. Objectives of the paper

2.3.1. General objective of the paper

The general objective of the study is to develop semi supervised method for medical image segmentation that can leverage the complementary information between CNN and Transformer when annotated data is scarce.

2.3.2. Specific objectives of the paper

The specific objectives of the paper:

- Gather the data set from publicly benchmark data set.
- Splite the data set into training and testing
- Frame work development, test:-Design and implement a novel framework that integrates CNNs and Transformers for semi-supervised medical image segmentation..
- Quantitatively evaluate segmentation result on dice and HD_{95} metrics
- Conduct Benchmarking:-Compare proposed method results against existing semi-supervised learning method.
- Conduct experiment on data gathered from different source.

2.4. Methods and Materials

In this section, I explain the methods and materials applied and utilized in the study under discussion. The training set consists of less labeled and more unlabeled data. So the entire training set is the union of the above two data sets. For label data, a commonly supervised loss function was used to update the parameters of the model.

This means the model parameters are updated based on the known ground truth of these labeled MRI medical images. For those unlabeled data sets, a cross-teaching strategy was applied. This strategy involves cross-supervision between CNN and the transformer.

The researchers in this study utilized resources and tools like the ACDC dataset, software PyTorch, and hardware Ubuntu desktops equipped with a GTX1080TI GPU.

2.4.1. Cross teaching between CNN and Transformer

CNN and transformers were trained using both labeled and unlabeled data. Both are supervised individually for labeled data, while for unlabeled data, they are supervised using cross-teaching.

Unlike the previous method, the proposed framework introduces perturbations at the learning paradigm and output level. For an input image x, the proposed framework produces two predictions:

1. Prediction of CNN and
2. Prediction of transformer.

Pseudo labels are generated based on the predictions of both models, and a bidirectional loss function, called cross-teaching loss, guides the training process. Unlike consistency regularization, this loss function doesn't explicitly enforce similar predictions but facilitates mutual learning between the models. The Transformer is used for complementary training, not for final predictions. CNN and

Transformer are different learning paradigms for vision recognition, where CNN relies on the local convolution operation and the Transformer is based on long-range self-attention, so these predictions have different properties essentially in the output level.

2.4.2. The overall objective function

The Overall training objective function is a joint loss with two parts a supervised loss on the labeled data and an unsupervised loss for the unlabeled data. The supervised loss comprises cross-entropy and dice losses, while the overall objective includes a weighted combination of supervised and unsupervised losses. The weight factor is determined by a time-dependent Gaussian warming up function based on the current iteration and total iteration.

They define objective function as follow:

$$L_{total}= L_{sup}+ \lambda Lct$$

2.4.3. Data set preparation and pre-processing

2.4.3.1 Data set preparation

The study utilizes the public benchmark ACDC data set contain 200 cardiac cine-MR images from 100 patients. Among them 140 from 70 patients,60 from 30 patients were randomly selected for training and validation respectively.

2.4.3.2. Data pre-processing

The researchers pre-process data set ass follows:

1. Resize all slice into 256X256 pixels.
2. Rescale intensity into[0,1] for normalize pixel values
3. Standard augmentation technique applied
4. Random cropping size and random flipping

2.5. Results

The proposed cross teaching framework for semi supervised medical image segmentation, incorporate both CNNs and transformers, it performs better than eight previous semi supervised methods on public benchmark data set. To assess the performance of the proposed cross teaching against other SSL method Dice(F1 score) and HD_{95} were employed. They are able to achieved 90.1% dice score coefficient and 11.2 mm $HD_{95.}$ The higher the dice the higher performance of the model and lower the HD95 indicate best performance of the model. Even though it outperforms the existing semi supervised method, it falls short of achieving comparable results to the state-of-the-art fully-supervised method’s performance.

2.6.Critiques (Strong and Weak sides)

2.6.1. Strong side

Strong side of the paper from proposed solution point of view:

- Transparency:- they put the source code on publicly accessible GitHub link which is good for reference and for improvement.
- The cross teaching between CNN and transformer is innovative.
- The method is simple to follow and implement. The underlying motivation of the cross-teaching is clear.
- The paper gave experimental results and comparisons with baselines and existing methods, demonstrating the effectiveness of the proposed approach.
- Both quantitative and qualitative results are presented.

2.6.2.Weak side

Weak side of the paper from proposed solution point of view:

- The results were only validated on one data set. Therefor it has lack of generalizability.
- The computational cost and training speed of the proposed method vs. existing methods were not mentioned in the paper.
- There is no the time and GPU memory of the network both in the training and test stage
- There is no information provided on the number of unlabeled images used in the training.
- Limitations and challenges of the proposed approach in real-world clinical work could be further discussed to provide a more comprehensive understanding of its applicability.

2.7. Contributions

- The first contribution is unlike previous studies that have focused solely on CNN or transformer, this research provides a simple yet efficient cross teaching scheme between CNN and transformer for semi supervised medical image segmentation.
- The second contribution is using of transformers to perform the semi-supervised medical image segmentation task and showed that it outperform eight existing semi-supervised methods on a public benchmark.
- Third it was validated with a public benchmark dataset, demon- strating State-of-the-Art performance. The source code of research work and all baseline methods are made public available.

3.Related Works

Tittle	Authors	Description	Year	Limitation/Gap
CNN-based segmentation of Medical Imaging Data	Barıs Kayalıbay, Grady Jensen, Patrick van der Smagt.	They discussed a CNN-based medical image segmentation method using three-dimensional filters and applying it to hand and brain MRI for bone and tumor segmentation tasks. They also explored a modified U-Net architecture, testing two modifications: combining multiple segmentation maps at different scales and element-wise summation of forward feature maps within the network stages.	2022	Despite their success, poor performance in learning global context and long-range spatial dependencies, which can severely impact the segmentation performance for challenging tasks of a network. Supervised methods like U-Net rely on manual image annotation, which is time-consuming, expensive, and dependent on expert skills.
Semi-supervised Medical Image Segmentation through Dual-task Consistency	Xiangde Luo, Jieneng Chen, Tao Song, Guotai Wang	They proposed a novel dual-task consistency semi-supervised framework for the first time. Concretely, we use a dual-task deep network that jointly predicts a pixel-wise segmentation map and a geometry-aware level set representation of the target	2020	Not considering disagreement among sub-networks can affect co-training and information extraction from unlabeled data, leading to reliance on labeled data or unreliable cross-supervision. Accuracy of 89% is good but many study exceed this achievement later.
Remixmatch: semi-supervised learning with distribution alignment and	David Berthelot, Nicholas Carlini, Ekin D. Cubuk,	This paper presents ReMixMatch an improved version of MixMatch. The main contributions are the distribution alignment and the augmentation anchoring.	2019	They only take into account the empirical ground-truth class distribution. This can be problematic when

augmentation anchoring.	Alex Kurakin, Han Zhang, Colin Raffel.	Distribution alignment rescales the predictions based on the difference between the model marginals and the ground truth running average estimation.		dealing with scarce labeled data, as the class distribution may be highly imbalanced or biased. Furthermore, estimating the labeled distributions can be computationally expensive, especially for dense prediction tasks like image segmentation. As suggestion they can maintain consistent distributions of both labeled and unlabeled data.
Semi-supervised semantic-segmentation	Chen, X., Yuan, Y., Zeng, G., Wang, J	They studied the semi-supervised semantic segmentation problem via exploring both labeled data and extra unlabeled data. We propose a novel consistency regularization approach, called cross pseudo supervision (CPS).	2021	CPS method, pseudo labels participate in supervision, which is not suitable for the small-scale medical data set because errors in the pseudo labels potentially mislead the networks when there are not enough training data
Swin-unet: Unet-like pure transformer for medical image segmentation	H. Cao, Y. Wang, J. Chen, D. Jiang, X. Zhang, Q. Tian, and M. Wang.	They were propose Swin-Unet, which is a Unet-like pure Transformer for medical image segmentation. The tokenized image patches are fed into the Transformer-based U-shaped Encoder-Decoder architecture with	2021	Their design is generally based on fully-supervised learning. Therefore, the performance improvements of fully-supervised.

		skip-connections for local global semantic feature learning. Specifically, we use hierarchical Swin Transformer with shifted windows as the encoder to extract context features. A		segmentation are restricted due to the limited number of labeled training samples.
UNETR: Transformers for 3D Medical Image Segmentation	Ali Hatamizadeh ,Yucheng Tang ,Vishwesh Nath , Dong Yang ,Andriy Myronenko ,Bennett Landman ,Holger R. Roth ,Daguang Xu	They introduced a novel architecture, Transformers (UNETR), that apply a transformer as the encoder to learn sequence representations of the input volume and effectively capture the global multi-scale information, while also following the successful “U-shaped” network design for the encoder and decoder. The transformer encoder is directly connected to a decoder via skip connections at different resolutions to compute final segmentation output	2022	This work achieved an accuracy of 97.24% but they need large amount of labeled data set to attain such accuracy.

4.Research gap

The following is the research gap:

- The accuracy of proposed model was good but still needs to be improved because medical image segmentation report is sensetive as it directly affect the patient, so there should be a model that can achieve higher accuracy mage segmentation and
- The model proposed has time complexity and memory demand. As a result there should be model that can reduce these resource demand.
- Lack of diversity in the training dataset of cardiac image to assure generalization and trust.
- The study found no advanced techniques that could further reduce the annotation cost for dense-annotation-based multi-organ segmentation tasks discrimination unwanted background.

4.1.Research questions

RQ1. How can proposed model improve the accuracy of semi-supervised medical image segmentation, while reducing the computational cost?

RQ2. How can I develop a more diverse and representative training dataset for cardiac image segmentation?

RQ3. What advanced techniques can be developed to lower the cost of dense-annotation-based multi-organ segmentation tasks?

5.Proposed solution

As discussed on section 4,the research has some potential gaps as it would not achieved current SOTA accuracy, computational cost and memory demands due to the fact that the transformer has quadratic computation, and its attention-based approach made itself lost in discriminating out background noise and highlighting target regions within the input data. In addition to those gaps, CNN and transformer failed to further reduce dense annotation in multi-organ medical image segmentation.

To bridge the mentioned gaps, I proposed semi-supervised medical image segmentation based on a combination of Visual Mamba and CNN (Unet++), which I dubbed semi-vmamba-Unet++. The combination of Vmamba and Unet++ is expected to resolve the identified gaps mentioned in Section 4 because Mamba has a linear computational cost, improved performance compared to transformers, and for strengthen discriminate against unwanted background I will use technique of pixel-level contrastive and pixel-level Cross-Supervised.

5.1.Objective of the proposed paper

5.1.1.General Objective

The general objective of this study is to develop a semi-supervised approach that combines the current state of the art Visual Mamba model and CNN (a-Unet++) model for medical image segmentation tasks. The goal is to enhance accuracy while also significantly reducing computational and memory demands.

5.1.2. Specific Objectives

The following are the specific objectives in my study that has to be done to achieve general objectives:

- Conduct review of previous work on deep learning image classification, medical image segmentation and other related topics.
- Gather sufficient amount of cardiac MRI data, including both label and unlabeled from different clinic or hospitals.
- Split the dataset into training and validation. The training set is used to train the model, while the validation set is used to evaluate the model's performance during training.

- Data cleaning and augmentation: It's necessary to clean and augment your dataset to improve the accuracy of the model. That can be achieved by resizing, rotating, flipping or adjusting the brightness and contrast of your images.
- Build a prototype of deep learning model which is powerful Visual mamba with improved Unet++ model and train the model with custom training dataset prepared.

5.2. Methodology

To achieve the objective of the research, there are different methods and approaches will be utilized.

5.2.1. Data Collection

Deep learning models are effective when the data set is large, as training on a few datasets can lead to lower performance and accuracy. Thus, the model should be trained on variety of sufficient datasets to detect or output with high accuracy when we apply the model to new data. To achieve the proposed solution for this study, sufficient cardiac medical image should be collected from different hospitals. Currently, hospitals like St. Paul's and Addis Cardiac Hospital offer CT scans and MRI-based diagnoses, offering patient diagnosis in ,so there will be a possibility to obtain dataset for my study.

5.2.2.Data Pre-processing

The quality of data affect the final result produced by the model so data preprocessing step plays crucial role in enhancing the quality of data and helps the suggested model to improve performance of its prediction so some common preprocessing like follow will be conducted :

- Data cleaning: These techniques, manual and automated, remove data.
- Noise reduction :-noise reduction techniques like Gaussian filtering will be applied.
- Resizing the input image: is all about resize the input data by keeping the original information to a specific size that is suitable for proposed model. This can be done using OpenCV or other python image processing.

5.2.3. Data Augmentation

Data augmentation is the process by which data quantity and complexity increase. Data augmentation is one of the techniques to enhance training set and preventing model. Techniques such as random cropping, flipping, and rotation will be utilized to enlarge the training data. I will apply Mix up data augmentation this study. Mix up is a data augmentation technique that combines two images and their labels to create a new training data or example.

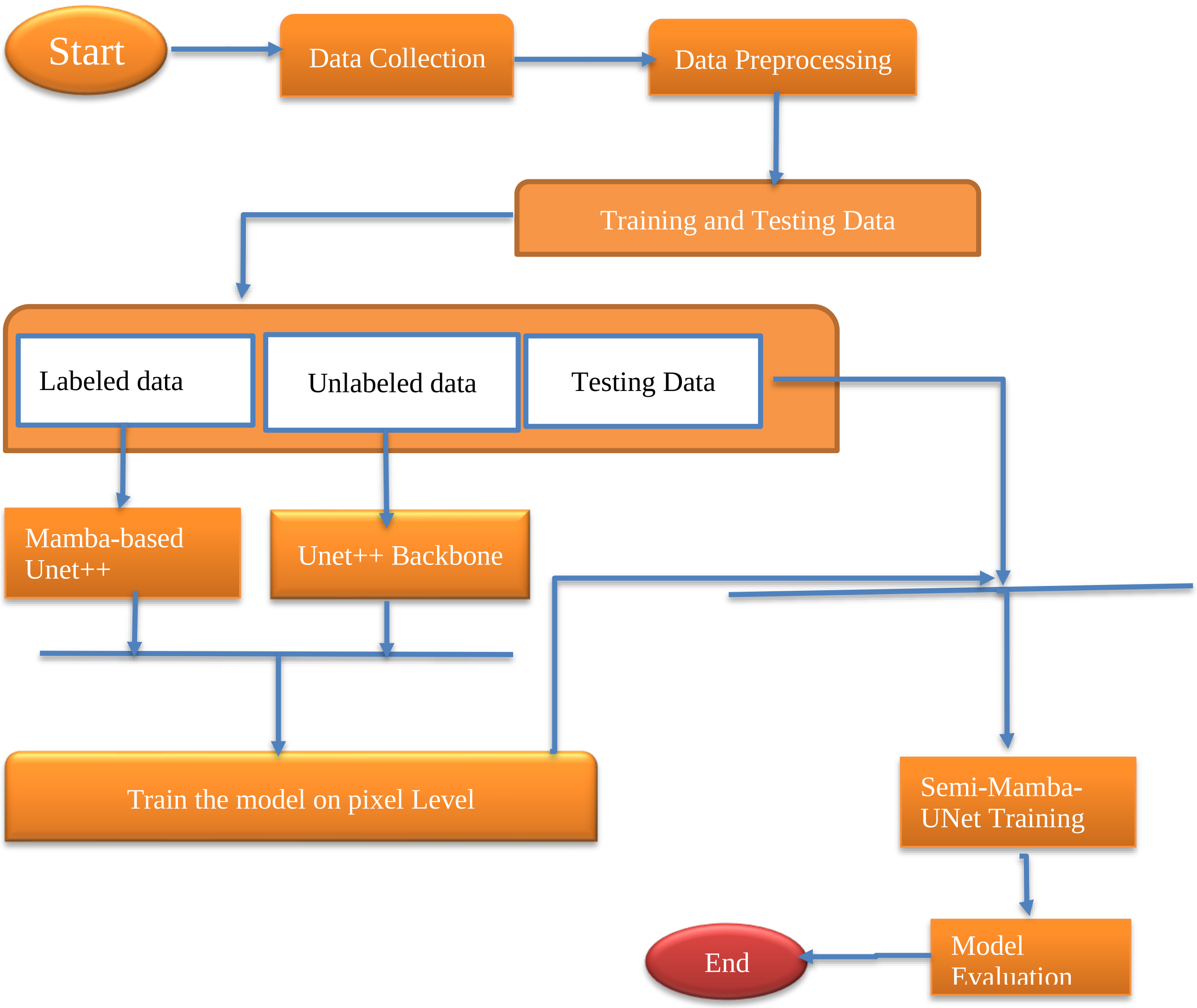
5.2.4.Normalization

Normalizing the input image helps in stabilizing and enhancing the efficiency of the training process which can achieve this by using standard image rescaling intensity into [0,1]. Normalizing the data implies reducing the load on memory and the need for process iterations during training the model. It makes data with different feature transform to similar scale.

5.2.5. Semi- VMamba-CNN(Unet++) Approach

I proposed Semi-VMamba-Unet++ for medical image segmentation. It combines the Visual Mamba architecture with the Unet++ architecture. Vmamba with a selective state space model can achieve state-of-the-art performance by monitoring training dynamics, rejecting inaccurate predictions, and modeling and predicting spatial data. Mamba is more powerful and efficient and can replace the complex attention and multi layer perception (MLP) blocks of transformers with a single, unified SSSM block. This reduces computational complexity and improve inference speed. Mamba is a new neural network that shows promising performance on in computer vision tasks, image classification, image segmentation long sequence data[12]. I will introduce pixel-level cross-supervised learning and pixel-level contrastive learning techniques for enabling the Visual Mamba and Unet++ to directly assist each other, discriminate unwanted background and the enhances the network's feature extraction capabilities

Workflow of the proposed system graphically



UNet++ is based on nested and dense skip connections. The skip connections have proved effective in recovering fine-grained details of the target objects with fine details even on complex background[12].

5.3.Expected outcomes

The expected outcomes of this study after its successful completion include:

A. Reduce computational cost and annotation Effort:- The proposed model has linear computational time so it will reduce both computational cost and memory due to its linearity.

B. Improved segmentation accuracy:The proposed approach is expected to enhance the accuracy of medical image segmentation beyond the current state of the art.

C. Reduce workload:- Save the efforts and time of the doctors or radiologist participate in the manual annotations of medical image because the proposed semi-supervised based approach consume less labeled data.

Overall, successful completion of this study will offer better model that result in high performance, speed up preprocessing time, speed up training time and better accuracy that outperforms the prior work for medical image segmentation.

6. Conclusion

In the course of this report we have discussed how medical image segmentation is an important task in health center practice, facilitating accurate diagnosis, treatment planning, and disease monitoring. Conducting thorough and reliable research in this field can advance the healthcare system. Semi-supervised based medical image segmentation has tackle the challenges were faced bay full supervised learning because of it consume few annotated data. Cross teaching between CNN(Unet) and transform(SWI-NET) has shown promising results on public benchmarks, but addressing the remaining challenges is crucial to achieve best performance and to be comparable with current SOTA.

Finding an alternative current quadratic computational time of transformer, foreground-background discrimination, attaining more dice, accuracy and advancements in medical image segmentation. My proposed solution, Semi-Vmamba-CNN(Unet++), aims to address these issues by incorporating advanced techniques in the process of semi-supervised medical image segmentation.

7.References

[1]. X. Liu, L. Song, S. Liu, and Y. Zhang, “A review of deep-learning-based medical image segmentation methods,” MDPI, https://www.mdpi.com/2071-1050/13/3/1224 (accessed May 12, 2024).

[2]. D. Karimi, H. Dou and A. Gholipour, "Medical Image Segmentation Using Transformer Networks," in IEEE Access, vol. 10, 2022,

[3].C.Wang,“Medical Image Segmentation with Deep Learning Medical Image Segmentation With Deep Learning,” 2020. Available: https://dc.uwm.edu/cgi/viewcontent.cgi?article=3439&context=etd

[4]. M. H. Hesamian, W. Jia, X. He, and P. Kennedy, “Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges,” Journal of Digital Imaging, vol. 32, no. 4, pp. 582–596, May 2019, doi: https://doi.org/10.1007/s10278-019-00227-x.

[5]. F. Milletari, N. Navab, and S.-A. Ahmadi, “V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation,” 2016 Fourth International Conference on 3D Vision (3DV), Oct. 2016, doi: https://doi.org/10.1109/3dv.2016.79.

[6]. Z. Wang and C. Ma, “Dual-Contrastive Dual-Consistency Dual-Transformer: A Semi-Supervised Approach to Medical Image Segmentation,” . 2023.

[7]. M. H. Hesamian, W. Jia, X. He, and P. Kennedy, “Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges,” *Journal of Digital Imaging*, vol. 32, no. 4, pp. 582–596, May 2019, doi: https://doi.org/10.1007/s10278-019-00227-x.

[8]. X.-X. Yin, L. Sun, Y. Fu, R. Lu, and Y. Zhang, “U-Net-Based Medical Image Segmentation,” Journal of Healthcare Engineering, vol. 2022, p. 4189781, Apr. 2022,doi: https://doi.org/10.1155/2022/4189781.

[9]. A.Hatamizadeh et al., “UNETR:Transformers for 3D Medical Image Segmentation,” openaccess.thecvf.com, 2022.

[10]. Christian S. Perone,J. Cohen “Deep Semi-supervised Segmentation with Weight-Averaged Consistency Targets”

[11].H.Zhang, Y.Zhu, D.Wang, L.Zhang,T.Chen, and Z. Ye, “A Survey on Visual Mamba,” *arXiv.org*, Apr. 26, 2024. https://arxiv.org/abs/2404.15956 (accessed May 10, 2024).

[12].Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, “UNet++: A Nested U-Net Architecture for Medical Image Segmentation,” Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, vol. 11045, pp. 3–11, 2018, doi: https://doi.org/10.1007/978-3-030-00889-5_1.

HARAMAYA UNIVERSITY

SCHOOL OF POSTGRADUATE STUDIES

COLLEGE OF COMPUTING AND INFORMATICS

Advisor Approval form for Seminar report 2024

1. **Student Information**

_________________________ _______________________ ___________________

Name Father's Name Grandfather's Name

2. **Seminar Advisor Name**

_________________________ _______________________ ___________________

Name Father's Name Grandfather's Name

3. **Student Signatures**

Signature: ______________________________ Date:_________________

4. **Advisor Signatures**

Signature: ______________________________ Date:__________________