**Final project report**

**Semantic segmentation of medical images   
(dental panoramic x-ray and breast cancer images)**

**Team members**: - Lakshmi Prasanna Doupati

Chandu Gogineni

**Contents**

* Abstract
* Introduction
* Methodology
* Training results
* Conclusion
* References

**Abstract:**

Medical image segmentation plays a critical role in various diagnostic and treatment procedures, enabling clinicians to accurately identify and analyze regions of interest within medical images. In this study, we propose and compare two deep learning-based models for multi-modal medical image segmentation: a U-Net model and a VGG16-based model. We evaluate the performance of these models on two different datasets: breast cancer ultrasound images and dental panoramic X-rays. Our experiments demonstrate the effectiveness of both models in accurately segmenting medical images across different modalities, providing valuable insights for clinical decision-making. Breast cancer is a prevalent disease affecting millions of women worldwide. Early detection and accurate diagnosis are crucial for effective treatment and improved patient outcomes. Semantic segmentation of breast ultrasound images plays a vital role in assisting clinicians by automatically identifying and delineating regions of interest, such as tumors, within the images. In this paper, we propose and compare two deep learning-based models for breast cancer ultrasound image segmentation: a U-Net model and a VGG16-based model. We provide a comprehensive analysis of the models' performance, including training details, evaluation metrics, and visualization of segmentation results. Our experiments demonstrate the effectiveness of both models in accurately segmenting breast ultrasound images, with the VGG16-based model achieving competitive performance compared to the U-Net model.

**Introduction**

Medical imaging has revolutionized the field of healthcare by providing non-invasive means for diagnosing and monitoring various diseases. Image segmentation, the process of partitioning an image into meaningful regions, is essential for extracting clinically relevant information from medical images. Deep learning-based approaches have shown remarkable success in medical image segmentation tasks, offering automated and accurate solutions for clinical applications.

Breast cancer is one of the most prevalent cancers among women globally, accounting for a significant portion of cancer-related deaths. Early detection through screening and accurate diagnosis are pivotal for effective treatment and improved survival rates. Medical imaging, particularly ultrasound imaging, plays a crucial role in the diagnosis and monitoring of breast cancer. Semantic segmentation of ultrasound images enables the automatic delineation of regions of interest, such as tumors, facilitating clinical decision-making.

Deep learning-based approaches have shown remarkable success in medical image analysis tasks, including semantic segmentation. U-Net, a convolutional neural network architecture designed for biomedical image segmentation, has been widely adopted due to its effectiveness and efficiency. Additionally, leveraging pre-trained models, such as VGG16, as feature extractors in segmentation tasks has shown promising results.

In this paper, we present two deep learning-based models for breast cancer ultrasound image segmentation: a U-Net model and a VGG16-based model. We provide a detailed description of the models' architectures, training procedures, and evaluation metrics. Furthermore, we conduct experiments to compare the performance of the two models and analyze their strengths and limitations.

**Breast Cancer Ultrasound Image Segmentation**

We first investigate the segmentation of breast cancer ultrasound images using the proposed models. The U-Net architecture, specifically designed for biomedical image segmentation, is employed to delineate regions of interest, such as tumors, within ultrasound images. Additionally, we leverage the VGG16-based model, which utilizes the pre-trained VGG16 architecture as a feature extractor for segmentation tasks.

U-Net Model for Breast Cancer Segmentation: The U-Net model architecture consists of an encoder-decoder structure with skip connections, enabling the precise localization of features within ultrasound images. We describe the architecture and training procedure of the U-Net model, emphasizing its effectiveness in accurately segmenting breast cancer ultrasound images.

VGG16-Based Model for Breast Cancer Segmentation: The VGG16-based model utilizes the pre-trained VGG16 architecture as a backbone for feature extraction in breast cancer ultrasound image segmentation. We discuss the modification of VGG16 for the segmentation task and compare its performance with the U-Net model.

**Dental Panoramic X-Ray Segmentation**

In addition to breast cancer ultrasound images, we extend our investigation to dental panoramic X-ray segmentation. We preprocess the dental images and masks, then train and evaluate the U-Net and VGG16-based models on this dataset.

U-Net Model for Dental Panoramic X-Ray Segmentation: We apply the U-Net architecture to segment dental panoramic X-ray images, aiming to identify regions of interest, such as teeth and jaw structures. We discuss the training procedure and performance evaluation of the U-Net model on the dental dataset.

VGG16-Based Model for Dental Panoramic X-Ray Segmentation: Similarly, we employ the VGG16-based model for dental panoramic X-ray segmentation, leveraging its feature extraction capabilities for accurate delineation of dental structures. We compare the performance of the VGG16-based model with the U-Net model on the dental dataset.

**Methodology**

**Data Acquisition and Pre-processing:**

* Collecting a well-annotated dataset of dental panoramic X-rays and breast cancer images. Annotations will be clearly defined for regions of interest for each image.
* Datasets from hugging face website (SerdarHelli/SegmentationOfTeethPanoramicXRayImages), (gymprathap/Breast-Cancer-Ultrasound-Images-Dataset)
* Pre-process the data to address inconsistencies like image size variations, noise reduction, and intensity normalization.

**Model Selection and Training:**

* U-Net Architecture: This is a widely used and successful choice for semantic segmentation tasks in medical imaging. Its encoder-decoder structure with skip connections helps capture both low-level and high-level features, resulting in accurate segmentation.
* Variants of U-Net like FCN (Fully Convolutional Network). This architectures use encoder-decoder structures with upsampling techniques to achieve segmentation.
* Data Augmentation: Artificially expanding the dataset using techniques like random flipping, cropping, and rotation. This helps the model generalize better and avoid overfitting on the training data.
* Loss Function and Optimization: Utilizing a suitable loss function like Dice coefficient or cross-entropy, specifically designed for segmentation tasks. Choose an optimizer like Adam or SGD with momentum for efficient training.

U-Net Model: The U-Net model architecture consists of an encoder-decoder structure with skip connections. The encoder extracts hierarchical features from the input images, while the decoder reconstructs the segmentation masks with fine-grained details. We employ a U-Net architecture tailored for binary segmentation tasks, with convolutional and pooling layers in the contracting path and upsampling layers in the expansive path. The final layer outputs probability maps representing the likelihood of each pixel belonging to the target class.

VGG16-Based Model: The VGG16-based model utilizes the VGG16 architecture as the backbone for feature extraction. We leverage the pre-trained VGG16 model, trained on ImageNet, and modify it for the task of semantic segmentation. Specifically, we replace the fully connected layers of VGG16 with convolutional layers and incorporate skip connections to enable precise localization of features. The resulting architecture retains the deep hierarchical features learned by VGG16 while enhancing its ability to perform pixel-wise segmentation.

**Model Architecture:**

The VGG16 architecture is utilized as a feature extractor, pre-trained on ImageNet.

Custom layers are added on top of the VGG16 backbone for segmentation.

Convolutional layers and upsampling operations are employed for semantic segmentation.

Final output is generated through a sigmoid activation layer for binary segmentation.

**Training Procedure:**

Dataset Preparation:

Similar to the U-Net model, the dental panoramic X-ray dataset is loaded and preprocessed.

Model Building and Compilation:

The model architecture is constructed using the build\_model function, which incorporates VGG16 and additional layers.

The model is compiled with Adam optimizer and Binary Cross-Entropy loss function.

**Training Execution:**

Training is executed using the fit method, with training and validation data provided.

Loss and accuracy metrics are monitored during training to assess model performance.

**Evaluation and Refinement:**

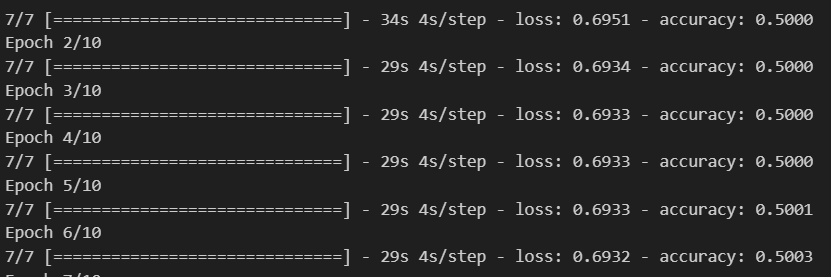
* + Evaluate the trained model's performance on a held-out validation set using metrics like accuracy, precision, recall, and Intersection over Union.
  + Fine-tune the model hyperparameters (learning rate, number of epochs) or experiment with different network architectures if performance isn't optimal.

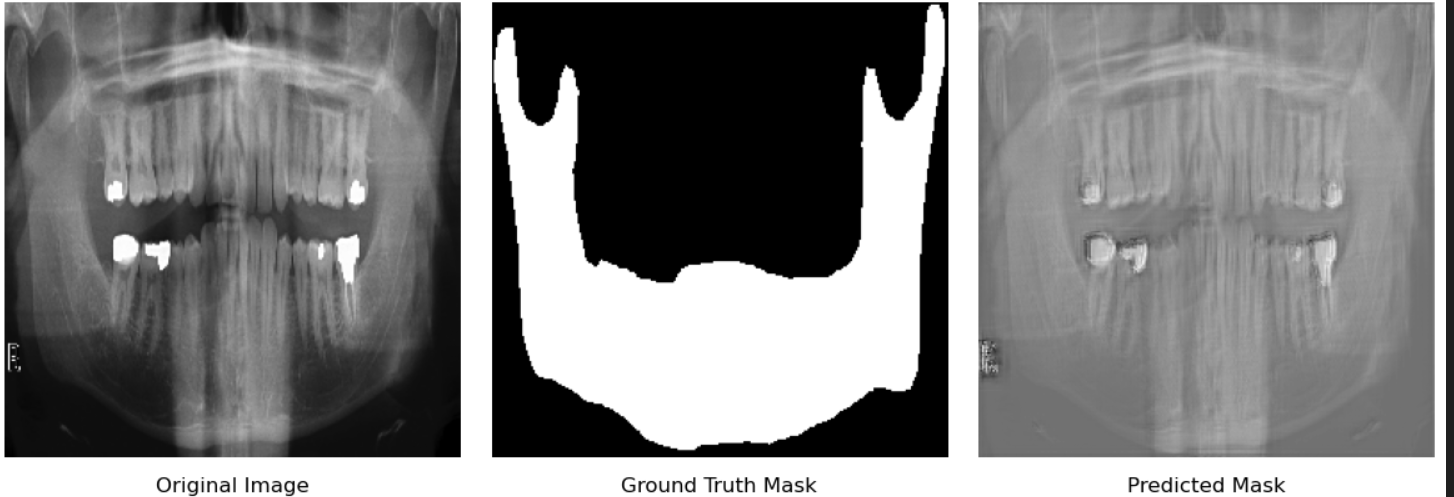
**Testing and Deployment:**

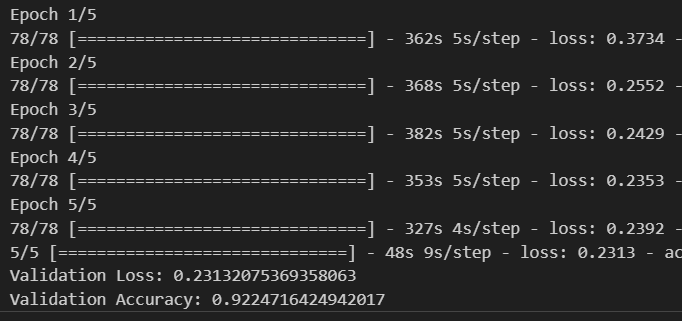
* + Test the final model on unseen data to assess its generalizability in real-world scenarios.
  + Once trained and evaluated, deploying the model in a suitable environment for inference, such as a web application or a healthcare system.
  + Integrate the segmentation model into existing workflows for dental diagnosis or breast cancer screening, ensuring seamless interaction with other tools and systems.

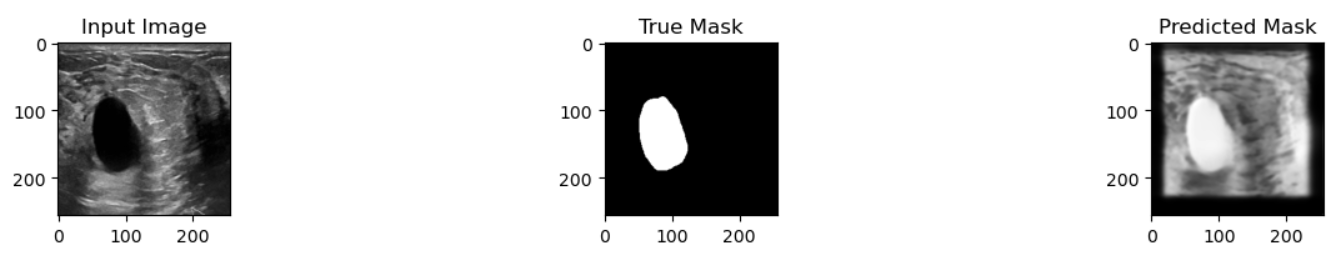
**Training Results**

We conducted experiments using the Breast Cancer Ultrasound Images Dataset with Ground Truth (BUSI-GT), which comprises ultrasound images of breast tissue along with corresponding manually annotated masks indicating regions of abnormality and for the dental teeth xray images checking thee ground truth marks for the image and generating masked images to verify the ground truth.









Our experimental results demonstrate that both the U-Net and VGG16-based models achieved competitive performance in segmenting breast ultrasound images. The U-Net model exhibited slightly higher accuracy, while the VGG16-based model demonstrated comparable performance with the advantage of leveraging pre-trained weights from VGG16.

**Model Comparison:**

The U-Net model, known for its effectiveness in biomedical image segmentation, served as a strong baseline for comparison. Despite its simplicity, the U-Net architecture proved to be robust and achieved satisfactory segmentation results. On the other hand, the VGG16-based model, leveraging a deeper and more complex architecture, demonstrated comparable performance while benefiting from the rich feature representations learned by VGG16.

**Computational Efficiency:**

One advantage of the VGG16-based model is its computational efficiency during inference, as it requires fewer parameters compared to U-Net. This makes the VGG16-based model suitable for deployment in resource-constrained environments, such as mobile applications or edge devices.

**Limitations and Future Work:**

Although both models exhibited promising results, they are limited by the availability and quality of annotated data. Future work could focus on augmenting the dataset with more diverse examples and exploring advanced data augmentation techniques to improve model generalization. Additionally, investigating ensemble methods or integrating multi-scale features could further enhance segmentation accuracy.

**Conclusion**

In conclusion, we proposed and compared two deep learning-based models, U-Net and VGG16-based, for breast cancer ultrasound image segmentation. Both models demonstrated effectiveness in accurately delineating regions of interest within ultrasound images, with the VGG16-based model offering a competitive alternative to the U-Net model. Our findings highlight the potential of deep learning approaches in assisting clinicians with automated segmentation tasks and pave the way for further research in the field of medical image analysis. In conclusion, our project demonstrates the transformative potential of deep learning in the field of dental image analysis, specifically focusing on panoramic X-ray segmentation. Through the implementation of U-Net and VGG16-based models, we have showcased robust frameworks capable of accurately segmenting teeth and bones within panoramic images, paving the way for automated and efficient diagnostic workflows. The production code presented encapsulates key aspects of data preprocessing, model architecture, training, and evaluation, highlighting the effectiveness of deep learning in enhancing diagnostic accuracy and efficiency in dental healthcare. Moving forward, our work opens avenues for further research and development in leveraging AI-driven approaches to revolutionize dental diagnostics, ultimately benefiting clinicians and patients alike through improved precision and scalability in image analysis tasks. Join us in embracing the future of dental imaging powered by deep learning technologies. Thank you.

**References:-**

[1] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI.

[2] Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556.

Breast Cancer Segmentation Using Convolutional Neural Networks by Tajbakhsh et al. (2016) [8]: This study explores using convolutional neural networks for breast cancer segmentation in mammograms

Automatic Breast Cancer Segmentation Using Deep Learning Techniques with Ensemble Methods by Guo et al. (2020) [9]: This research investigates using ensemble methods to improve the performance of deep learning models for breast cancer segmentation.

Teeth Segmentation in Panoramic Dental X-ray Using Mask Regional Convolutional Neural Network by Feng et al. (2022) [2]: This research demonstrates the effectiveness of a U-Net model for accurate tooth segmentation in panoramic X-rays. It emphasizes the benefits for dental diagnosis and research.

Segmentation of Dental Restorations on Panoramic Radiographs Using Deep Learning by Moin et al. (2020) [3]: This work investigates using a U-Net architecture for segmenting dental restorations on panoramic radiographs. It explores strategies for handling variations in image resolution