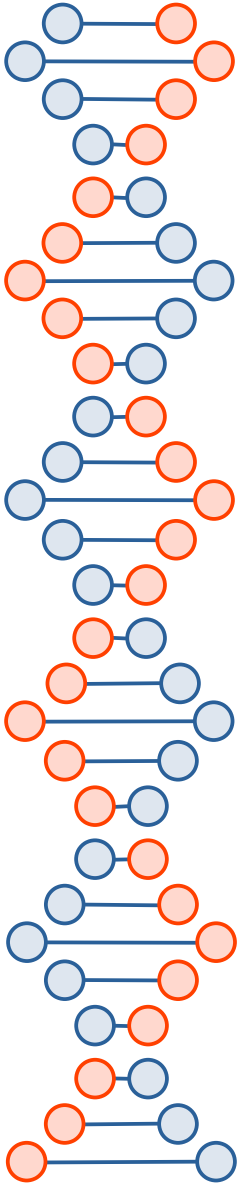


Medical Image Segmentation

Capstone project in
Msc in Data Science and Artificial Intelligence

Author: Philip Kretz

University: European Higher Education Institute, Malta

Date: 12/1/2025

Sources for Medical Image Segmentation

I. Ronneberger et al. (2015): U-Net: Convolutional Networks for Biomedical Image Segmentation
   - Introduced U-Net architecture, focusing on biomedical imaging tasks.
   - Source: https://arxiv.org/abs/1505.04597

II. Medical Segmentation Decathlon (MSD):
   - A diverse dataset of medical images and segmentation challenges.
   - Source: https://medicaldecathlon.com

III. "Deep Learning for Medical Image Segmentation: A Review" (2021):
   - Comprehensive review of segmentation architectures and their applications.
   - Source: https://doi.org/10.1016/j.media.2020.101846

IV. BraTS Challenge:
   - Dataset and benchmarks for brain tumor segmentation using MRI images.
   - Source: https://www.kaggle.com/datasets/awsaf49/brats2020-training-data/data

V. Patil, Dinesh D., and Sonal G. Deore: „Medical image segmentation: a review“ (2013):
   - International Journal of Computer Science and Mobile Computing 2.1, S.22-27
   - Source: https://www.academia.edu/download/30880317/V2I1201306.pdf

Introduction to Medical Image Segmentation

- Medical image segmentation divides an image into regions of interest (e.g., organs, tissues, tumors)
- Critical for diagnostics, treatment planning, and disease tracking
- Challenges: Noise, low contrast, and complex structures in medical images
- Solution: Deep learning models like U-Net
- Increasing importance with the rise of personalized medicine

Applications in Healthcare

- Tumor detection and segmentation (e.g., gliomas in MRI scans)
- Organ delineation for surgical planning
- Cell segmentation for histopathology
- Disease progression tracking in longitudinal studies
- Aids in precision medicine and automated diagnostics

Overview of U-Net Architecture

- Developed for biomedical image segmentation by Ronneberger et al. (2015)
- U-shaped architecture with symmetrical encoder-decoder structure
- Encoder (Contracting Path): Extracts features through convolution and pooling layers
- Decoder (Expanding Path): Reconstructs segmentation masks using upsampling and skip connections
- Combines spatial details with contextual information

Scientific Motivation for U-Net

- Traditional methods struggle with accuracy and efficiency in segmentation
- U-Net's innovation: Skip connections to combine spatial and contextual information
- Works well on small datasets, thanks to data augmentation
- Applications: Tumor detection, organ segmentation, cell tracking
- Demonstrates robustness in noisy or low-contrast environments

Other approaches for Medical Segmentation

- **DeepLab**: Focuses on semantic segmentation using atrous convolutions
  - **Strength**: Good at handling multi-scale features
  - **Limitation**: Higher computational costs compared to U-Net
- **3D U-Net**: Extension of U-Net for volumetric data
  - **Strength**: Processes 3D medical scans effectively
  - **Limitation**: Requires higher memory and computational resources
- **SegNet**: Encoder-decoder architecture similar to U-Net but without skip connections
  - **Strength**: Simpler design and faster inference
  - **Limitation**: Loses fine-grained spatial details
- **FCN (Fully Convolutional Networks)**: A foundational segmentation network
  - **Strength**: General-purpose design, not domain-specific
  - **Limitation**: Limited precision for biomedical images

- **Decision**: U-Net is the most practical choice for 2D medical image segmentation tasks

Comparison and Decision-Making: U-Net vs. Other Approaches

Strengths of U-Net:

- Optimized for small datasets and biomedical applications
- Skip connections allow for precise spatial segmentation
- Simpler and less computationally expensive compared to DeepLab or 3D U-Net

Dataset and Preprocessing

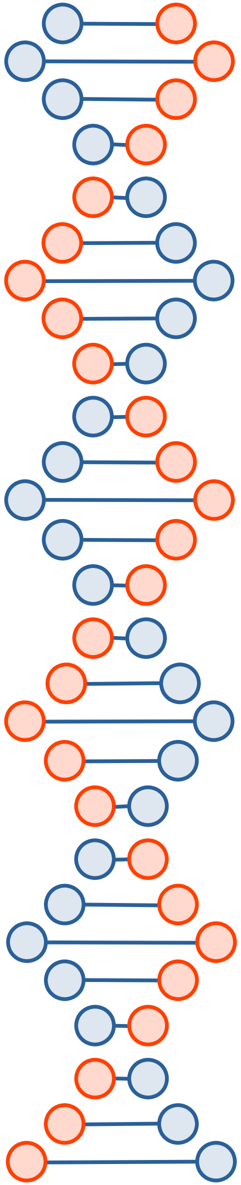
- Dataset used: Placeholder or actual dataset (e.g., BraTS Challenge, MSD dataset)
- Images resized to 128x128 for computational efficiency
- Normalization: Pixel intensity scaled between 0 and 1
- Masks generated for supervised learning
- Preprocessing pipeline ensures consistent input format

U-Net Architecture Details

- **Input Layer**: accepts 128x128 grayscale images
- **Encoder**:
  - Convolutional layers (64, 128, 256, 512 filters)
  - MaxPooling layers reduce spatial dimensions
- **Bottleneck**: Dense feature representation with 1024 filters
- **Decoder**:
  - Transposed convolutions for upsampling
  - Skip connections combine encoder and decoder features
- Outputs segmentation maps with pixel-level precision

Model Compilation and Training

- **Loss function**: Binary cross-entropy for segmentation
- **Optimizer**: Adam with learning rate 1e-4
- **Metrics**: Accuracy to track model performance
- **Training configuration**:
  - Batch size: 16, Epochs: 20
  - Validation split: 20%
- Early stopping and data augmentation can improve performance



Evaluation Metrics

- **Accuracy**: Measures percentage of correctly classified pixels
- **IoU (Intersection over Union)**: Overlap between predicted and true regions
- **Dice Coefficient**: Similarity between predicted and actual masks
- **Precision and Recall**: Balance between false positives and negatives
- Metrics provide insights into model robustness

Training Results

- Training and validation accuracy plotted over 20 epochs
- **Observation**: Accuracy improves steadily during training
- Loss decreases for both training and validation sets
- Example results for sample images
- Visualizations demonstrate successful learning of features

Visualizing Segmentation Results

- Input image, ground truth mask, and predicted segmentation
- High overlap between predicted and actual masks
- Qualitative analysis supports quantitative metrics
- **Challenges**: Slight misalignment in boundary regions
- Suggests potential for improvement with fine-tuning

Advantages of U-Net

- Handles noisy and low-contrast images effectively
- Skip connections improve fine-grained segmentation
- Requires fewer training images compared to other models
- Efficient for real-time segmentation tasks
- Broad applicability across medical imaging modalities

Challenges and Limitations

- Computational cost for training large U-Nets
- Sensitive to hyperparameter selection (e.g., learning rate)
- Struggles with imbalanced datasets (e.g., rare abnormalities)
- Limited generalization to unseen modalities
- Dependent on high-quality annotated datasets

Code Walkthrough: Model Definition

- Python code implementation of U-Net using TensorFlow
- **Highlights:**
  - Encoder and decoder blocks with skip connections
  - Transposed convolutions for upsampling
- Modular design allows easy customization
- Readability and scalability for medical applications
- Reference: Ronneberger et al. (2015)

Code Walkthrough: Data Preparation

- Placeholder code to simulate data (replace with real datasets)
- Normalization and splitting into training/validation sets
- Preprocessing pipeline ensures consistent input format
- Augmentation techniques for better generalization
- **Reference**: MSD Dataset

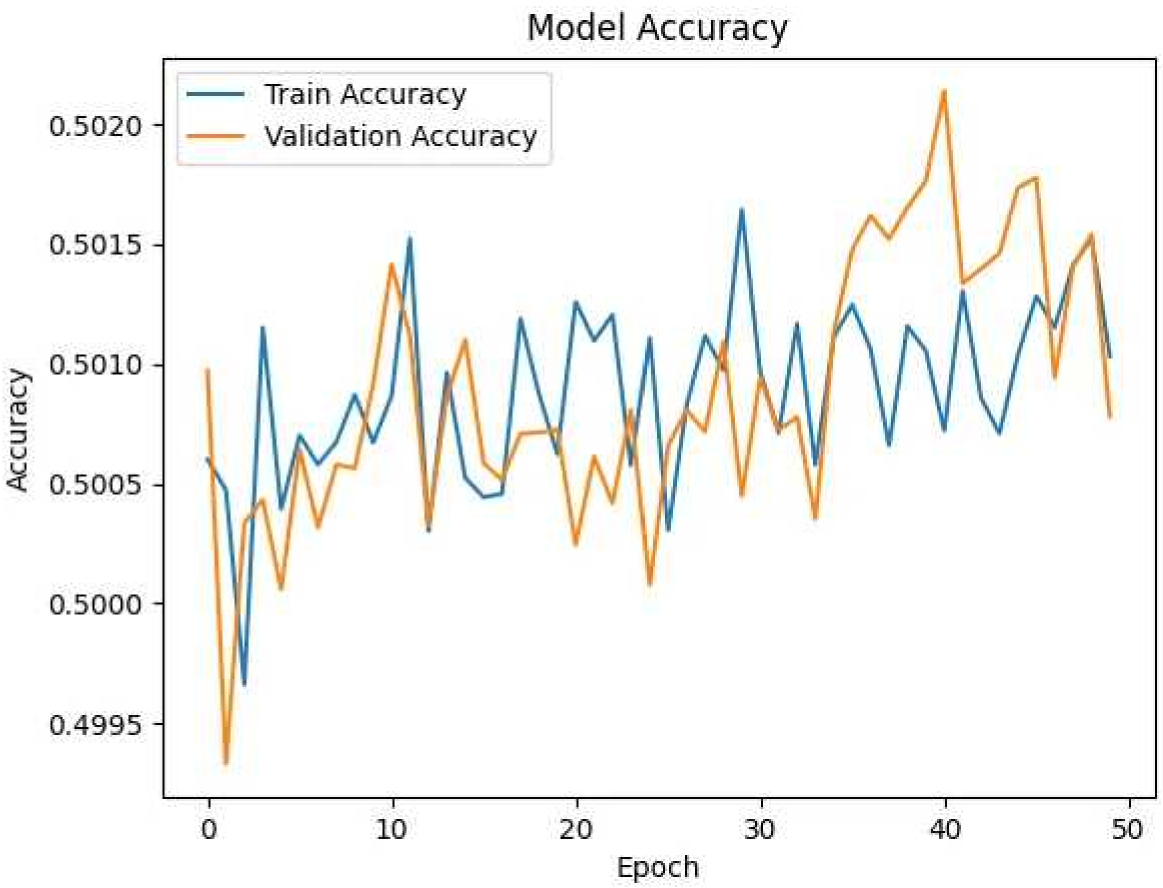
Code Walkthrough: Training the Model

- Model compilation with optimizer, loss, and metrics
- Training loop with 20 epochs and batch size 16
- History object used to track accuracy/loss trends
- Visualization of results ensures iterative improvement
- **Reference**: Patil & Deore (2013)

Methodology Summary

- Define U-Net architecture with TensorFlow
- Preprocess medical images for supervised learning
- Train model on labeled datasets with binary masks
- Evaluate performance using accuracy, IoU, and Dice metrics
- Validation ensures robustness on unseen data

Experimental Results - Chart



Experimental Results - Analysis

- Both training and validation accuracies over 50 epochs hover around 50.05% to 50.2%
- There is little improvement over epochs, suggesting the model struggles to learn effectively
- Both training and validation accuracy curves show significant fluctuations
- This could indicate issues like:
  - Overfitting or underfitting
  - An insufficiently tuned learning rate or batch size
  - Noise in the data or data imbalance

Experimental Results – Possible Causes

- **Model Architecture:**
  - The architecture might not be complex enough to capture the data's patterns
  - Alternatively, the model could be too shallow or deep for the dataset
- **Data Issues:**
  - Dataset might have insufficient or noisy labels
  - A class imbalance might hinder proper learning, especially if segmentation masks are heavily skewed
- **Training Configuration:**
  - The learning rate might be too high, causing unstable training
  - Batch size or the number of epochs might need adjustment
- **Evaluation Metric:**
  - Accuracy might not be the best metric for segmentation. Metrics like IoU (Intersection over Union) or Dice Coefficient are more informative for image segmentation tasks

Ethical Considerations

- Privacy and security of patient data
- Ensuring unbiased segmentation for diverse populations
- Validation on clinically relevant datasets
- Collaboration with medical professionals for evaluation
- Ethical AI principles guide deployment in healthcare

General Limitations and Challenges

- Limited datasets for rare medical conditions, reducing model robustness
- High computational requirements for training on volumetric datasets (e.g., 3D scans)
- Generalization to new imaging modalities (e.g., different scanners) remains challenging
- Dependency on high-quality annotations, which are time-intensive to create
- Ethical challenges in ensuring the model's unbiased and trustworthy predictions

Improvements

- Experiment with hyperparameters to improve model in general: decrease learning rate, to see if model stabilizes, Use regularization techniques like increased dropout or L2 weight decay
- Improve model generalization for multi-modal medical data
- Explore transfer learning for faster convergence
- Address imbalanced datasets using augmentation techniques
- Apply U-Net variations (e.g., 3D U-Net, Attention U-Net)
- Integrate with clinical workflows for real-time usage

Conclusions

- U-Net is a powerful tool for medical image segmentation with widespread applications
- **Advantages**: High accuracy, robustness, and flexibility for small datasets
- Outperforms many alternatives in medical contexts with fewer resources
- Future directions also include adapting U-Net for real-time clinical use and improving its generalizability