Course Name: Computer Vision

Weekly Report: 5

Group Name: Plain

Vanilla Ice-cream

Submitted to faculty:

Mehul Raval

Date of Submission:

29 March 2025

Student Details

Roll No.	Name of the Student	Name of the Program
AU2240	Raj Koticha	B.Tech in CSE
AU2240	Dhruv Premani	B.Tech in CSE
AU2240085	Hariohm Bhatt	B.Tech in CSE

Table of Contents.

WORK DONE THIS WEEK

This week, our efforts centered on implementing and critically evaluating a U-Net architecture as a baseline for retinal image segmentation. We recreated the U-Net model from scratch to gain a deeper understanding of the fundamental components and to benchmark our segmentation performance against more advanced architectures.

Detailed Implementation and Analysis:

1. **U-Net Architecture Development:**

    - **Encoder-Decoder Structure:**
      We implemented a classic U-Net design featuring an encoder that progressively reduces spatial dimensions while capturing context, and a decoder that upsamples the feature maps to restore spatial resolution. The design includes skip connections that bridge the encoder and decoder layers, ensuring the preservation of fine-grained details—an essential feature for segmenting intricate lesion structures.

    - **Activation and Normalization:**
      We used ReLU activations throughout the network, coupled with batch normalization layers to improve training stability and convergence. This combination was key in enhancing the network's ability to learn complex features from the retinal images.

    - **Loss Function:**
      To address the imbalance between lesion and non-lesion pixels, we applied a weighted binary cross-entropy loss. This approach was critical in penalizing misclassifications of smaller lesions, which often appear less prominently in the images.

2. **Dataset Preparation and Augmentation:**

    - **Data Preprocessing:**
      The retinal images were preprocessed by standardizing image sizes and normalizing pixel intensities. This ensured consistency across the training set and facilitated effective model training.

    - **Augmentation Techniques:**
      We employed data augmentation methods such as rotations, flips, and brightness adjustments to artificially expand the dataset. This step was essential for improving the robustness of the U-Net model, particularly in handling variations present in clinical imaging scenarios.

3. **Baseline Evaluation and Insights:**

    - **Segmentation Performance:**
      The U-Net model was evaluated on a curated subset of retinal images. Visual inspections showed that the model accurately identified and segmented larger lesion regions. However, challenges were noted in detecting smaller, more subtle lesions—a critical insight that underscores the need for further architectural enhancements.

    - **Quantitative Metrics:**
      We computed key performance metrics including the Dice coefficient and Intersection over Union (IoU). These metrics provided quantitative evidence of the model's strengths and limitations. While the Dice scores were promising for larger lesion areas, the performance drop on smaller regions highlighted a gap that future iterations of the DRG-Net architecture must address.

4. **Rationale for Recreating U-Net:**

    - **Benchmarking for Improvement:**
      Recreating the U-Net architecture served as a fundamental exercise to establish a clear baseline for our segmentation tasks. By understanding the performance and limitations of this well-known model, we are better positioned to appreciate the enhancements offered by the advanced segmentation modules (S-Net, PD-Net, AD-Net) within the DRG-Net framework.

    - **Identification of Critical Areas:**
      The shortcomings observed in our U-Net implementation, especially in detecting fine lesion details, validate the necessity for incorporating domain adaptation and attention mechanisms. These advanced components are essential for overcoming challenges such as domain shifts and annotation scarcity in real-world clinical data.

    - **Proof of Concept:**
      The experiment provided us with valuable visual and quantitative proof of concept. Overlays of the segmented masks on the original retinal images clearly demonstrated areas of accurate detection as well as regions needing further improvement. This evidence reinforces our strategic direction towards enhancing the segmentation process.

Overall, this week's work on the U-Net implementation not only deepened our understanding of segmentation fundamentals but also clearly illustrated the need for more sophisticated techniques in the DRG-Net pipeline. The insights gained from this baseline exercise will directly

inform the subsequent enhancements aimed at achieving higher diagnostic accuracy in diabetic retinopathy detection.

WORK TO BE DONE NEXT WEEK

1. Complete implementation of U-Net and testing on the IDRiD Dataset.
2. Variations of U-Net to create an efficient but different pipeline similar to DRG-Net.