WiDS Healthcare AI & Vision: Learning Journey & Experimental Report

Jennifer Esbel Mary

January 2026

Contents

1 Introduction

This report documents the learnings, experiments, and projects completed during the WiDS Healthcare AI & Vision program. The curriculum progressed from the fundamentals of digital image processing to advanced deep learning architectures, culminating in a real-world case study on Breast Cancer Pathology (TIGER Grand Challenge).

2 Week 1: The Science Behind Images

2.1 Conceptual Learnings

The program began with an exploration of the fundamentals of computer vision. Key learnings included:

- **Human vs. Computer Vision:** Understanding how digital images are represented as matrices of pixel intensities compared to human biological perception.
- **Coordinate Systems & Transforms:** Learned about image coordinates, resizing techniques, and color spaces (RGB, HSV) through the lectures of Joseph Redmon (creator of YOLO).
- **Traditional Algorithms:** Studied foundational techniques such as kernels, convolutions, and filters before moving to deep learning.

2.2 Experimental Tasks

I implemented a Python-based image processing pipeline to manipulate natural and medical images.

- **Preprocessing:** Performed grayscale conversion and histogram visualization to understand intensity distributions.
- **Filtering:** Applied Gaussian filters for noise removal and experimented with edge detection kernels to identify structural boundaries.
- **Outcome:** These exercises established a strong understanding of how raw pixel data is manipulated prior to feeding it into neural networks.

3 Week 2: Neural Networks & Medical Datasets

3.1 Deep Learning Fundamentals

This week focused on the transition from manual feature extraction to automated representation learning.

- **Neural Networks:** Studied the work of Geoffrey Hinton on backpropagation, learning how networks minimize loss through gradient descent.
- **CNN Architectures:** Learned how Convolutional Neural Networks (CNNs) preserve spatial hierarchies using convolutional layers, pooling, and activation functions.
- **Medical Application:** Explored how CNNs are applied in healthcare for tasks like tumor detection in X-rays and MRI scans.

3.2 Dataset Exploration

I analyzed open-source medical imaging datasets (e.g., NIH, Kaggle) to understand data-centric challenges.

- **Class Imbalance:** Observed significant disparities between "Normal" and "Disease" classes, a common issue in medical data.
- **Annotation Quality:** Noted challenges with noise and varying label quality which necessitates rigorous cleaning and preprocessing.

4 Week 3: Modern Architectures & YOLO

4.1 Advanced Architectures

We examined state-of-the-art architectures that serve as backbones for modern vision systems:

- **Evolution of CNNs:** Traced the development from AlexNet and VGG to ResNet (solving vanishing gradients) and EfficientNet (optimizing parameter efficiency).
- **Darknet Framework:** Explored the Darknet framework, known for its speed and efficiency in real-time detection.

4.2 YOLO Experimentation

I utilized the Ultralytics ecosystem to train a YOLO (You Only Look Once) model.

- **Task:** Classification on a skin lesion dataset (Roboflow).
- **Workflow:**
  1. Data split into Train/Val/Test sets.
  2. Model training using YOLOv8/v11 on GPU.
  3. Evaluation using confusion matrices and loss curves.
- **Outcome:** Gained hands-on experience with the *Train-Val-Predict* workflow and interpreting model metrics.

5 Week 4: Case Study - TIGER Grand Challenge

5.1 Problem Statement

The final project focused on the **Automated Assessment of Tumor-Infiltrating Lymphocytes (TILs)** in breast cancer pathology.

- **Clinical Relevance:** TILs are a vital biomarker for predicting survival in Triple Negative and HER2-positive breast cancers.
- **Current Limitation:** Manual scoring by pathologists is subjective, labor-intensive, and prone to inter-observer variability.

5.2 Methodology

I developed a Deep Learning pipeline to segment lymphocyte nuclei from H&E stained histopathology slides.

5.2.1 Data Preprocessing

- **Standardization:** ROIs were resized to 512×512 pixels.
- **Augmentation:** Applied heavy augmentations using Albumentations, including Random Rotations and RGB Shift, to handle stain heterogeneity (variations in lab coloring).

5.2.2 Model Architecture

- **Architecture:** U-Net, chosen for its biomedical segmentation dominance.
- **Encoder:** EfficientNet-B0 pre-trained on ImageNet to leverage transfer learning while maintaining computational efficiency.
- **Loss Function:** A composite loss strategy was used to handle the "Small Object Dilemma":

$$Loss = DiceLoss + FocalLoss \tag{1}$$

  Focal loss focused learning on hard-to-classify examples, while Dice loss handled the class imbalance.

5.3 Results

The model was trained for 5 epochs using the AdamW optimizer ($lr = 1e^{-4}$).

Metric	Value
Validation IoU	0.4789
Validation Loss	0.6366

Table 1: Performance Baseline (5 Epochs)

Clinical Metric (TIGER Score): Post-processing was applied to calculate the TILs Density:

$$Density = \frac{TotalLymphocytePixels}{TotalTissueArea} \times 100 \tag{2}$$

The model successfully identified lymphocyte clusters while ignoring larger tumor nuclei and stromal tissue.

5.4 Challenges & Conclusion

- **Small Objects:** Lymphocytes occupy $< 1\%$ of pixels. U-Net skip connections were critical for recovering these fine details.
- **Confluent Clusters:** Semantic segmentation treats clumped cells as a single region. Future work should explore Instance Segmentation (Mask R-CNN) for accurate cell counting.

In conclusion, this project demonstrated that deep learning can effectively automate TIL scoring, offering a pathway toward reproducible precision medicine.