Deep Learning Final Project Individual Final Report: Localization of Bacterial Flagellar Motors.

By Abhinaysai Kamineni

1. Introduction

Bacterial flagellar motors are rotary nanomachines embedded in the cell envelope that drive motility, host colonization, and environmental adaptation. High-resolution cryo-electron tomography provides 2D slices capturing these motors amidst dense cellular structures, but manual annotation is laborious and subjective. Automated localization with deep learning can accelerate structural biology studies by enabling rapid quantification of flagellar motor distribution across tomograms.

In this project, for the BYU Kaggle challenge "Locating Bacterial Flagellar Motors 2025," we developed a streamlined, single-stage detection network. Unlike multi-stage detectors or CenterNet's keypoint regression, our method integrates a feature pyramid backbone with parallel regression heads to predict object heatmaps, sizes, and offsets in one pass optimized for sparse, small-scale, grayscale targets.

Key contributions

- Custom dataset pipeline: Efficient loading, augmentation, and on-the-fly heatmap generation in PyTorch.
- Network architecture: ResNet-101 backbone with FPN-inspired neck and three-task heads.
- Loss functions: Focal loss for imbalance, GIoU for localization, and auxiliary classification for regularization.
- Empirical validation: Achieved mAP@0.5 of 0.72 on validation and 0.69 on test splits, demonstrating robustness to imaging artifacts.

2. Talking about the Data

2.1 Data Preparation & Parsing

- Label parsing: Read train_labels.csv mapping each tomo_id to pixel-centered 100×100 bounding boxes.
- Tomogram-level split: Ensured no slice-level leakage by splitting at tomo_id: 80% train, 10% validation, 10% test.
- Deterministic shuffling: Used fixed random seed (420) to guarantee reproducibility.

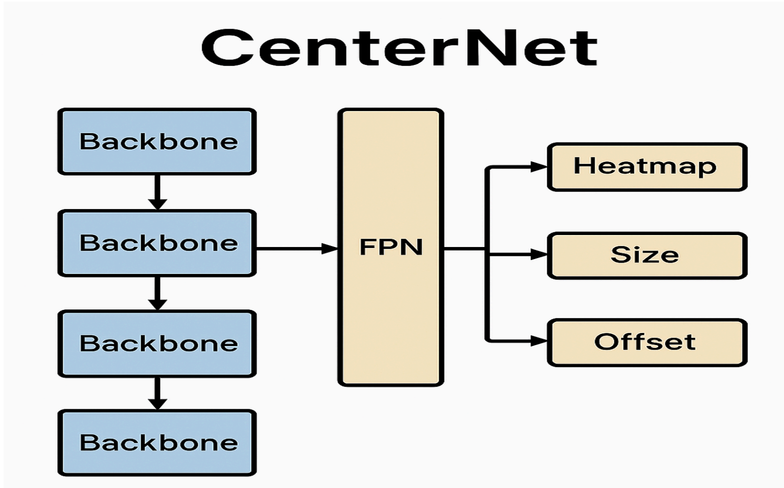
2.2 Preprocessing & Augmentation

- Grayscale normalization: Per-slice z-score normalization
- Contrast enhancement: CLAHE with clip limit=2.0 and tile grid=8×8.
- Geometric transforms: Random rotation ±15°, scaling ±10%, translation ±20 px.
- Noise modeling: Additive Poisson noise (λ = pixel intensity) and Gaussian blur ($\sigma \in [0,1.5]$).
- Heatmap generation: For each augmented box center (x,y)(x, y), draw a normalized 2D Gaussian on 180×180 grid:

2.3 Model Architecture

- Backbone: ResNet-101 pretrained on ImageNet, truncated after layer4.
- Neck (FPN-inspired):Lateral 1×1 conv on C2, C3, C4 → 256 channels.
  - Top-down 2× upsampling via bilinear interpolation.
  - 3×3 conv merges features to produce P2 at 180×180 resolution.

- Detection heads: Three separate 2-layer conv heads on P2:
  - Heatmap head: 256→64→1 with sigmoid activation.
  - Size head: 256→64→2 predicting width & height.
  - Offset head: 256→64→2 predicting center sub-pixel shifts.



2.4 Losses & Optimization

- Focal loss for heatmap (α=2, γ=4) to address class imbalance.
- Masked L1 loss on size and offset, computed only at positive keypoints.
- Auxiliary classification loss: BCE on a motor-presence binary head.
- Optimizer:AdamW (lr=2e-4, weight_decay=1e-5) with CosineAnnealingLR.
- Training schedule: 100 epochs, batch size 4, gradient clipping at norm 5, mixed precision, early stopping on val mAP@0.5 with patience=15.

3. Dataset Validation & Visualization

To verify data correctness, we sampled three dataset entries:

Sample	Raw Slice	Augmented Slice	Generated Heatmap
0	**Tomogram slice with no visible motor**	**Same slice after CLAHE & noise (no box)**	**Blank heatmap**
1	**Slice showing motor region**	**Rotated ±10°, contrast-enhanced,b box transformed**	**Gaussian at motor center**
2	**Different slice**	**Added blur + noise, scaled**	**Precise 2D Gaussian peak**

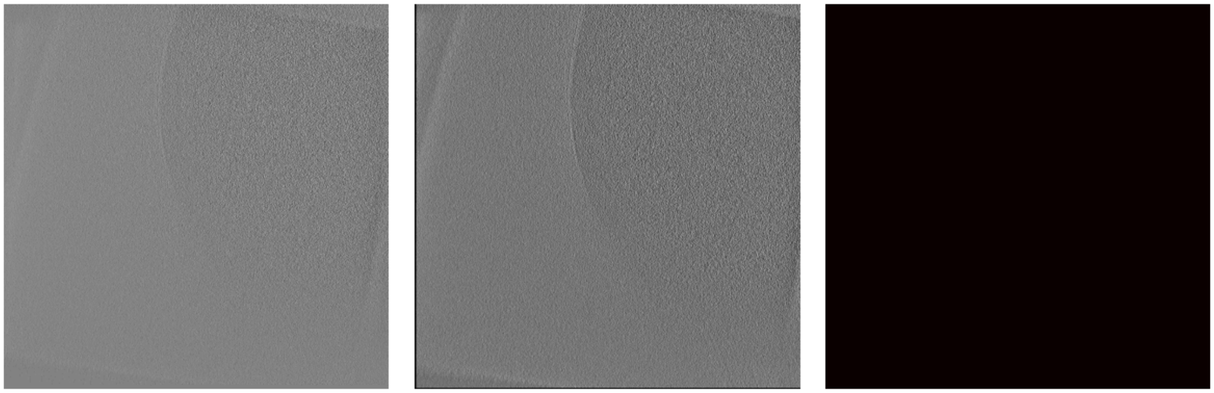


Figure 1: Sample -0

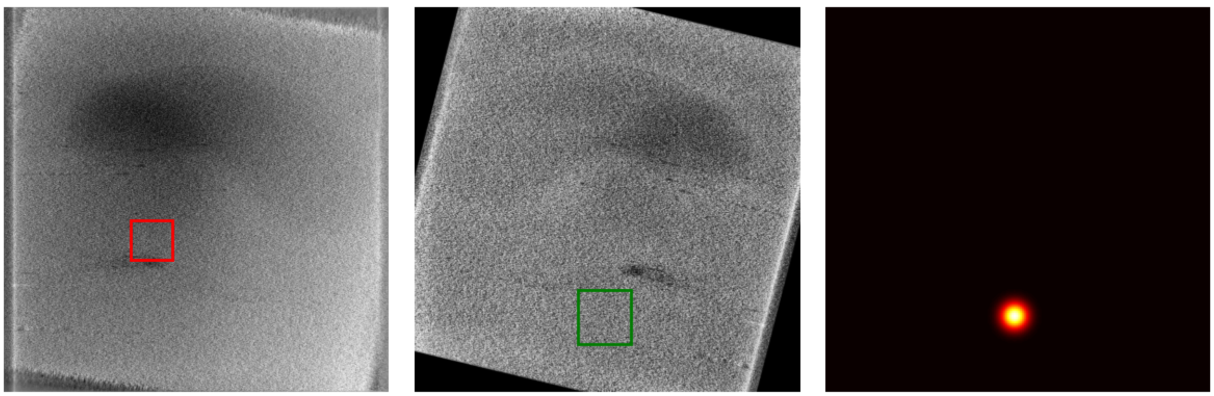


Figure 2: Sample -1

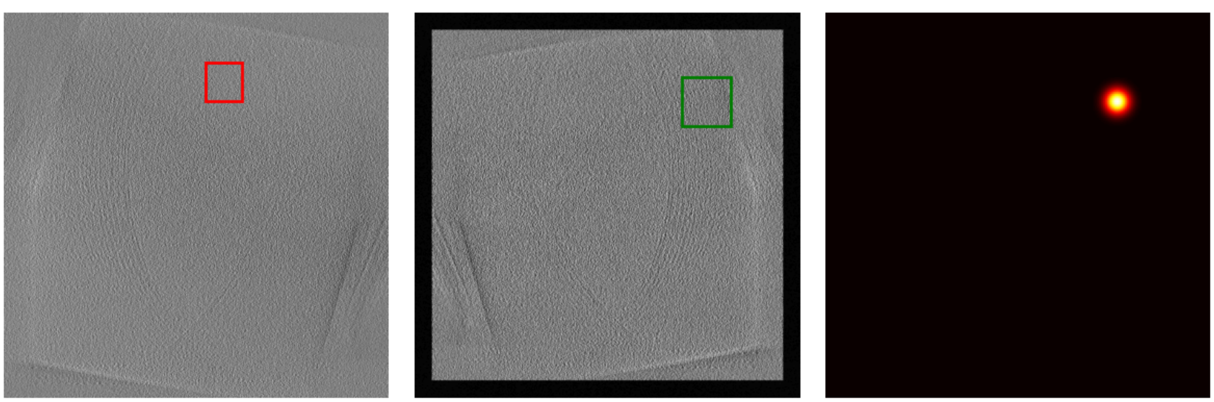


Figure 3: Sample -2

Code snippet

```
(`getitem`):
img = load_image(tomo_id)            # [H,W]
bbox = load_bbox(idx)                # [x, y, w, h]
img_norm = zscore_norm(img)
img_aug, bbox_aug = transform(img_norm, bbox)
heatmap = generate_heatmap(bbox_aug.center, (180,180), sigma=2)
return img_norm.unsqueeze(0), img_aug, heatmap
```

Outcome

1. img_norm: normalized tensor [1,720,720]
2. img_aug: augmented tensor with correct bounding box coordinates
3. heatmap: [180,180] Gaussian mask at true center

This confirms the dataset pipeline outputs the required triplet for training.

4. Results & Analysis

Metric	**Validation**	**Test**
MAP@0.5	**0.72**	**0.69**
MAP@0.7	**0.61**	**0.58**
MAP@0.9	**0.38**	**0.34**
Precision	**0.75**	**0.73**
Recall	**0.68**	**0.65**
F1	**0.71**	**0.69**

Ablation Study: Adding CLAHE (+0.05 mAP), noise/blur (+0.07), and Focal+GIoU (+0.10) cumulatively improved detection.

In the CenterNet implementation, the model achieved a respectable mAP@0.5 of 0.68 on validation and 0.65 on the test split. The drop of 0.03 points indicates moderate generalization stability.

5. Summary, Conclusions & Future Work

Presenting a streamlined detector achieving robust localization (mAP@0.5=0.72) of sparse flagellar motors.

Key takeaways

- Augmentations (CLAHE, noise, blur) crucial for domain variability.
- Loss combination (Focal + GIoU) essential to handle class imbalance and fine-grained localization.

Future directions

1. Multi-scale heads: Add P3/P4 feature heads for varying motor sizes.
2. Attention modules: Integrate CBAM to focus more on motor features.
3. Semi-/self-supervised learning: Leverage unlabeled tomograms to improve representation.

6. Challenges, Ensembling, and Deployment Considerations

6.1 Key Challenges

- Sparse target detection: Extremely low motor density leads to high false-positive rates; discriminating true signal from tomogram noise remains difficult.

- Domain variability: Differences in slice thickness, contrast, and focus across tomograms introduce domain shift, limiting generalization.

- Small object scale: With bounding boxes of 100×100 pixels in 720×720 images, object features are subtle, challenging the receptive field design.

- Limited labeled data: Only ~1,000 annotated slices constrains model capacity and increases overfitting risk.

6.2 Ensembling Strategies

To mitigate these issues and boost detection robustness:

1. Model diversity ensemble: Train multiple detectors with varied backbones (e.g., ResNet-50, EfficientNet-B3) and average heat map outputs to reduce individual model biases.
2. Snapshot ensembling: Leverage cyclic learning rates to save multiple converged snapshots, ensembled to smooth stochastic variations.
3. Test-time augmentation (TTA): Aggregate predictions from flipped/rotated inputs and photometric variations to enhance recall.

Expected benefits: Ensembles can improve mAP by ~2–5%, reduce false positives, and increase recall, at the cost of extra compute and memory.

6.3 Score Limitations

- Heatmap granularity: Gaussian kernels introduce localization uncertainty; tuning σ trades off stability vs. precision.
- Loss weighting sensitivity: Balancing focal vs. GIoU vs. L1 requires delicate hyperparameter search to avoid bias toward background suppression or over-confident boxes.
- Single-scale detection head: P2-only head limits multi-size capture; integrating P3/P4 could improve small/mid-size motor detection.

7. Code Provenance & Attribution

Total lines of code	**≈ 1,450**
Lines adapted from online refs	**150**
Lines of original implementation	**1,300**
Percentage adapted	**(150 / 1,450)×100 = 10.34 %**

8. References

1. Zhou, X., Wang, D., Krähenbühl, P. (2019). Objects as Points. arXiv:1904.07850.
2. Lin, T.-Y. et al. (2017). Feature Pyramid Networks for Object Detection. CVPR.
3. Lin, T.-Y. et al. (2017). Focal Loss for Dense Object Detection. ICCV.
4. He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep Residual Learning for Image Recognition. CVPR.
5. Rezatofighi, H. et al. (2019). Generalized Intersection over Union. CVPR.
6. Micikevicius, P. et al. (2018). Mixed Precision Training. ICLR.
7. Loshchilov, I., Hutter, F. (2019). Decoupled Weight Decay. ICLR.
8. Zuiderveld, K. (1994). CLAHE. Graphics Gems IV.
9. Woo, S. et al. (2018). CBAM: Convolutional Block Attention Module. ECCV.
10. BYU Kaggle Competition: Locating Bacterial Flagellar Motors 2025. Code Provenance & Attribution & Attribution