Fetal Ultrasound Biometry Landmark Detection

Problem Statement

Development of algorithms to identify biparietal diameter (BPD) and occipitofrontal diameter (OFD) landmark points in fetal axial ultrasound images. These measurements are crucial for:

- Estimating gestational age
- Assessing fetal growth
- Monitoring neurodevelopment

Approaches Implemented

Part A: Landmark Detection-Based Approach

Direct regression of 4 landmark points (2 for BPD, 2 for OFD) using deep learning models with heatmap regression.

Part B: Segmentation-Based Approach

Segmentation of fetal cranium followed by ellipse fitting and geometric calculation of biometry points.

Dataset Structure

- **Total Images**: 624 fetal ultrasound images
- **Annotations**: CSV file with 8 landmark coordinates (x, y) per image
  - `ofd_1_x, ofd_1_y` : First OFD landmark
  - `ofd_2_x, ofd_2_y` : Second OFD landmark
  - `bpd_1_x, bpd_1_y` : First BPD landmark
  - `bpd_2_x, bpd_2_y` : Second BPD landmark

Project Structure

```
.
├── README.md
├── requirements.txt
├── config.py
├── data/
│   ├── dataset.py
│   ├── preprocessing.py
│   └── augmentation.py
├── models/
│   ├── landmark_detection/
│   │   ├── heatmap_model.py
│   │   ├── coordinate_regression.py
│   │   └── attention_pyramid.py
│   └── segmentation/
│       ├── unet.py
│       ├── attention_unet.py
│       └── deeplabv3.py
├── utils/
│   ├── metrics.py
│   ├── visualization.py
│   └── ellipse_fitting.py
├── train_landmark.py
├── train_segmentation.py
└── inference.py
```

Methodology & Thought Process

1. Data Understanding

- Ultrasound images have inherent challenges: noise, speckle, variable contrast
- Landmark points form an ellipse (cranium) with BPD and OFD as perpendicular diameters
- Dataset size (624 images) requires careful augmentation strategy

2. Preprocessing Strategy

Rationale:

- **Normalization**: Ultrasound images have variable intensity ranges
- **CLAHE (Contrast Limited Adaptive Histogram Equalization)**: Enhances local contrast in ultrasound
- **Denoising**: Reduces speckle noise characteristic of ultrasound imaging
- **Resizing**: Standardizes input dimensions while preserving aspect ratio

3. Augmentation Strategy

Rationale:

- **Geometric Transformations** (rotation, flipping, scaling):
    - Accounts for different fetal head orientations
    - Increases dataset diversity (5-10x effective size)
    - Must transform both images and landmark coordinates
- **Intensity Augmentations** (brightness, contrast, gamma):
    - Simulates different ultrasound machine settings
    - Improves model robustness to varying image quality
- **Elastic Deformations**:
    - Simulates natural anatomical variations
    - Particularly important for medical imaging
- **Conservative Augmentation**:
    - Avoid extreme transformations that create unrealistic anatomy
    - Preserve clinical validity of measurements

4. Model Variations Explored

Part A: Landmark Detection (3+ Approaches)

Approach 1: Heatmap Regression with CNN

- **Architecture**: ResNet50 backbone + decoder for heatmap generation
- **Hypothesis**: Heatmaps provide spatial probability distribution, more robust than direct coordinate regression
- **Loss**: MSE on heatmaps + coordinate extraction via argmax
- **Pros**: Implicit spatial reasoning, handles ambiguity
- **Cons**: Computational overhead, resolution dependent

Approach 2: Direct Coordinate Regression

- **Architecture**: EfficientNet-B3 backbone + fully connected layers
- **Hypothesis**: Direct regression is simpler and faster for well-defined landmarks
- **Loss**: Smooth L1 loss on normalized coordinates
- **Pros**: Fast inference, simple architecture
- **Cons**: Less robust to occlusions, no spatial context

Approach 3: Attention-Based Feature Pyramid

- **Architecture**: Feature Pyramid Network + Attention mechanisms
- **Hypothesis**: Multi-scale features capture both fine details and global context
- **Loss**: Wing loss (better for long-tailed distribution of errors)
- **Pros**: Handles scale variations, attention focuses on relevant regions
- **Cons**: More complex, requires more data

Approach 4: Transformer-Based Landmark Detection

- **Architecture**: Vision Transformer (ViT) with landmark-specific tokens
- **Hypothesis**: Self-attention captures long-range dependencies between landmarks
- **Loss**: Combined coordinate loss + landmark relationship constraints
- **Pros**: Models geometric relationships, state-of-the-art performance
- **Cons**: Requires more training data, computationally expensive

Part B: Segmentation-Based (3+ Approaches)

Approach 1: U-Net

- **Architecture**: Classic U-Net with skip connections
- **Hypothesis**: Standard for medical image segmentation, proven effectiveness
- **Loss**: Dice loss + Binary Cross Entropy
- **Pros**: Well-established, good with small datasets
- **Cons**: Limited receptive field

Approach 2: Attention U-Net

- **Architecture**: U-Net + Attention gates
- **Hypothesis**: Attention gates suppress irrelevant features, focus on cranium
- **Loss**: Focal Dice loss (handles class imbalance)

- **Pros**: Better boundary localization, fewer false positives
- **Cons**: Slightly slower training

Approach 3: DeepLabV3+

- **Architecture**: Atrous Spatial Pyramid Pooling (ASPP) + encoder-decoder
- **Hypothesis**: Multi-scale context through dilated convolutions
- **Loss**: Combined Dice + Focal loss
- **Pros**: Captures multi-scale features, good for varying cranium sizes
- **Cons**: Memory intensive

Approach 4: Post-Processing Pipeline

- After segmentation: Ellipse fitting (OpenCV) → Extract major/minor axes → Calculate BPD/OFD points
- **Techniques**: RANSAC for robust ellipse fitting, morphological operations

5. Loss Functions

For Landmark Detection:

- **MSE/Smooth L1**: Standard for coordinate regression
- **Wing Loss**: Better handles outliers, focuses on difficult samples
- **Adaptive Wing Loss**: Dynamically adjusts to error magnitude
- **Geometric Constraint Loss**: Enforces perpendicularity of BPD/OFD

For Segmentation:

- **Dice Loss**: Handles class imbalance (background vs cranium)
- **Focal Loss**: Focuses on hard-to-segment regions
- **Boundary Loss**: Emphasizes accurate edge delineation

6. Training Strategy

- **Cross-validation**: 5-fold stratified split
- **Transfer Learning**: ImageNet pre-trained backbones
- **Learning Rate Scheduling**: Cosine annealing with warm restarts
- **Regularization**: Dropout, weight decay, early stopping
- **Mixed Precision Training**: Faster training, reduced memory

7. Evaluation Metrics

- **Mean Radial Error (MRE)**: Average Euclidean distance from predicted to ground truth
- **Successful Detection Rate (SDR)**: % of landmarks within threshold (2mm, 2.5mm, 3mm)
- **Dice Coefficient**: For segmentation masks
- **Clinical Metrics**: BPD/OFD measurement error (mm)

Installation

```
pip install -r requirements.txt
```

Usage

Training Landmark Detection Model

```
python train_landmark.py --model heatmap --epochs 100 --batch_size 16
```

Training Segmentation Model

```
python train_segmentation.py --model unet --epochs 150 --batch_size 8
```

Inference

```
python inference.py --model_path checkpoints/best_model.pth --image_path test_image.png
```

Key Insights & Decisions

1. **Why Both Approaches?**
    - Landmark detection: Direct, but requires precise annotations
    - Segmentation: More robust to annotation noise, provides additional context (cranium outline)
2. **Data Preprocessing Choices**

- CLAHE instead of global histogram equalization: Preserves local features
- Minimal cropping: Maintains anatomical context around cranium

3. **Model Selection Rationale**

    - Started with simpler models (U-Net, ResNet) for baseline
    - Progressive complexity (attention, transformers) to improve performance
    - Ensemble potential: Combine predictions from multiple models

4. **Clinical Considerations**

    - Measurements must be reproducible (low inter-observer variability)
    - False negatives are critical: Missing measurements can delay diagnosis
    - Explainability: Heatmaps provide interpretable confidence regions

References

1. Regressing Heatmaps for Multiple Landmark Localization using CNNs
2. Cephalometric Landmark Detection by Attentive Feature Pyramid Fusion
3. U-Net: Convolutional Networks for Biomedical Image Segmentation
4. Attention U-Net: Learning Where to Look for the Pancreas
5. DeepLabV3+: Encoder-Decoder with Atrous Separable Convolution

Future Improvements

- **Domain Adaptation**: Fine-tune on specific ultrasound machines
- **Uncertainty Estimation**: Bayesian deep learning for confidence intervals
- **Active Learning**: Prioritize difficult cases for annotation
- **Multi-task Learning**: Simultaneous landmark detection + segmentation
- **3D Extension**: Utilize volumetric ultrasound data if available