Complete Data Processing Pipeline

Original Image (Any size) → **YOLO Detection (Best.pt finetuned)** → **Confidence-based Masking**

Resize to 128×128

Data Augmentation (Stronger/Minimal)

Keypoint Heatmap Generation

HybridKeypointNet (YOLO + ViT)

Predicted Heatmap (1, 128, 128)
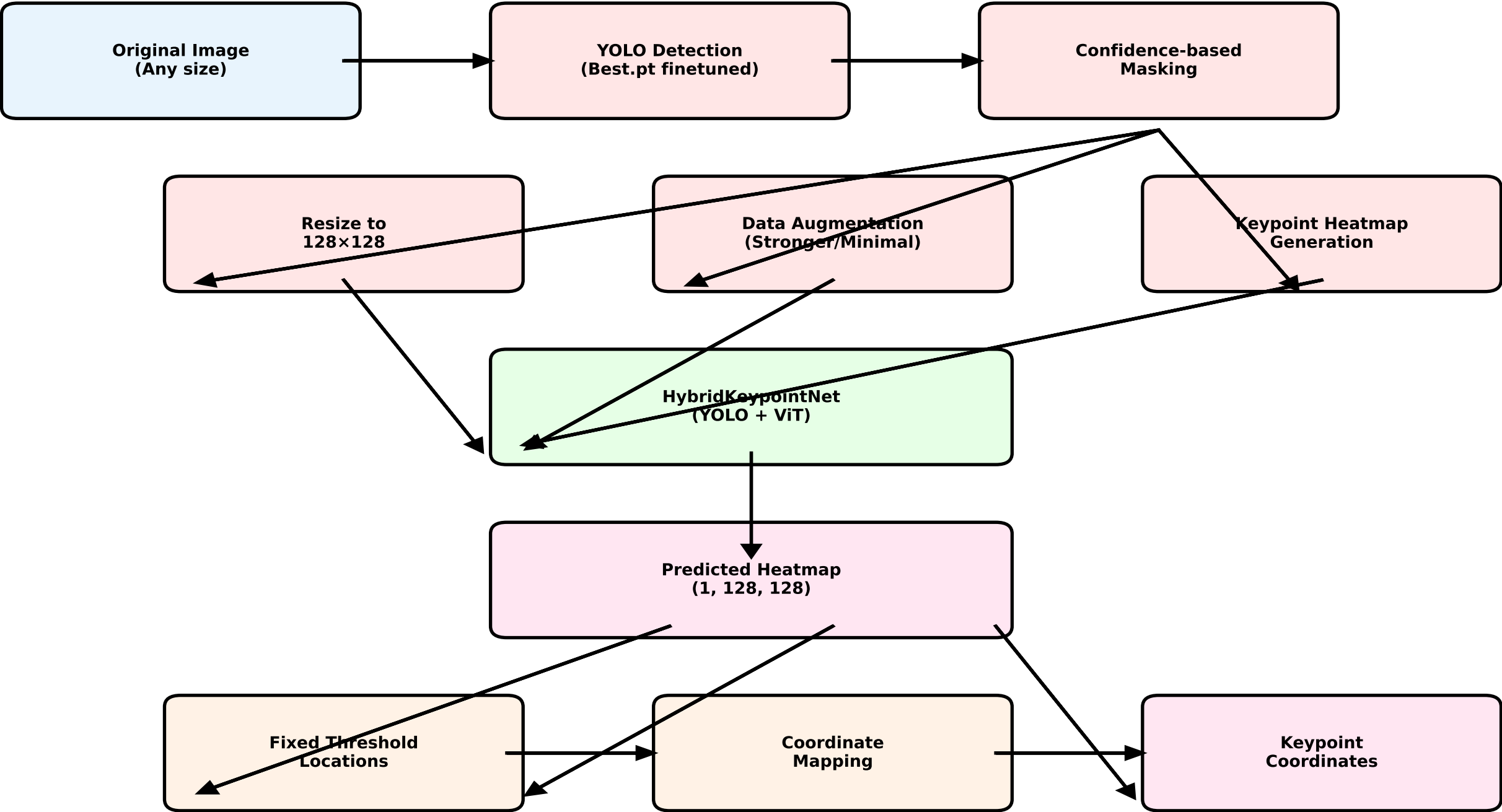
Fixed Threshold Locations → **Coordinate Mapping** → **Keypoint Coordinates**

Key Optimizations:
• Confidence-based masking refinement
• Fixed coordinate system mapping
• Single peak detection per keypoint
• Proper train/eval data separation
• FP16 mixed precision training
• Gradient accumulation & early stopping