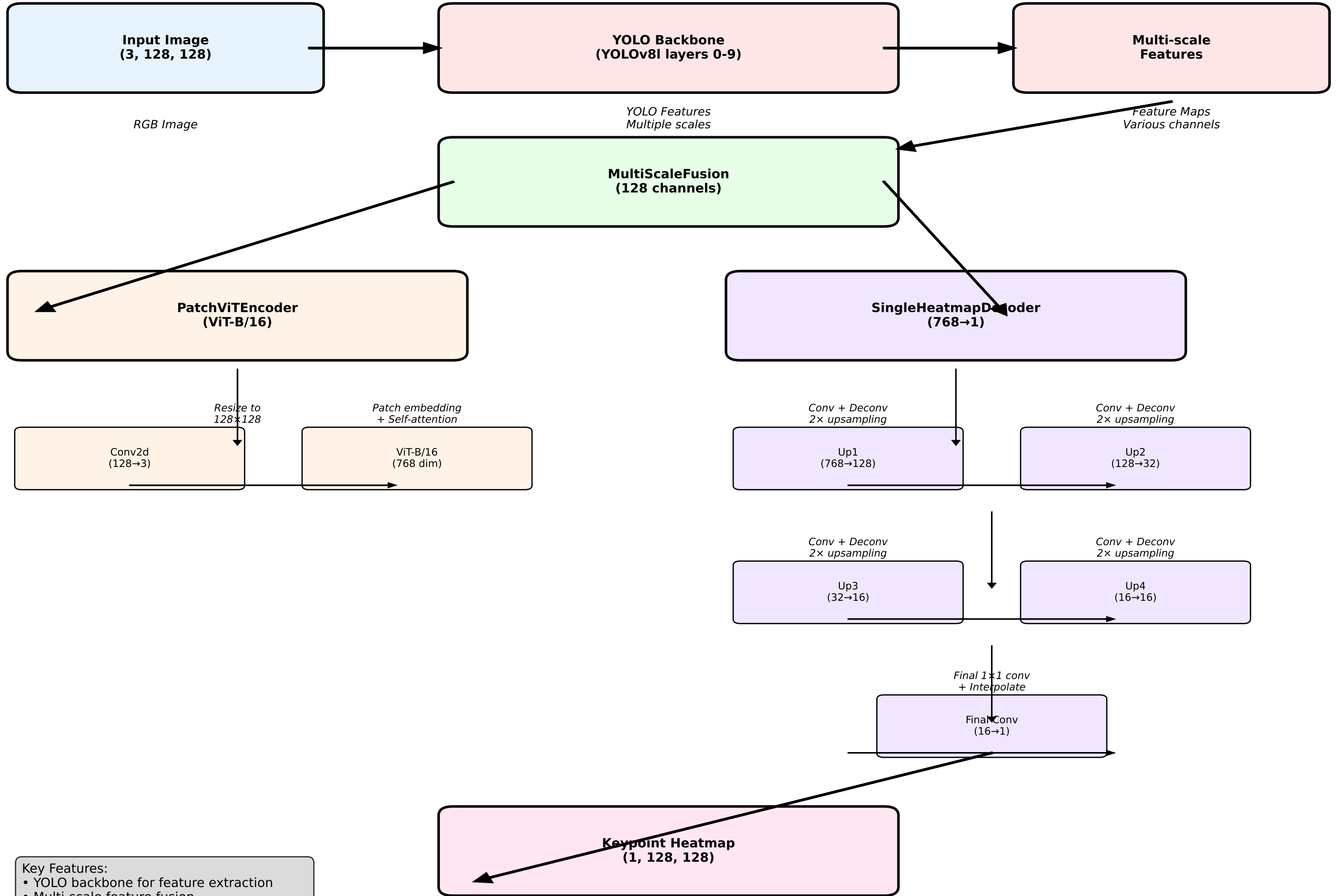


HybridKeypointNet Architecture

YOLO + Vision Transformer Hybrid Model for Keypoint Detection



Key Features:

- YOLO backbone for feature extraction
- Multi-scale feature fusion
- Vision Transformer for global context
- U-Net style decoder for precise localization
- Spatial softmax for keypoint probability
- Single heatmap per keypoint type