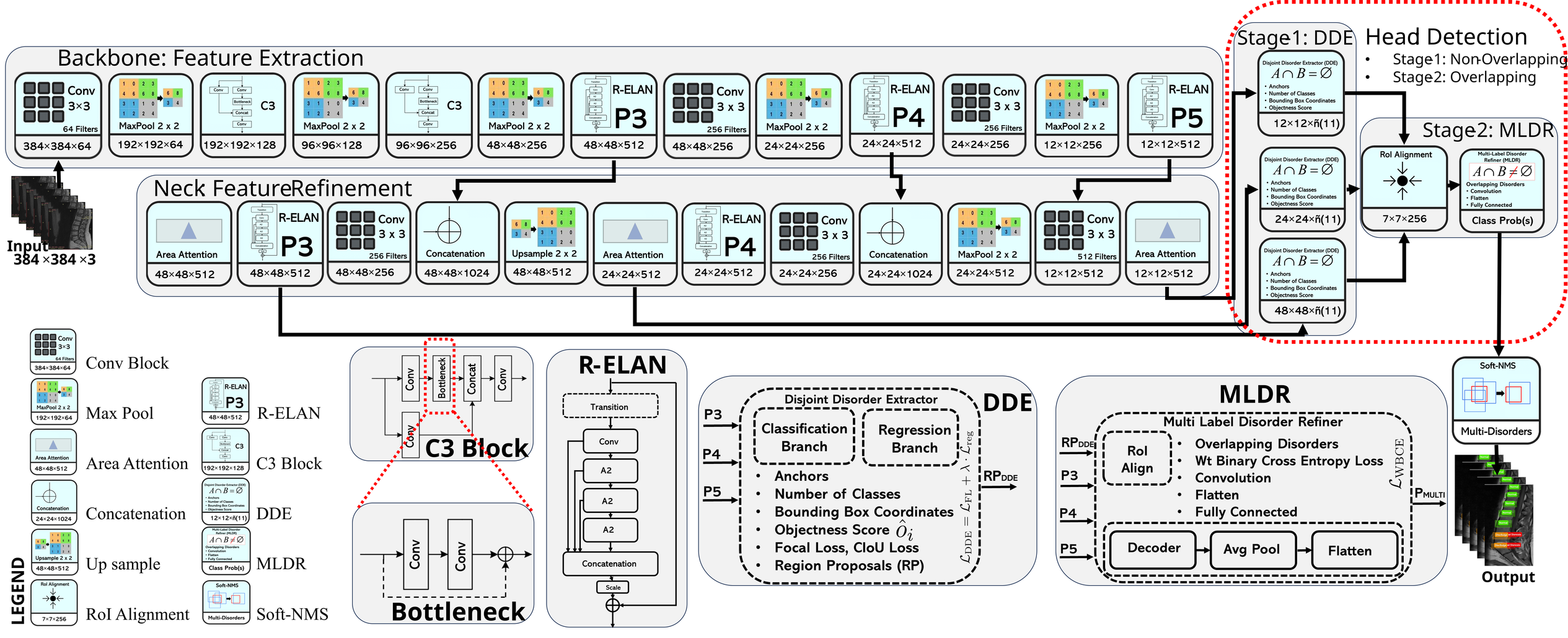
Backbone: Feature Extraction
Conv 3×3 64 Filters 384×384×64
MaxPool 2 x 2 192×192×64
C3 192×192×128
MaxPool 2 x 2 96×96×128
C3 96×96×256
MaxPool 2 x 2 48×48×256
R-ELAN P3 48×48×512
Conv 3 x 3 256 Filters 48×48×256
MaxPool 2 x 2 24×24×256
R-ELAN P4 24×24×512
Conv 3 x 3 256 Filters 24×24×256
MaxPool 2 x 2 12×12×256
R-ELAN P5 12×12×512
Input 384 ×384 ×3
Neck FeatureRefinement
Area Attention 48×48×512
R-ELAN P3 48×48×512
Conv 3 x 3 256 Filters 48×48×256
Concatenation 48×48×1024
Upsample 2 x 2 48×48×512
Area Attention 24×24×512
R-ELAN P4 24×24×512
Conv 3 x 3 256 Filters 24×24×256
Concatenation 24×24×1024
MaxPool 2 x 2 24×24×512
Conv 3 x 3 512 Filters 12×12×512
Area Attention 12×12×512
Stage1: DDE
Disjoint Disorder Extractor (DDE) $A \cap B = \varnothing$
• Anchors
• Number of Classes
• Bounding Box Coordinates
• Objectness Score
12×12×ñ(11)
24×24×ñ(11)
48×48×ñ(11)
Head Detection
• Stage1: Non-Overlapping
• Stage2: Overlapping
Stage2: MLDR
RoI Alignment 7×7×256
Multi-Label Disorder Refiner (MLDR) $A \cap B \neq \varnothing$
Overlapping Disorders
• Convolution
• Flatten
• Fully Connected
Class Prob(s)
Soft-NMS
Multi-Disorders
Output
LEGEND
Conv Block
Max Pool
Area Attention
Concatenation
Up sample
RoI Alignment
R-ELAN
C3 Block
DDE
MLDR
Soft-NMS
C3 Block
Bottleneck
R-ELAN
Transition
Conv
A2
A2
A2
Concatenation
Scale
DDE
Disjoint Disorder Extractor
Classification Branch
Regression Branch
• Anchors
• Number of Classes
• Bounding Box Coordinates
• Objectness Score $\hat{O}_i$
• Focal Loss, CIoU Loss
• Region Proposals (RP)
$\mathcal{L}_{DDE} = \mathcal{L}_{FL} + \lambda \cdot \mathcal{L}_{reg}$
P3
P4
P5
RP_DDE
MLDR
Multi Label Disorder Refiner
RoI Align
• Overlapping Disorders
• Wt Binary Cross Entropy Loss
• Convolution
• Flatten
• Fully Connected
$\mathcal{L}_{WBCE}$
Decoder
Avg Pool
Flatten
P_MULTI