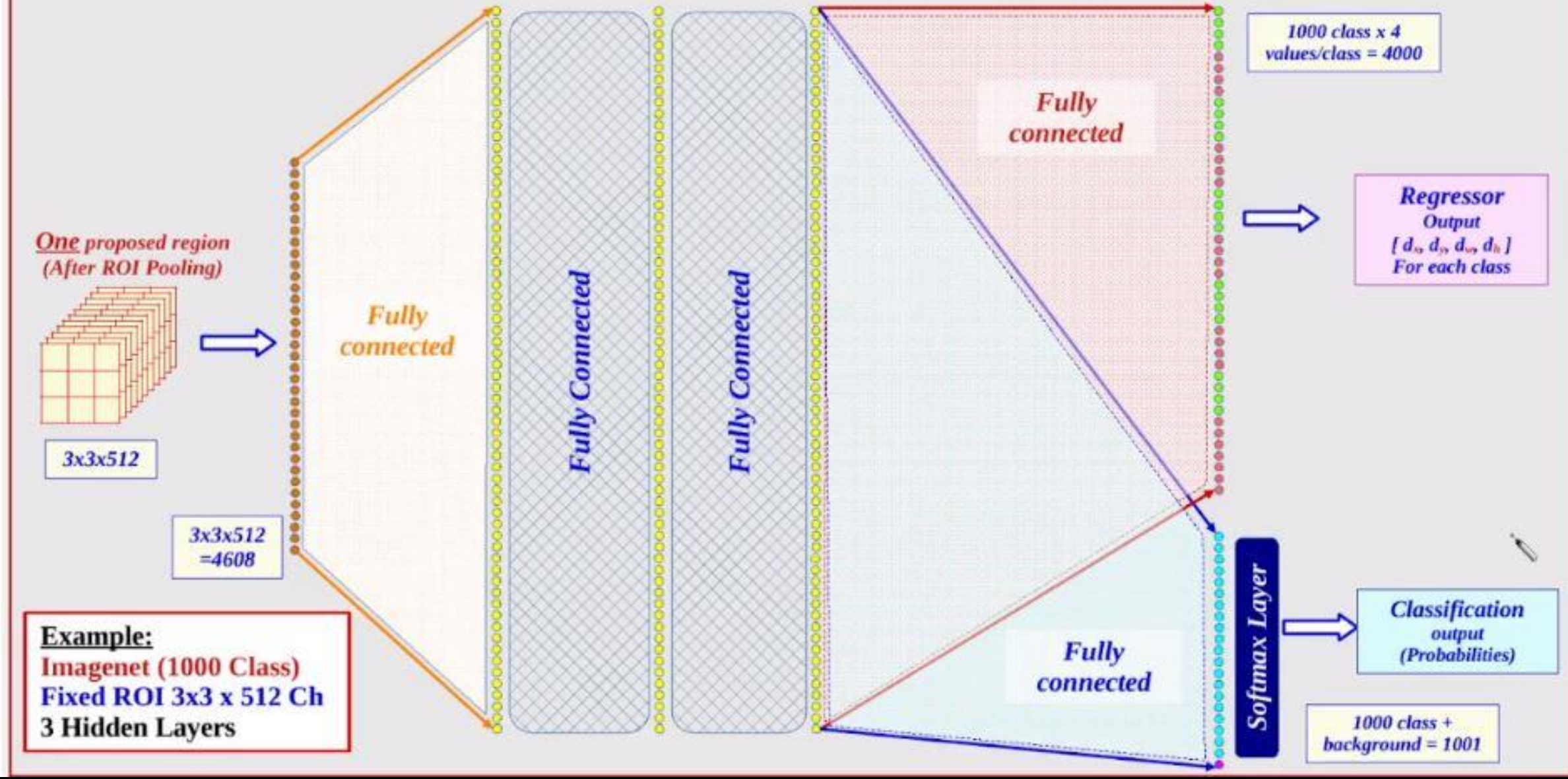
Training of Faster RCNN Detector

Architecture of Detector Network *(for Fast RCNN, Faster RCNN)*



CNN

Feature Map

RPN +
(non-max suppression, max count,..)

Proposed Boxes Vectors

Objectness Vectors

Proposed region

IOU with all Ground Truth (GT) Boxes

If < 0.5 → BG

Else Get GT box with highest IOU

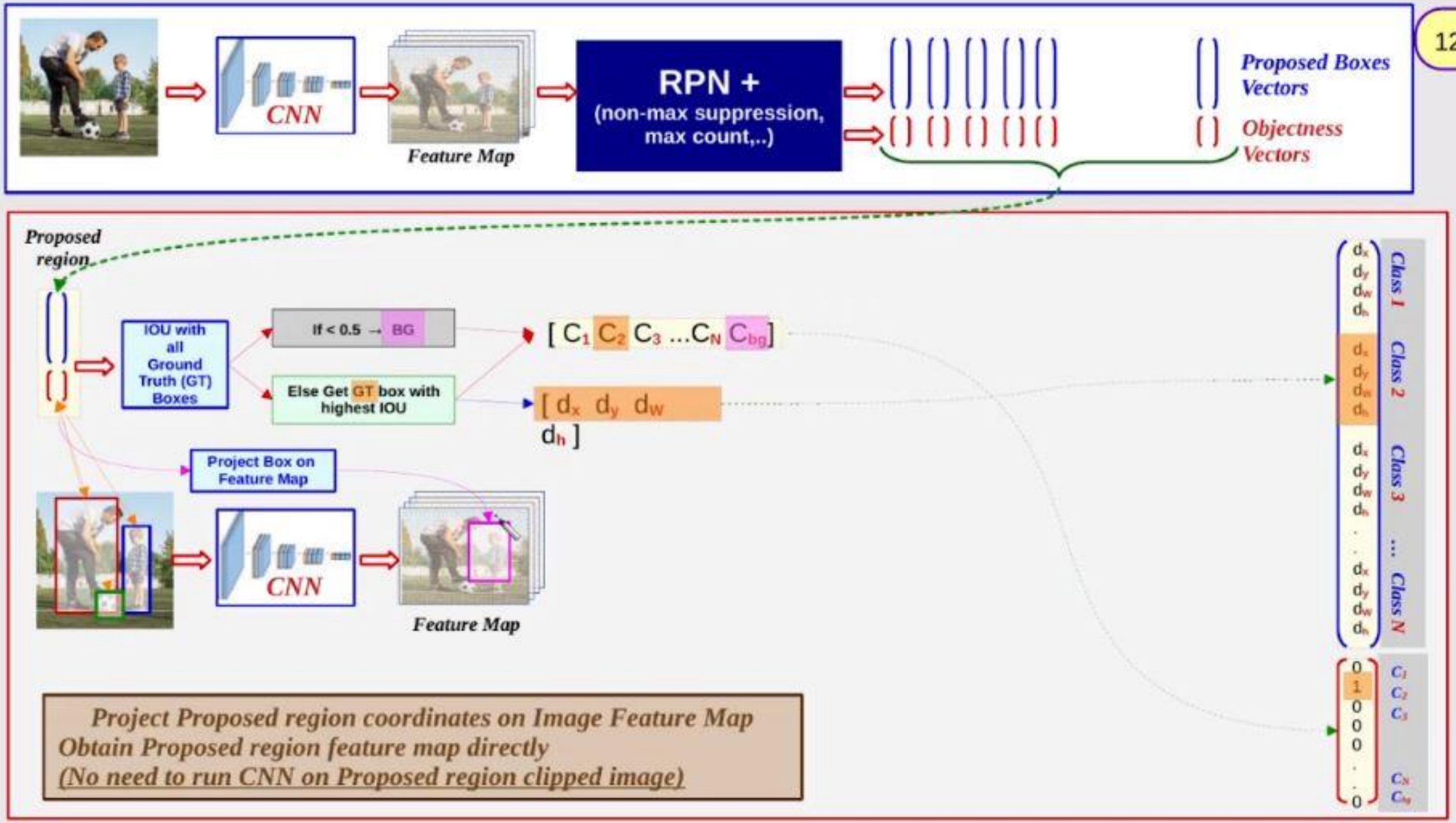
$[\ C_1\ C_2\ C_3\ \ldots C_N\ C_{bg}]$

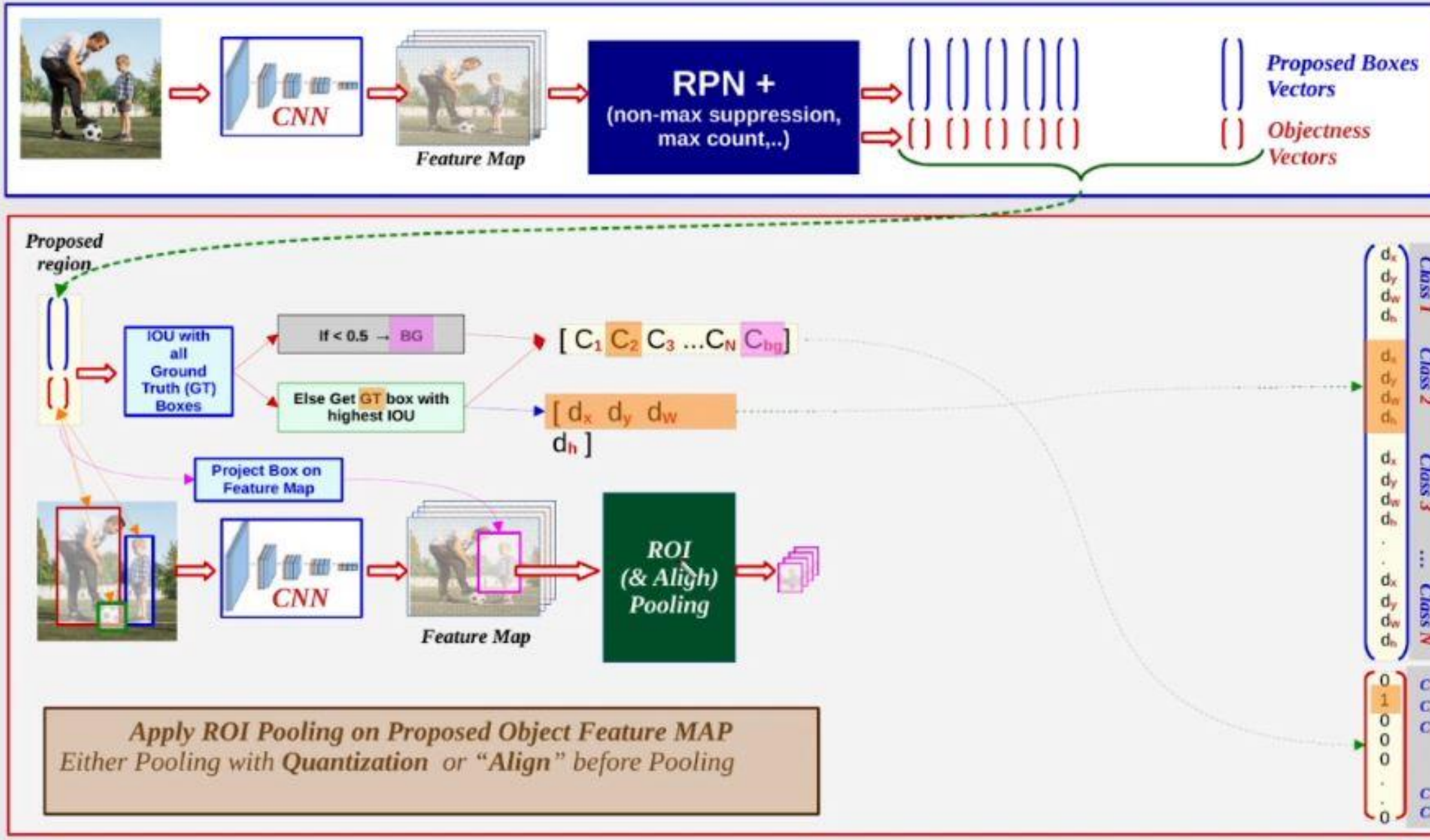
$[\ d_x\ d_y\ d_w\ d_h\ ]$

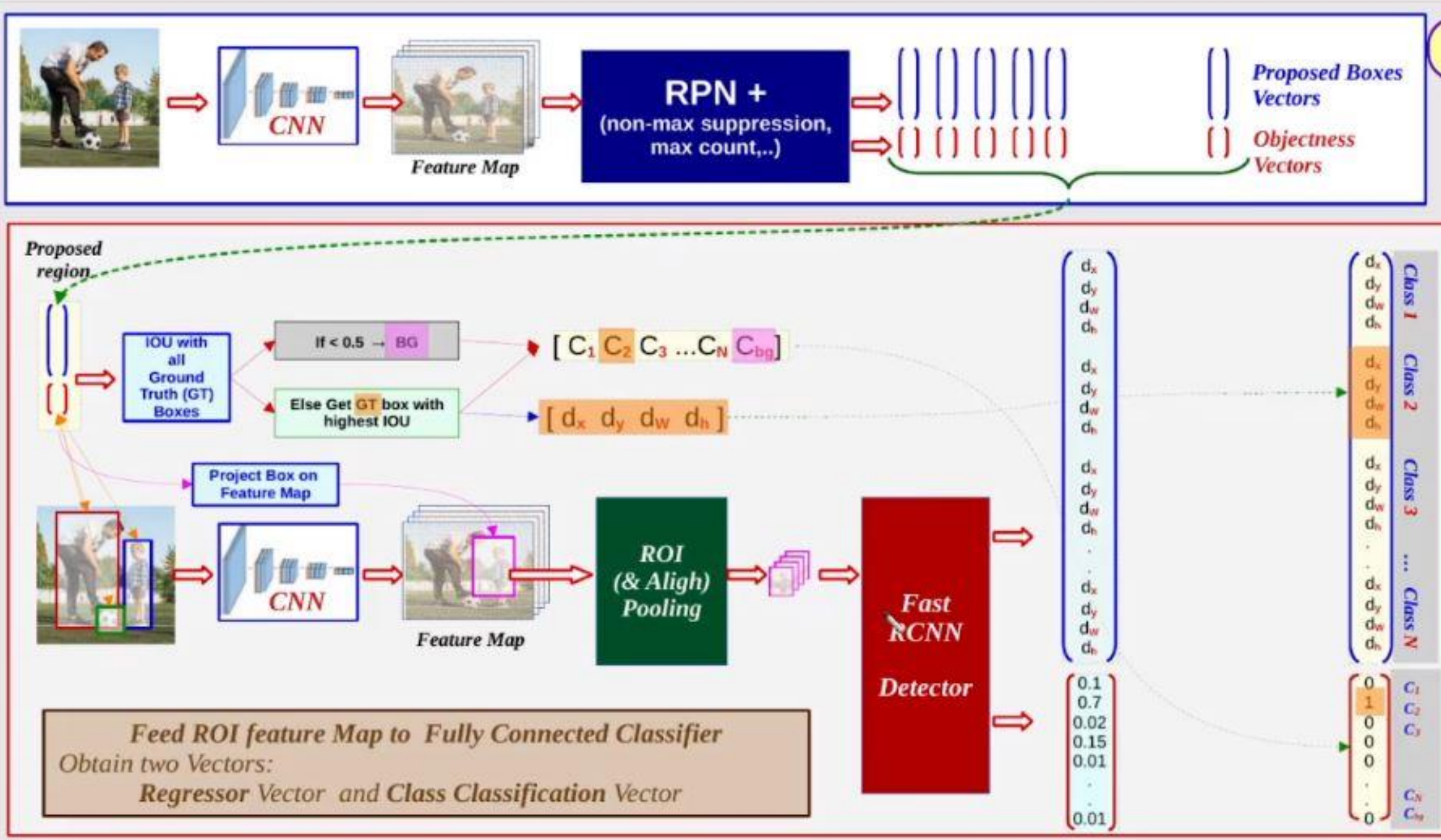
$d_x\ d_y\ d_w\ d_h$ . . $d_x\ d_y\ d_w\ d_h$

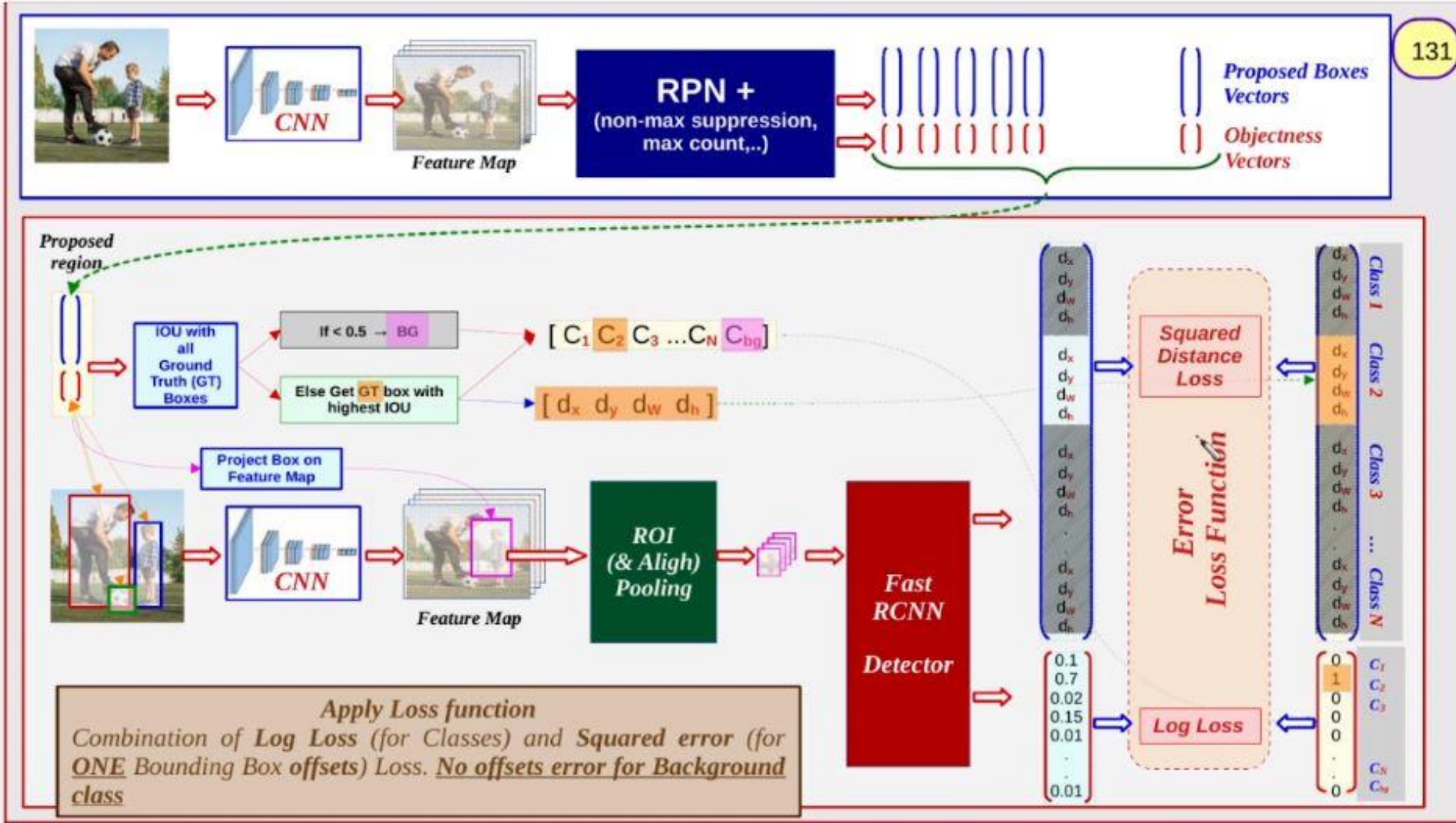
0 1 0 0 0 . . 0

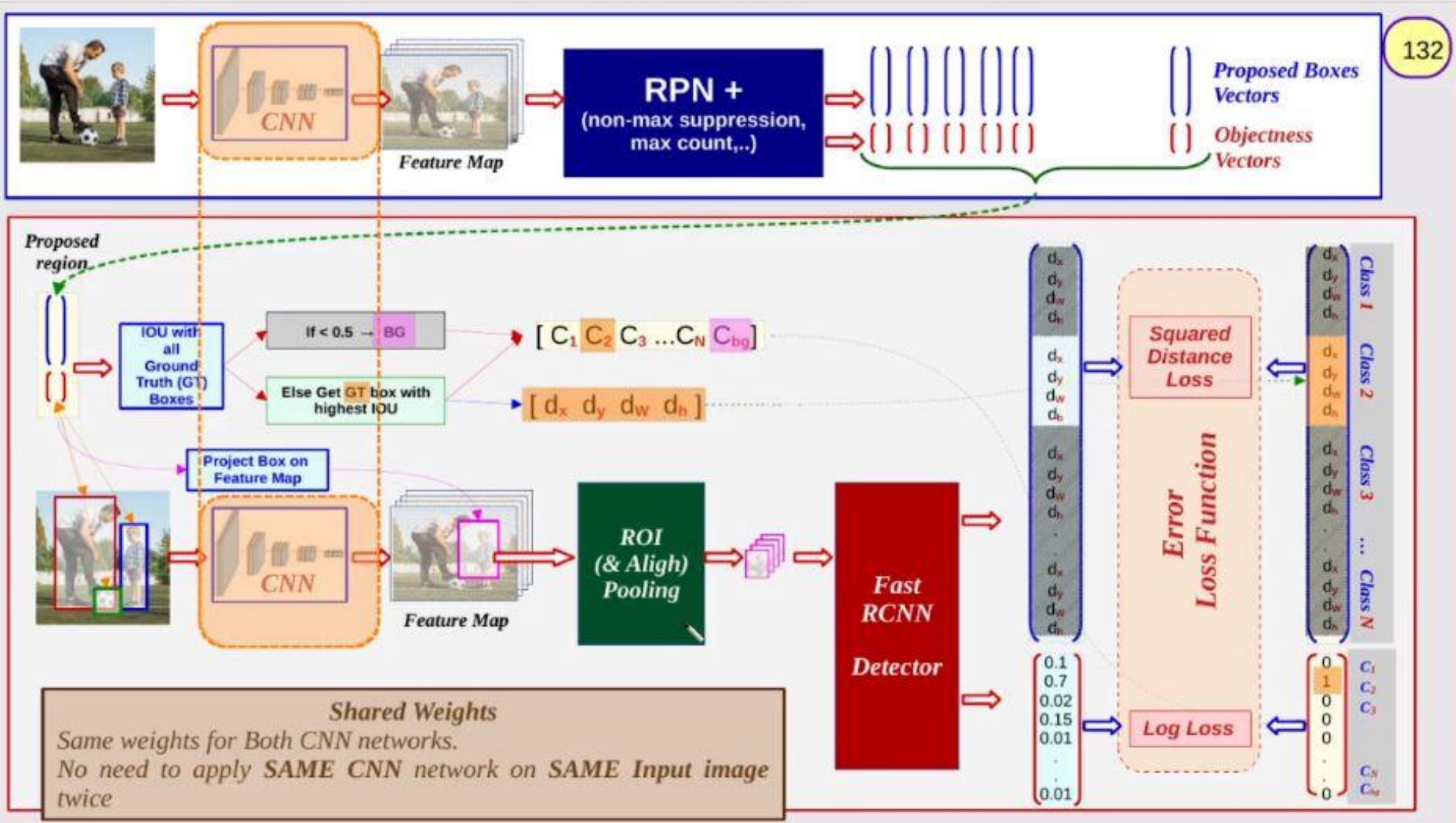
Prepare Ground Truth Labels for Proposed Region
Classify Proposed region *as enclosing Object / Background / Mix*
Get Regressor Vector *(w.r.t. heights IOU ground truth Object)*

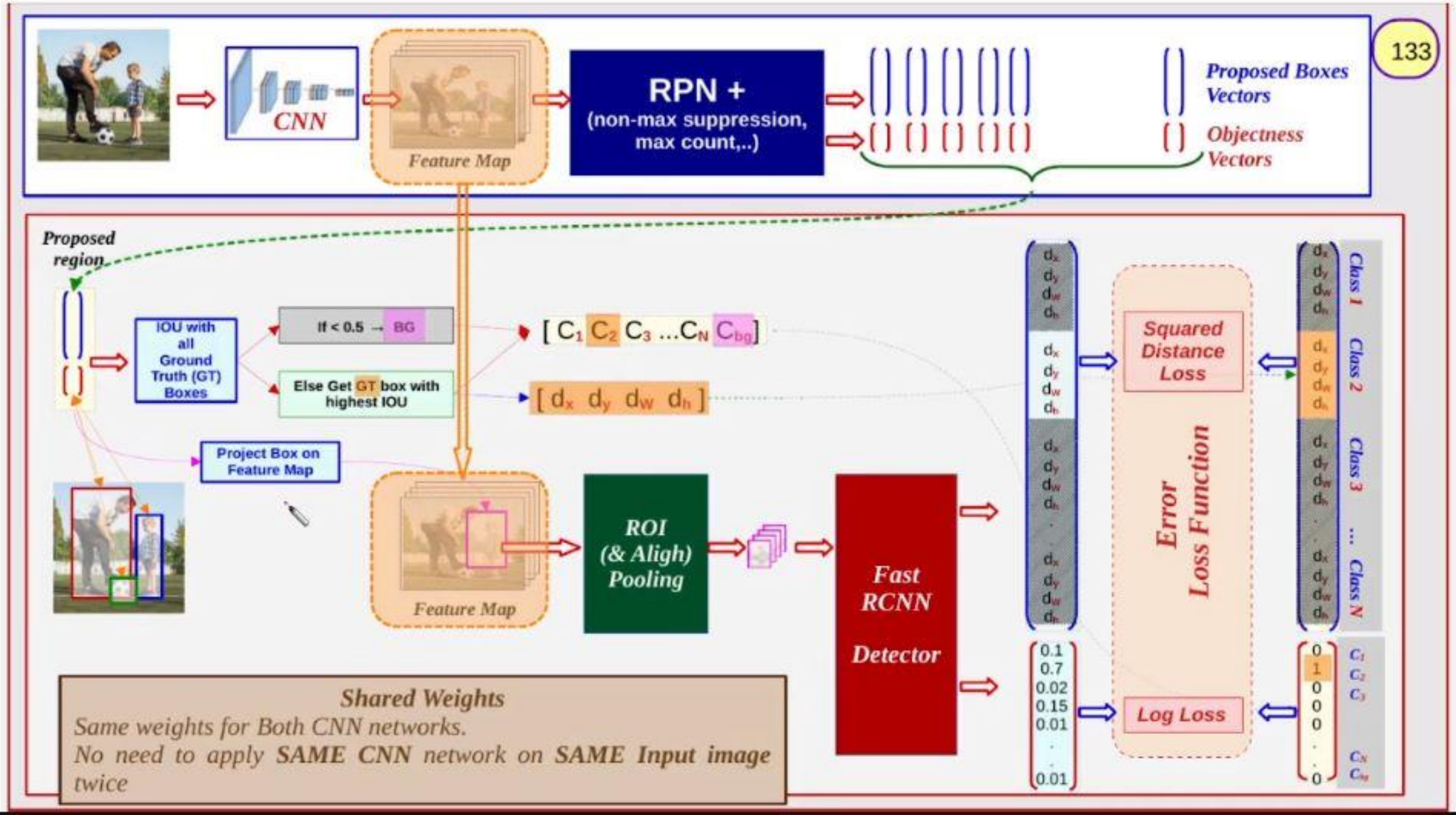

CNN
Feature Map
RPN +
(non-max suppression, max count,..)
Proposed Boxes Vectors
Objectness Vectors
Proposed region
IOU with all Ground Truth (GT) Boxes
If < 0.5 → BG
Else Get GT box with highest IOU
[C1 C2 C3 ...CN Cbg]
[dx dy dw dh]
Project Box on Feature Map
CNN
Feature Map
Class 1
Class 2
Class 3
...
Class N
C1
C2
C3
CN
Cbg
Project Proposed region coordinates on Image Feature Map
Obtain Proposed region feature map directly
(No need to run CNN on Proposed region clipped image)


CNN
Feature Map
RPN +
(non-max suppression, max count,..)
Proposed Boxes Vectors
Objectness Vectors
Proposed region
IOU with all Ground Truth (GT) Boxes
If < 0.5 → BG
Else Get GT box with highest IOU
[C1 C2 C3 ...CN Cbg]
[dx dy dw dh]
Project Box on Feature Map
CNN
Feature Map
ROI (& Align) Pooling
Class 1
Class 2
Class 3
... Class N
Apply ROI Pooling on Proposed Object Feature MAP
Either Pooling with Quantization or "Align" before Pooling

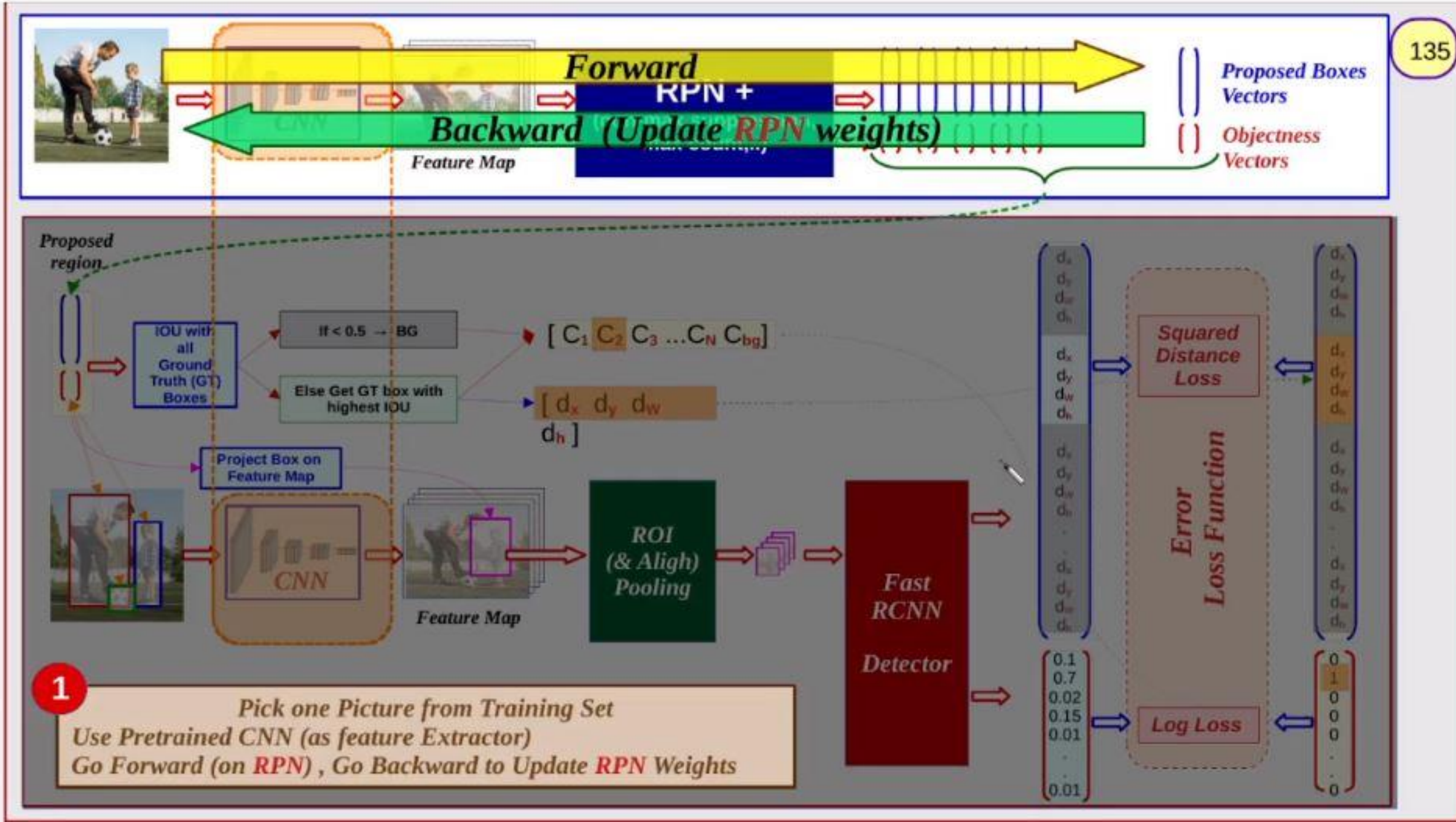

CNN
Feature Map
RPN +
(non-max suppression, max count,..)
Proposed Boxes Vectors
Objectness Vectors
Proposed region
IOU with all Ground Truth (GT) Boxes
If < 0.5 → BG
Else Get GT box with highest IOU
[C1 C2 C3 ...CN Cbg]
[dx dy dw dh]
Project Box on Feature Map
CNN
Feature Map
ROI (& Aligh) Pooling
Fast RCNN Detector
Class 1
Class 2
Class 3
... Class N
Feed ROI feature Map to Fully Connected Classifier
Obtain two Vectors:
Regressor Vector and Class Classification Vector

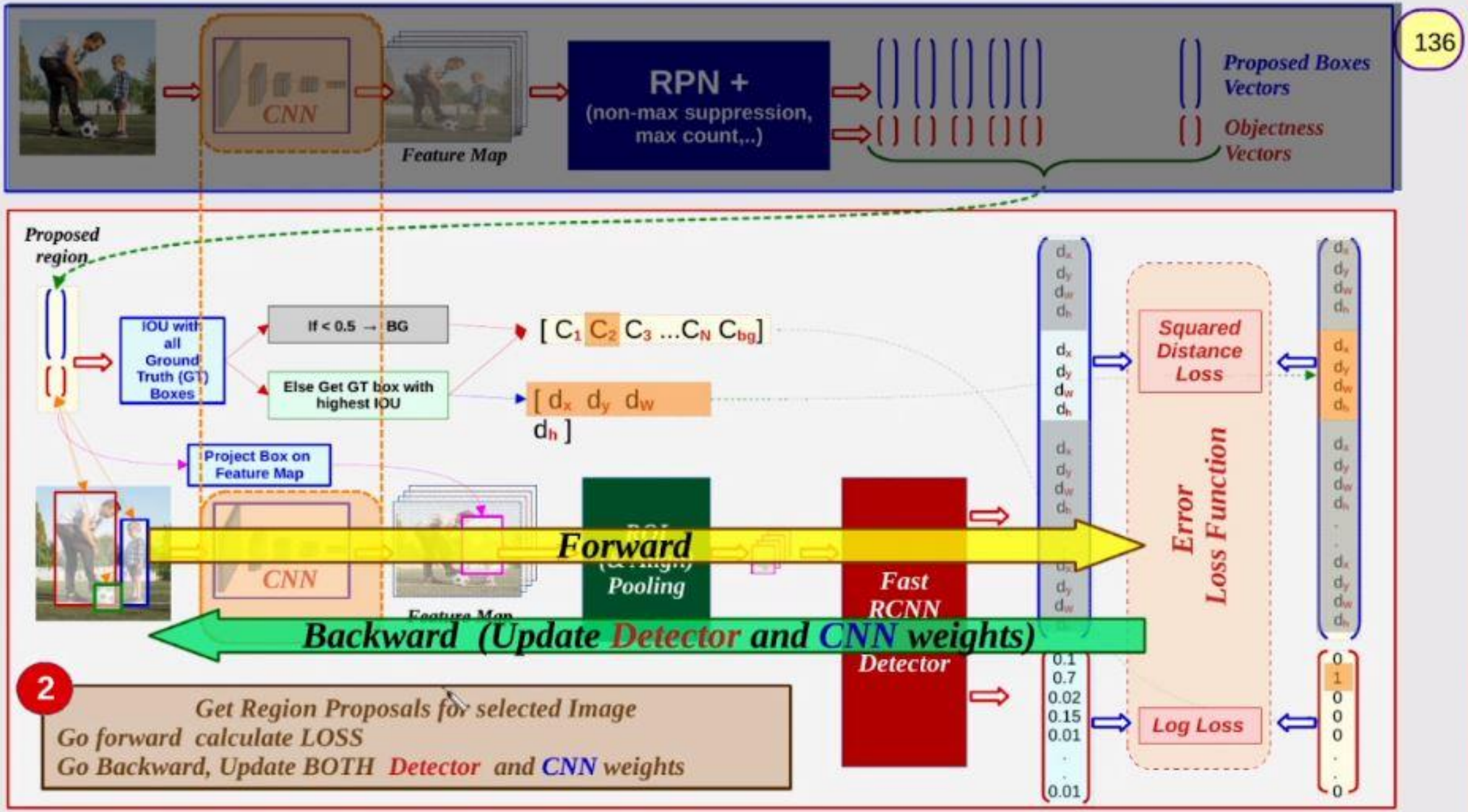

CNN
Feature Map
RPN +
(non-max suppression, max count,..)
Proposed Boxes Vectors
Objectness Vectors
Proposed region
IOU with all Ground Truth (GT) Boxes
If < 0.5 → BG
Else Get GT box with highest IOU
[C1 C2 C3 ...CN Cbg]
[dx dy dw dh]
Project Box on Feature Map
CNN
Feature Map
ROI (& Aligh) Pooling
Fast RCNN Detector
Squared Distance Loss
Error Loss Function
Log Loss
Class 1
Class 2
Class 3
...
Class N
Apply Loss function
Combination of Log Loss (for Classes) and Squared error (for ONE Bounding Box offsets) Loss. No offsets error for Background class

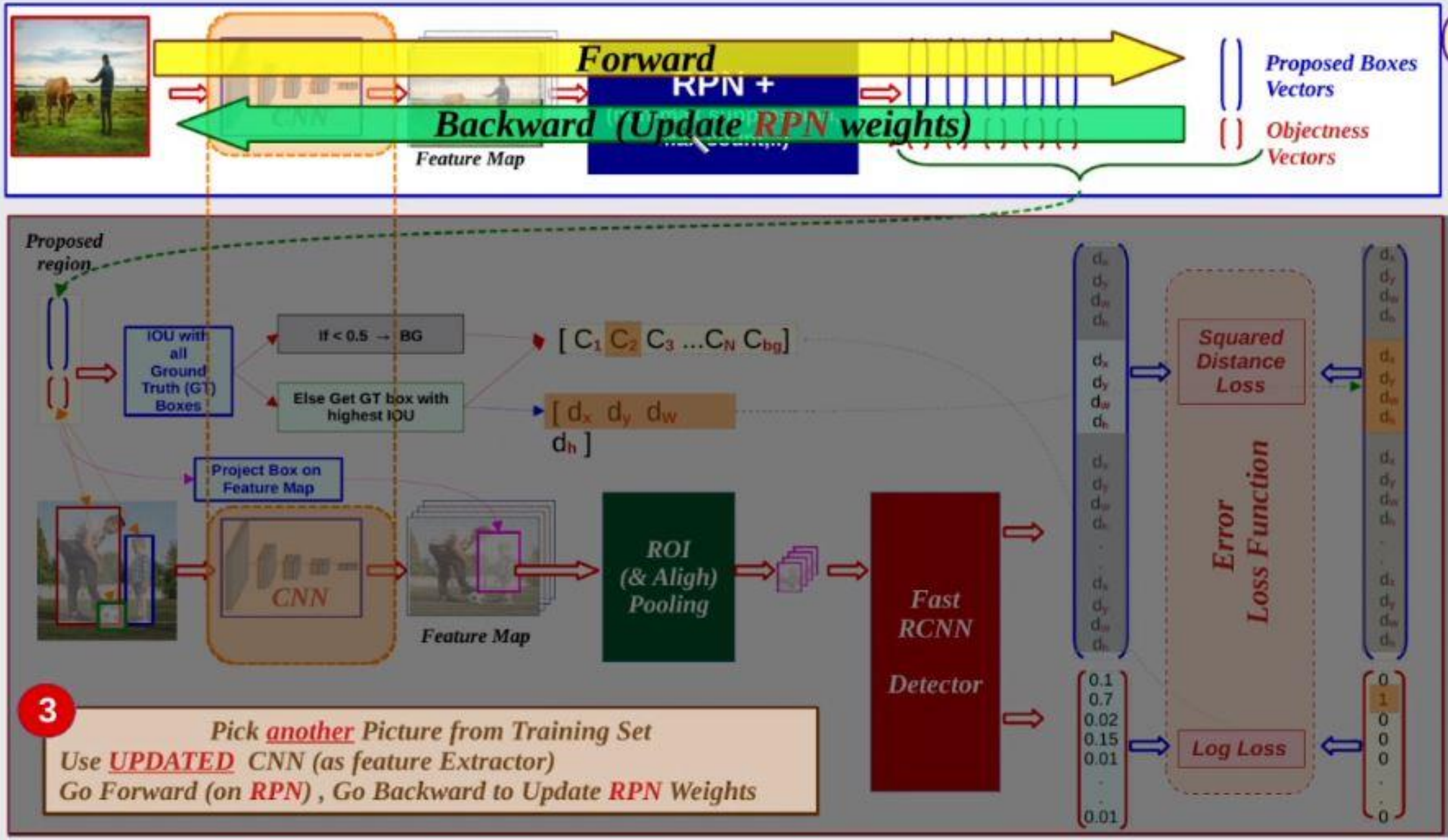

CNN
Feature Map
RPN +
(non-max suppression, max count,..)
Proposed Boxes Vectors
Objectness Vectors
Proposed region
IOU with all Ground Truth (GT) Boxes
If < 0.5 → BG
Else Get GT box with highest IOU
[C1 C2 C3 ...CN Cbg]
[dx dy dw dh]
Project Box on Feature Map
CNN
Feature Map
ROI (& Aligh) Pooling
Fast RCNN Detector
Squared Distance Loss
Error Loss Function
Log Loss
Class 1
Class 2
Class 3
... Class N
0.1
0.7
0.02
0.15
0.01
0.01
Shared Weights
Same weights for Both CNN networks.
No need to apply SAME CNN network on SAME Input image twice

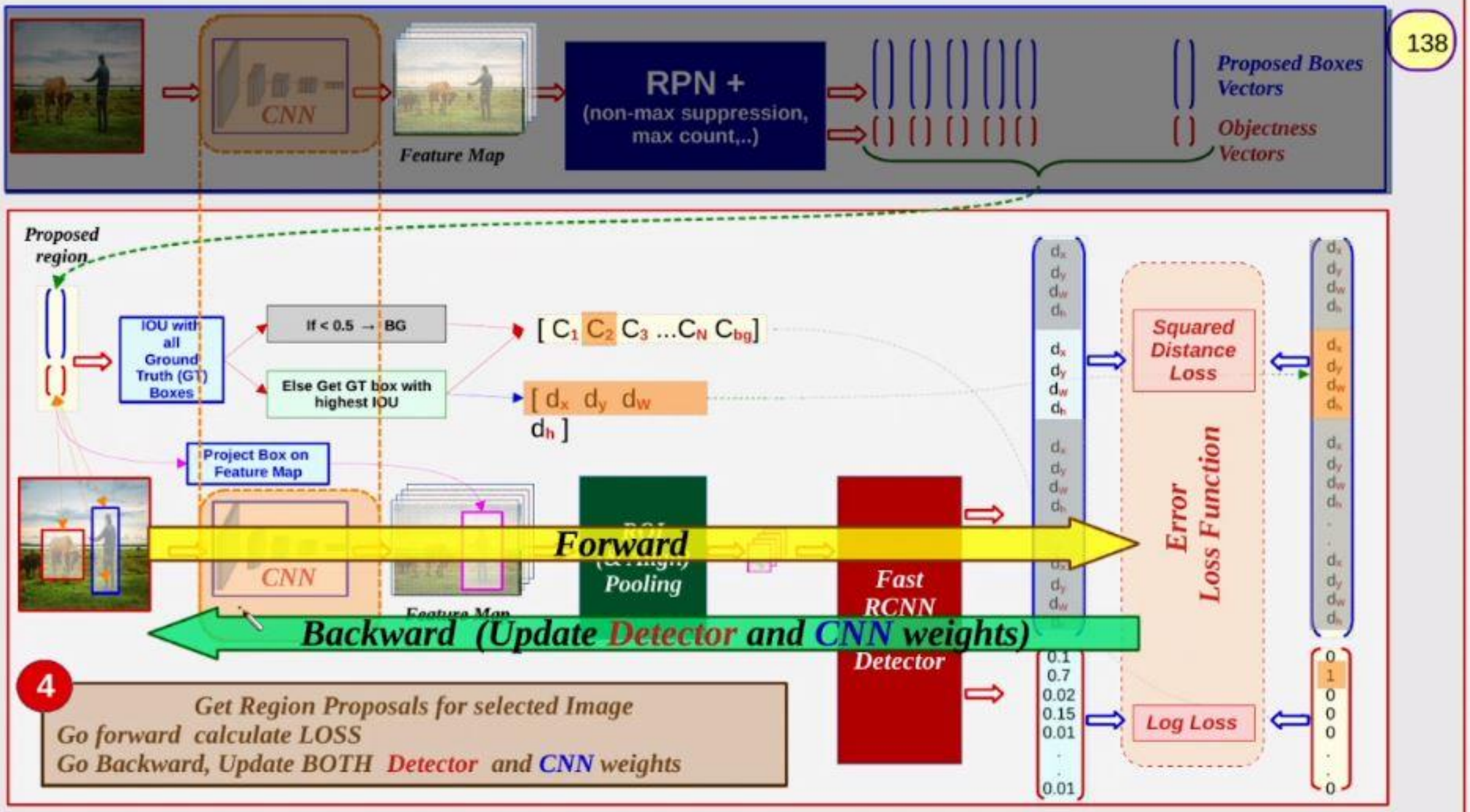

CNN
Feature Map
RPN +
(non-max suppression, max count,..)
Proposed Boxes Vectors
Objectness Vectors
Proposed region
IOU with all Ground Truth (GT) Boxes
If < 0.5 → BG
Else Get GT box with highest IOU
[C1 C2 C3 ...CN Cbg]
[dx dy dw dh]
Project Box on Feature Map
Feature Map
ROI (& Aligh) Pooling
Fast RCNN Detector
Squared Distance Loss
Error Loss Function
Log Loss
Class 1
Class 2
Class 3
... Class N
Shared Weights
Same weights for Both CNN networks.
No need to apply SAME CNN network on SAME Input image twice

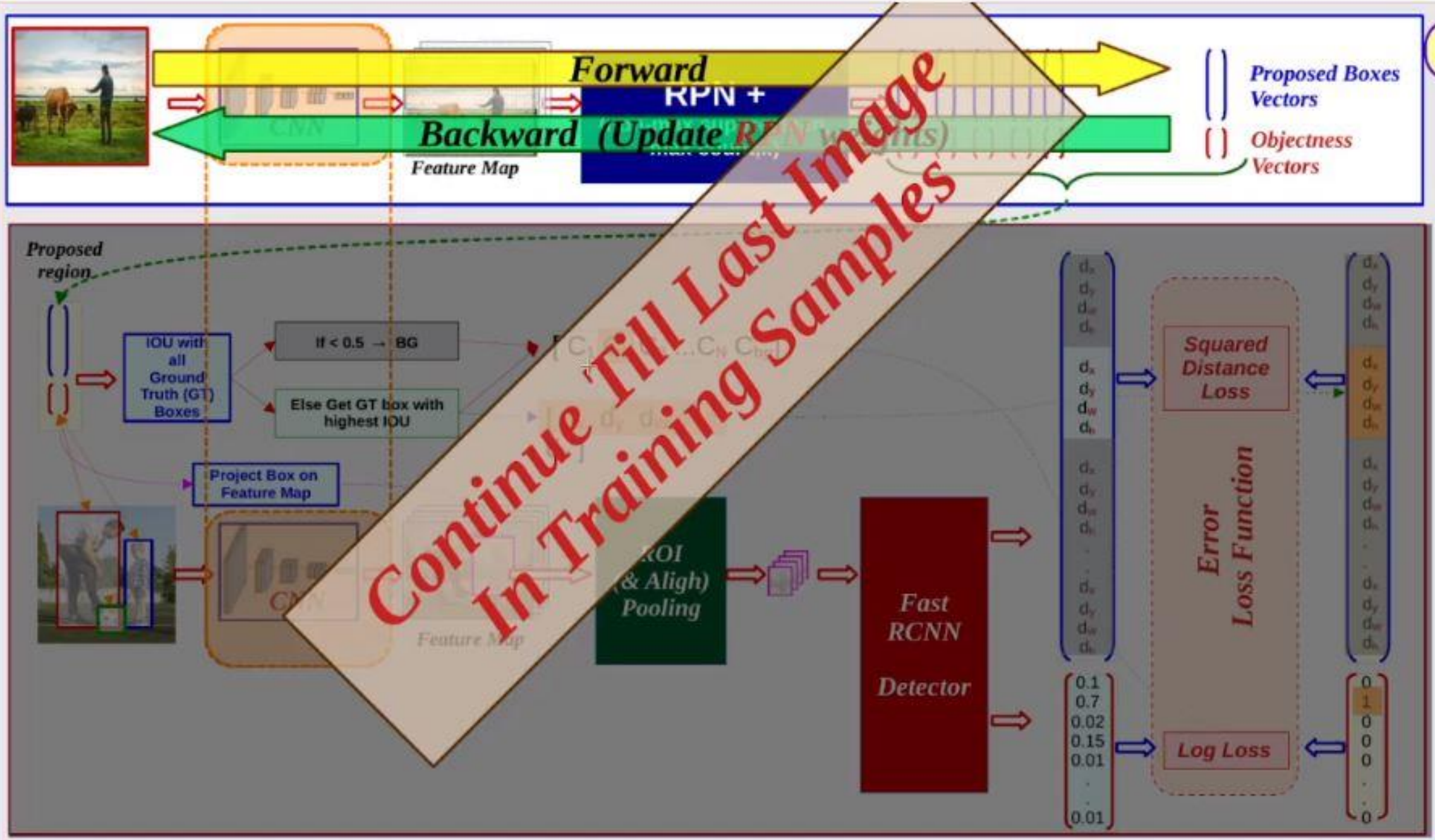
Alternate Training for RPN and Faster RCNN Detector

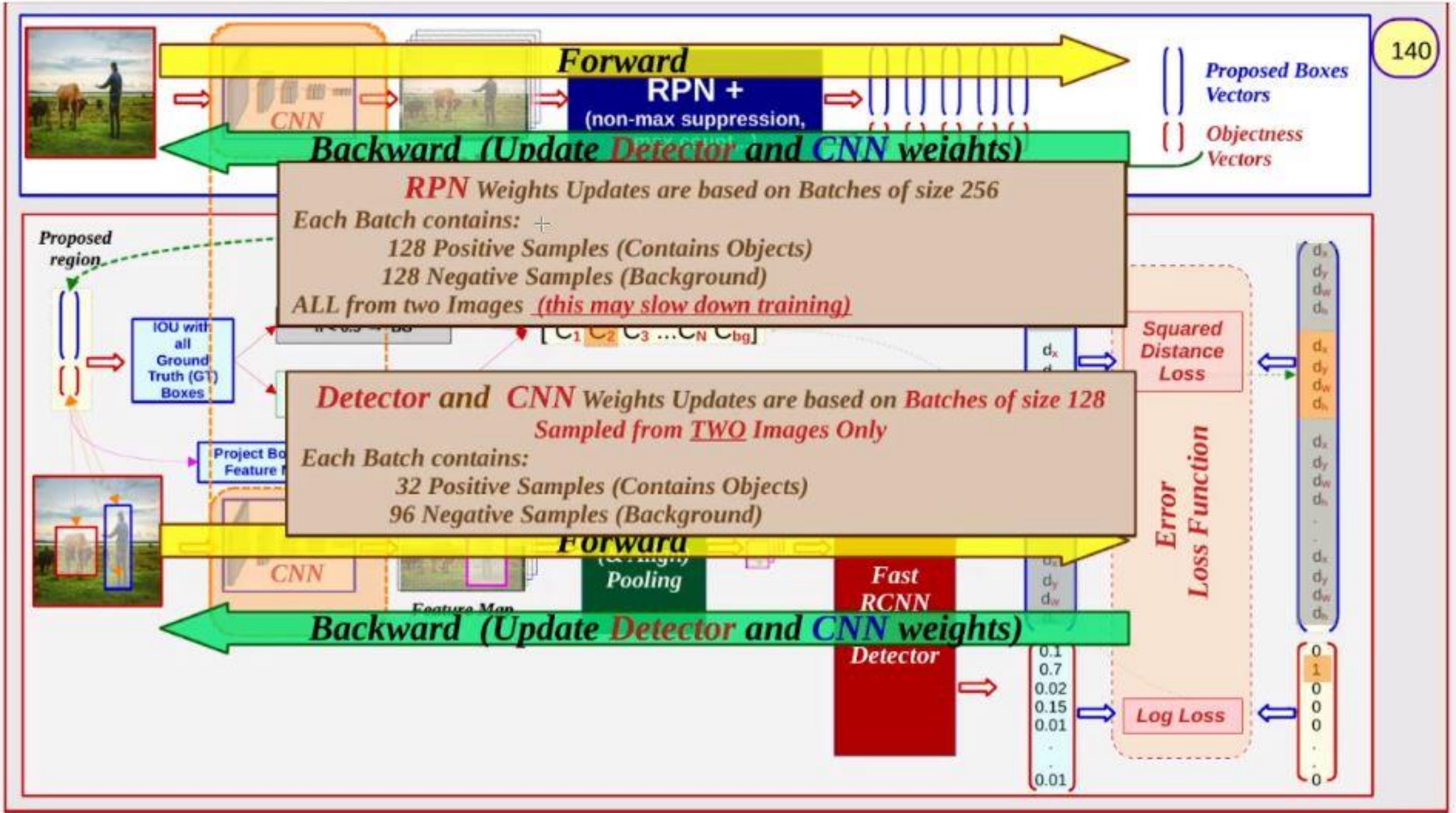

Forward
RPN +
Backward (Update RPN weights)
Feature Map
Proposed Boxes Vectors
Objectness Vectors
Proposed region
IOU with all Ground Truth (GT) Boxes
If < 0.5 → BG
Else Get GT box with highest IOU
[C1 C2 C3 ...CN Cbg]
[dx dy dw dh]
Project Box on Feature Map
CNN
Feature Map
ROI (& Aligh) Pooling
Fast RCNN Detector
Squared Distance Loss
Error Loss Function
Log Loss
1
Pick one Picture from Training Set
Use Pretrained CNN (as feature Extractor)
Go Forward (on RPN) , Go Backward to Update RPN Weights

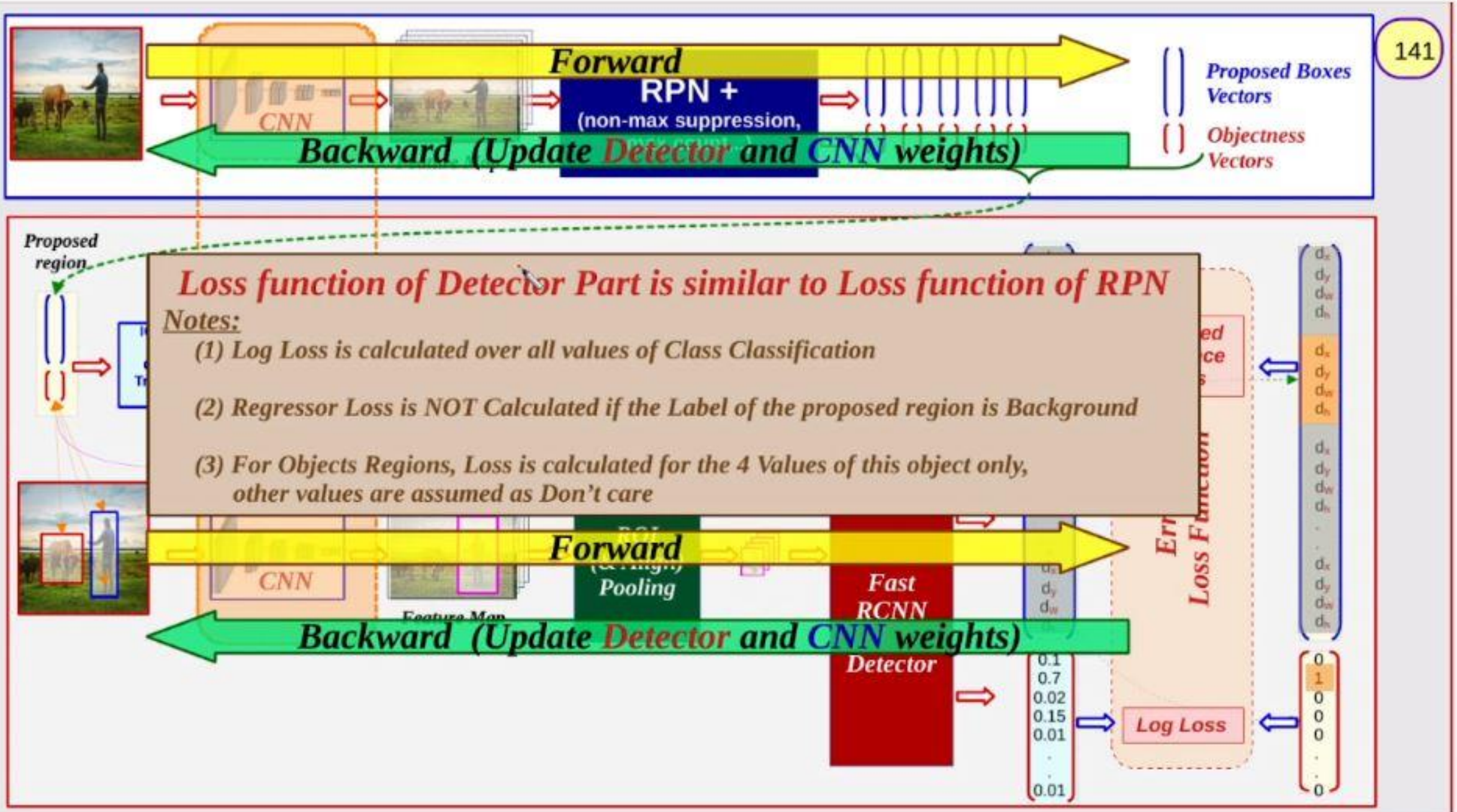

CNN
Feature Map
RPN +
(non-max suppression, max count,..)
Proposed Boxes Vectors
Objectness Vectors
Proposed region
IOU with all Ground Truth (GT) Boxes
If < 0.5 → BG
Else Get GT box with highest IOU
[C1 C2 C3 ...CN Cbg]
[dx dy dw dh]
Project Box on Feature Map
CNN
Feature Map
RoI (& Align) Pooling
Fast RCNN Detector
Forward
Backward (Update Detector and CNN weights)
Squared Distance Loss
Error Loss Function
Log Loss
0.1
0.7
0.02
0.15
0.01
0.01
2
Get Region Proposals for selected Image
Go forward calculate LOSS
Go Backward, Update BOTH Detector and CNN weights


Forward
Backward (Update RPN weights)
CNN
Feature Map
RPN +
Proposed Boxes Vectors
Objectness Vectors
Proposed region
IOU with all Ground Truth (GT) Boxes
If < 0.5 → BG
Else Get GT box with highest IOU
[C1 C2 C3 ...CN Cbg]
[dx dy dw dh]
Project Box on Feature Map
CNN
Feature Map
ROI (& Aligh) Pooling
Fast RCNN Detector
Squared Distance Loss
Error Loss Function
Log Loss
3
Pick another Picture from Training Set
Use UPDATED CNN (as feature Extractor)
Go Forward (on RPN) , Go Backward to Update RPN Weights


CNN
Feature Map
RPN +
(non-max suppression, max count,..)
Proposed Boxes Vectors
Objectness Vectors
Proposed region
IOU with all Ground Truth (GT) Boxes
If < 0.5 → BG
Else Get GT box with highest IOU
[C1 C2 C3 ...CN Cbg]
[dx dy dw dh]
Project Box on Feature Map
CNN
Feature Map
Pooling
Fast RCNN Detector
Forward
Backward (Update Detector and CNN weights)
Squared Distance Loss
Log Loss
Error Loss Function
0.1
0.7
0.02
0.15
0.01
0.01
4
Get Region Proposals for selected Image
Go forward calculate LOSS
Go Backward, Update BOTH Detector and CNN weights


Forward
Backward (Update RPN weights)
RPN +
Feature Map
Proposed Boxes Vectors
Objectness Vectors
Proposed region
IOU with all Ground Truth (GT) Boxes
If < 0.5 → BG
Else Get GT box with highest IOU
Project Box on Feature Map
CNN
Feature Map
ROI (& Aligh) Pooling
Fast RCNN Detector
Squared Distance Loss
Error Loss Function
Log Loss
Continue Till Last Image In Training Samples


Forward
RPN +
(non-max suppression,
CNN
Proposed Boxes Vectors
Objectness Vectors
Backward (Update Detector and CNN weights)
RPN Weights Updates are based on Batches of size 256
Each Batch contains:
128 Positive Samples (Contains Objects)
128 Negative Samples (Background)
ALL from two Images (this may slow down training)
Proposed region
IOU with all Ground Truth (GT) Boxes
Squared Distance Loss
Detector and CNN Weights Updates are based on Batches of size 128
Sampled from TWO Images Only
Each Batch contains:
32 Positive Samples (Contains Objects)
96 Negative Samples (Background)
Forward
CNN
Pooling
Fast RCNN Detector
Error Loss Function
Log Loss
Backward (Update Detector and CNN weights)

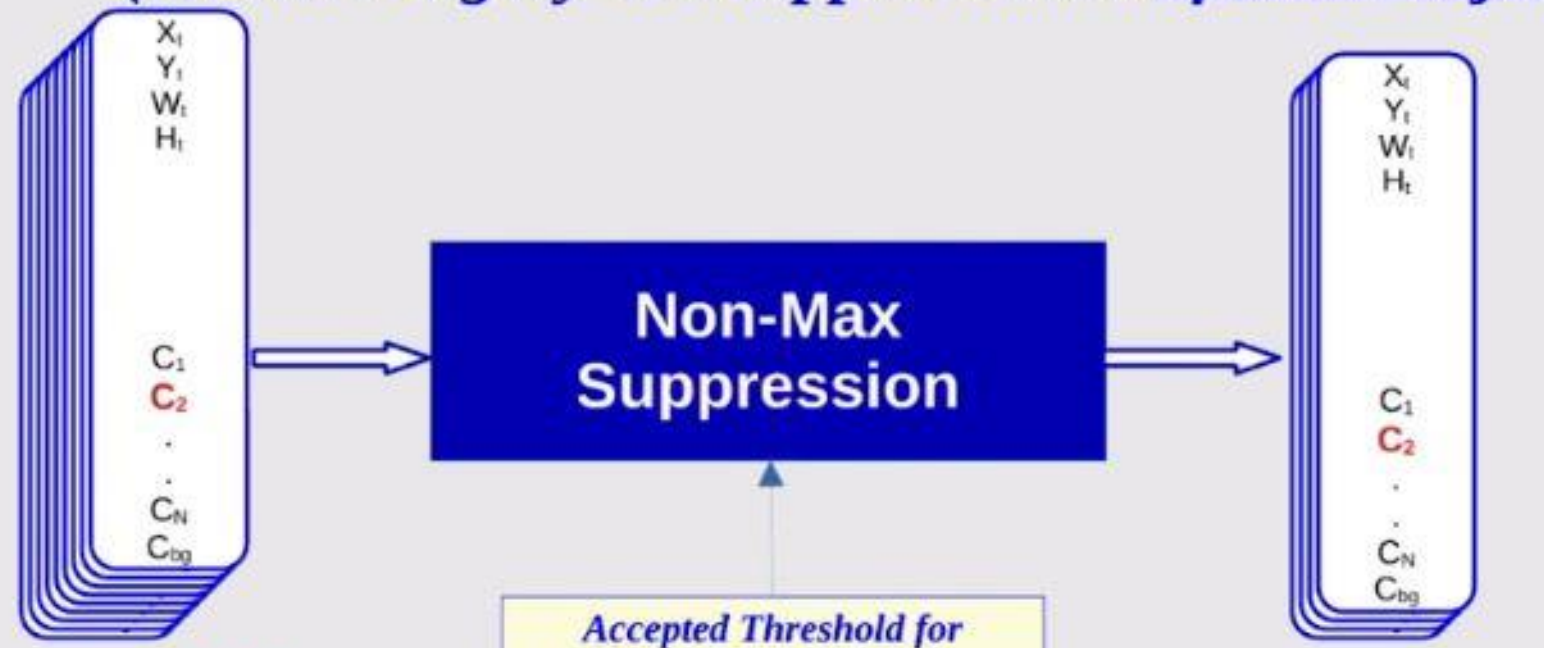

Forward
RPN +
(non-max suppression,
CNN
Backward (Update Detector and CNN weights)
Proposed Boxes Vectors
Objectness Vectors
Proposed region
Loss function of Detector Part is similar to Loss function of RPN
Notes:
(1) Log Loss is calculated over all values of Class Classification
(2) Regressor Loss is NOT Calculated if the Label of the proposed region is Background
(3) For Objects Regions, Loss is calculated for the 4 Values of this object only, other values are assumed as Don't care
Forward
CNN
Pooling
Fast RCNN Detector
Backward (Update Detector and CNN weights)
Log Loss
0.1
0.7
0.02
0.15
0.01
0.01
0
1
0
0
0
0

Final Detected Objects

using Detector and Non-Max Suppression

Filter Boxes of detected Objects (within same Object Class)

(Remove highly overlapped Boxes of same object)



Repeat for All Classes