BRAIN TUMOR DETECTION USING THE RETINANET MODEL

-- JESSE ANNAN

INTRODUCTION TO OBJECT DETECTION

- <u>Computer Vision Task:</u>
  - Image Classification
  - **Object Detection**
  - Image Segmentation

- Object Detection is a computer vision technique for classifying an object in an image and localizing the object.

- <u>Applications</u>: Surveillance, autonomous vehicles, image retrieval, medical imaging, etc

THE RISE OF DEEP LEARNING METHODS

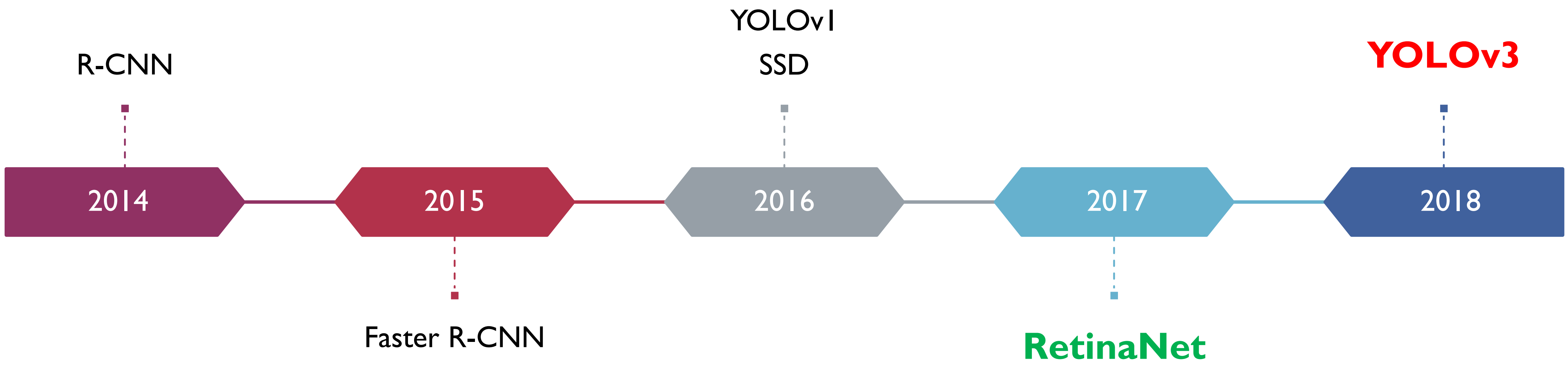
- Early Methods: Traditional Computer Vision techniques:

1. Histogram of Oriented Gradients (HOG)
2. Scale-Invariant Feature Transform (SIFT)

- Rise of Deep Learning (2014 - Present):

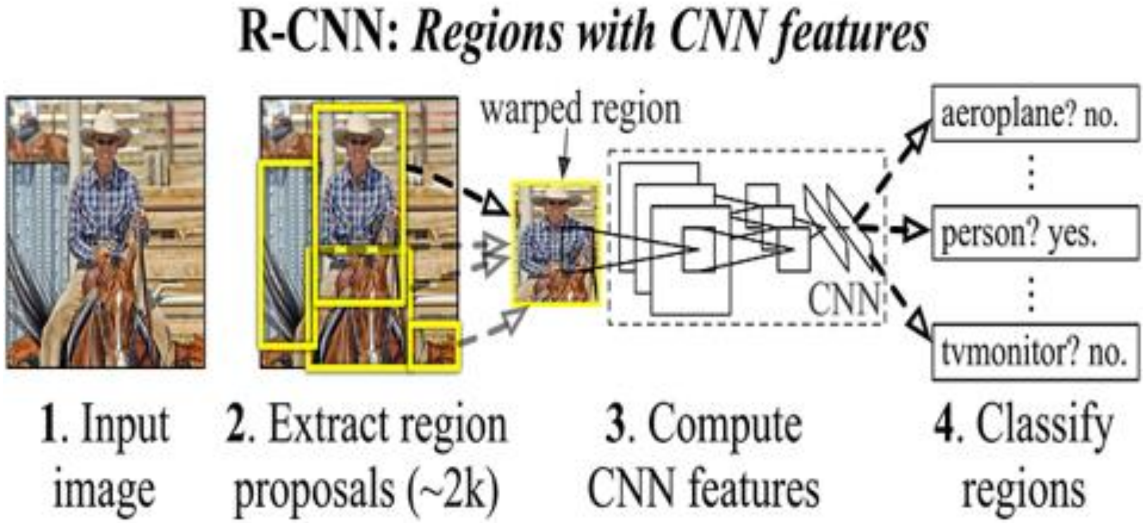
1. R-CNN
2. Faster R-CNN
3. SSD
4. YOLOv1
5. RetinaNet
6. YOLOv3

OBJECT DETECTION TIMELINE



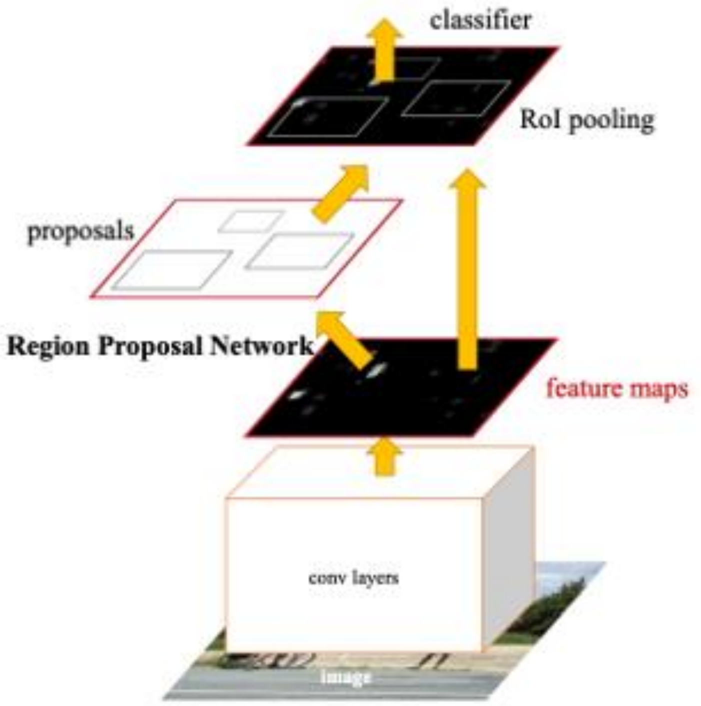
REGIONS WITH CNN (R-CNN)

- **Developed by:** Ross Girshick et al.
- **Process:**
  - Extract region proposals using selective search
  - Warp and resize each region
  - Classify per class region with SVMs using extracted features from a CNN
- **Pros/Impact:** Inspired research into deep learning-based object detection frameworks.



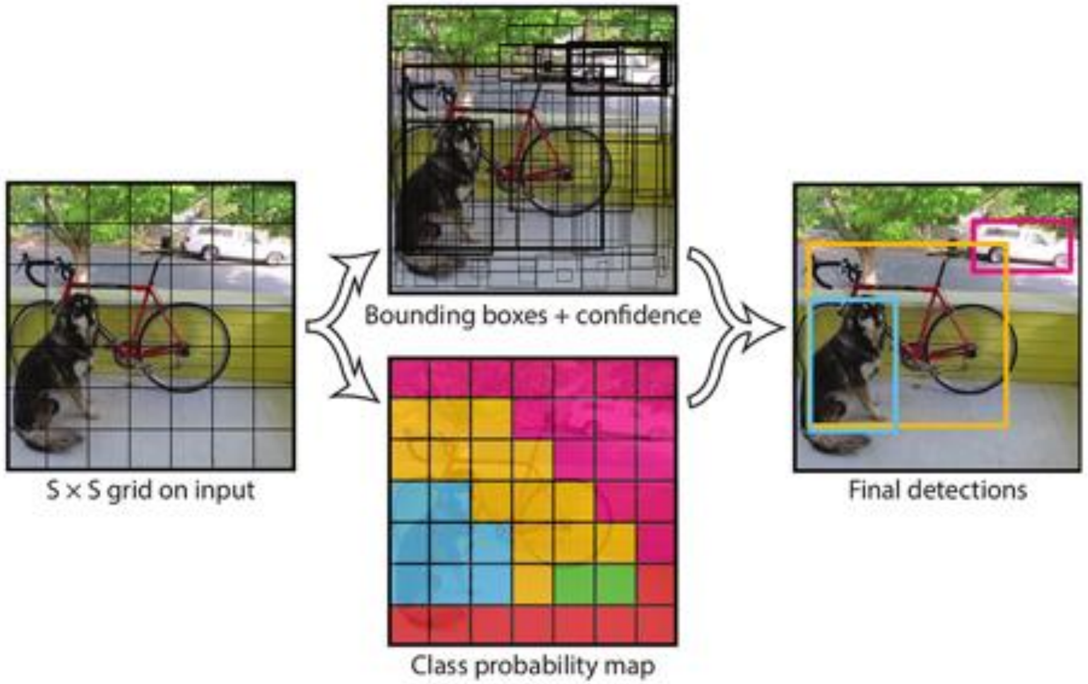
FASTER R-CNN

- **Developed by:** Shaoqing Ren et al.
- **Process:**
  - Use RPN for region proposals
  - Region of Interest (RoI) pooling layer
  - Final Classification and bounding box regression
- **Pros/Impact:** Faster than R-CNN plus they introduced Regio Proposal Network (RPN) and an end-to-end training.



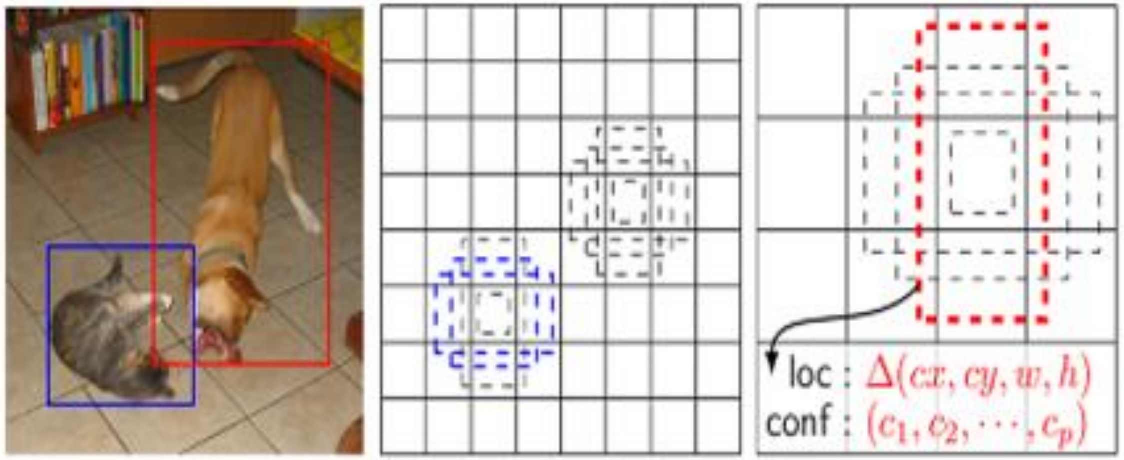
YOU ONLY LOOK ONCE (YOLO)

- **Developed by:** Joseph Redmon et al.
- **Process:**
  - Divide the input into an S x S grid
  - Each grid cell predicts B bounding boxes. Each box has coordinates (x, y, width, height) and a confidence score.
  - Each grid cell also predicts a class probability for each class
- **Pros/Impact:** Introduced a single-stage object detector, extremely fast



SINGLE SHOT MULTIBOX DETECTOR (SSD)

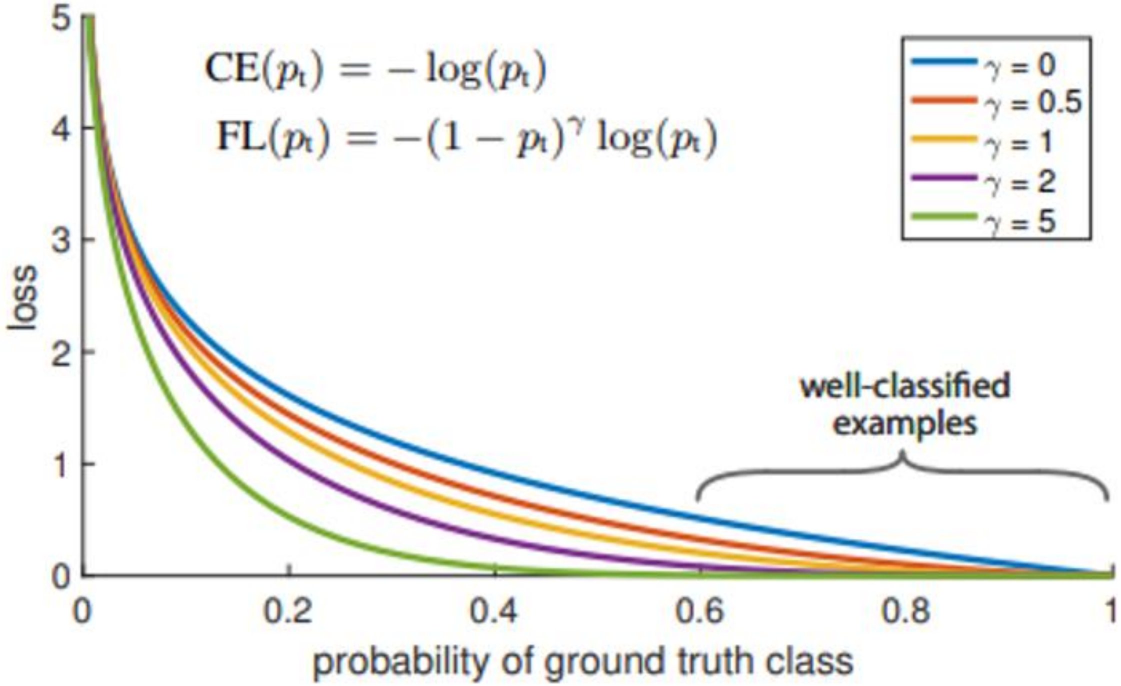
- **Developed by:** Wei Liu et al.
- **Process:**
  - Generate feature maps at multiple scales for detecting objects of different sizes
  - Predefined default boxes of various aspect ratios and scales at each feature map location
  - Predict offsets relative to default boxes, their confidence score, and class probabilities for each default box
- **Pros/Impact:** Faster than YOLOv1 and suitable for real-time detection



(a) Image with GT boxes (b) 8×8 feature map (c) 4×4 feature map

RETINANET

- **Developed by:** Tsung-Yi Lin et al.
- **Process:**
  - Focal Loss (FL)
  - Addresses class imbalance
  - Down-weights easy examples, focuses on hard examples.
- **Pros/Impact:** Feature Pyramid Network and A new loss function (Focal Loss). Good for small object detection



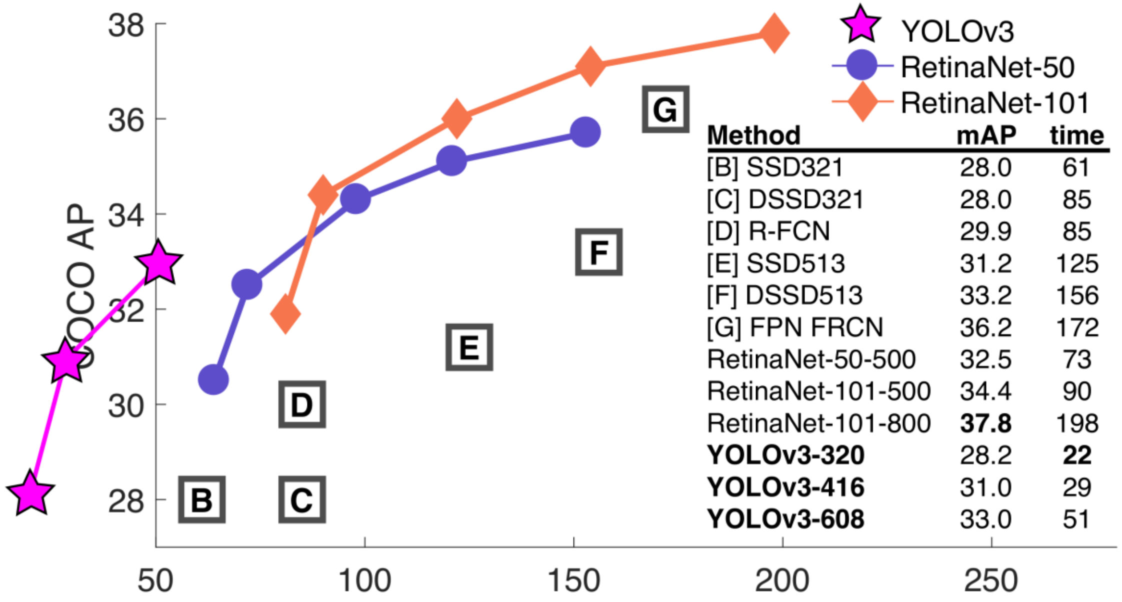
YOLO V3

- **Developed by:** Joseph Redmon et al.
- **Process:**
  - Improved over YOLOv1 and v2
  - Introduction of a multiscale prediction
  - Better performance on small objects
  - Uses Darknet-53 which competes with deeper netwoks: ResNet-101 and ResNet-152
- **Pros/Impact:** Faster than previous methods and suitable for real-time detection

	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1×	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
2×	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
8×	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
8×	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
4×	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

THE SHIFT FROM YOLO V3 TO RETINANET

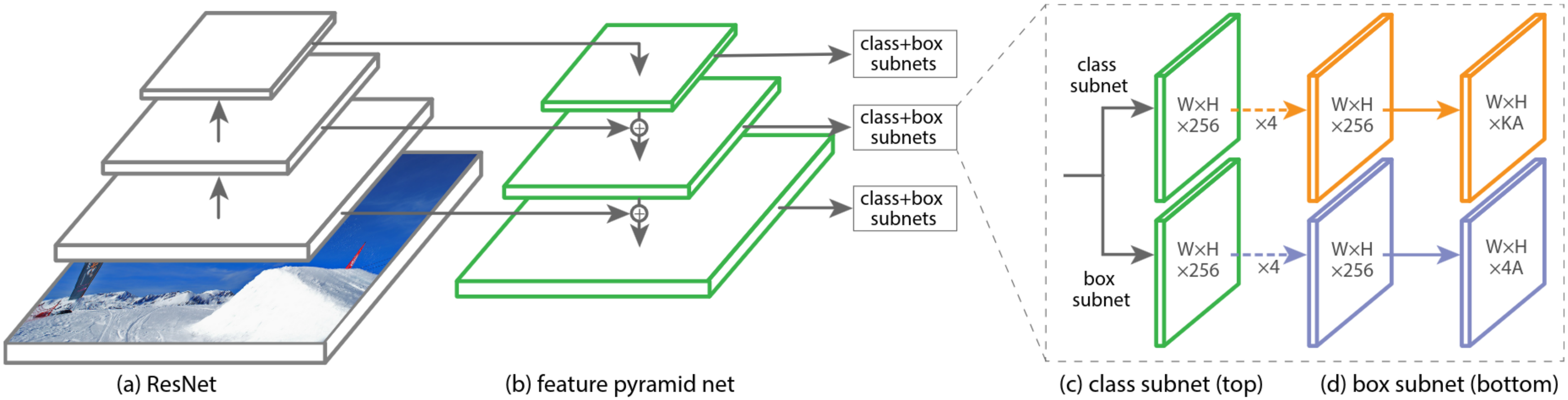
Performance Comparison



Method	mAP	time
[B] SSD321	28.0	61
[C] DSSD321	28.0	85
[D] R-FCN	29.9	85
[E] SSD513	31.2	125
[F] DSSD513	33.2	156
[G] FPN FRCN	36.2	172
RetinaNet-50-500	32.5	73
RetinaNet-101-500	34.4	90
RetinaNet-101-800	**37.8**	198
YOLOv3-320	28.2	**22**
YOLOv3-416	31.0	29
YOLOv3-608	33.0	51

Key Issues Included

- Difficulty in adapting an existing dataloader to our brain tumor dataset
- Complications in implementing the loss function
- Limited understanding of the complete object detection pipeline, hindering our ability to debug effectively



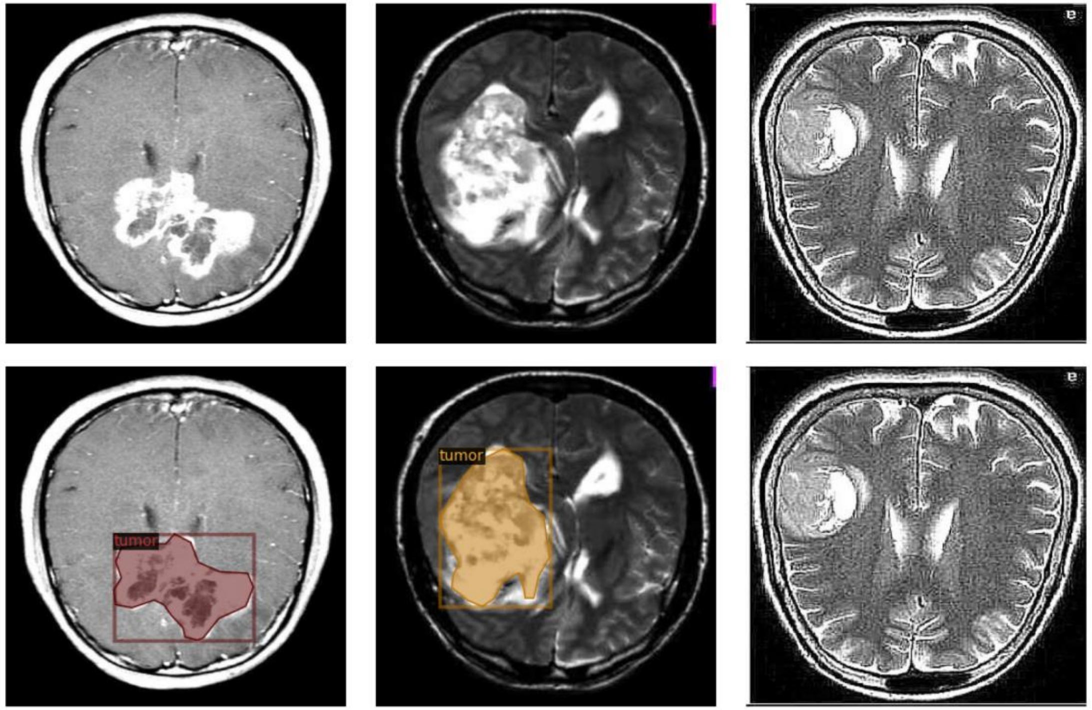
RETINANET ARCHITECTURE

$$Focal\ Loss\ (FL) \coloneqq FL(p_t) = -\alpha(1-p_t)^{\lambda}\log(p_t)$$

OBJECT DETECTION IN MEDICAL IMAGING

- Ezhilarasi et al. used Faster R-CNN for tumor detection and classification. They also apply transfer learning by fine-tuning the pre-trained AlexNet. Their study, conducted on a dataset of 50 MRI brain images, classifies tumors into four types: benign, malignant, glial, and astrocytoma
- Dipu et al. provide a comprehensive study comparing seven different object detection algorithms for brain tumor detection. The authors emphasize the importance of dataset quality, preprocessing steps, and data augmentation techniques in achieving high accuracy. Their study concludes that YOLOv5 (95.07% mAP) shows promising results for real-time brain tumor detection.
- Mercaldo et al. employ the YOLOv8 (small) object detection model for both detecting and localizing brain cancer in MRI images. Using a dataset of 300 brain MRIs, their method achieved impressive results with a precision of 0.943, recall of 0.923, and mAP of 0.941 at an IoU threshold of 0.5. These results demonstrate the effectiveness of modern object detection algorithms in medical imaging tasks.

THE DATASET



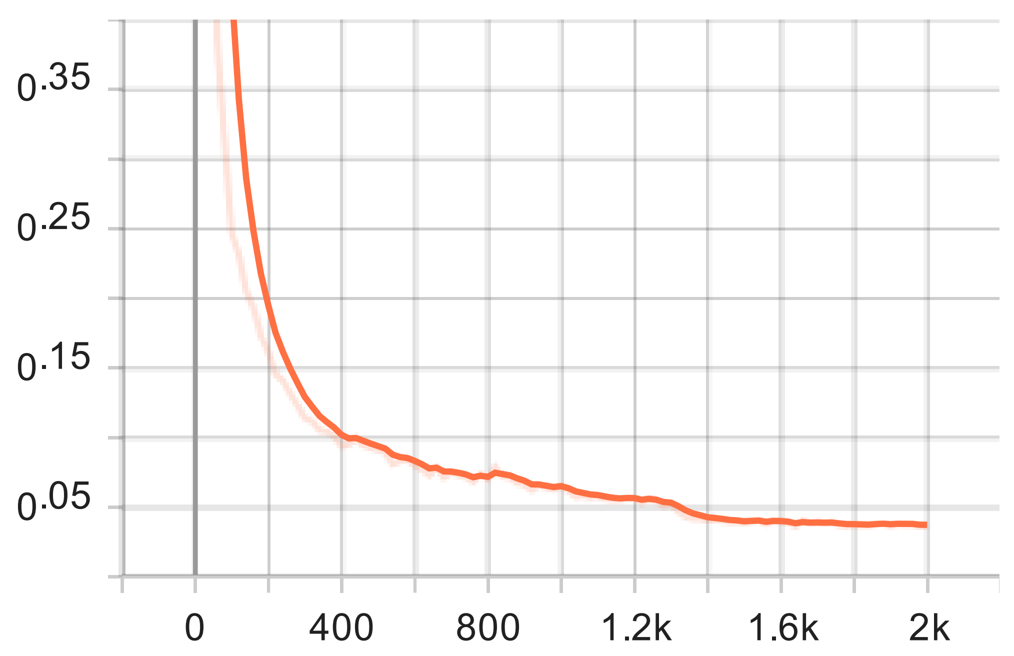
- The original dataset consists of 1,229 images
- 922 training images, 184 validation images, 123 test images
- Auto-Orient, Resize, 3 Outputs per image, Grayscale, Blur, Noise, Flip *(2766 training images)*

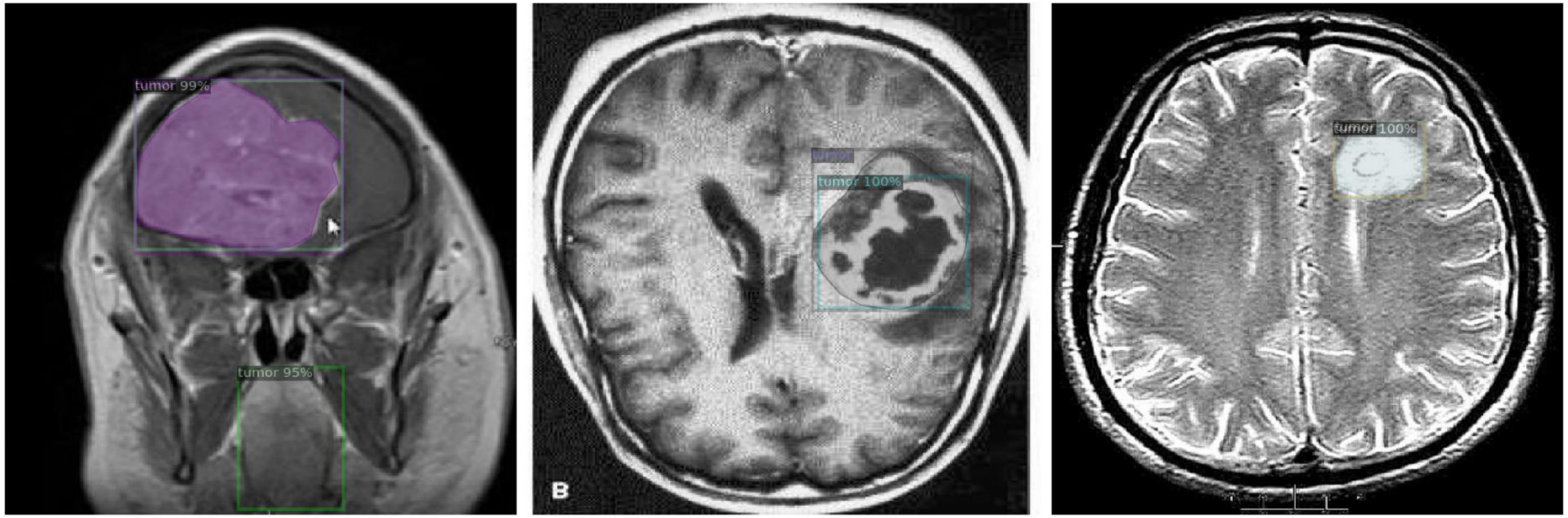
RETINANET_50_FPN_3X TRAINING

Configuration

- Google Colab A100 GPU with 40 GB RAM
- cfg.SOLVER.IMS_PER_BATCH = 64
- cfg.SOLVER.BASE_LR = 0.01
- cfg.SOLVER.MAX_ITER = 2000
- cfg.SOLVER.STEPS = (1300, 1700)
- cfg.MODEL.RETINANET.NUM_CLASSES = 2
- cfg.TEST.EVAL_PERIOD = 150
- @ IoU [0.5 : 0.95] mAP ≈ 80% Recall ≈ 86%

Loss Curve





SAMPLE TEST PREDICTIONS

(Left) The model correctly identifies a tumor but also incorrectly predicts a tumor in a non-tumor region (green, 95% confidence), illustrating a false positive case. **(Middle)** The model accurately detects the presence of a tumor, but the predicted bounding box (teal) does not fully encompass the entire extent of the tumor visible in the image. **(Right)** An example of a perfect prediction, where the model correctly identifies and localizes the tumor with a single bounding box.In other cases, two overlapping bounding boxes cover the tumor region, indicating that the NMS did not fully suppress redundant detections

CONCLUSION

Limitation of current approach:

- Imbalanced representation in the training data
- Challenges in distinguishing medium tumors from surrounding tissues
- Potential bias towards larger tumors
- Limited testing on diverse MRI sequences
- Runtime shutdonws and reduced training iterations due to limited compute units on Google Colab

Future work could focus on:

- Fine-tuning the model to improve detection of medium-sized tumors
- Exploring data augmentation techniques to address size imbalance
- Testing on a more diverse dataset to ensure generalizability
- Investigating the integration of 3D information from MRI sequences

THANKS FOR YOUR ATTENTION

QUESTIONS & ANSWERS