Recent Advances in Computer Vision

Sanghyuck Lee

tkdgur658@gmail.com

16th Nov, 2022

Contents

1. 개인 연구 (Recent Advances in Computer Vision)
2. TAO 연구 (CAU Hospital & NVIDIA)
3. Orbit structures segmentation 논문 작업
4. 졸업 논문 SCIE 투고 작업

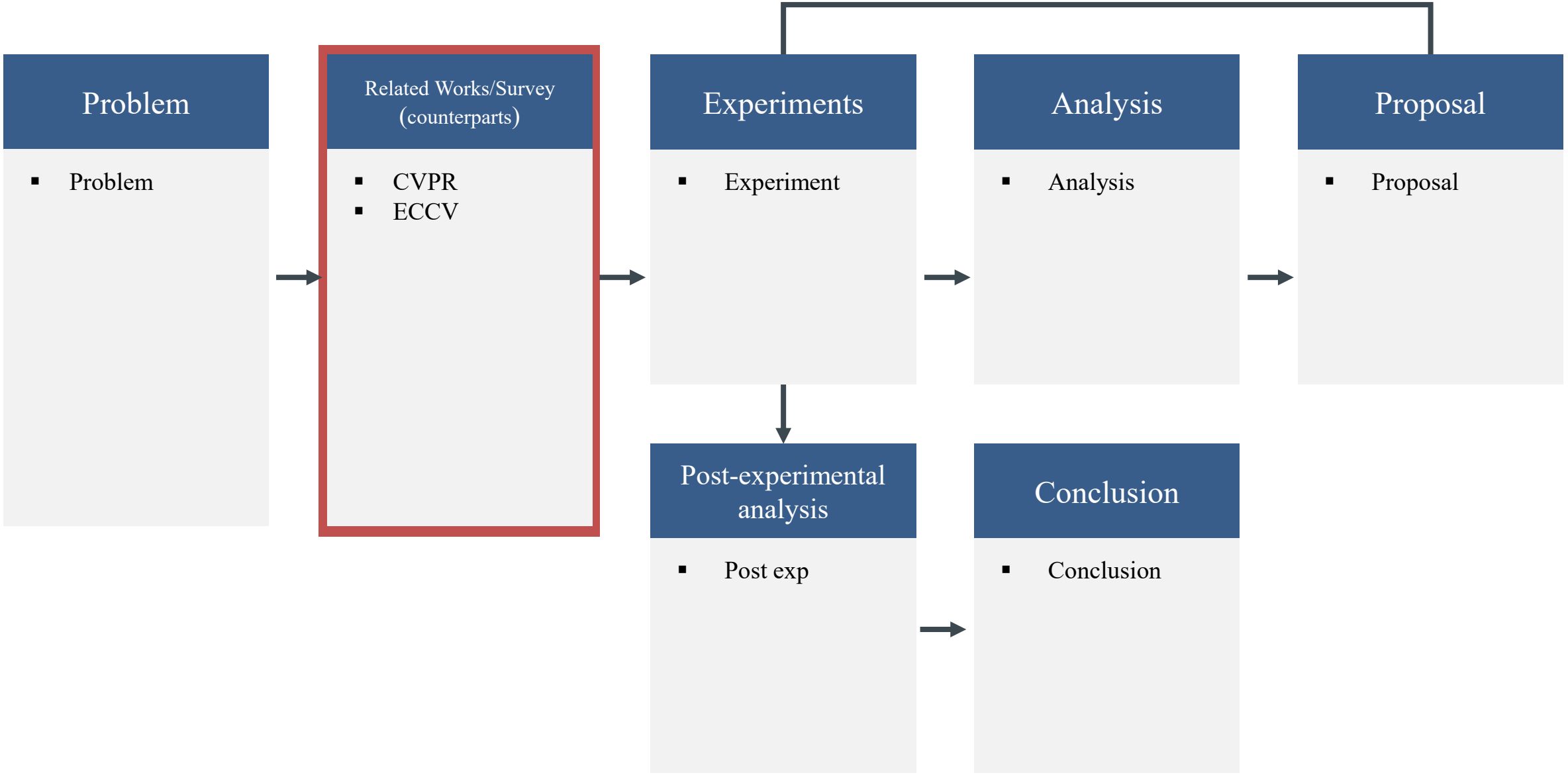
To Do List from Previous Seminar

22.11.16

- 개인 연구 : 논문 읽기 진행
- ~~SW중심대학지원사업 : 결과보고서 제출~~
- TAO 연구 : Activity 논문 및 ~~Label 수신 대기~~ => **수신 완료**
- ~~중견 연구 제안서 : 온라인 설명회 시청, 제안서 작성~~
- 졸업논문 : 후속 절차 완료 => **+SCIE 투고**

Overall Progress

Paper submission deadline
CVPR 2024 : TBD (대략 '23.11. 중순)
ECCV 2024 : TBD (대략 '24.03. 초)



Related Work

CVPR 2022's research categories

MENU	Select Primary Subject Area	# papers	Oral	Poster
1	Recognition: detection, categorization, retrieval	177	25	152
2	Image and video synthesis and generation	157	26	131
3	3D from multi-view and sensors	137	26	111
4	Low-level vision	110	19	91
5	Vision + language	105	20	85
6	Segmentation, grouping and shape analysis	99	16	83
7	Transfer/ low-shot/ long-tail learning	86	15	71
8	Deep learning architectures and techniques	85	20	65
9	Self-& semi-& meta- & unsupervised learning	84	7	77
10	Video analysis and understanding	77	15	62
11	Pose estimation and tracking	62	14	48
12	Representation learning	61	11	50
13	3D from single images	60	10	50
14	Scene analysis and understanding	56	9	47
15	Face and gestures	54	7	47
16	Computational photography	53	10	43
17	Motion and tracking	53	8	45
18	Adversarial attack and defense	52	10	42
19	Datasets and evaluation	52	7	45
20	Machine learning	41	7	34
21	Action and event recognition	40	8	32
22	Efficient learning and inferences	40	3	37
23	Medical, biological and cell microscopy	37	5	32
24	Vision applications and systems	32	4	28
25	Navigation and autonomous driving	31	2	29
26	Vision + graphics	24	6	18
27	Privacy and federated learning	21	3	18
28	Vision + X	20	4	16
29	Physics-based vision and shape-from-X	16	2	14
30	Robot vision	16	3	
31	Explainable computer vision	15	2	13
32	Demo	15		15
33	Optimization methods	14	2	12
34	Transparency, fairness, accountability, privacy and ..	14	4	
35	Document analysis and understanding	12		12
36	Biometrics	11	2	9

Source: https://blog.roboflow.com/cvpr-2022-best-papers-and-highlights/

Top 5 Google scholar citation papers for each CVPR & ECCV 2020~2022

Conf.	GS Citations	Title	Authors
ECCV 2022	143	Bytetrack: Multi-object tracking by associating every detection box	Zhang et al.
ECCV 2022	80	Motr: End-to-end multiple-object tracking with transformer	Zeng et al.
ECCV 2022	60	Slip: Self-supervision meets language-image pre-training	Mu et al.
ECCV 2022	42	Detecting twenty-thousand classes using image-level supervision	Zhou et al.
ECCV 2022	36	Vqgan-clip: Open domain image generation and editing with natural language guidance	Crowson et al.
CVPR 2022	854	Resnest: Split-attention networks	Zhang et al.
CVPR 2022	761	Masked autoencoders are scalable vision learners	He et al.
CVPR 2022	451	A convnet for the 2020s	Liu et al.
CVPR 2022	293	Video swin transformer	Liu et al.
CVPR 2022	263	Scaling vision transformers	Zhai et al.
CVPR 2021	1355	Exploring simple siamese representation learning	Chen et al.
CVPR 2021	973	Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers	Zheng et al.
CVPR 2021	575	Coordinate attention for efficient mobile network design	Hou et al.
CVPR 2021	568	Scaled-yolov4: Scaling cross stage partial network	Wang et al.
CVPR 2021	559	Pre-trained image processing transformer	Chen et al.
ECCV 2020	3954	End-to-end object detection with transformers	Carion et al.
ECCV 2020	1350	Contrastive multiview coding	Tian et al.
ECCV 2020	812	Oscar: Object-Semantics Aligned Pre-training for Vision-Language Tasks	Li et al.
ECCV 2020	799	Uniter: Universal image-text representation learning	Chen et al.
ECCV 2020	710	Raft: Recurrent all-pairs field transforms for optical flow	Teed et al.
CVPR 2020	5060	Momentum contrast for unsupervised visual representation learning	He et al.
CVPR 2020	2620	Efficientdet: Scalable and efficient object detection	Tan et al.
CVPR 2020	2605	Analyzing and improving the image quality of stylegan	Karras et al.
CVPR 2020	2038	nuscenes: A multimodal dataset for autonomous driving	Caesar et al.
CVPR 2020	1513	CSPNet: A new backbone that can enhance learning capability of CNN	Wang et al.

TAO 연구 (CAU Hospital & NVIDIA)

Activity & Treatment

- Activity: 논문 진행 대기중
- Treatment
  - CT & Label 데이터 수신 완료
  - Pilot 실험 진행 예정

Orbit structures segmentation 논문 작업

연구실 기존 논문(Chae et al., 2021) 구조 파악 및 모방

- 파악
  - ✓ 기존 음악 태그 분류 연구들이 다양한 Dataset, Feature Extractor, Feature Selection, Classifier 사용으로 공정한 비교가 어려우니 이들을 잘 일치시킨후, Feature Selection 방식에 대해서 비교를 진행함.
- 모방
  - ✓ 현 Orbit structures segmentation 연구의 논문 작업 Material들에 Proposed method가 포함되어 있지 않으니, **연구 Survey 쪽으로 확장하여 논문 작업 진행**함.
  - ✓ **다양한 Segmentation 연구들을 추가적으로 조사**하고, 이들을 잘 그루핑하여 **현 실험 및 분석 결과를 잘 사용할 수 있도록** 논문을 작성함.
- 계획
  - ✓ IEEE Access, MTAP 투고 예정 => MTAP의 저자 가이드라인 읽는 중
  - ✓ 최신 알고리즘을 비교 실험에 조금 추가하면 어떨지

Chae, Jonghoon, et al. "Toward a fair evaluation and analysis of feature selection for music tag classification." *IEEE Access* 9 (2021): 147717-147731.

졸업 논문 SCIE 투고 작업

내용 보강

- 비교 모델이었던 TransMed에 Proposed Idea인 Group SE Block을 추가하는 방식으로 실험 예정
  - ✓ 최신 알고리즘에 아이디어를 추가하는 것이니, 리뷰어들이 더 신뢰하면서 긍정적으로 볼 것으로 기대함.

To Do List

- 개인 연구 : 논문 읽기 진행
- TAO 연구 : Treatment Pilot 실험 진행
- Orbit structures segmentation 논문 작업 : 작업 진행
- 졸업논문 SCIE 투고 작업 : 보강

Thank you

Eyeball Volume Estimation

Research Briefing

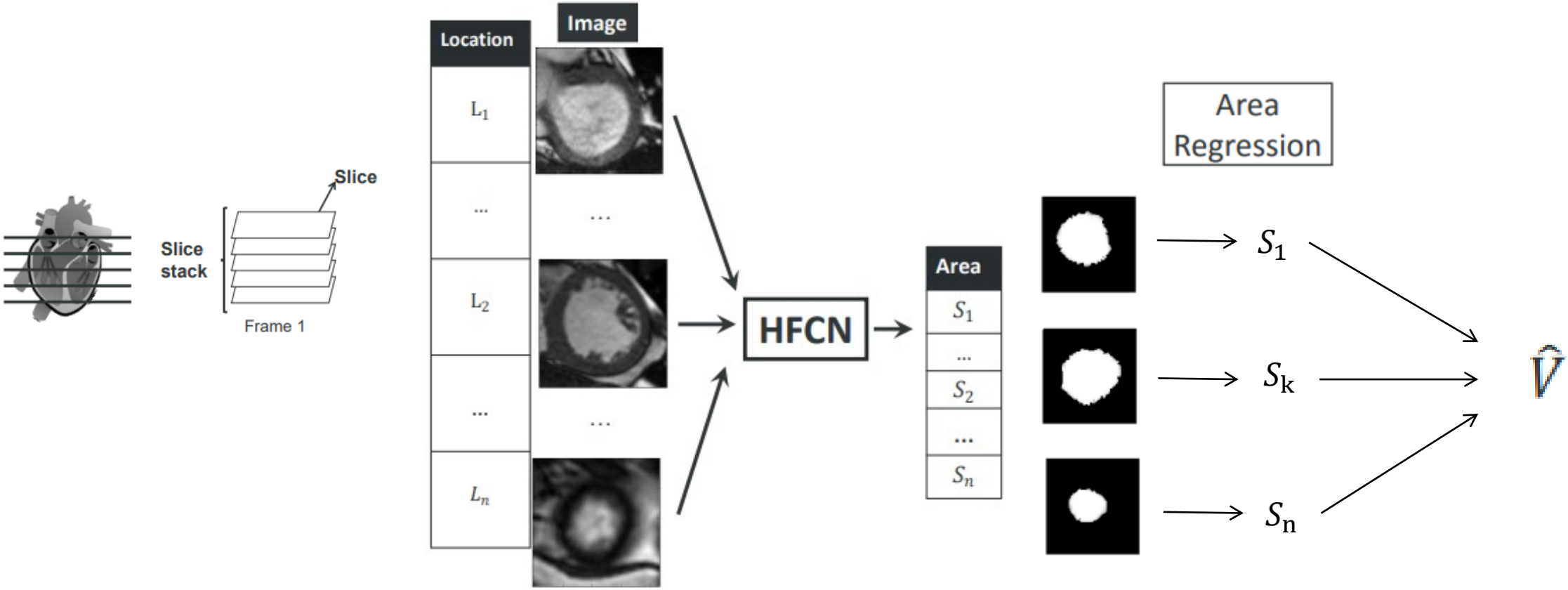
Sujeong Han

0001entropy@gmail.com

14th Dec, 2022

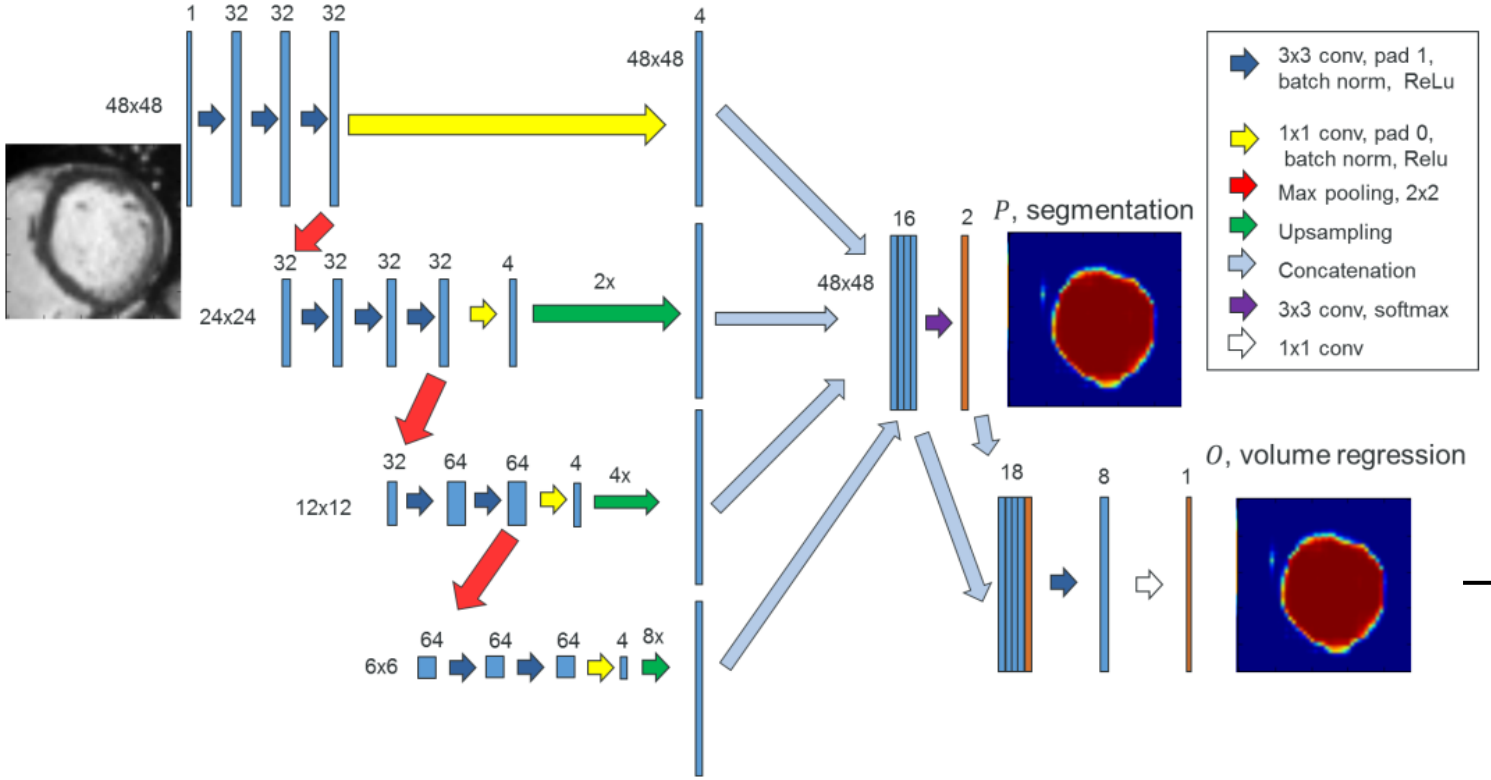
Related Work

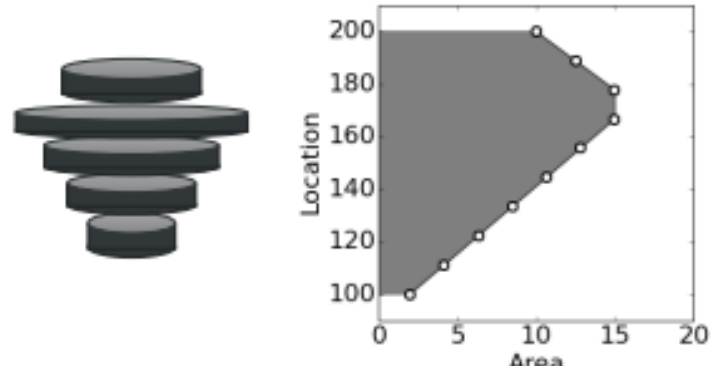
Estimating the volume of the left ventricle from MRI images using deep neural networks (IEEE, 2019)



Related Work

Estimating the volume of the left ventricle from MRI images using deep neural networks (IEEE, 2019)





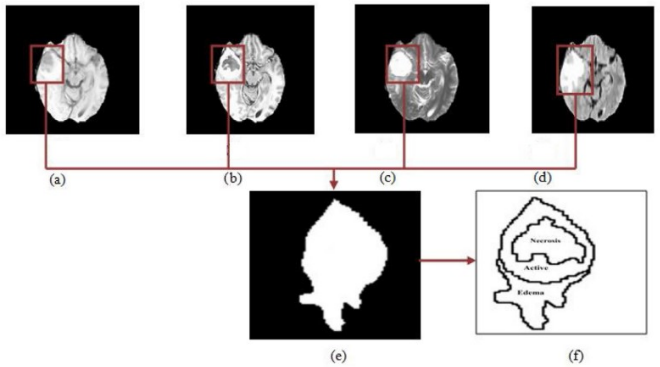
$$\hat{V} = \sum (S_i + S_{i+1}) \times (L_{i+1} - L_i)/2 \rightarrow Loss = |V - \hat{V}|$$

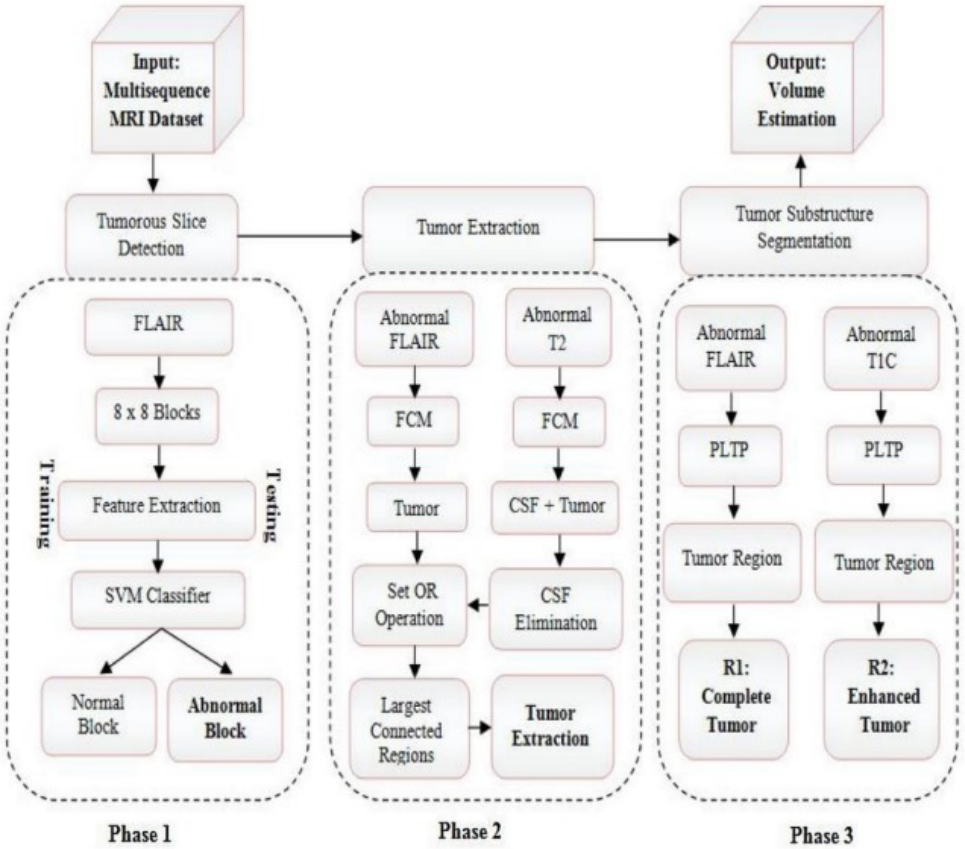
Volume

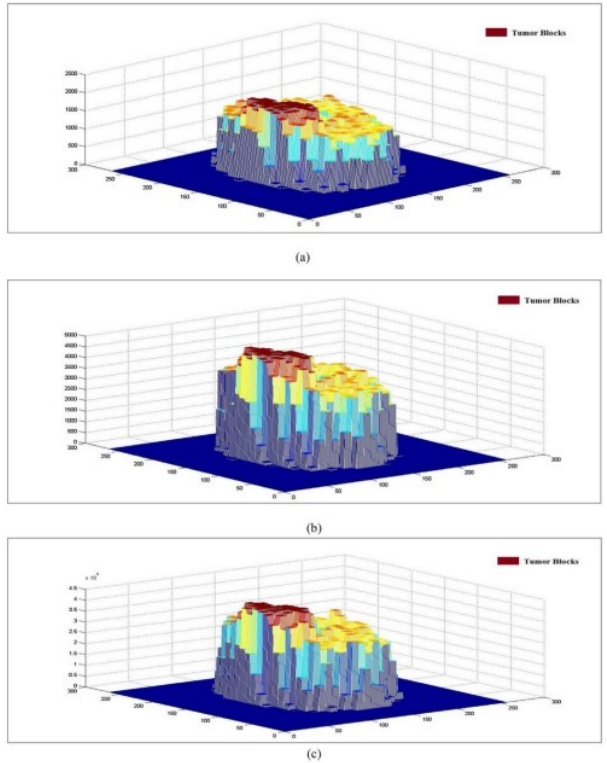
V

Related Work

Modified local ternary patterns technique for braintumour segmentation and volume estimation from MRI multi-sequence scans with GPU CUDA machine (Biocybernetics and Biomedical Engineering, 2019)

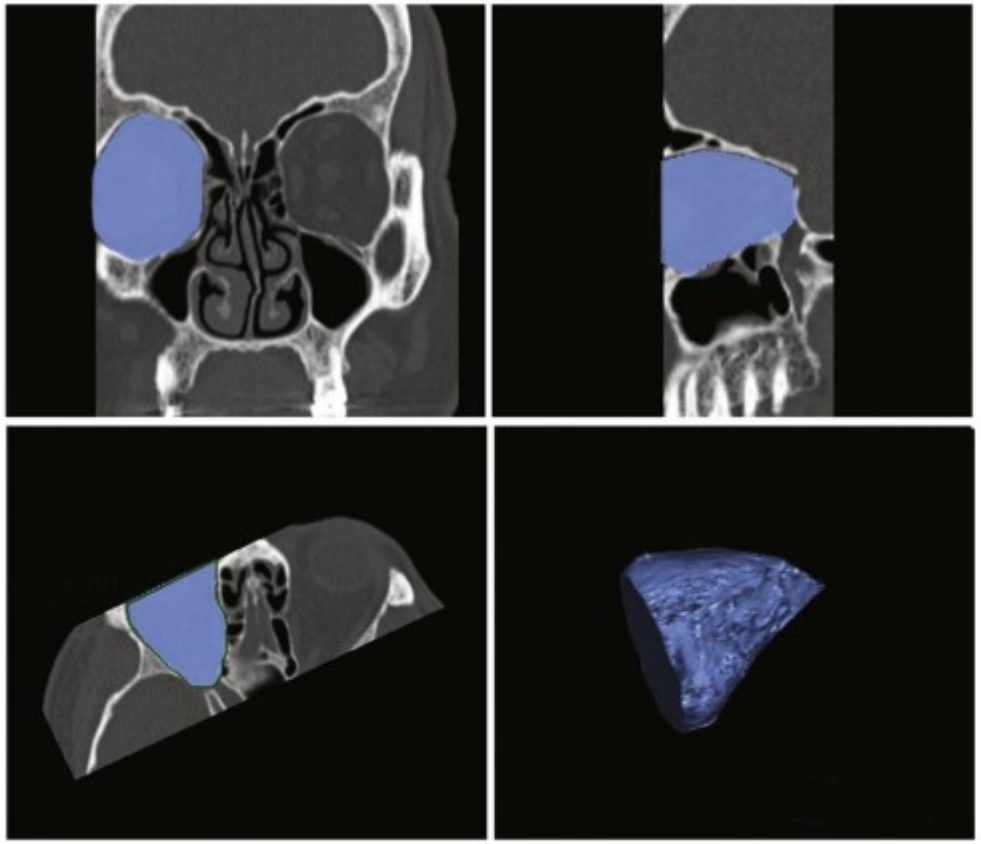


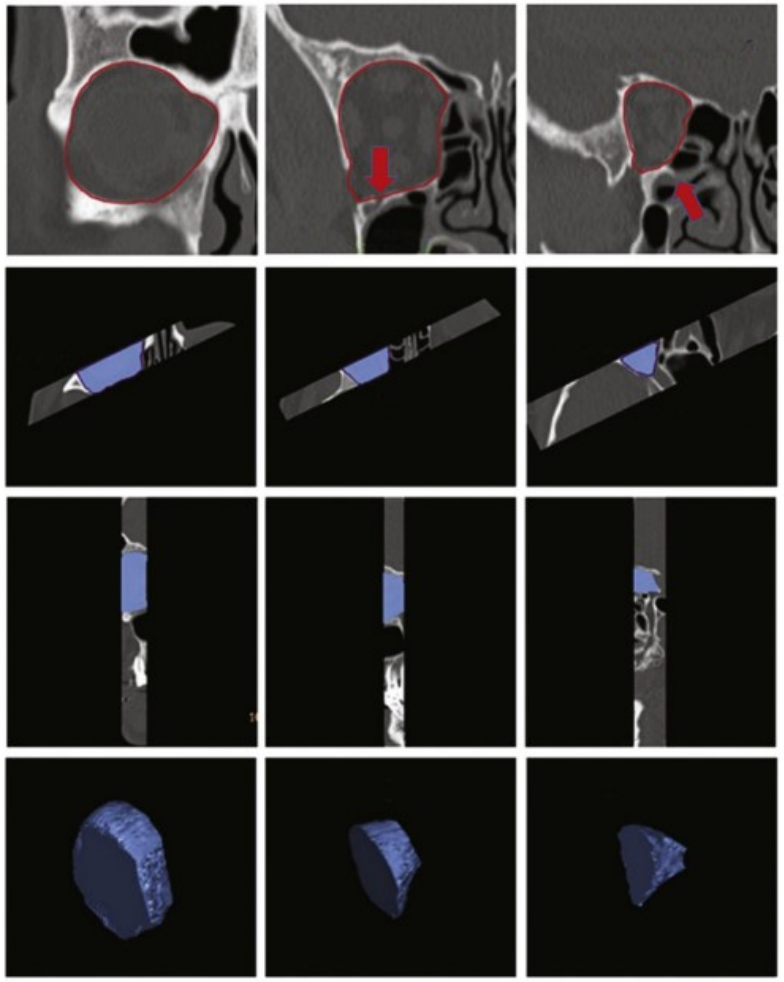




Related Work

Characterization of the orbital volume in normal population (Journal of cranio-maxillofacial surgery, 2018)





Thank you

Individual Study on Natural Language Processing

Chaelyn Lee

mylynchae@gmail.com

10th Nov, 2022

Contents

1. To Do List from Previous Seminar
2. Individual study
3. To Do List

To Do List from Previous Seminar

- ✓ Paper survey for selection of research topics
- ✓ Proceed in the order of finding the task first and studying the model in depth.

individual study

Natural Language Processing Task

- Classification (문장 분류)
- Sentiment Analysis (감정 분석)
- Summarization (요약)
- Machine Translation (기계 번역)
- Question Answering (질문 응답)
- 그 외에 정보 추출, 용어 추출, 주제 모델링 등

individual study

Sentiment Analysis ? Emotion Recognition ?

- machine learning + natural language processing
- Many studies work on classifying sentences as positive, neutral, or negative.

  -> emotion recognition in text (mainly many studies on speech emotion recognition)
- SA - I think the event was **boring** (*boring:* -1 [SenticNet], negative 1, positive 0 [NRC])
- ED - I am **disappointed** because I think I have wasted time. (*disappointed:* anger 1, anticipation 0, disgust 1, fear 0, joy 0, sadness 1, surprise 0, trust 0 [NRC])

individual study

- Reading a survey paper on text sentiment analysis

  (Sentiment analysis using deep learning architectures: a review, *Artificial Intelligence Review* 53.6 (2020))

  (Sentiment analysis based on deep learning: A comparative study., *Electronics* 9.3 (2020))

  (A survey of state-of-the-art approaches for emotion recognition in text, Knowledge and Information Systems (2020))

individual study

- Studying books related to GPT -> Curious about zero-shot running
- --> Zero-shot Text Classification via Reinforced Self-training (acl, 2020)

To Do List

- ✓ Decide on the task as quickly as possible
- ✓ Read task-related survey papers

Thank you

Appendix

개인 연구 리뷰

Goeun Kim

goeun678@gmail.com

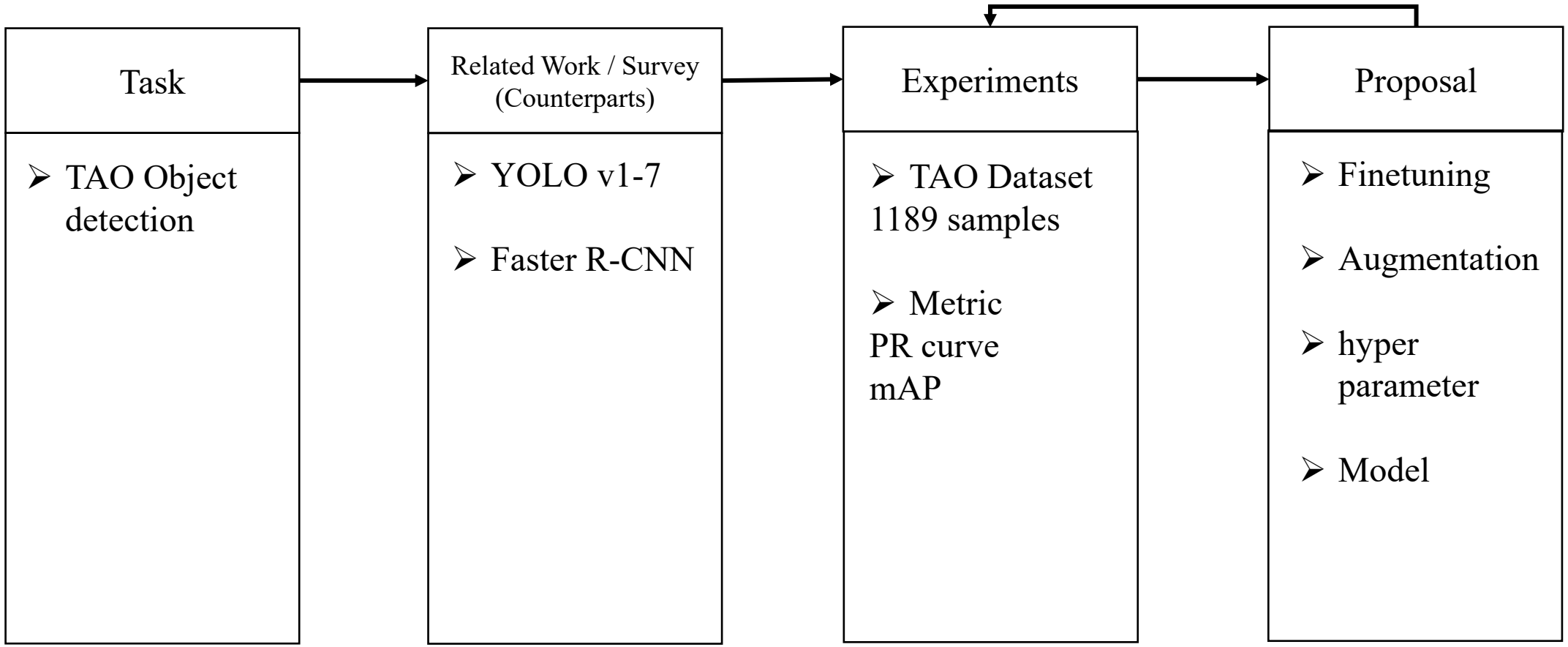
14th Dec, 2022

Contents

TAO Object Detection

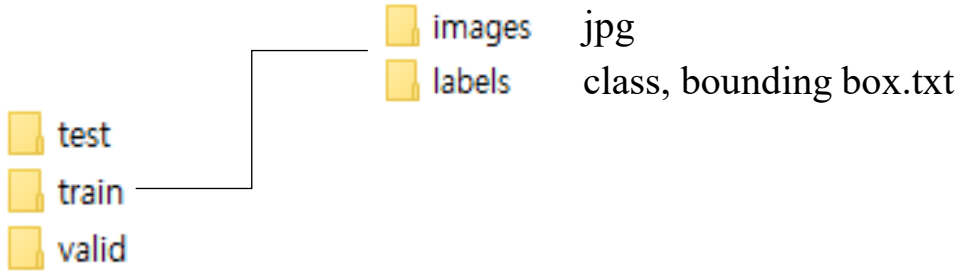
1. Overall Progress
2. Introduction
3. Experiment
4. Conclusion

Overall Progress



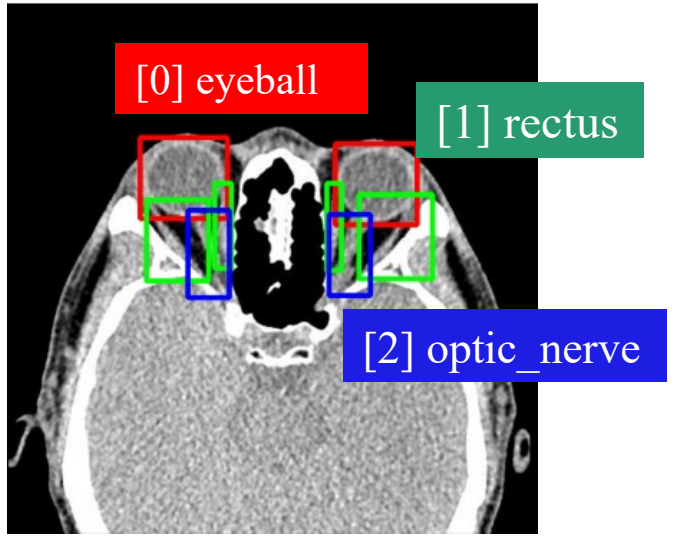
TAO **t**hyroid **a**ssociated **o**rbitopathy (or thyroid-associated ophthalmopathy)

- ✓ 일명 그레이브스 병이라고 불리는 갑상선 관련 안와병증(TAO)
- ✓ 안와의 주변 조직, 갑상선, 드물게 전전공 피부나 손가락에 영향을 줄 수 있는 질병으로 자가면역과정의 일부.



< forder structure >

< case 1: CT axial >

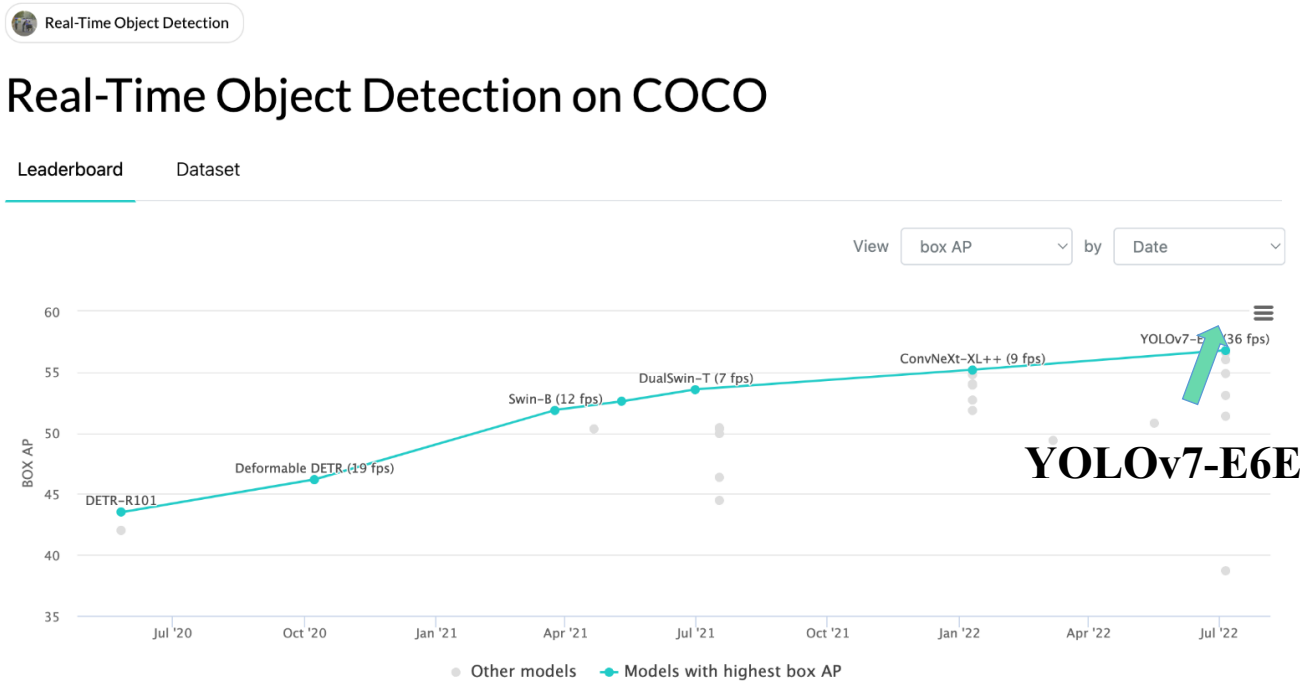


< case 1: CT axial with label >

Introduction

Paper : https://arxiv.org/pdf/2207.02696.pdf
GitHub : https://github.com/WongKinYiu/yolov7

YOLOv7: Trainble bag-of-freebies sets new state-of-the-art for real-time object detectors



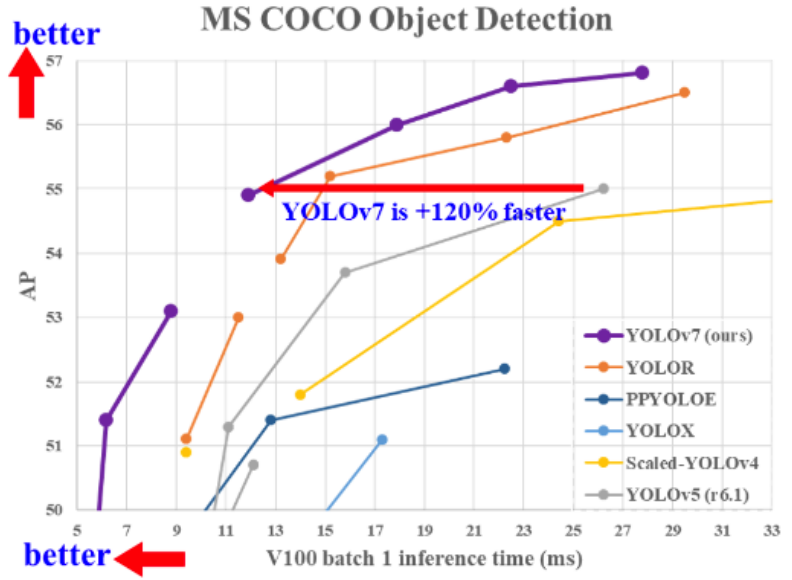


Figure 1: Comparison with other real-time object detectors, our proposed methods achieve state-of-the-arts performance.

Experiment

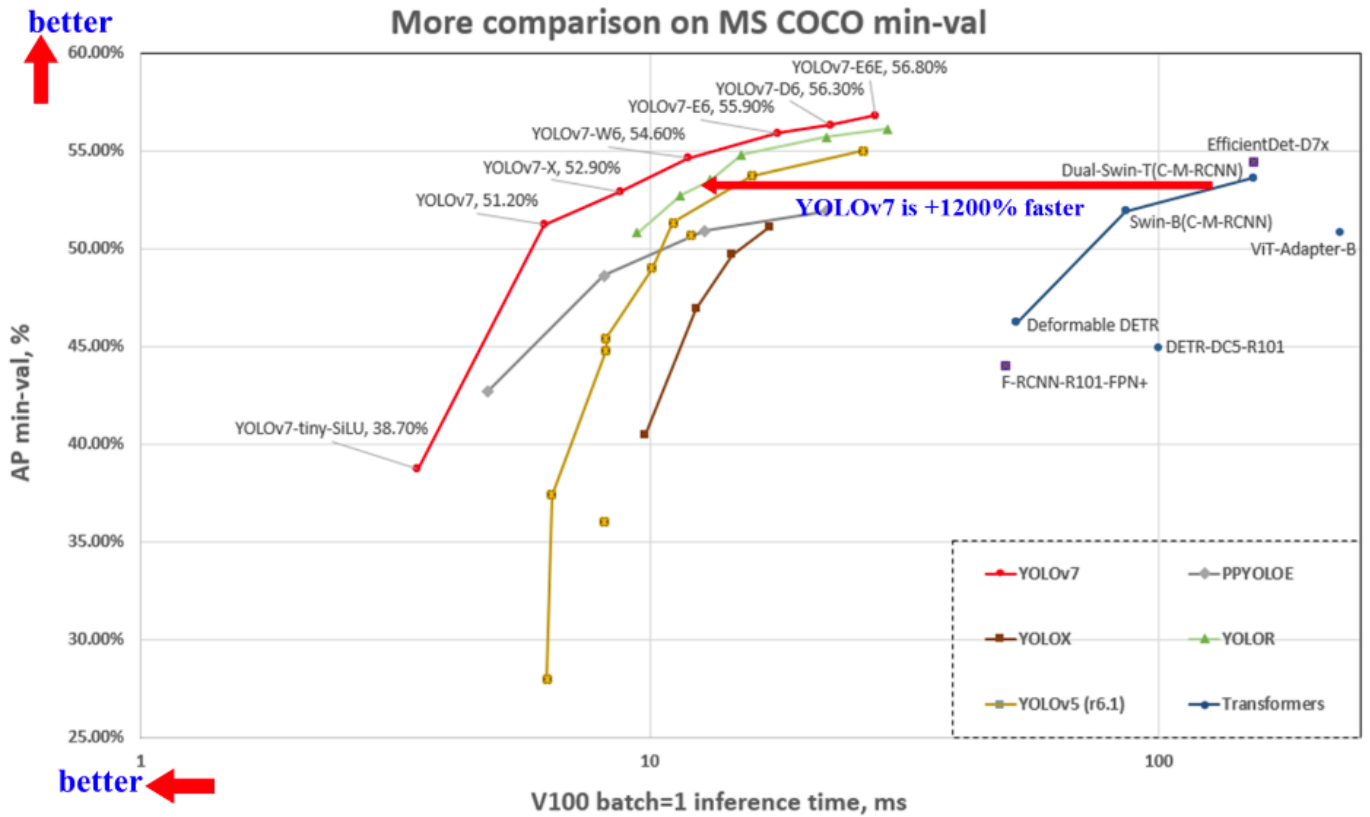
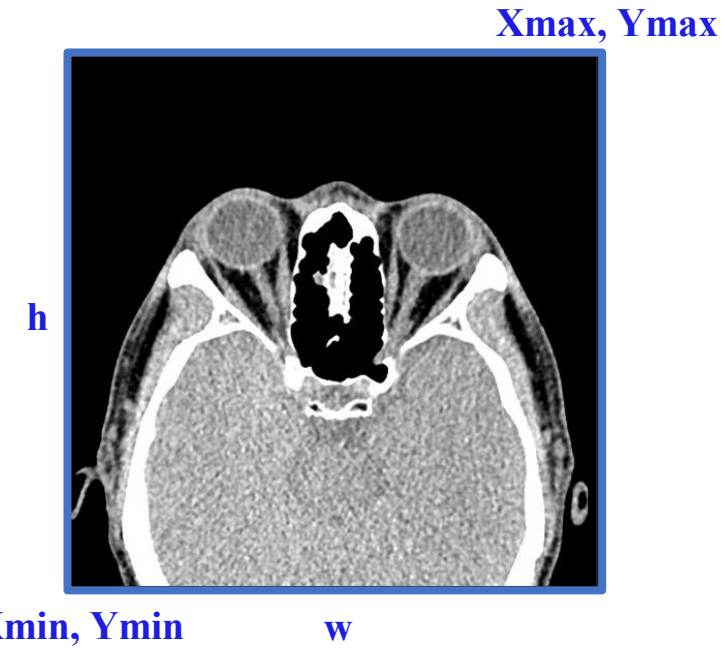


Figure 9: Comparison with other object detectors.

Preprocessing

- ✓ Data format : jpg
- ✓ Input Size : 512 x512 ➞ 640 x 640
- ✓ Annotation Normalization



Normalization Formula

- ✓ Normalized(Xmin) = (Xmin+w/2)/Image_Width
- ✓ Normalized(Ymin) = (Ymin+h/2)/Image_Height
- ✓ Normalized(w) = w/Image_Width
- ✓ Normalized(h) = h/Image_Height

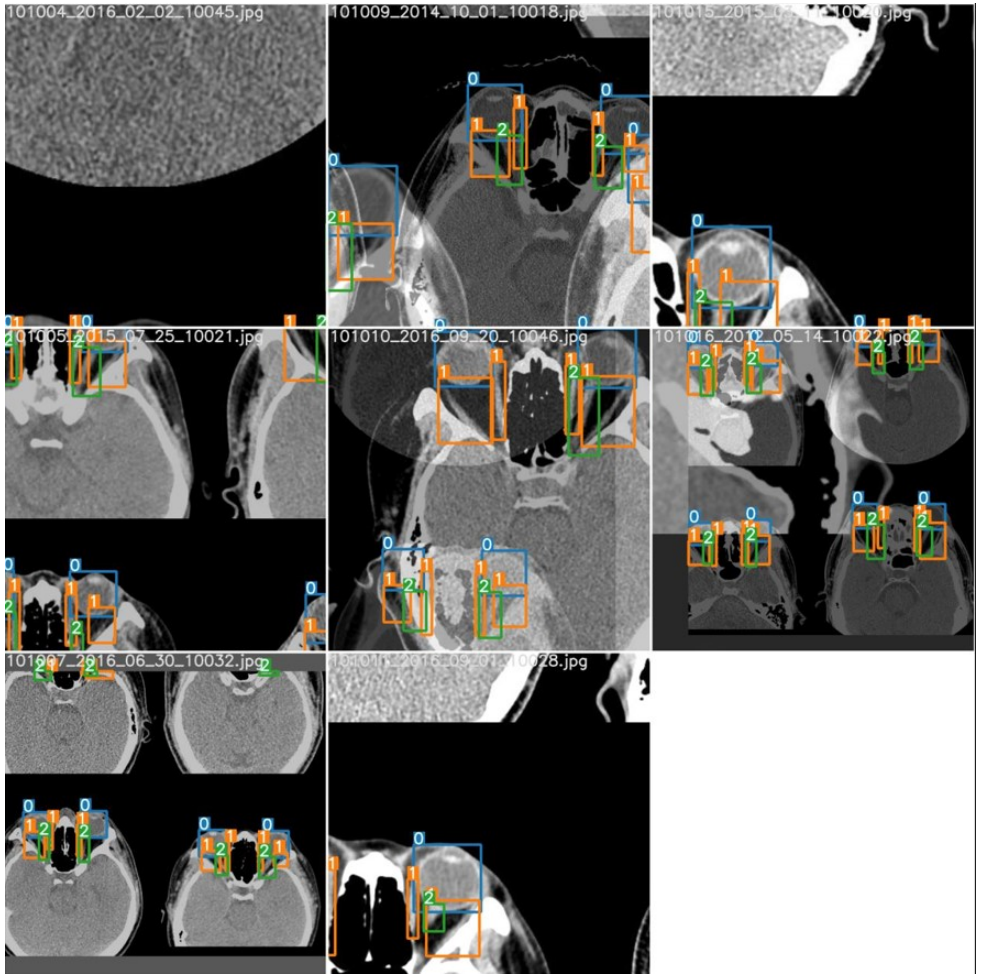
label index, N(Xmin), N(Ymin), N(w), N(h)

```
0 0.33203125 0.330078125 0.1640625 0.15234375
0 0.6943359375 0.341796875 0.154296875 0.15234375
1 0.40625 0.419921875 0.03125 0.16015625
1 0.615234375 0.421875 0.02734375 0.1640625
1 0.3203125 0.447265625 0.1171875 0.15234375
1 0.7314453125 0.4404296875 0.138671875 0.158203125
2 0.37890625 0.47265625 0.078125 0.1640625
2 0.64453125 0.474609375 0.078125 0.15234375
```

Augmentation

- ✓ hsv_h : 0.015 image HSV-Hue augmentation (fraction)
- ✓ hsv_s : 0.7 image HSV-Saturation augmentation (fraction)
- ✓ hsv_v: 0.4 image HSV-Value augmentation (fraction)
- ✓ translate : 0.2 image translation (+/- fraction)
- ✓ scale : 0.9 image scale (+/- gain)
- ✓ fliplr : 0.5 image flip left-right (probability)
- ✓ mosaic : 1.0 image mixup (probability)
- ✓ mixup : 0.15 image mixup (probability)
- ✓ paste_in : 0.15 segment copy-paste (probability)

< training batch images >



Experiment Setting

✓ Datasets (TAO)

Number of Samples: 1,189

train: 832

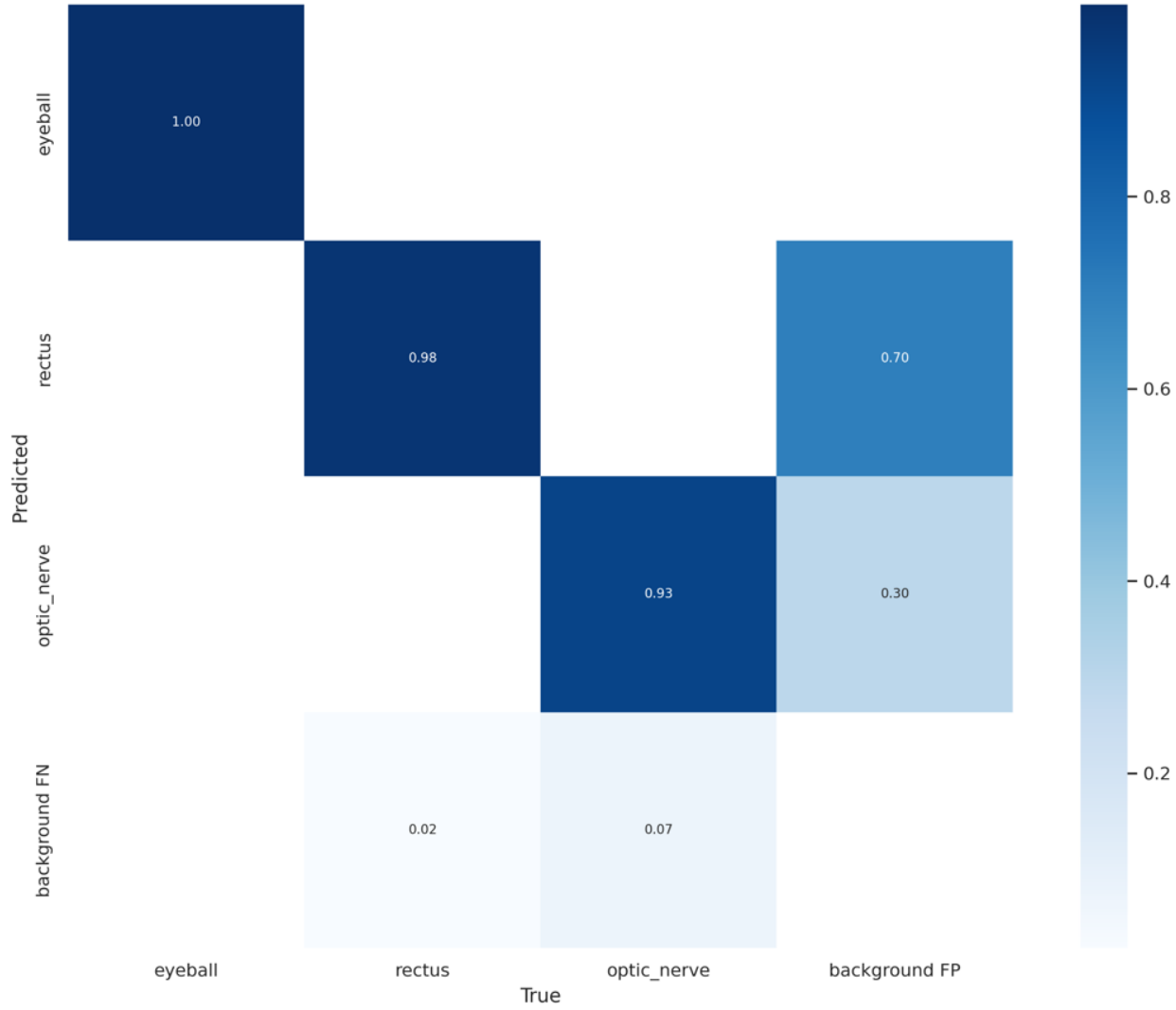
valid: 178

test: 179

✓ Hyper Parameter Setting

Pretrained weight	Model	Loss function	Optimizer	Epochs	Batch_size
Imagenet	YOLOv7	binary CE	SGD	50	8

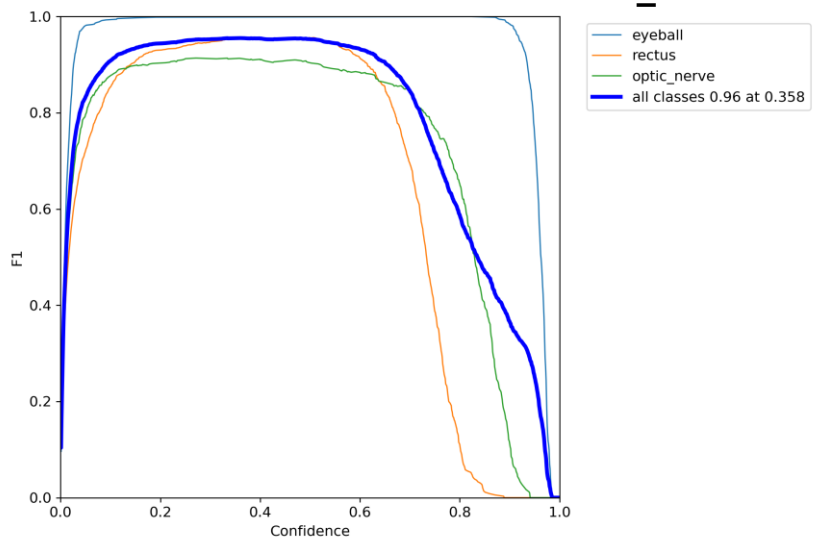
Experimental Results



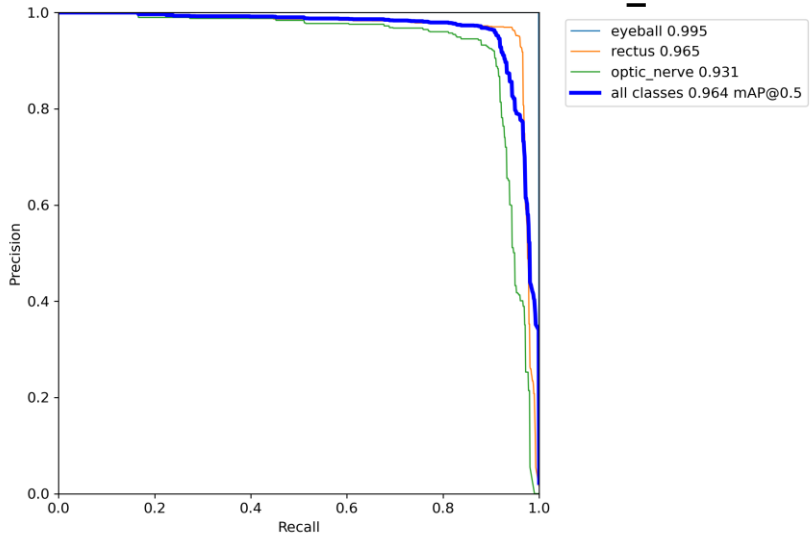
Confusion matrix

Experimental Results

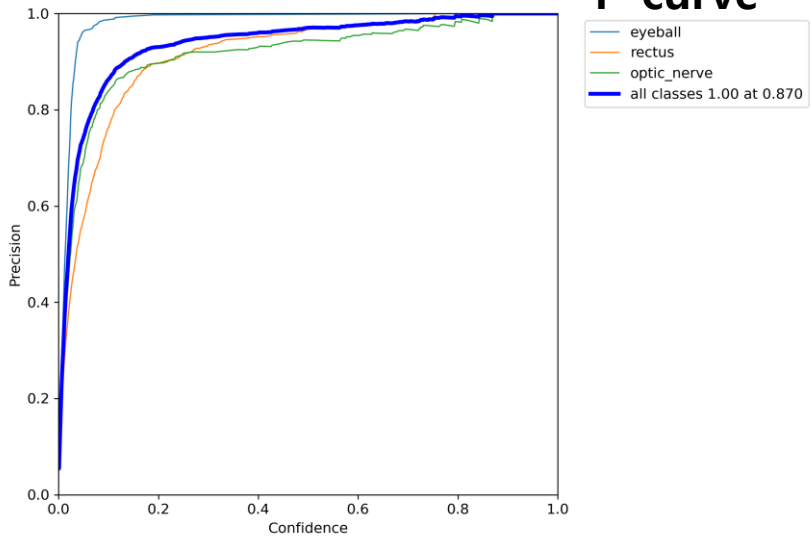
F1_curve



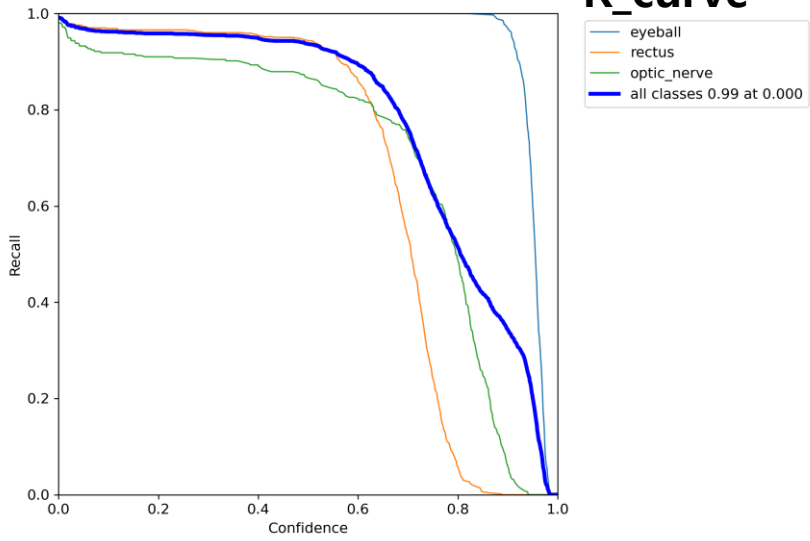
PR_curve



P-curve

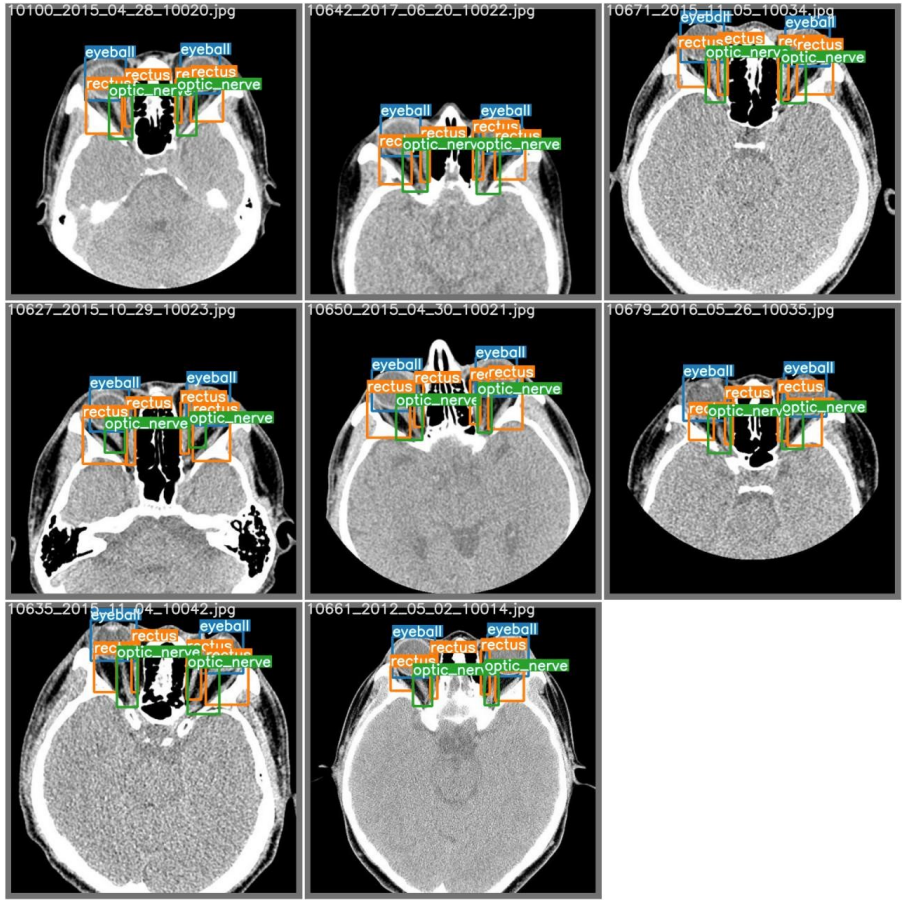


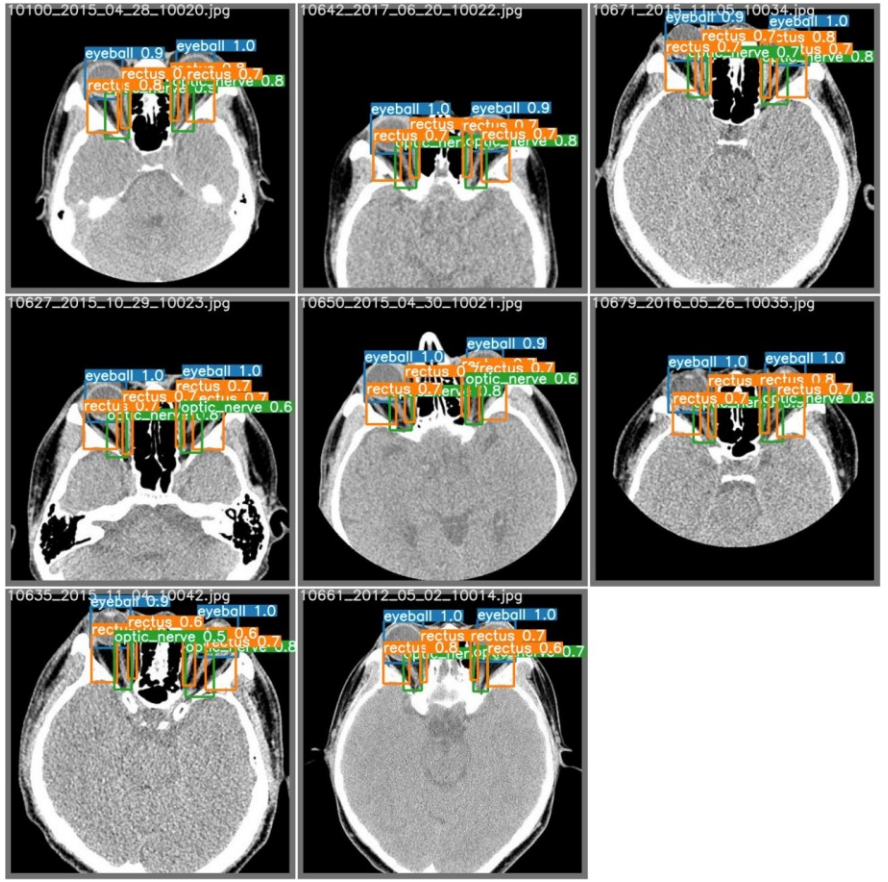
R_curve



Experimental Results

labels & pred





Proposal

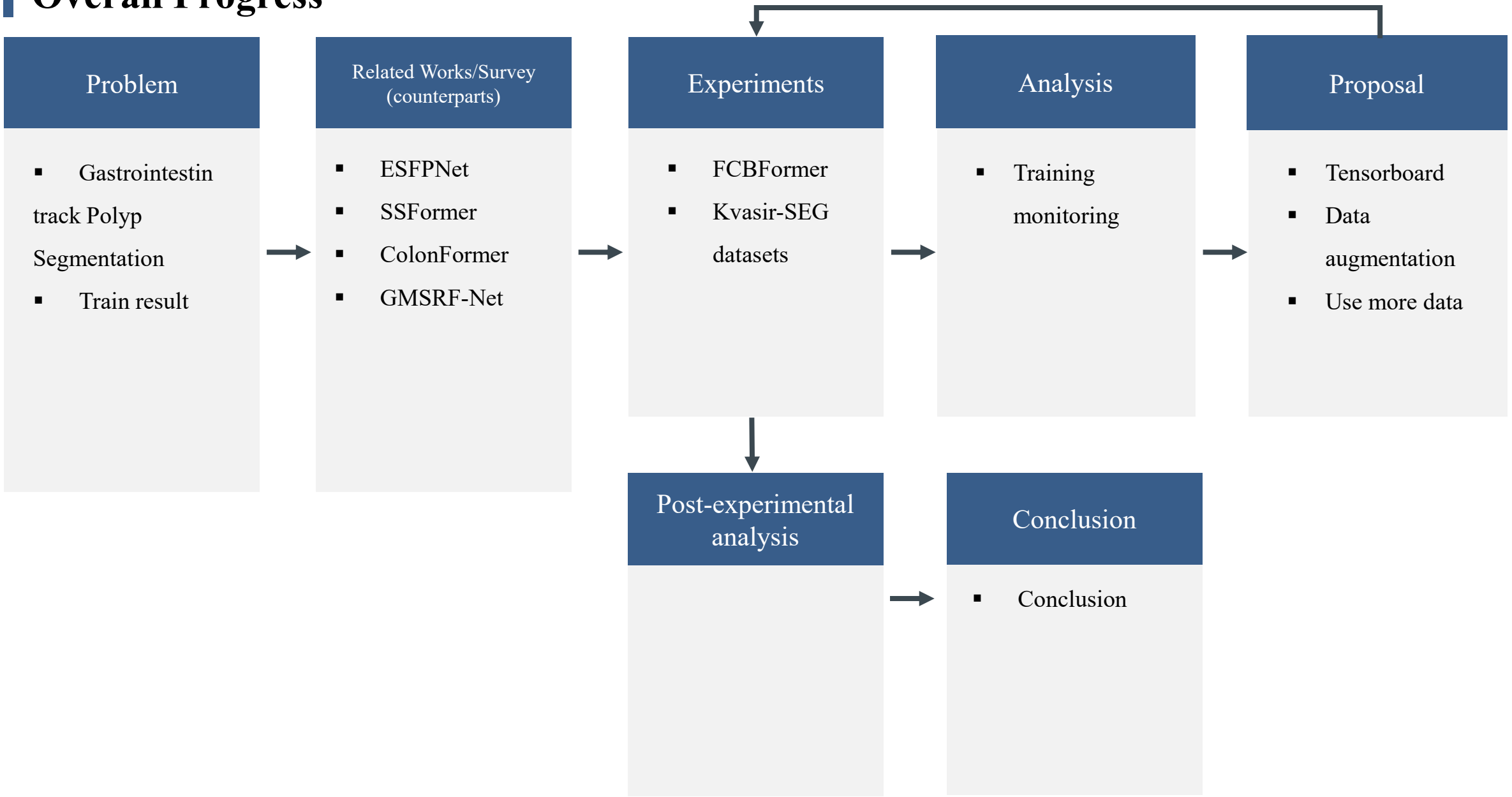
- ✓ Augmentation
- ✓ Fine tuning
- ✓ Optimizer (Hyper parameter)
- ✓ Model

Model	Test Size	AP^{test}	AP_{50}^{test}	AP_{75}^{test}	batch 1 fps	batch 32 average time
YOLOv7	640	**51.4%**	**69.7%**	**55.9%**	161 *fps*	2.8 *ms*
YOLOv7-X	640	**53.1%**	**71.2%**	**57.8%**	114 *fps*	4.3 *ms*
YOLOv7-W6	1280	**54.9%**	**72.6%**	**60.1%**	84 *fps*	7.6 *ms*
YOLOv7-E6	1280	**56.0%**	**73.5%**	**61.2%**	56 *fps*	12.3 *ms*
YOLOv7-D6	1280	**56.6%**	**74.0%**	**61.8%**	44 *fps*	15.0 *ms*
YOLOv7-E6E	1280	**56.8%**	**74.4%**	**62.1%**	36 *fps*	18.7 *ms*

GI Tract Segmentation

1. Overall Progress
2. Introduction
3. Experiment
4. Conclusion

Overall Progress



Introduction

Gastrointestinal track Polyp Segmentation (GI tract Polyp Segmentation)

GI track Polyp : 위창자관(위장관) 용종

➢ **딥러닝의 필요성**

기존 검사법	Deep Learning	효과
Colonoscopy	Segmentation model	정확도&속도 -> polyp 검출율 -> 검출 간과율 -> 대장암 위험

➢ **데이터셋**

Kvasir-SEG, CVC-ClinicDB, CVC-ColonDB, ETIS-LARIBPOLYPDB, Synapse multi-organ CT, MoNuSeg, 2018 Data Science Bowl, Automatic Cardiac Diagnosis Challenge(ACDC), GlaS, Medical Segmentation Decathlon

Introduction

Kvasir-SEG: A Segmented Polyp Dataset

Kvasir SEG

Segmented Polyp Dataset for Computer Aided Gastrointestinal Disease Detection.

- 위장관 용종 이미지와 해당 segmentation mask의 개방형 액세스 데이터 세트
- Contribution of the dataset

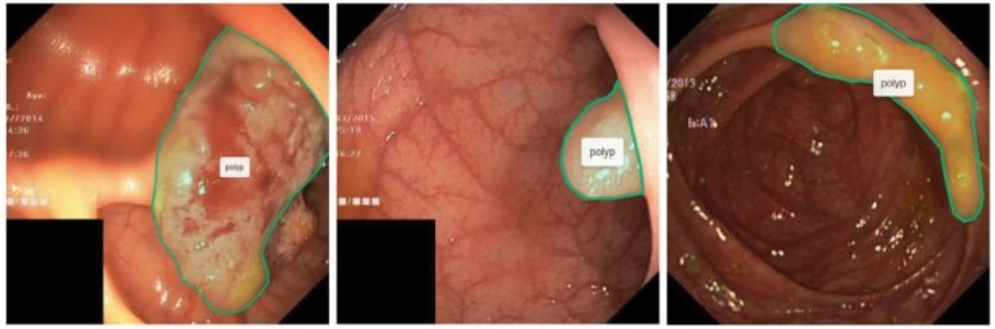
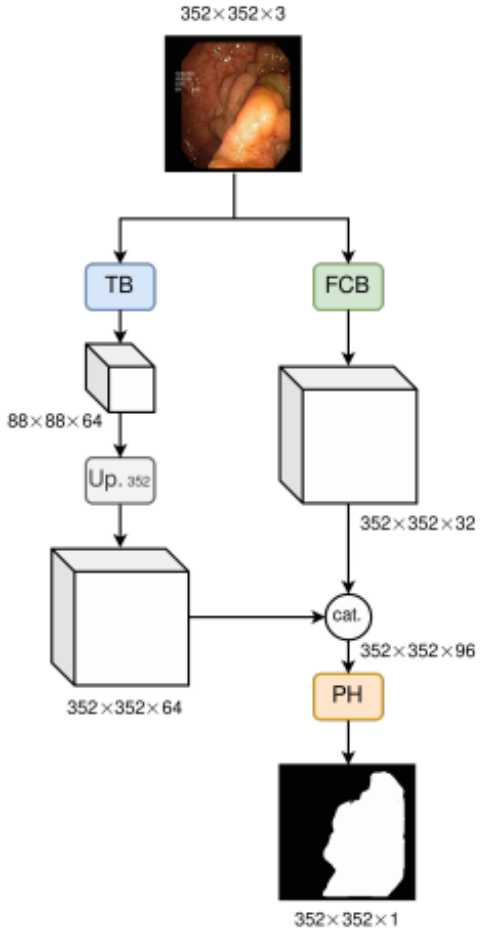


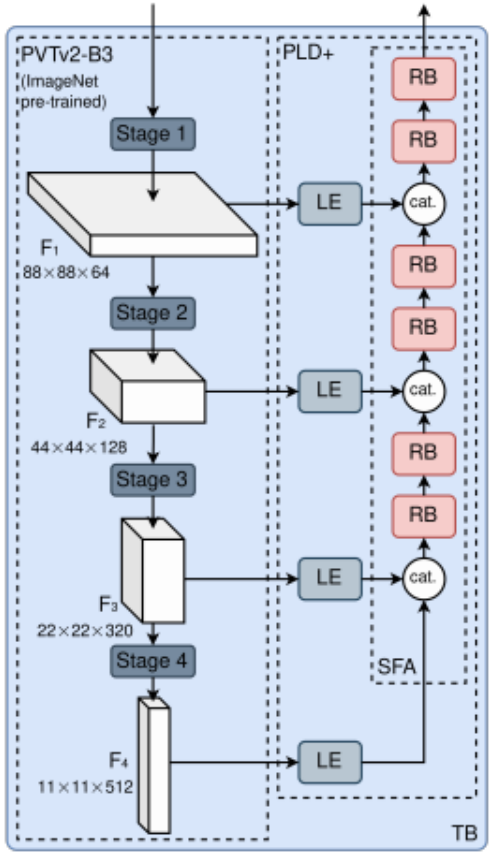
Fig. 1: Example frames from the Kvasir dataset where we additionally have marked the polyp tissue with green outlines.

Introduction

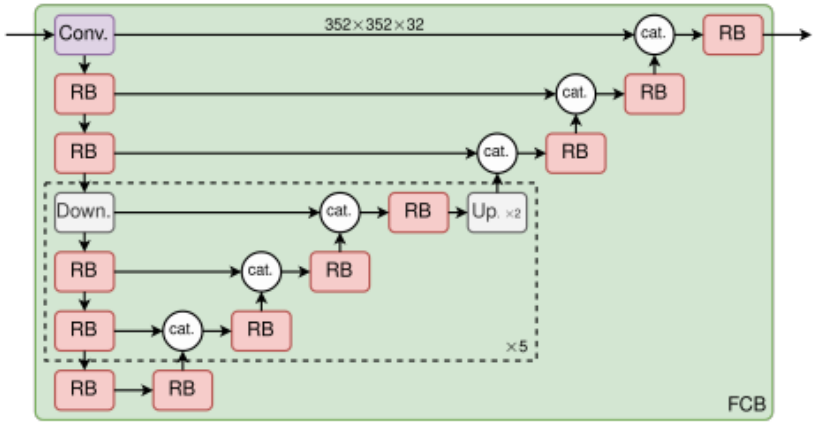
FCN-Transformer Feature for Polyp Segmentation



The architectures of FCBFormer



TB (Transformer Branch)



FCB (Fully convolutional Branch)

The architecture, named the **Fully Convolutional BranchTransFormer (FCBFormer)** uses two parallel branches which both start from a "h × w" input image:

- **Encoder**: **transformer branch** (**TB**) which returns reduced-size (h/4 × w/4) feature maps.(use the PVTv2 [34] for the image encoder)
- **Decoder**: **a fully convolutional branch (FCB)** which returns full-size (h×w) feature maps.

FCN-Transformer Feature Fusion for Polyp Segmentation (Image and Video Processing (eess.IV), Computer Vision and Pattern Recognition (cs.CV), Machine Learning(cs.LG)2022)

Introduction

Experimental Results

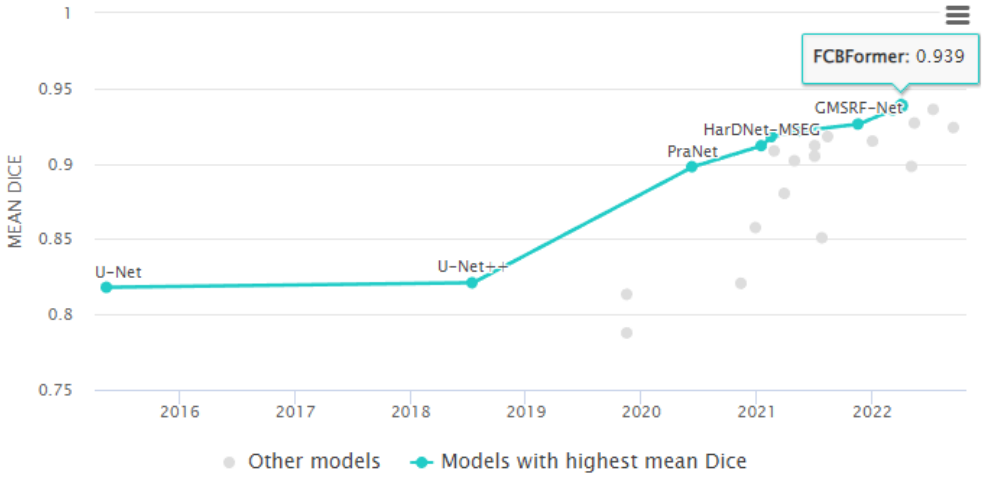


Figure 1. Experimental Results for Kvasir-SEG Dataset

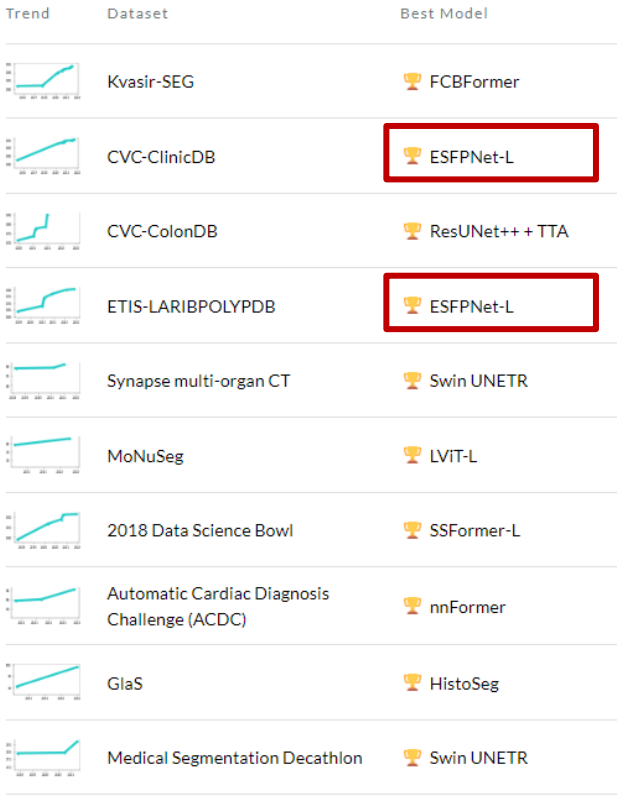


Figure 2. Best Model for Other Medical Image Segmentation dataset

Table 1. Experimental Results for Kvasir-SEG Dataset

Algorithms	mean Dice	mIoU
FCBFormer	**0.9385**	**0.8903**
ESFPNet-L	0.936	0.891
SSFormer-L	0.9357	0.8905
ColonFormer	0.927	0.877
GMSRF-Net	0.9263	0.8843

Experimental Settings

- Datasets
  - Kvasir Dataset
  - Number of Samples: 1,000 (train: 800, test: 100, validation:100)

- Model : FCN-Transformer (FCBFormer)

- Hyper Parameter Setting:

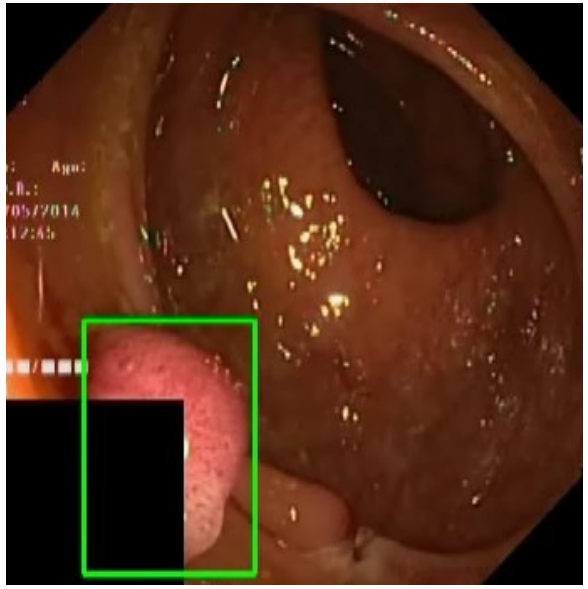
Model	Optimizer	loss	Epochs	Batch size	Learning rate
FCN-Transformer	AdamW	Dice_loss, BCE_loss	200	8	1e-3

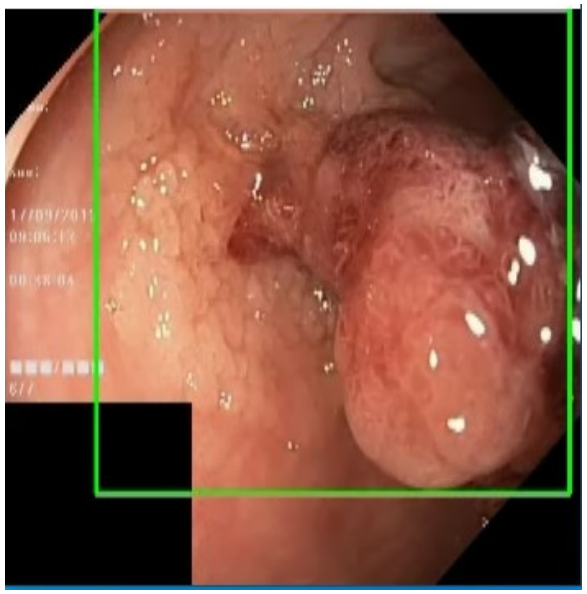
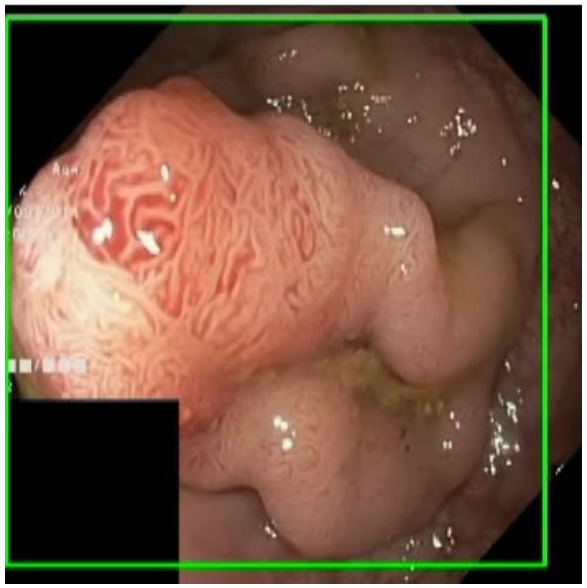
To Do List

- ✓ To do list 1
  - ➢ CVC-ClinicDB, CVC-ColonDB, ETIS-LARIBPOLYPDB 등의 위장계 polyp 검출용 데이터를 사용하여 FCN model 학습
    - ➢ Train monitoring
    - ➢ Data Augmentation
    - ➢ Use more data
    - ➢ Train, Test result visualization & save result

Thank you

Appendix



To Do List

- ✓ To do list
  - ➢ Hyper parmater 실험
  - ➢ Data Augmentation
  - ➢ Prediction result 분석, heatmap 분석
  - ➢ Train, Test result visualization (pytorch tensorboard) & save result