Individual Studies

Taekyung Song

songteagyong@gmail.com

2nd Nov, 2022

Contents

To Do List from Previous Seminar

Contents

- ✓ Individual Study
  - ➢ Sementic Segmentation using localized information

- ✓ 갑상선 연구 – 눈꺼풀 길이 추론
  - ➢ 눈의 중앙으로부터 눈꺼풀까지의 길이 정규화

Overall Progress

Problem

- Sementic Segmentation Accuracy

Related Works/Survey (counterparts)

- Using localization information

Experiments

Analysis

Proposal

Individual Study

Sementic Segmentation using localized information

Search keyword
- visual localization
- small object

cf		
2013	*Journal of mathematical imaging and vision*	Image Segmentation Using Local GMM in a Variational Framework
2018	*IEEE conference on computer vision and pattern*	Semantic visual localization
2019	*IEEE reviews in biomedical engineering*	A Review of Automated Methods for the Detection of Sickle Cell Disease

갑상선 연구 – 눈꺼풀 길이 추론

눈의 모양을 평가하는 metric

- 기존 시도의 문제

  1. 중앙에서 1°씩 $y = ax + b$ 그래프를 그렸을 때

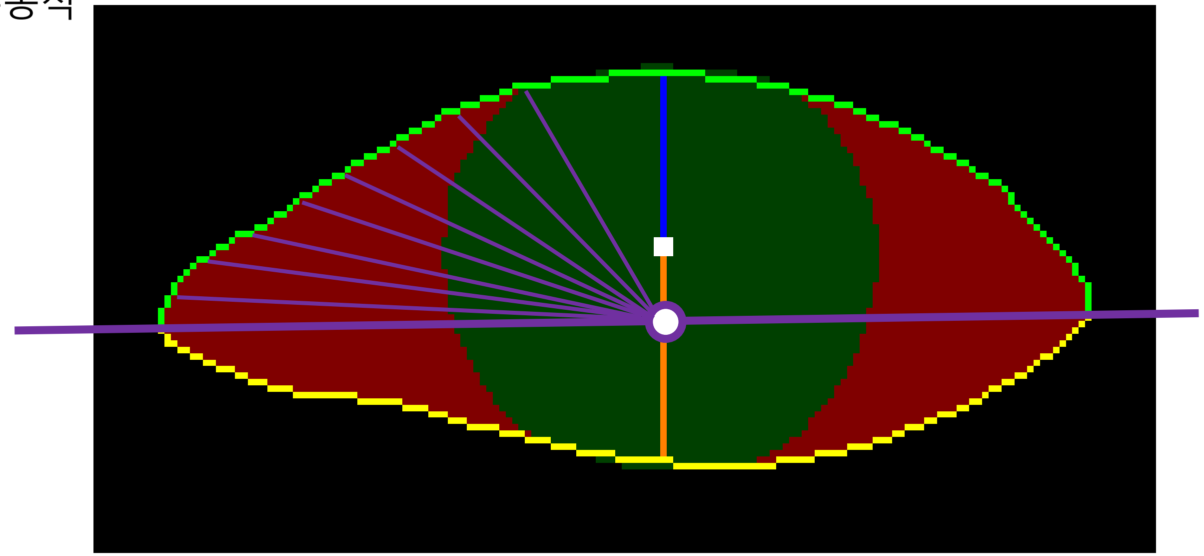
     - 그래프 위에 정확히 일치하는 픽셀이 존재하지 않을 경우

     ➡ sclera의 테두리 픽셀 개수를 360개로 정규화

  2. 눈꺼풀 픽셀 개수의 문제

     - 촬영 거리에 따라, 환자의 눈 크기에 따라 픽셀 개수가 유동적

     ➡ sclera의 테두리 픽셀 개수를 360개로 정규화



갑상선 연구 – 눈꺼풀 길이 추론

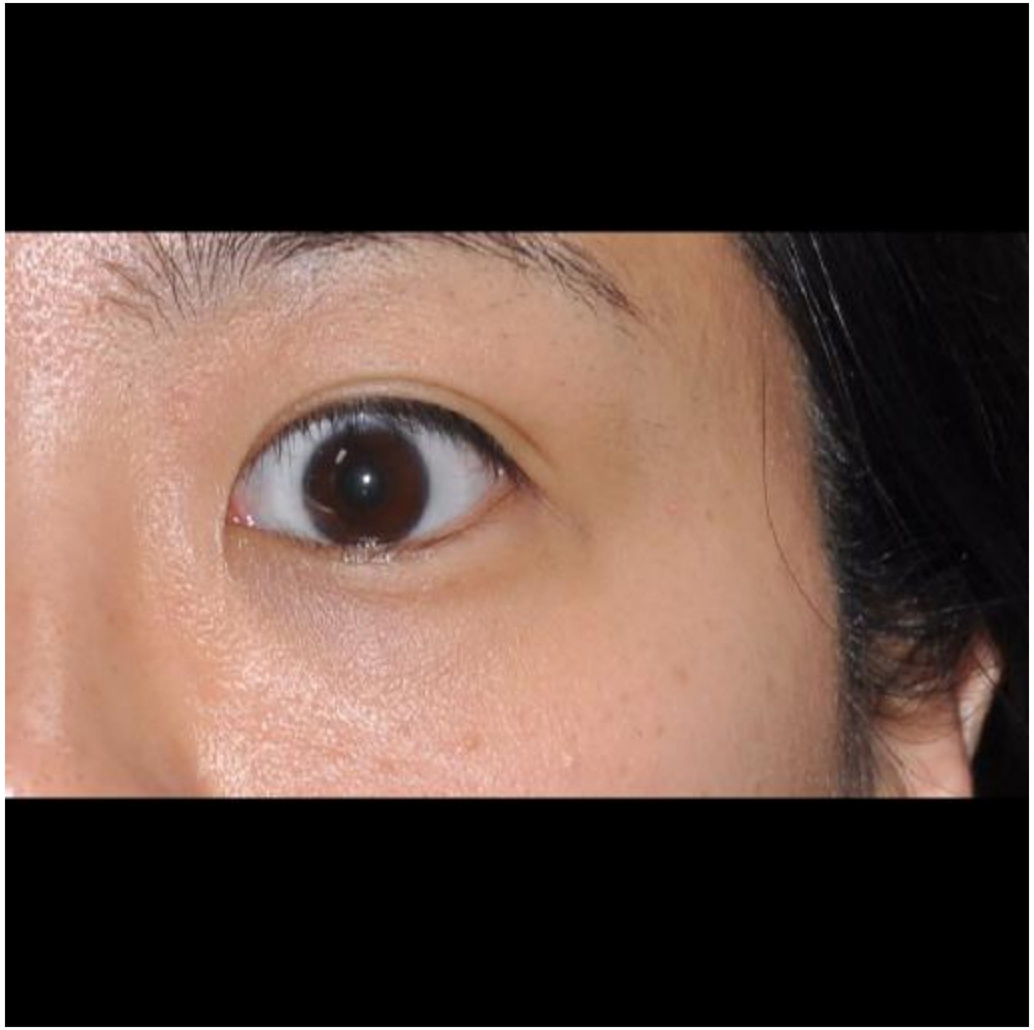
눈의 모양을 평가하는 metric

- 시도 방법
    1. 1도마다 변하는 $y = ax + b$ 계산하여 좌표 계산
    2. Linear regression을 이용한 좌표값 혹은 길이 추정
    3. 길이를 계산하여 360개로 확대 혹은 축소하는 보간

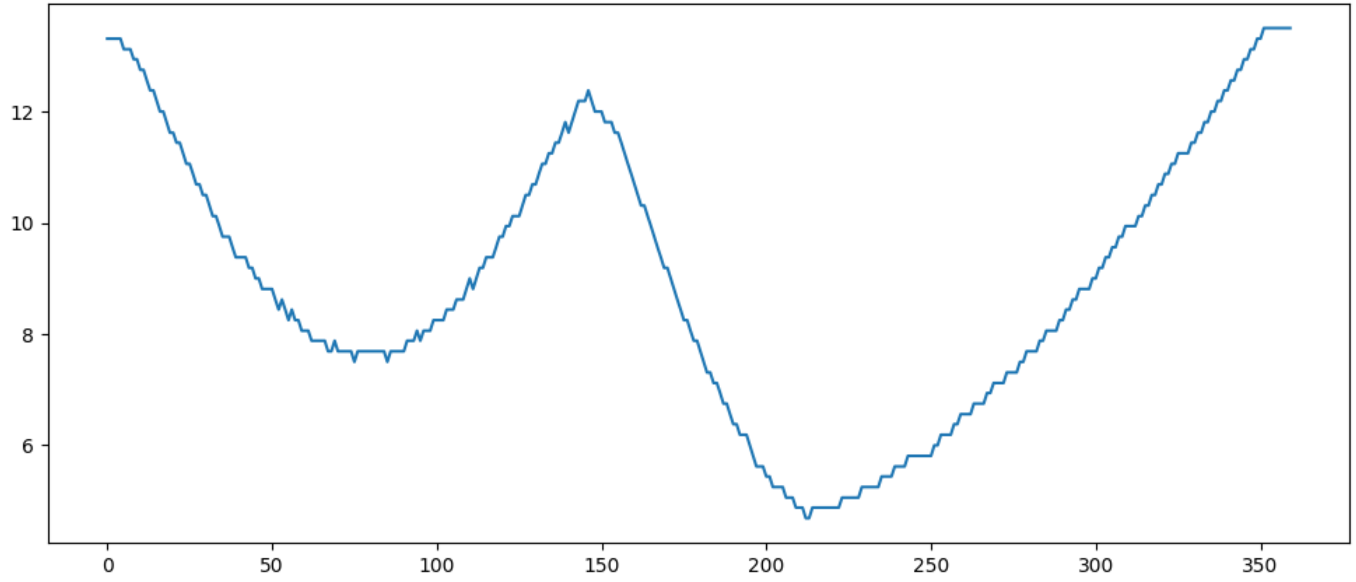
- 추가 진행사항
    - 개별 데이터가 아닌 전체 데이터에 대한 정규화 진행

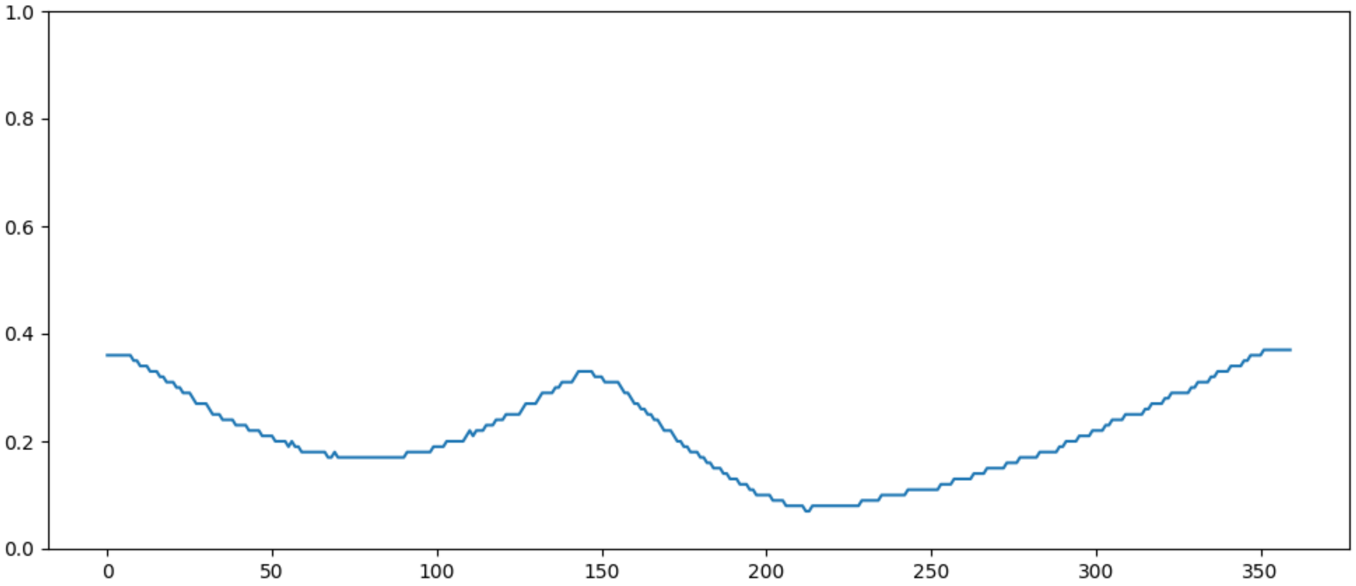
※ Reflex 미관측자 및 일부 sclera 측정이 잘 이루어지지 않은 경우 제외

갑상선 연구 – 눈꺼풀 길이 추론



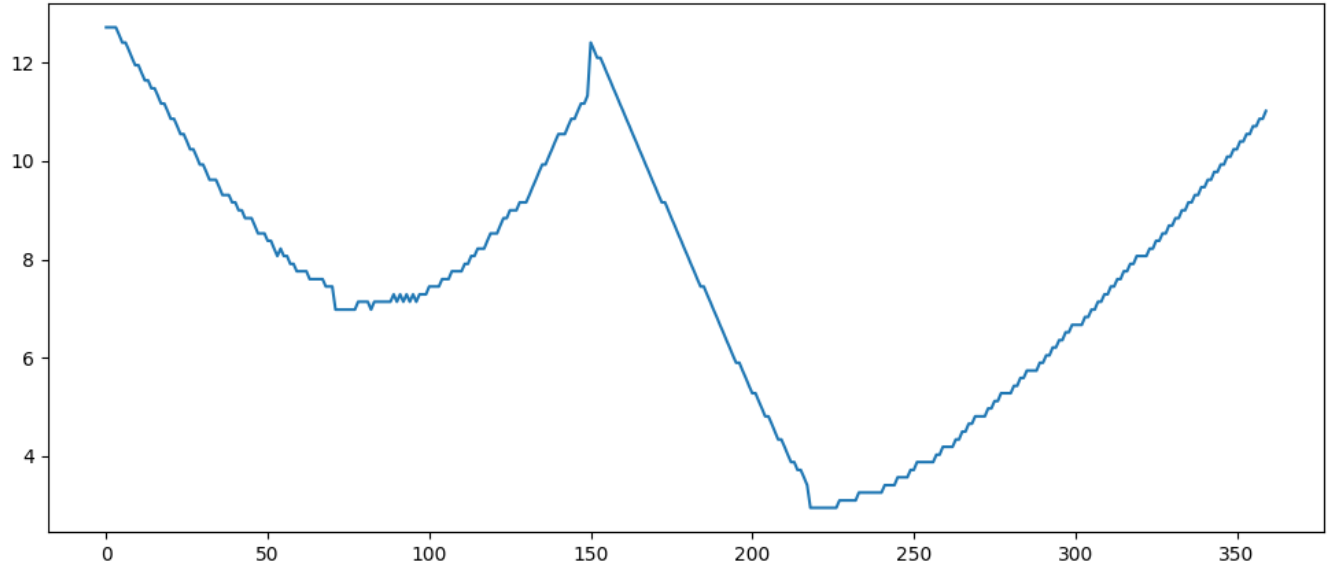
<101026_L>

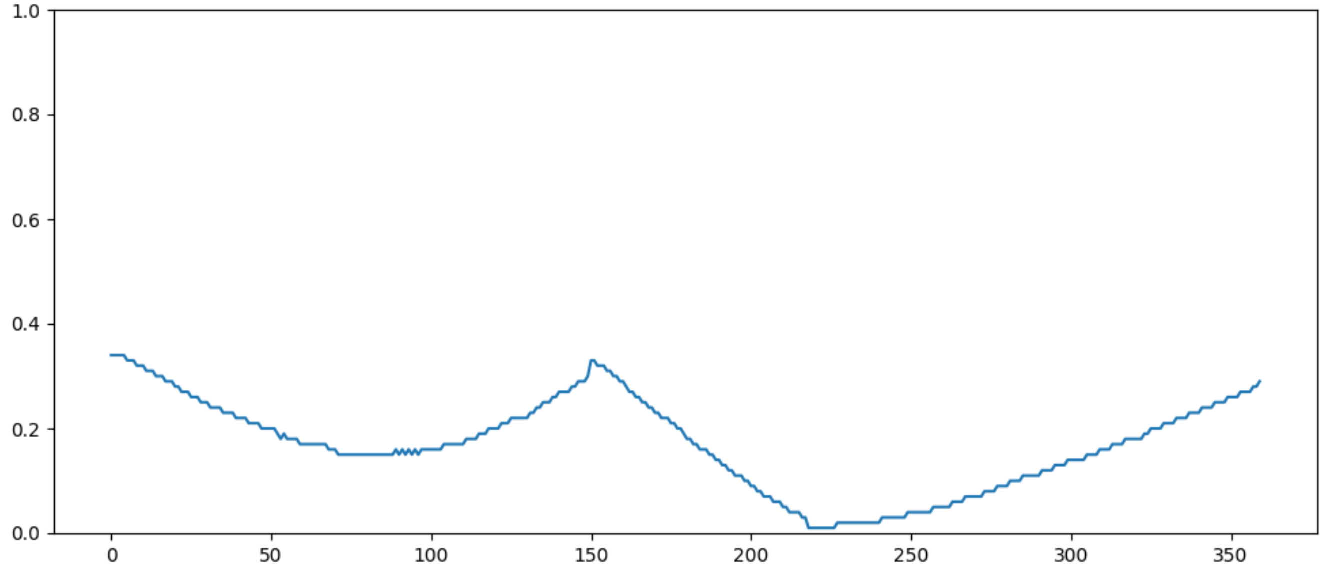




갑상선 연구 – 눈꺼풀 길이 추론

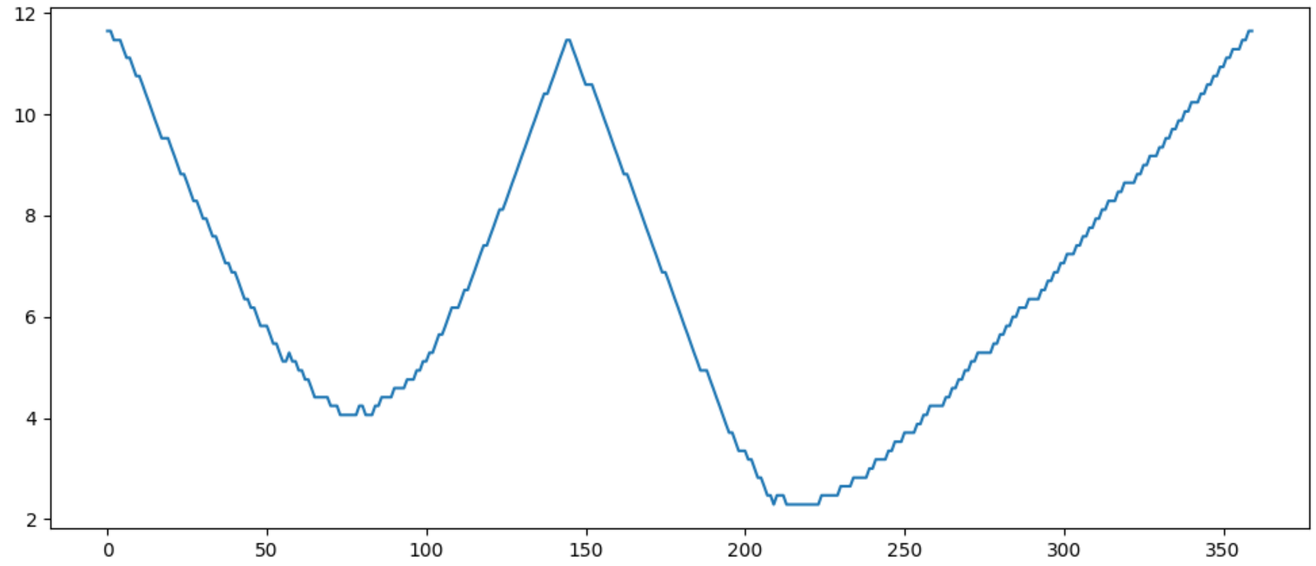
<10312_R>

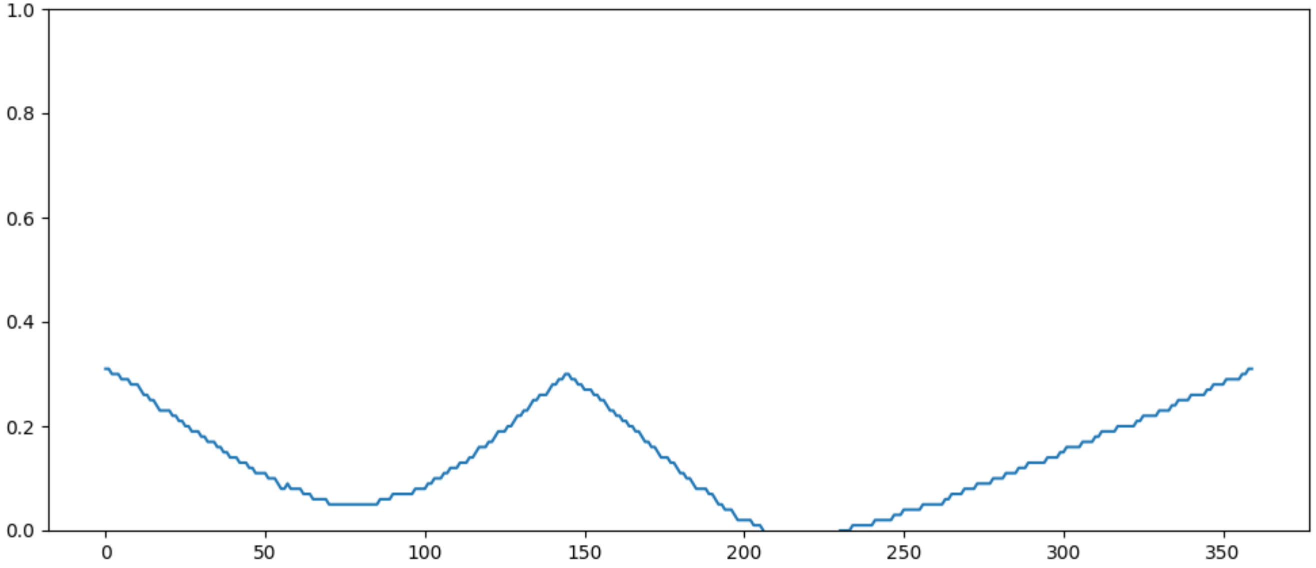




갑상선 연구 – 눈꺼풀 길이 추론

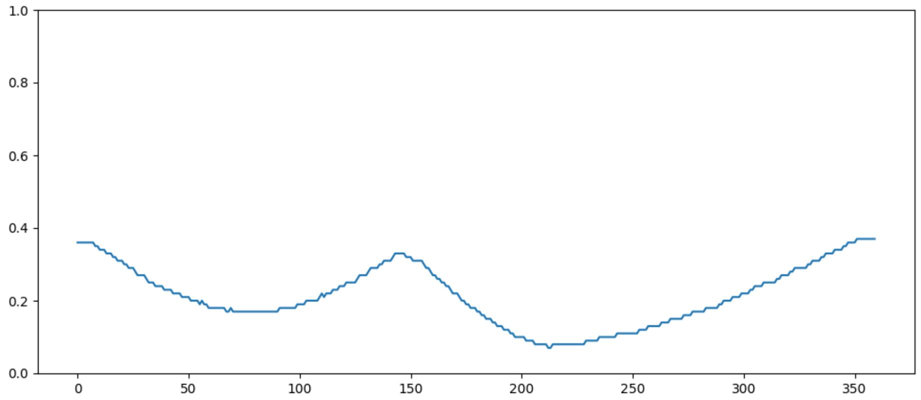
<101129_R>





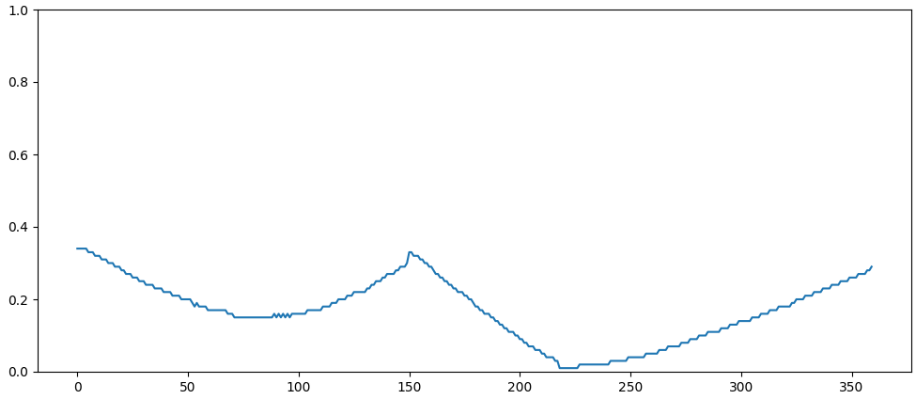
갑상선 연구 – 눈꺼풀 길이 추론

<101026_L>



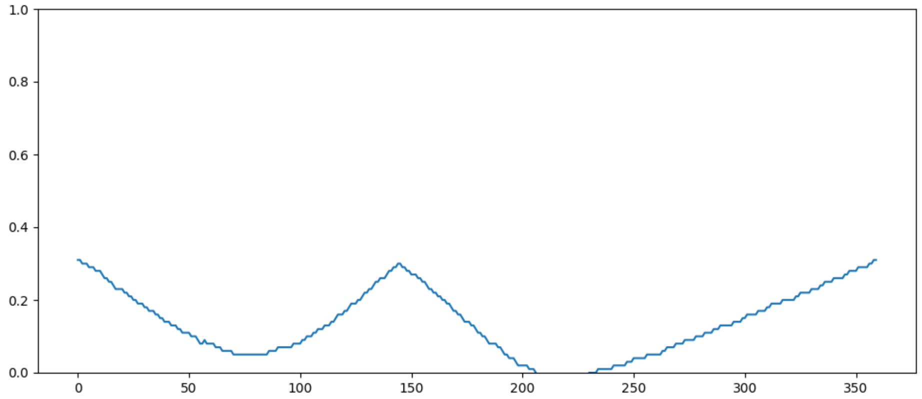
MRD1	MRD2	Upper	Lower
6.56	6	33.06	30.58

<10312_R>



MRD1	MRD2	Upper	Lower
5.43	4.5	31.73	28.34

<101129_R>



MRD1	MRD2	Upper	Lower
1.24	5.12	27.57	26.11

ICEIC Overall Progress

Problem
- Inaccurate Eyelid Length Measurement

→

Related Works/Survey (counterparts)
- Analysis of eyelid position using digital image processing software
- Sementic Segmentation

→

Experiments
- TAO dataset
- Compare model performance
  - U-Net, SegNet, Deeplab V3+
- Metric
  - mIoU

→

Analysis
- Single-task Multi-task performance comparison
- Length measurement using segmented images

→

Proposal
- Automated length measurement using deep learning models

To do list

Next progress

- 개인연구 주제 관련 Research

- 두산 관련 머신러닝 알고리즘 고도화 진행

- 갑상선 미팅 위한 준비(~ 10.20)

Employing Multi-branch Convolution Structure in 1D ResNet for Music Skip Prediction

Sanghyeong Jin

jinsanghyeong@gmail.com

2022. 11. 10.

Contents

To Do List

Experimental Results Discussion

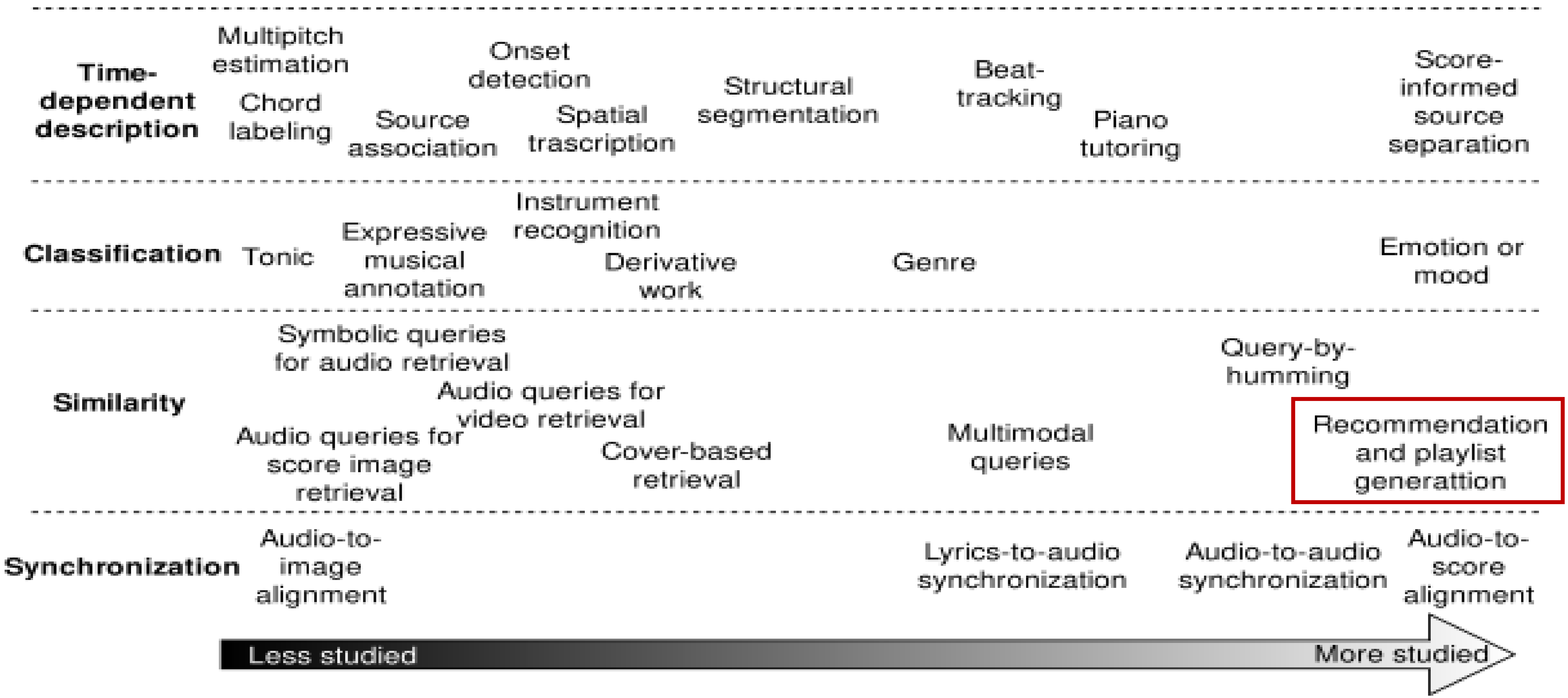
- Feature map 관련

Track feature에서 First / Second Half 분기의 필요성

- Track / Session 분기 관련 재실험

Introduction

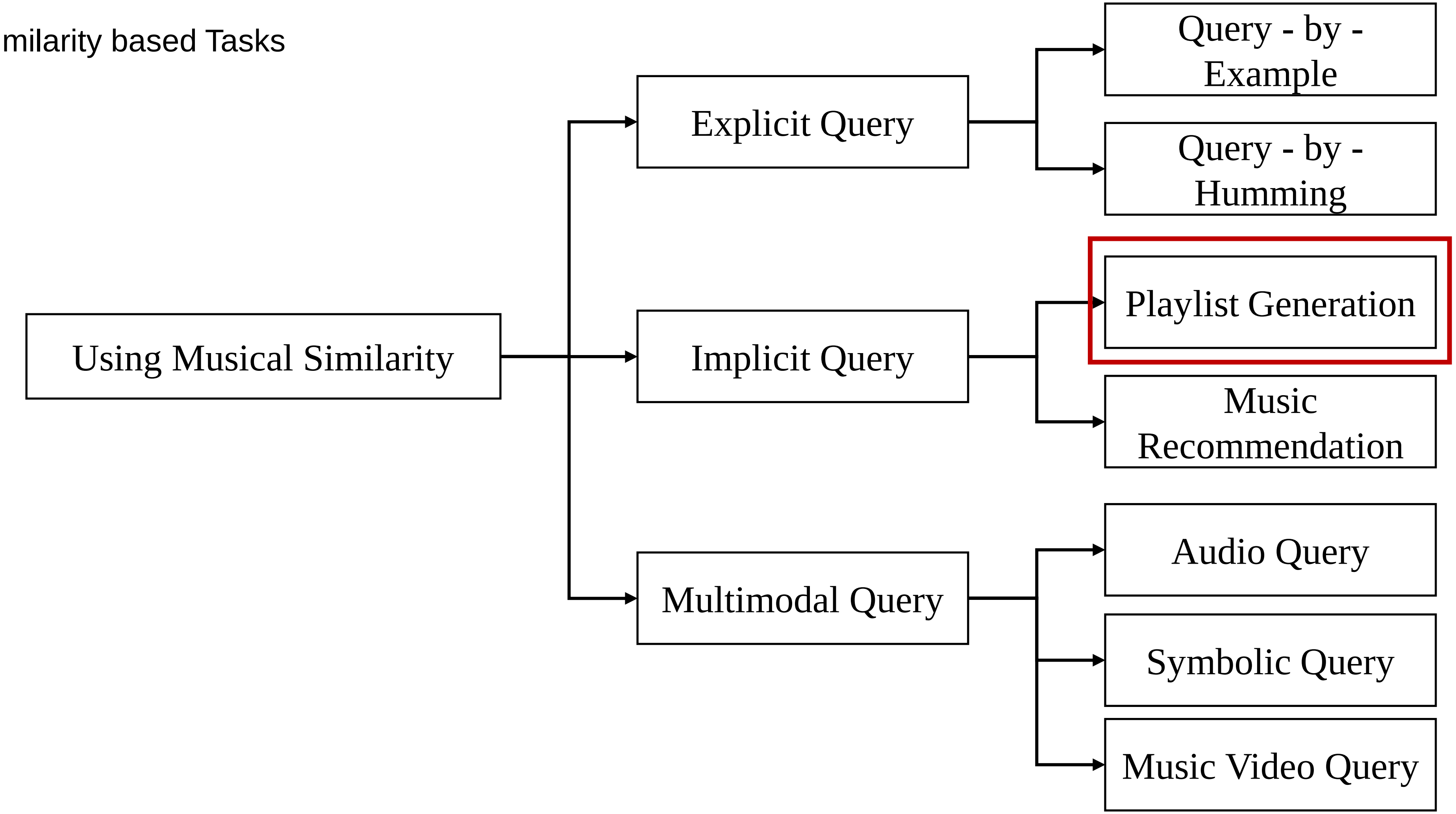
Music Information Retrieval



Introduction

Music Information Retrieval

- Musical Similarity based Tasks



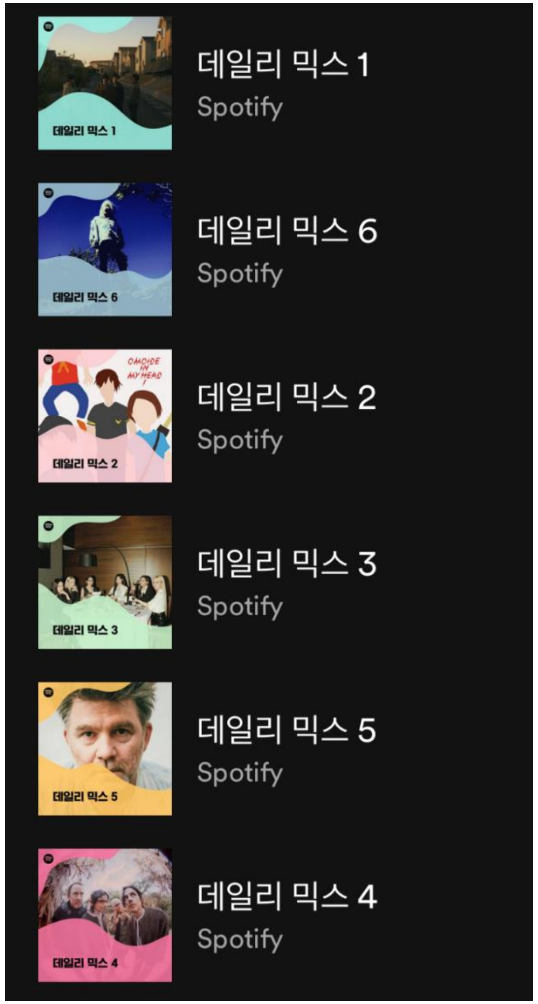
Introduction

Playlist Generation

Playlist Generation

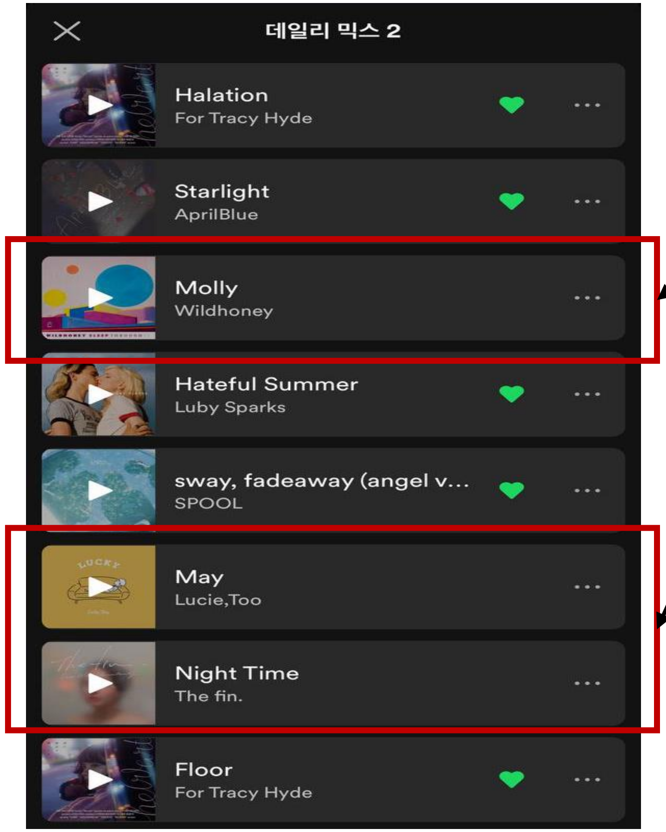
- Reinforcement Learning
- Frequent Pattern Mining
- Collaborative Filtering
- Statistical Model (Markov Model)

⋮



Introduction

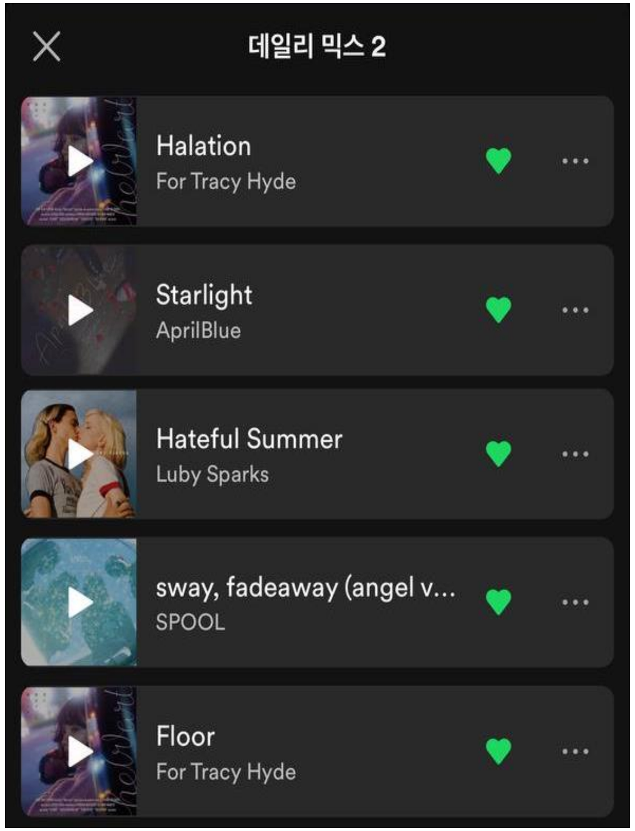
Music Track Skip Prediction



General mix playlist

Annoying to skip these tracks
at the end of each song

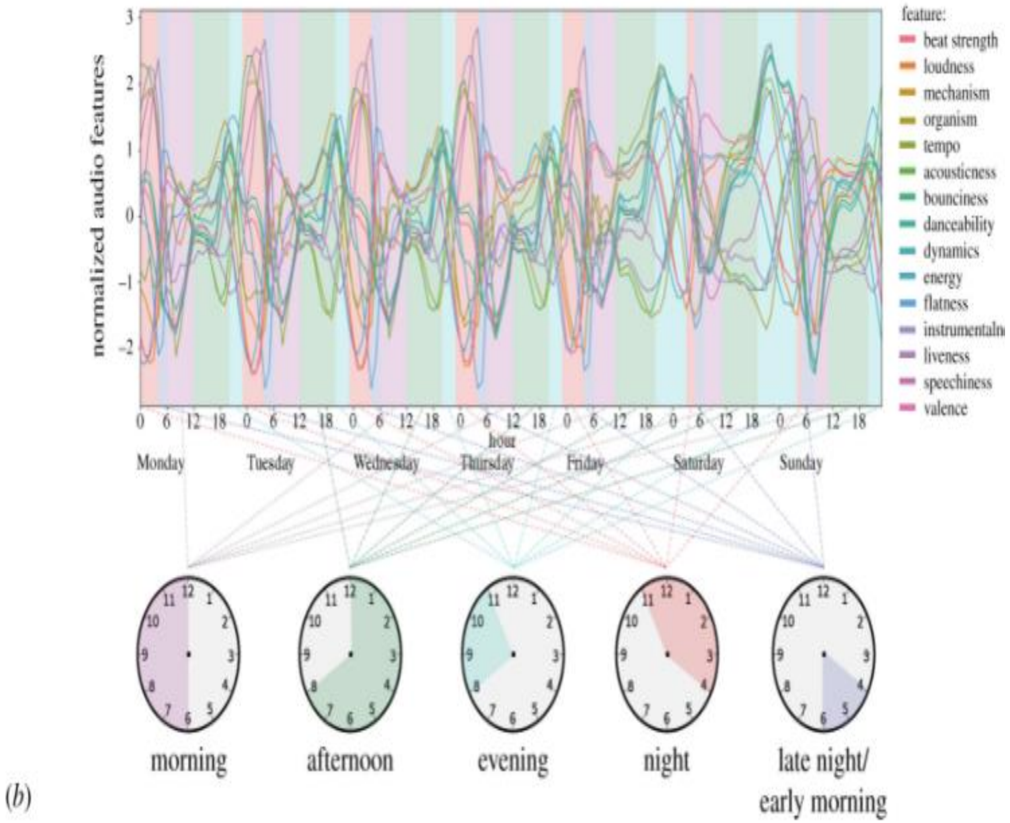
- Minimizing user skip intervention
- Smooth track sequence transition
- Generating user preferred playlist



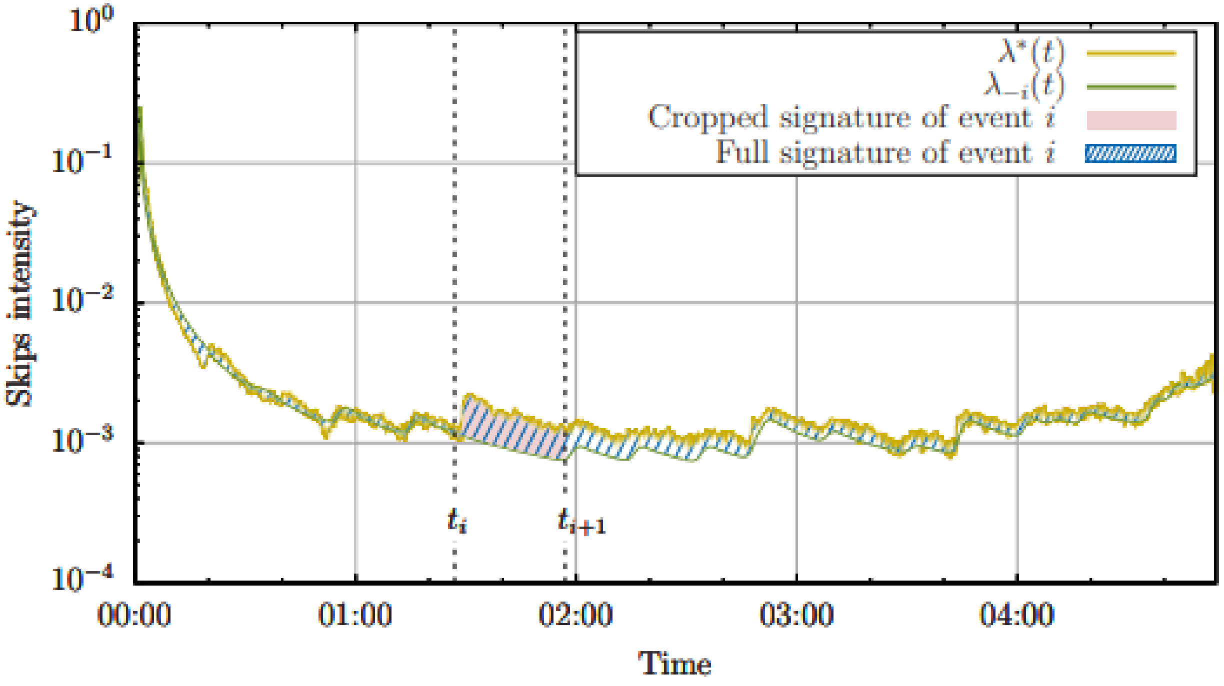
User preferred playlist

Related Work

Track Skip Behavior



Musical diurnal fluctuations

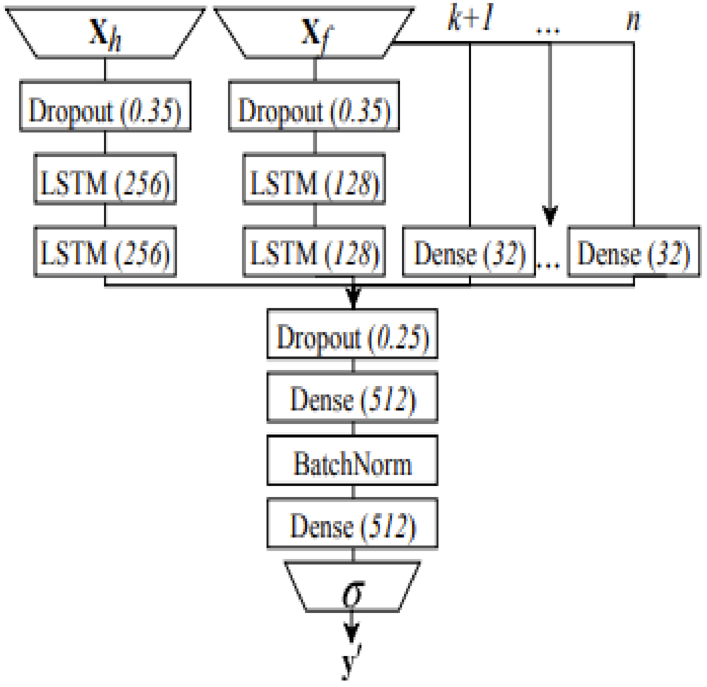


Skip rate based on time

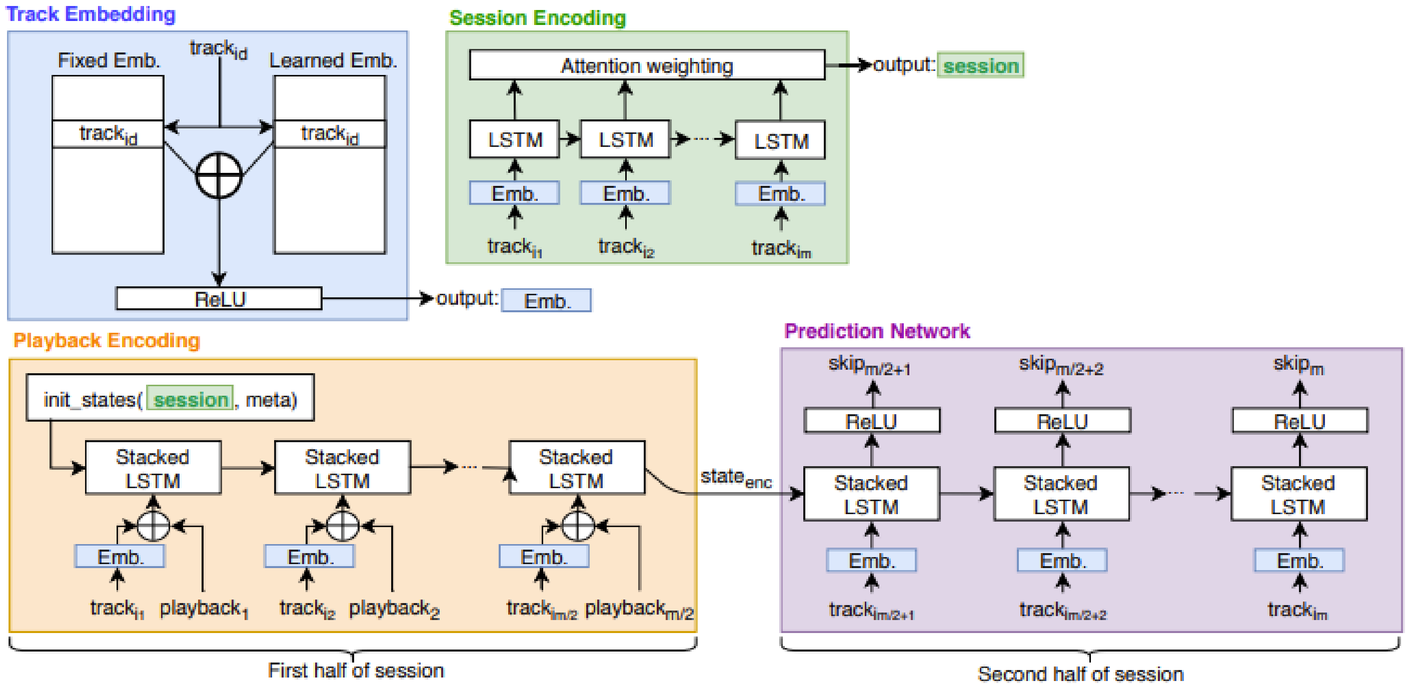
Related Work

Track Skip Prediction

- Mostly based on RNN Models



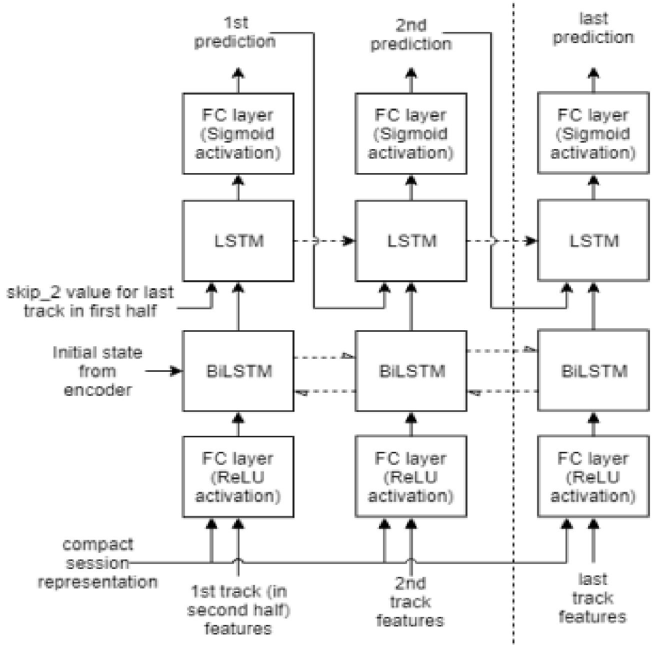
RNN-Dense



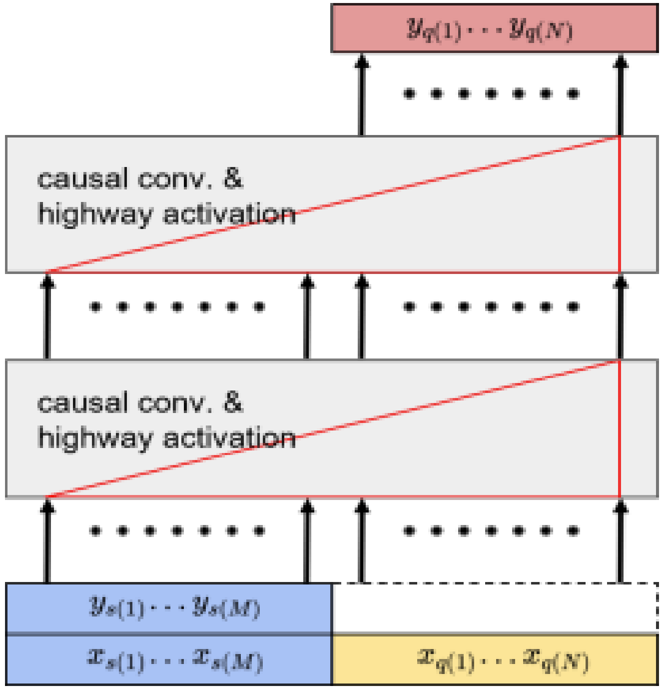
Multi-RNN

Related Work

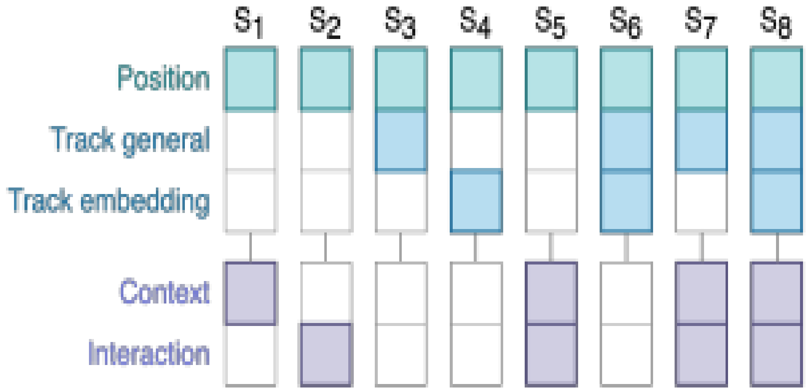
Track Skip Prediction



BiLSTM Encoder-Decoder



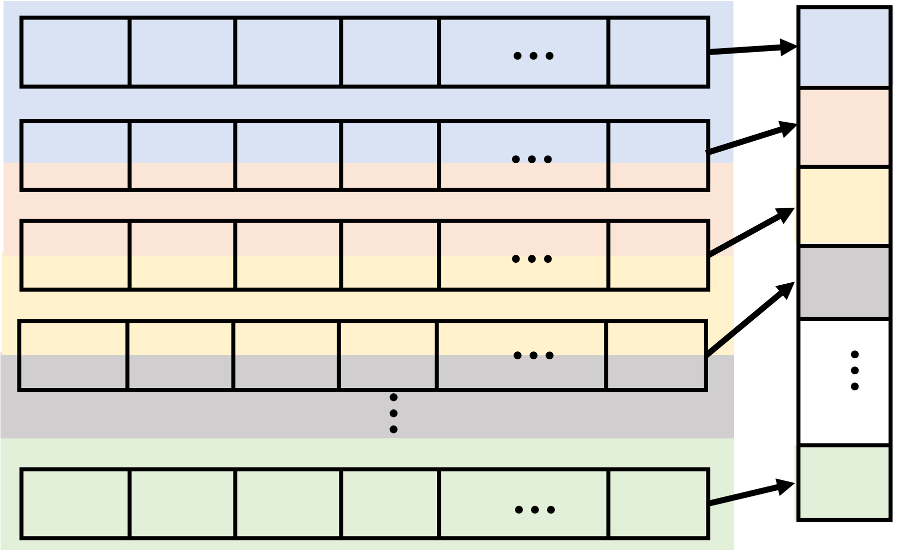
1D convolution with Highway Network



Interpretable Transformer

Related Work

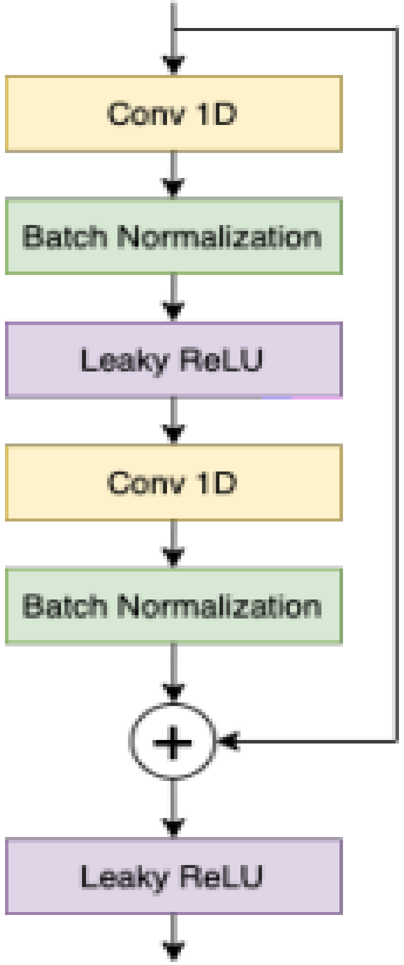
1D ResNet



1D Convolution

Variant	Functional Form
Plain	$H(\mathbf{x})$
Residual	$H(\mathbf{x}) + x$
T-Only	$H(\mathbf{x}) \cdot T(\mathbf{x}) + \mathbf{x}$
C-Only	$H(\mathbf{x}) + \mathbf{x} \cdot C(\mathbf{x})$
Coupled	$H(\mathbf{x}) \cdot T(\mathbf{x}) + \mathbf{x} \cdot (1 - T(\mathbf{x}))$
Full	$H(\mathbf{x}) \cdot T(\mathbf{x}) + \mathbf{x} \cdot C(\mathbf{x})$

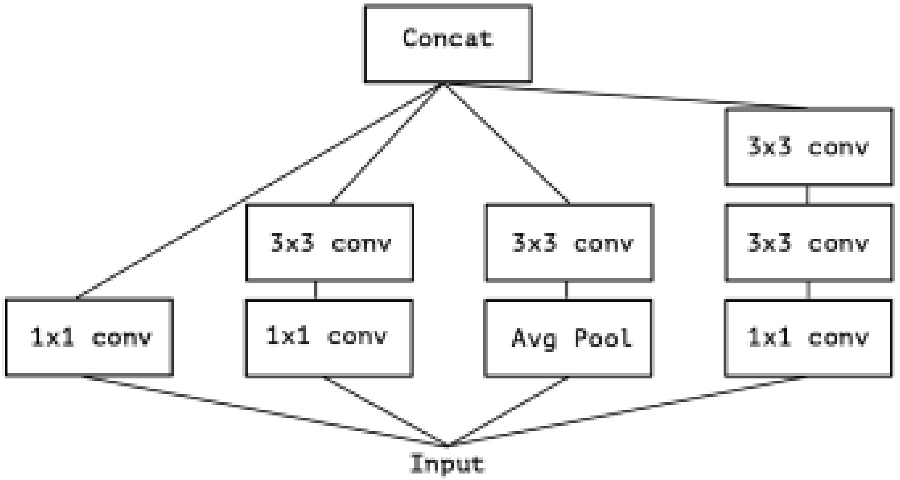
Residual Connection



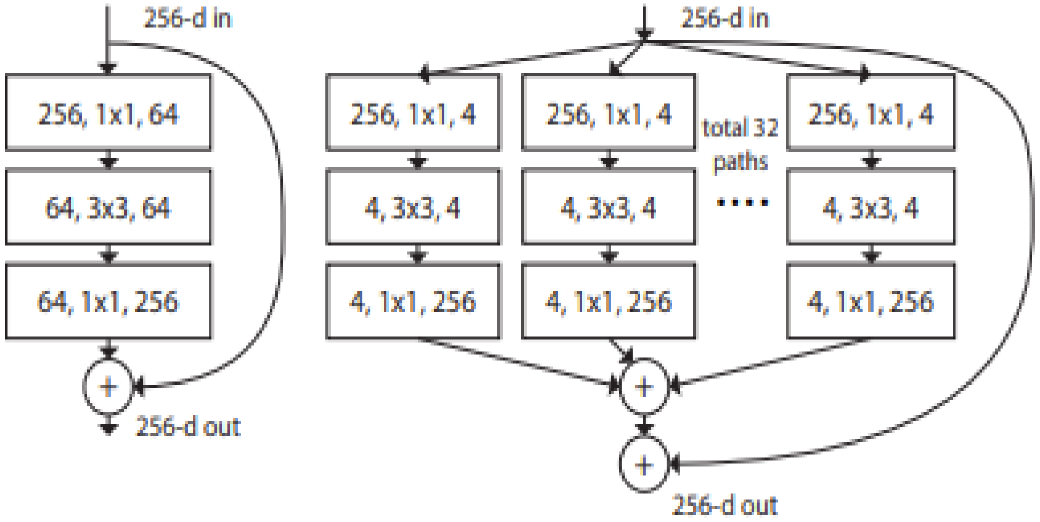
1D Residual Block

Related Work

Multi-branch Convolution



Inception V3



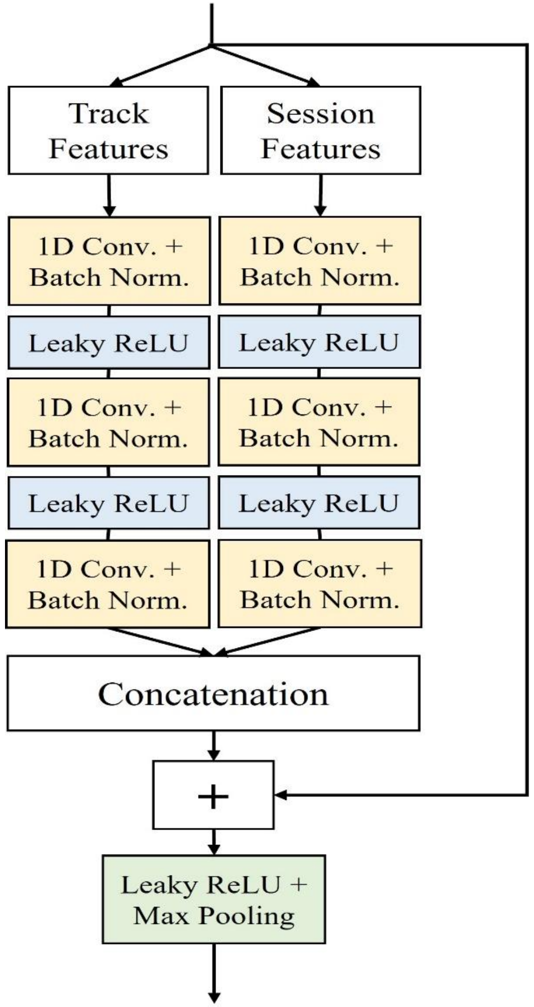
ResNeXt

Proposed Method

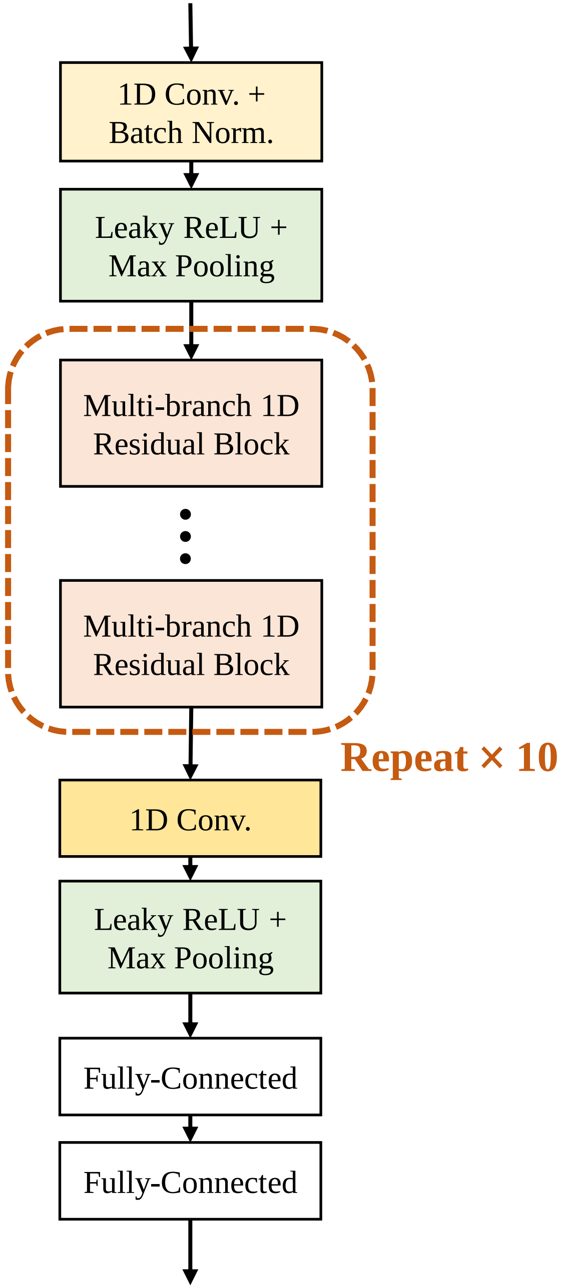
Multi-branch 1D Residual Block

B × (S + T) × L	
Track Features	Session Features
B × T × L	B × S × L
B × T/2 × L	B × S/2 × L
B × T × L	B × S × L
Concatenation	
B × (S + T) × L	

B : Batch size, S : Session size, T : Track size
L : Sequence Length



Multi-branch 1D ResNet



Experimental Settings

MAA (Mean Average Accuracy)

$$AA = \frac{\Sigma_{i=1}^{T} A(i)L(i)}{T}$$

T : Number of tracks to be predicted
A(i) : Accuracy to the *i*-th track in sequence
L(i) : Boolean indicator that if *i*-th prediction was correct (0, 1)

Ground Truth	Predicted	AA (Average Accuracy)
1, 1, 1, 1, 1	1, 1, 1, 1, 0	0.800
1, 1, 1, 1, 1	1, 1, 1, 0, 1	0.760
1, 1, 1, 1, 1	1, 1, 0, 1, 1	0.710
1, 1, 1, 1, 1	1, 0, 1, 1, 1	0.643
1, 1, 1, 1, 1	0, 1, 1, 1, 1	0.543

$$\frac{\left(\frac{1}{1} * 1 + \frac{2}{2} * 1 + \frac{3}{3} * 1 + \frac{4}{4} * 1 + \frac{4}{5} * 0\right)}{5} = 0.800$$

$$\frac{\left(\frac{1}{1} * 1 + \frac{2}{2} * 1 + \frac{2}{3} * 0 + \frac{3}{4} * 1 + \frac{4}{5} * 1\right)}{5} = 0.710$$

$$\frac{\left(\frac{0}{1} * 0 + \frac{1}{2} * 1 + \frac{2}{3} * 1 + \frac{3}{4} * 1 + \frac{4}{5} * 1\right)}{5} = 0.543$$

Experimental Settings

- Datasets
  - Music Streaming Sessions Dataset (MSSD)
  - Number of Samples: 167,880 session histories (10,000 unique session)
  - Number of Classes: 2 (Skip / Not Skip)
  - Input : (74, 20) - (Session & Track Features, Sequence Length) / Sequence that less than 20 includes zero padding
  - Output : (10, 1) - Prediction probability for the 10 tracks in second half
- Cross-Validation for Test Performance
  - 10 iterations / Group K-Fold (by Session id)
  - Data Split Ratio: (Train : Test) = (0.9: 0.1)
- Hyper Parameter Setting:

<table>
<tr><th>Model</th><th>Batch Size</th><th>Epochs</th><th>Loss Function</th><th>Optimizer</th><th>LR</th></tr>
<tr><td>Proposed</td><td rowspan="5">16</td><td rowspan="5">10</td><td rowspan="5">Binary Cross Entropy</td><td rowspan="2">Adam</td><td rowspan="3">1e-4</td></tr>
<tr><td>Highway 1D Conv.</td></tr>
<tr><td>Encoder-Decoder</td><td>RMSprop</td></tr>
<tr><td>Multi-RNN</td><td rowspan="2">Adam</td><td>1e-3</td></tr>
<tr><td>RNN-Dense</td><td>1e-4</td></tr>
</table>

Experimental Results

t-test Result

Model	MAA Mean ± Standard Deviation	Average Rank
Proposed	**0.6625 ± 0.0427**	**1.77**
Highway 1D Conv. [12]	0.5976 ± 0.0885 ▼	3.03
Encoder-Decoder [16]	0.5357 ± 0.0912 ▼	3.77
Multi-RNN [14]	0.6027 ± 0.0583 ▼	3
RNN-Dense [15]	0.5740 ± 0.0901 ▼	3.43

▼ indicates corresponding model is significantly worse than proposed model at significance level α=0.05

Experimental Results

Friedman-test Result

	Statistics	p-value
Proposed	27.573	1.52e-05

significance level α=0.05

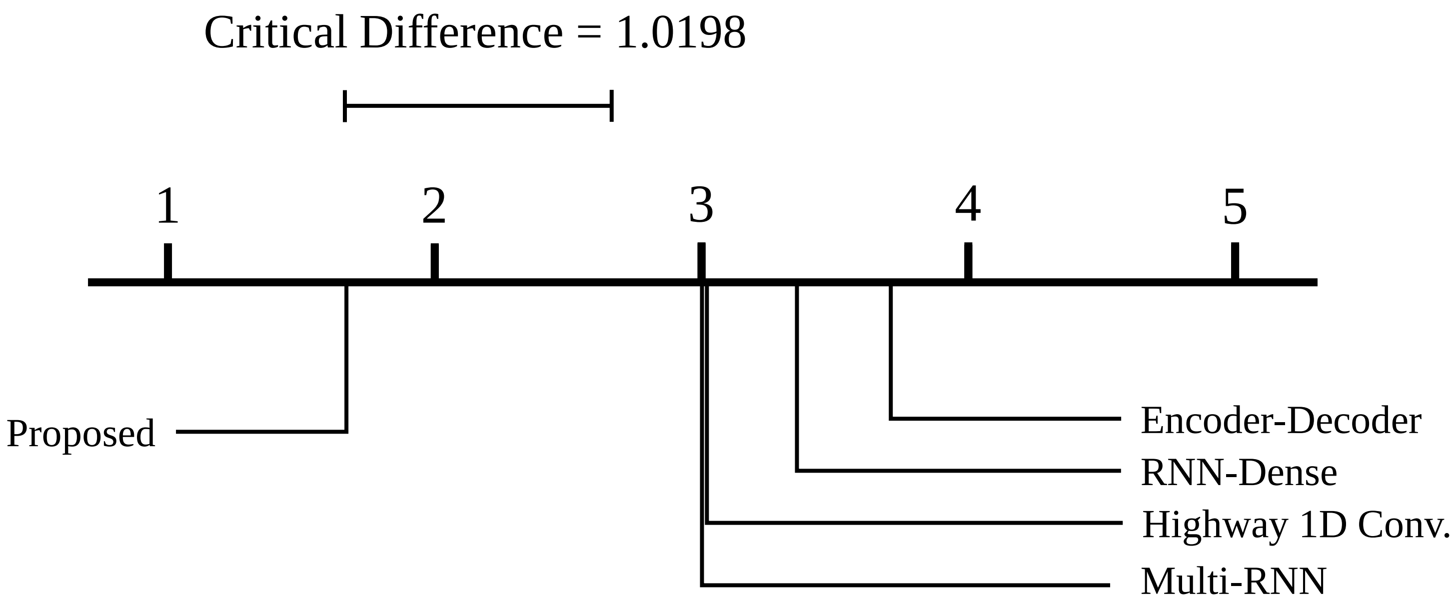
Wilcoxon signed-rank test Result

	Comparison models			
	Highway 1D Conv. [12]	Encoder-Decoder [16]	Multi-RNN [14]	RNN-Dense [15]
Proposed ***versus*** (p-value)	Win (8.31e-04)	Win (3.69e-06)	Win (5.29e-04)	Win (8.92e-05)

significance level α=0.05

Experimental Results

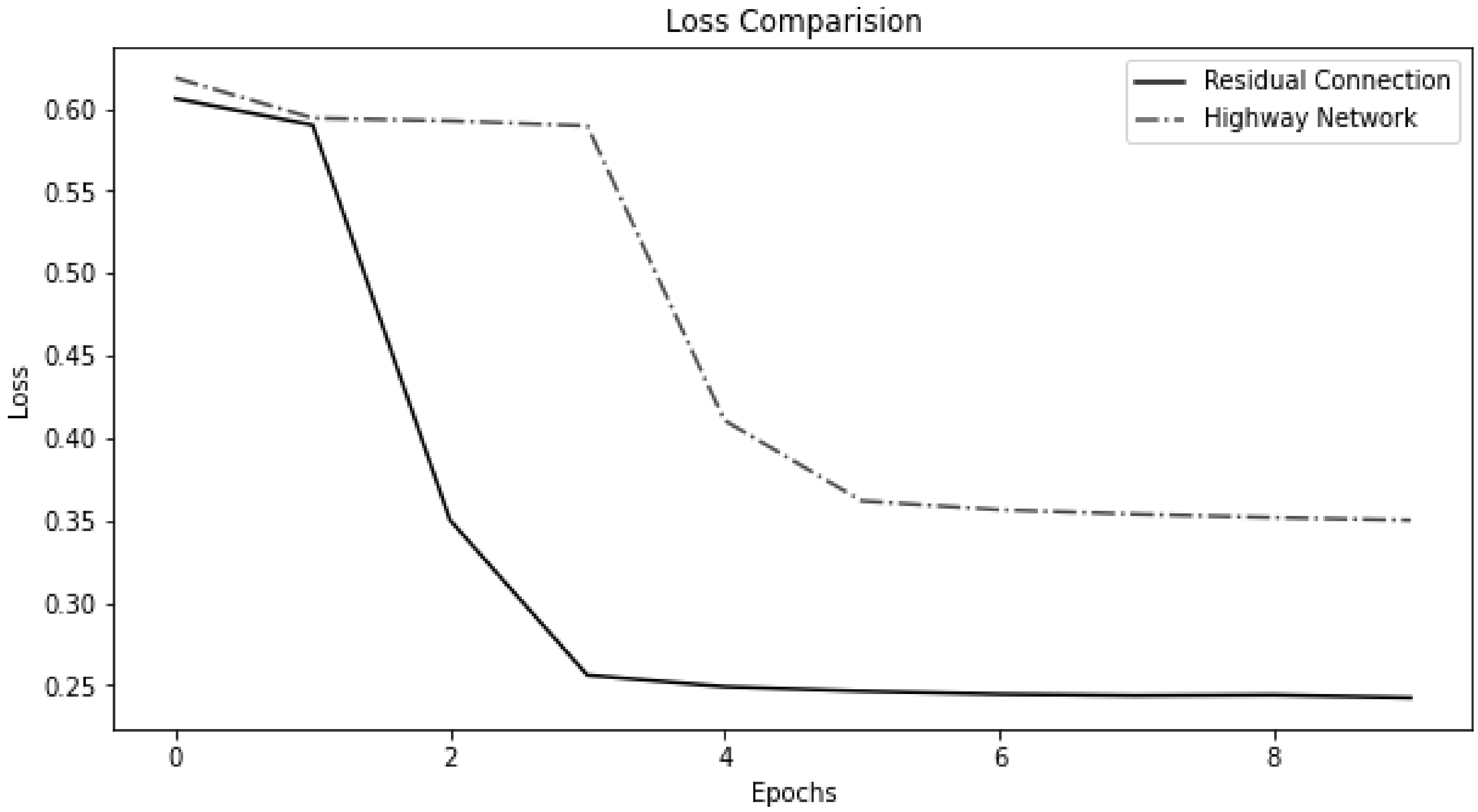
Bonferroni-Dunn test Result



Experimental Results

Discussion

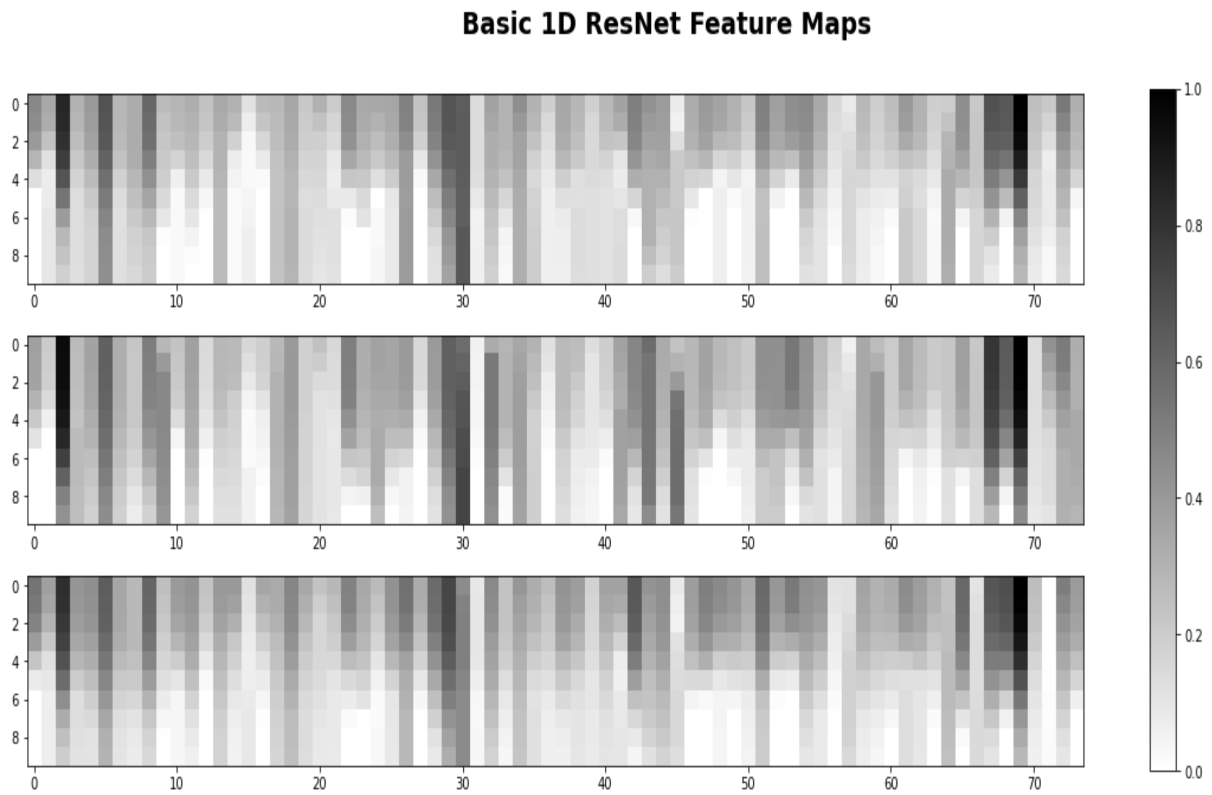
- Highway Network vs. Residual Connection

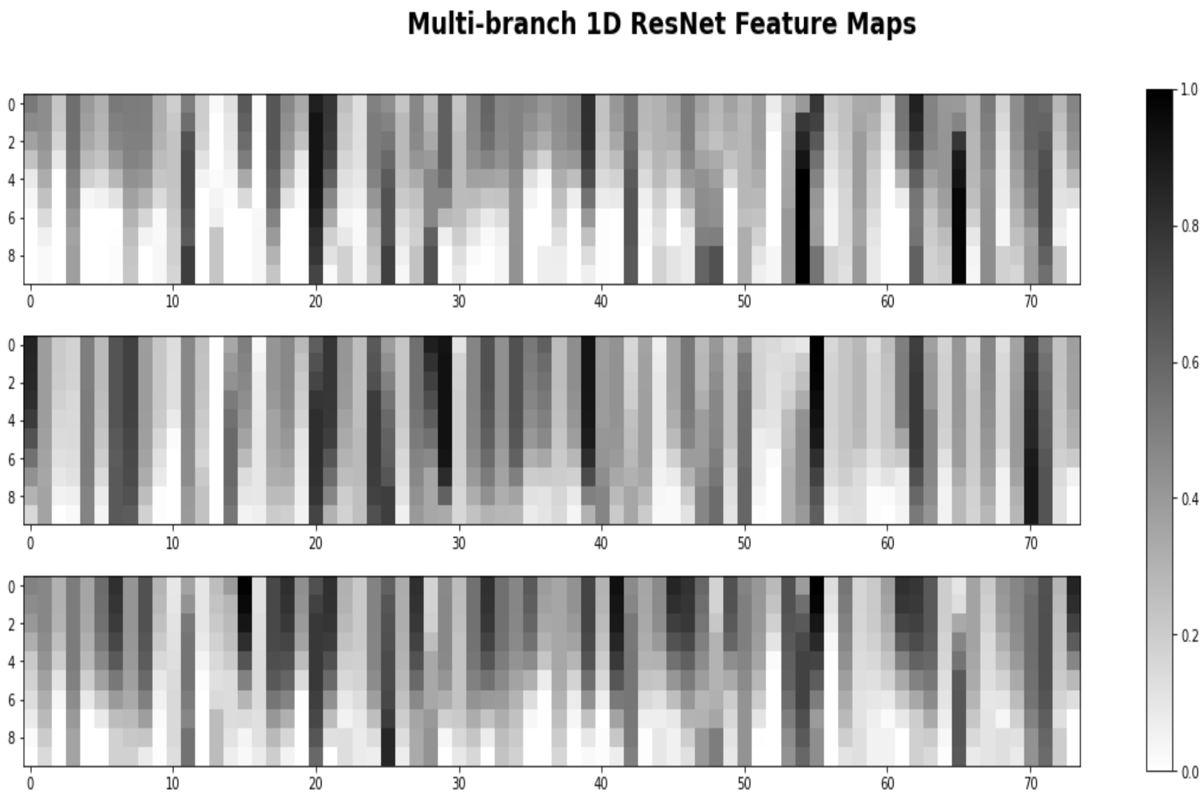


Experimental Results

Discussion

- 1D CNN Feature Maps Visualization





Conclusion

Summary

- Proposed multi-branch 1D residual block based on 1D convolution and 1D ResNet

- Proposed method perform significantly better than conventional RNN-based models

- Explicitly split multi-branch convolution can drive high-correlated feature learning

Future Works

- Using track more audio features

- Applying grouped convolution (separate inputs into same sizes)

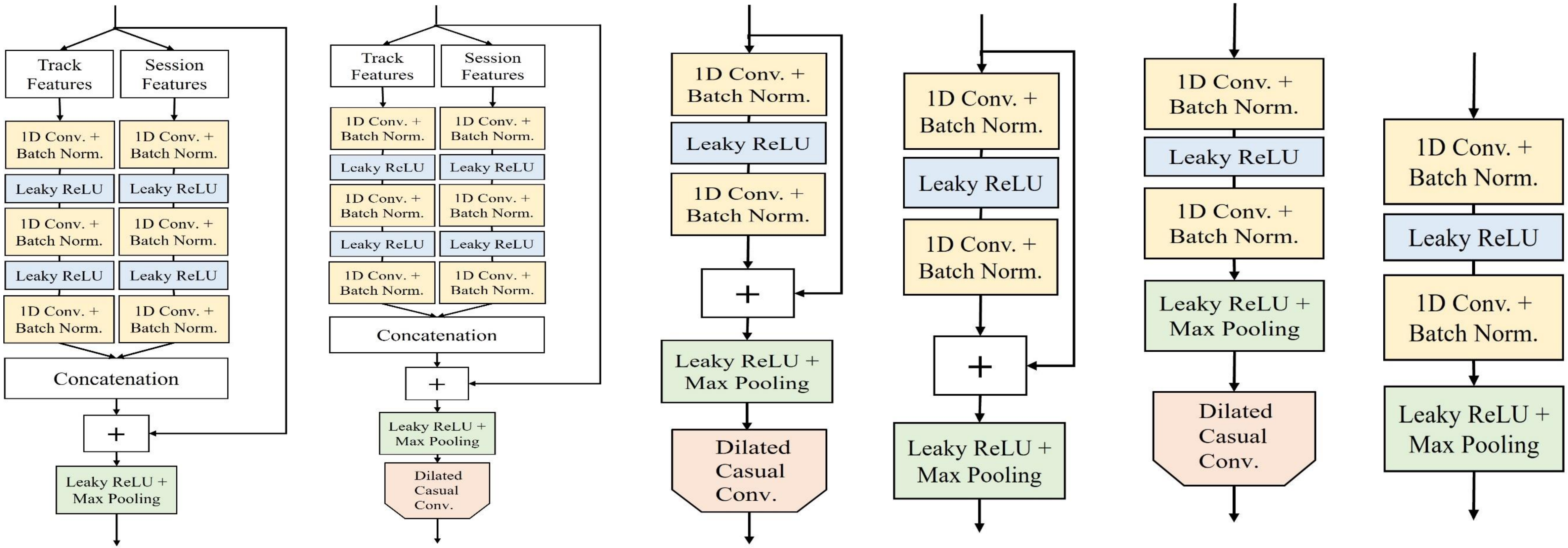
ICEIC 논문 관련

- 2~4 페이지
- 현재 진행 내용 간략히 정리
- Appendix의 variation 비교결과 사용?

Appendix

1D ResNet Variation Comparison

* DCC : Dilated Casual Convolution



1. Proposed 2. Multi-branch 1D ResNet + DCC 3. Basic 1D ResNet + DCC 4. Basic 1D ResNet 5. Residual Conn. Removed + DCC 6. Residual Conn. Removed

Appendix

t-test Result

Model	MAA Mean ± Standard Deviation	Average Rank
Proposed	**0.6625 ± 0.0427**	**1.6**
Multi-Branch 1D ResNet + DCC	0.6659 ± 0.0402	1.4
Basic 1D ResNet + DCC	0.5399 ± 0.0242 ▼	3.67
Basic 1D ResNet	0.5418 ± 0.0339 ▼	3.67
w/o Residual + DCC	0.4615 ± 0.0434 ▼	5.27
w/o Residual	0.4616 ± 0.0431 ▼	5.4

▼ indicates corresponding model is significantly worse than proposed model at significance level α=0.05

* DCC : Dilated Casual Convolution

Appendix

Friedman-test Result

	Statistics	p-value
Proposed	126.914	1.08e-25

significance level α=0.05

Wilcoxon signed-rank test Result

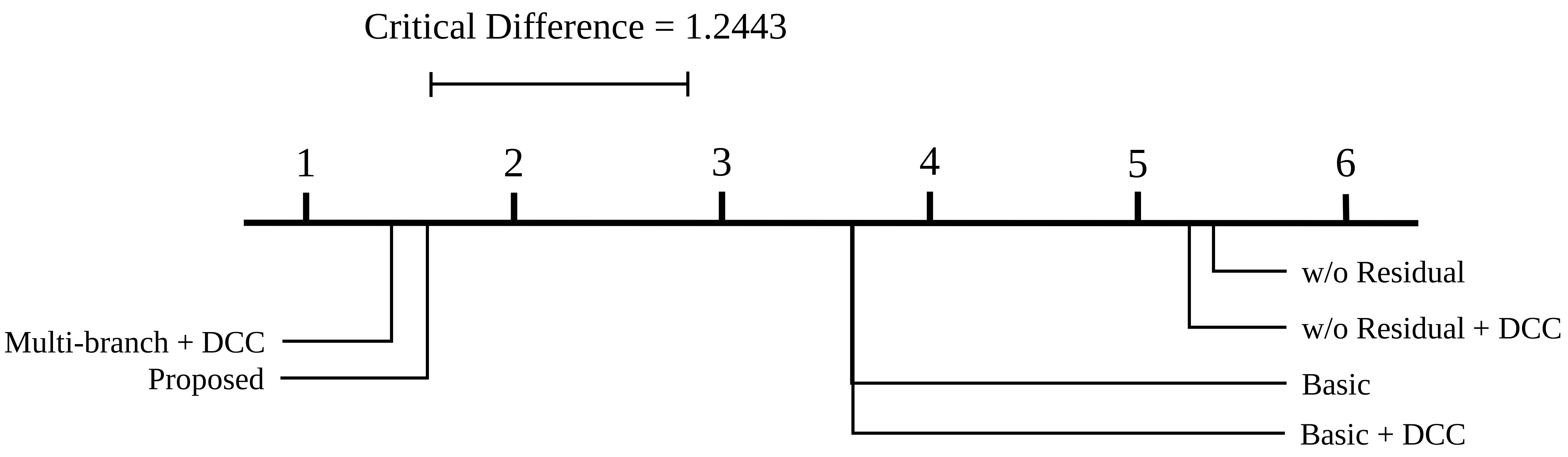
* DCC : Dilated Casual Convolution

	Comparison models				
	Multi-Branch 1D ResNet + DCC	Basic 1D ResNet + DCC	Basic 1D ResNet	w/o Residual + DCC	w/o Residual
Proposed ***versus*** (p-value)	Tie (0.5999)	Win (1.73e-06)	Win (1.73e-06)	Win (1.73e-06)	Win (1.73e-06)

significance level α=0.05

Appendix

Bonferroni-Dunn test Result



* DCC : Dilated Casual Convolution

References

[1] Brian Brost, Rishabh Mehrotra, Tristan Jehan. “The Music Streaming Sessions Dataset”, In Proceedings of the ACM Web Conference. 2019.

[2] Aaron Ng et al. "Investigating the Impact of Audio States & Transitions for Track Sequencing in Music Streaming Sessions." 14th ACM Conference on Recommender Systems. 2020.

[3] Ole Adrian Heggli et al. "Diurnal fluctuations in musical preference." Royal Society open science 8.11. 2021.

[4] Jonathan Donier, “The universality of skipping behaviours on music streaming platforms”, arXiv preprint arXiv:2005.06987, 2020.

[5] Serkan Kiranyaz et al. "1D convolutional neural networks and applications: A survey." Mechanical systems and signal processing 151. 2021.

[6] Federico Simonetta et al. "Multimodal music information processing and retrieval: Survey and future challenges." 2019 international workshop on multilayer music representation and processing (MMRP). IEEE, 2019.

[7] Francesco Meggetto, et al. "On skipping behaviour types in music streaming sessions." Proceedings of the 30th ACM International Conference on Information & Knowledge Management. 2021.

[8] Christian Szegedy et al. "Going deeper with convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

[9] Francois Chollet. "Xception: Deep learning with depthwise separable convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[10] Saining Xie et al. "Aggregated residual transformations for deep neural networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[11] Casper Hansen et al. "Contextual and sequential user embeddings for large-scale music recommendation.“14th ACM conference on recommender systems. 2020.

[12] Klaus Greff, Rupesh K. Srivastava, Jurgen Schmidhuber. “Highway and Residual Networks learn Unrolled Iterative Estimation”, ICLR 2017 Poster. 2017.

[14]] A. Oord, S. Dieleman, H. Zen, et al. “Wavenet: A generative model for raw audio”. arXiv:1609.03499, 2016.

[15] Sakurai, Keigo, et al. "Music playlist generation based on reinforcement learning using acoustic feature map." 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE). IEEE, 2020.

[16]] Darius Afchar, and Romain Hennequin. "Making neural networks interpretable with attribution: application to implicit signals prediction." 14th ACM Conference on Recommender Systems. 2020.

[17] Christian Hansen, Casper Hansen, et al. “Modelling sequential music track skips using a Multi-RNN approach”, 12th ACM International Conference on Web Search and Data Mining. 2019.

[18] Olivier Jeunen, Bart Goethals. “Predicting Sequential User Behavior with Session-based Recurrent Neural Networks”, 12th ACM International Conference on Web Search and Data Mining. 2019.

[19] Sainath Adapa. “Sequential modeling of Sessions using Recurrent Neural Networks for Skip Prediction”, 12th ACM International Conference on Web Search and Data Mining. 2019.

[20] Geoffray Bonnin, and Dietmar Jannach. "Automated generation of music playlists: Survey and experiments." ACM Computing Surveys (CSUR). 2014

Gastrointestinal Track Polyp Segmentation

Goeun Kim

goeun678@gmail.com

19th Oct, 2022

Contents

To Do List from Previous Seminar

- ✓ **To do list 1**
  - ➢ GI tract polyp segmentation Public dataset 조사
- ✓ **To do list 2**
  - ➢ Medical Image polyp segmentation Deep Learning model 조사 , Paper review
- ✓ **To do list 3**
  - ➢ GI Tract Segmentation model: FCBFormer 학습

Overall Progress

Problem

- Gastrointestin track Polyp Segmentation
- Train result
- Low accuracy

Related Works/Survey (counterparts)

- ESFPNet
- SSFormer
- ColonFormer
- GMSRF-Net

Experiments

- FCBFormer
- Kvasir-SEG datasets

Analysis

- Training monitoring
- Wrong location

Proposal

- Tensorboard
- Data augmentation
- Use more data

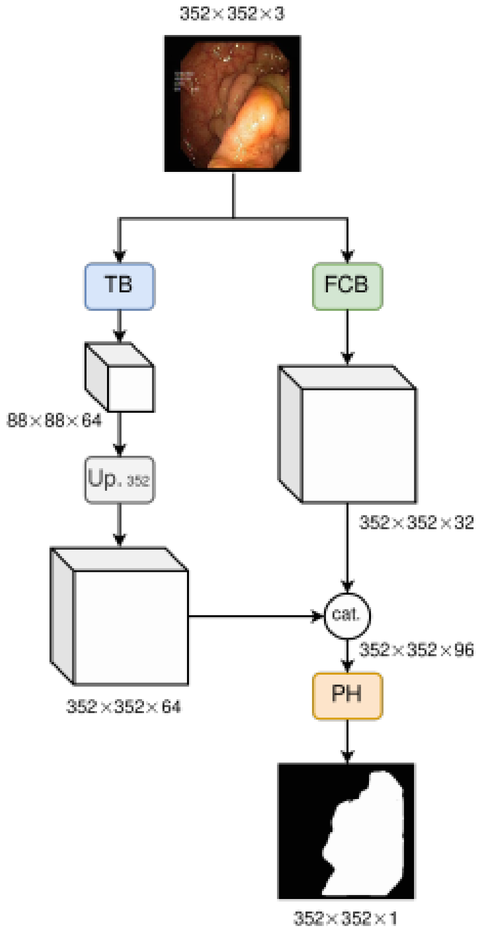
Post-experimental analysis

Conclusion

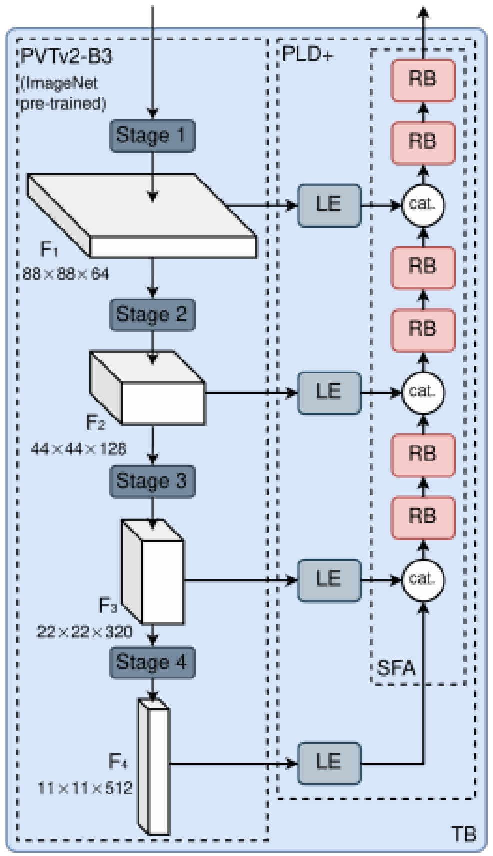
- Conclusion

Introduction

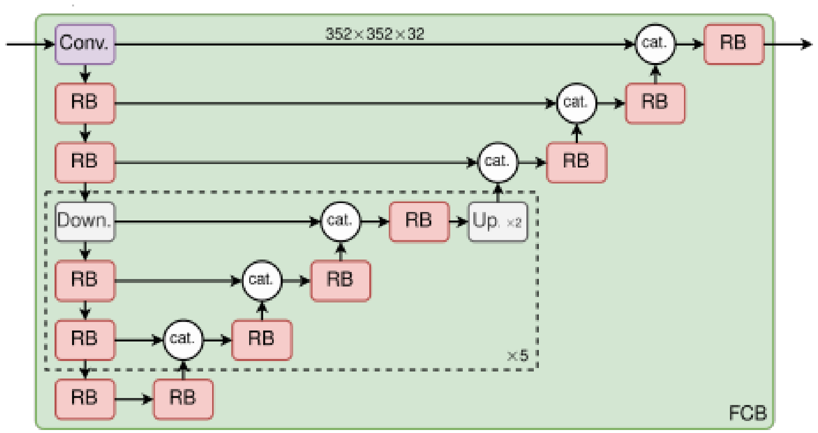
FCN-Transformer Feature for Polyp Segmentation



The architectures of FCBFormer



TB (Transformer Branch)



FCB (Fully convolutional Branch)

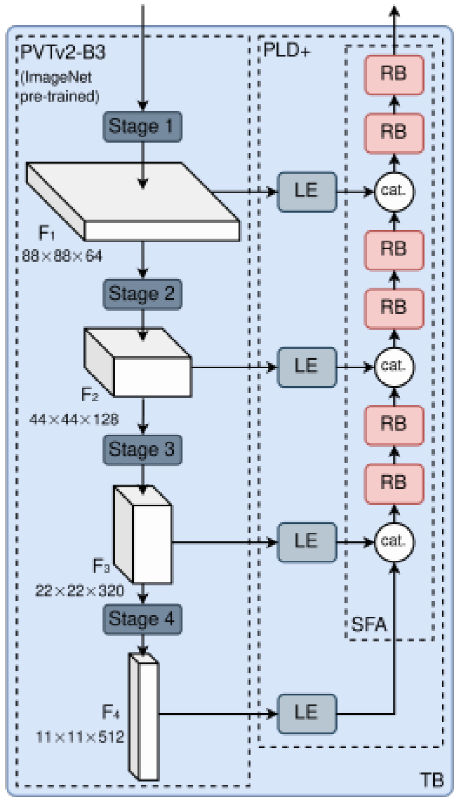
The architecture, named the **Fully Convolutional BranchTransFormer (FCBFormer)** uses two parallel branches which both start from a "h × w" input image:

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

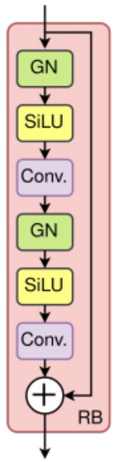
FCN-Transformer Feature Fusion for Polyp Segmentation (Image and Video Processing (eess.IV), Computer Vision and Pattern Recognition (cs.CV), Machine Learning(cs.LG)2022)

Introduction

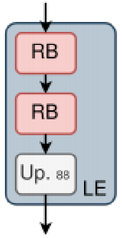
FCN-Transformer Feature for Polyp Segmentation



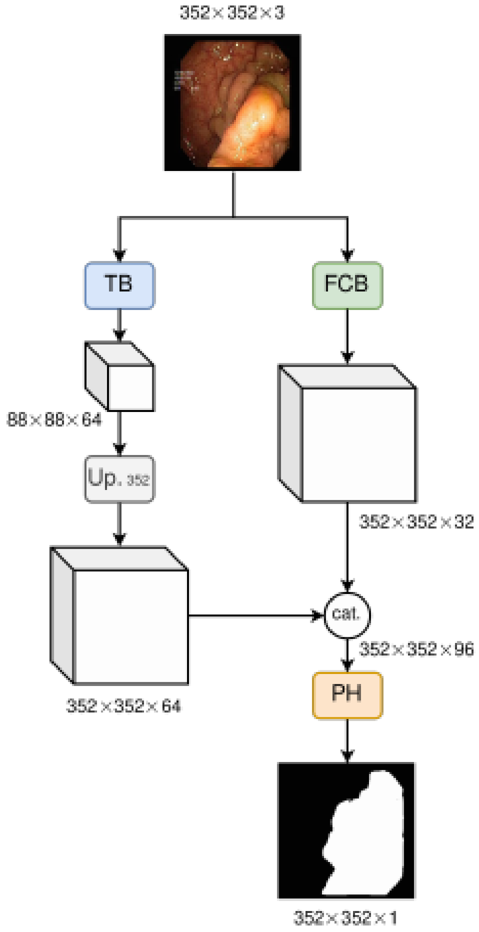
TB (Transformer Branch)



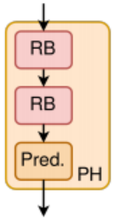
RB (Residual Block)



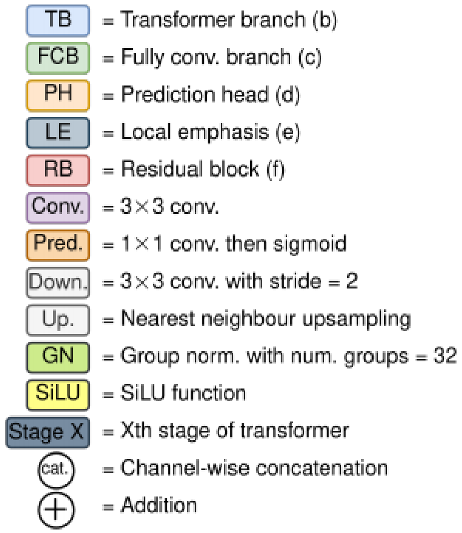
LE (Improved Local Emphasis)



The architectures of FCBFormer



PH (Prediction Head)



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)

Related Work

Experimental Results

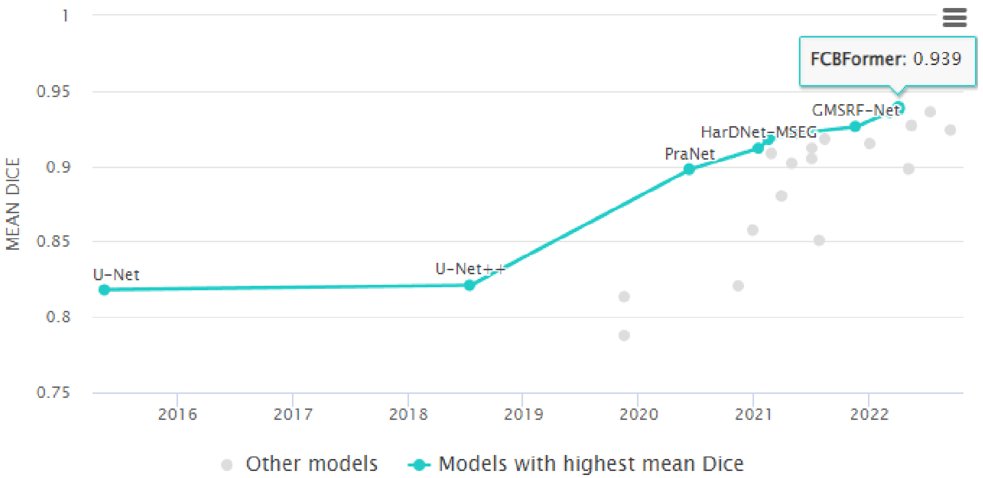


Figure 1. Experimental Results for Kvasir-SEG Dataset

Trend	Dataset	Best Model
	Kvasir-SEG	FCBFormer
	CVC-ClinicDB	ESFPNet-L
	CVC-ColonDB	ResUNet++ + TTA
	ETIS-LARIBPOLYPDB	ESFPNet-L
	Synapse multi-organ CT	Swin UNETR
	MoNuSeg	LViT-L
	2018 Data Science Bowl	SSFormer-L
	Automatic Cardiac Diagnosis Challenge (ACDC)	nnFormer
	GlaS	HistoSeg
	Medical Segmentation Decathlon	Swin UNETR

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

Related Work

Medical Image Segmentation Model Survey

	저자	연도	저널	제목	요약	별칭
1	Edward Sa	2022.08	Recognitio n (cs.CV), Machine Learning (cs.LG)	Feature Fusion for Polyp Segmentat ion	we propose a new architecture for full-size segmentation which leverages the strengths of a transformer in extracting the most important features for segmentation in a primary branch, while compensating for its limitations in full-size prediction with a	FCBFormer
	Qi Chang 외	2022.07	Image and	autofluore scence bronchosc opic video	We propose a framework for synchronizing multiple bronchoscopic videos to enable interactive multimodal analysis of bronchial lesions. Our methods first register the video streams to a reference 3D chest computed-tomography (CT) scan to produce	ESFPNet-L
3		2022.06	Pattern Recognitio n (cs.C	Local Guides Global	we propose a new State-Of-The-Art model for medical image segmentation, the SSFormer, which uses a pyramid Transformer encoder to improve the generalization ability of models. Specifically, our proposed Progressive Locality Decoder can be adap	SSFormer-L
4	Nguyen Th	2022.06	Computer Vision and Pattern Recognitio n (cs.CV)	Method for Colon Polyp Segmentat ion	This paper proposes a novel network, namely ColonFormer, to address these limitations. ColonFormer is an encoder-decoder architecture capable of modeling long-range semantic information at both encoder and decoder branches. The encoder is a lig	ColonFormer
5	Abhishek	2021.11	Computer Vision and Pattern Recognitio n (cs.CV)	fusion network for polyp segmentati on	Our proposed network maintains high-resolution representations while performing multi-scale fusion operations for all resolution scales. To further leverage scale information, we design cross multi-scale attention (CMSA) and multi-scale feature selecti	GMSRF-Net

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