기계학습자동화 연구실
주간 정기 세미나

이상혁, 조성현, 한수정, 송태경, 오해성

2022. 03. 08.

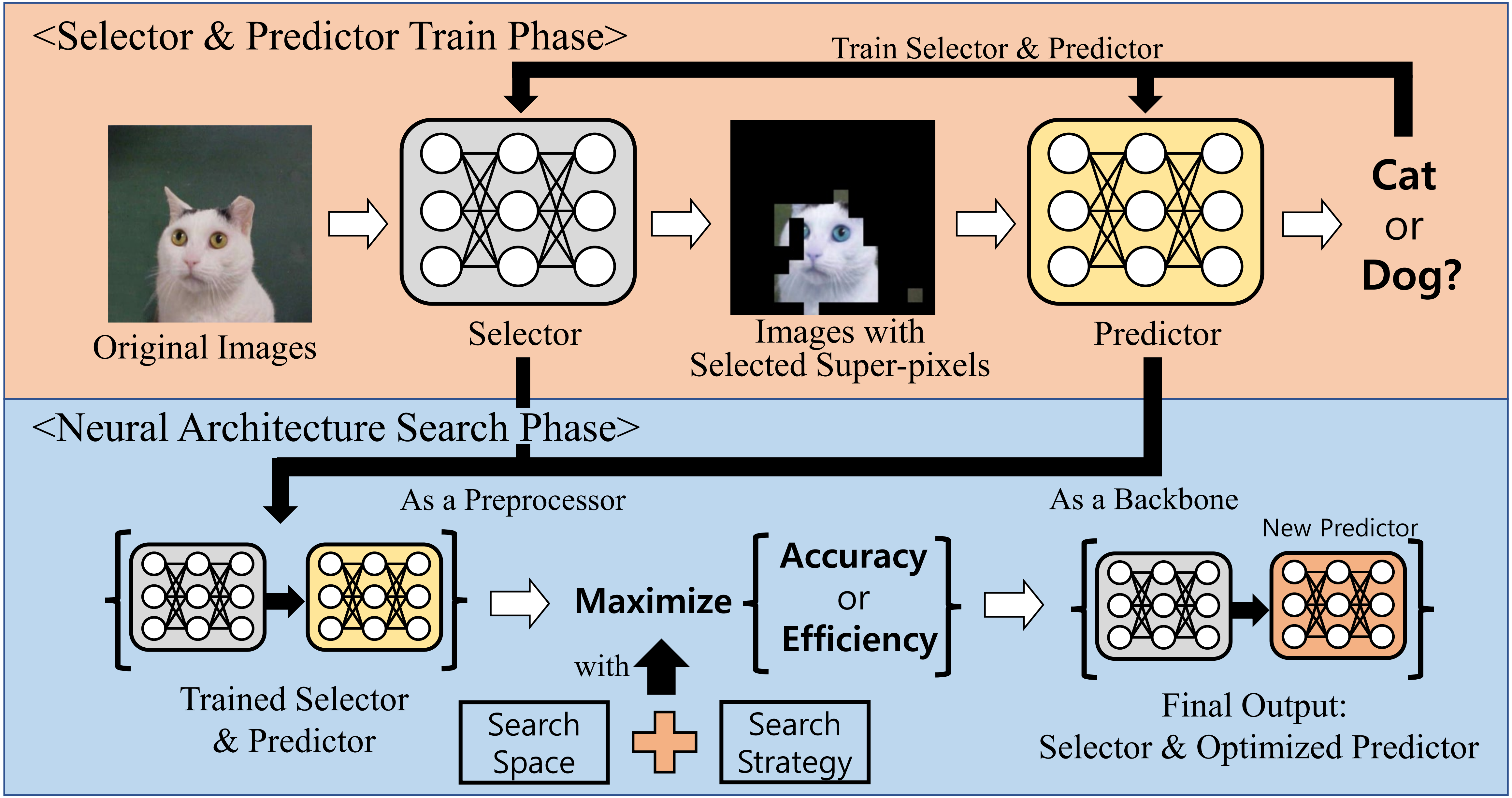
Outline

- **Individual Studies**
  - Sanghyuck Lee
  - Sung-Hyun Cho
  - Su-Jeong Han
  - Taekyung Song
  - Haesung Oh
- **Government-funded Research Projects**
  - 갑상선 연구(이상혁, 한수정, 송태경, 오해성)
    - ➢ 지난 보고 : CT 영상 기반 Activity 추론 연구 실험 최종 보고 및 시각화 보고
    - ➢ 이번 보고 : 사진 기반 CAS 점수 추론 연구 Baseline 실험 보고
  - 산학협력프로젝트(연구실 전체)
    - ➢ 지난 보고 : 간단한 진행 내용 보고
    - ➢ 이번 보고 : 간단한 진행 내용 보고

Individual Studies:

Sanghyuck Lee

Neural Architecture Search based on Super-pixel Selection



Selector Method Comparison

- Feature Selection Methods: INVASE, SCM, Real-X
- Dataset: CIFAR-10
- Predictor: VGG16
- 7 Iterations

Table 3. Metrics with 95% confidence interval (▼/△ indicates that the corresponding method is **significantly worse** than the best model written in bold based on paired t-test at 5% significance level).

FS Method	Predictor AUROC			Accruracy		
	Average	**95% CI Lower Upper**	**p-value**	**Average**	**95% CI Lower Upper**	**p-value**
Real-X	0.946	0.944 0.948		0.697	0.690 0.704	
SCM	0.799	0.759 0.839	0 (**▼)	0.41	0.347 0.473	0 (**▼)
INVASE	0.764	0.578 0.949	0.052 (*▼)	0.478	0.174 0.782	0.126
CI, confidence interval; *: p-value < 0.05; **: p-value < 0.01						

Qualitative Comparison

- INVASE:
  - Long time
  - Poor implementation detail for image data in the paper
  - FS does not work often (In many cases, select all features or all features)

- SCM:
  - Low Performance, need detail hyper-parameter setting
  - Fixed number of features

- Real-X:
  - Top Performance, but need to check experiment setting mistake
  - Tensorflow 2 implementation (others are Pytorch)

Next Step

- Approach 1: Convert to Pytorch

  - High understanding of the code and algorithm

  - Overall research delay

- Approach 2: Not convert, but keep Tensorflow

  - Apply FS to various datasets as the next step directly

  - Difficulty in application with PyTorch-based code

Individual Studies:

Sung-Hyun Cho

Progress

- Last Report
  - Comparison result
  - Dataset Integrity

- This Report
  - Comparison result

Related Work 1

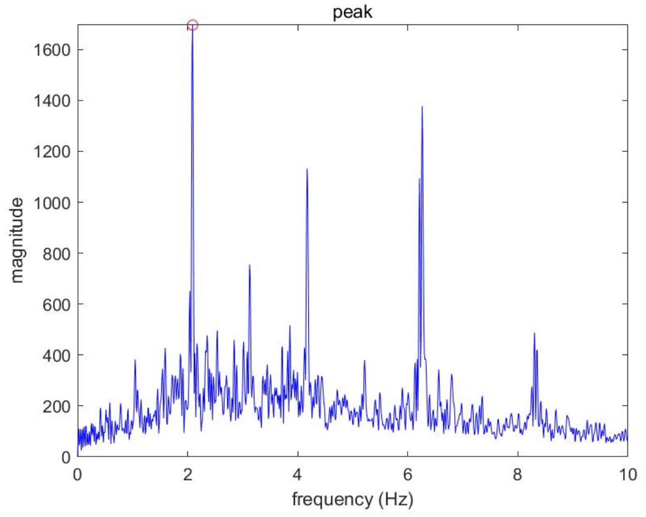
Author	Year	Neural Network	# of employed datasets	Employed Dataset	Input Type	Evaluation Metric	Ref.
Massoudi	2021	CNN	1	Urban Sound Classification Challenge	MFCC	Accuracy	[1]
Ghildiyal	2020	CNN	1	GTZAN	Mel-Spectrogram	Accuracy	[2]
Palanisamy	2020	CNN	3	ESC-50, GTZAN, UrbanSound8K	Log Mel-Spectrogram	Accuracy	[3]
Yu	2019	CNN	2	GTZAN, Extended Ballroom	STFT Spectrogram	Accuracy	[4]
Won	2020	CNN	3	MagnaTagATune, Million Song Dataset, MTG-Jamendo	Mel-Spectrogram	ROC-AUC, PR-AUC	[5]
Bisharad	2019	CNN	1	GTZAN	Mel-Spectrogram	Precision, Recall, F1 score	[6]
Zhang	2019	RNN	1	GTZAN	MFCC, chromagram	Accuracy	[7]
Chillara	2019	CNN	1	FMA-SMALL	Spectrogram	Accuracy	[8]
Zhang	2016	CNN	1	GTZAN	STFT Spectrogram	Accuracy	[9]
Choi	2017	CRNN	1	Million Song Dataset	Log Mel-Spectrogram	ROC-AUC	[10]
Su	2019	CNN	1	UrbanSound8k	Log Mel-Spectrogram, MFCC	Accuracy	[11]
Choi	2016	CNN	2	MagnaTagATune, Million Song Dataset	Mel-Spectrogram, MFCC, STFT	ROC-AUC	[12]
Li	2018	CNN	3	ESC-10, ESC-50, UrbanSound8k	Log Mel-Spectrogram	Accuracy	[13]
ASHRAF	2020	CNN-RNN	2	GTZAN, FMA-SMALL	Mel-Spectrogram	Accuracy, Precision, Recall, F1-Score	[14]

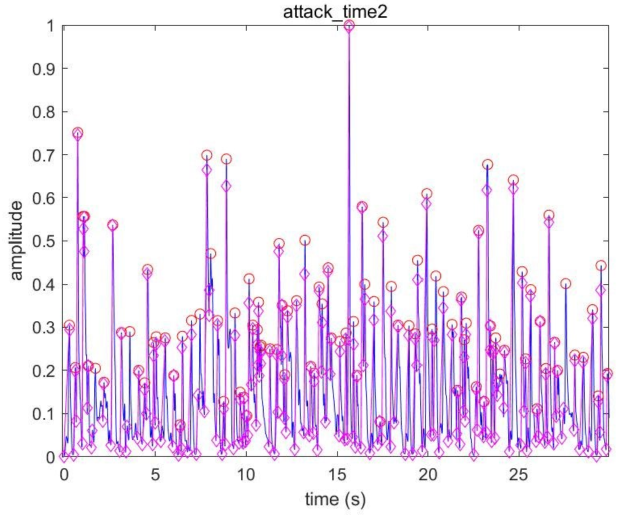
Related Work 2

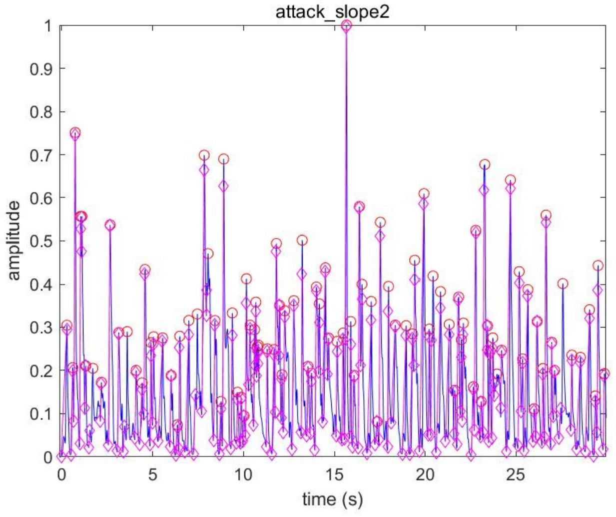
Author	Year	Neural Network	# of employed datasets	Employed Dataset	Input Type	Evaluation Metric	Ref.
Li	2021	CNN	4	FMA, GTZAN5, GTZAN10, EMA	STFT Spectrogram	Accuracy	[15]
Elbir	2018	CNN	1	GTZAN	Raw Audio, STFT Spectrogram, MFCC	Accuracy, Precision, Recall, F1-Score	[16]

MIRToolbox

X_axis = frequency	X_axis = time
Fluctuation	Rhythm



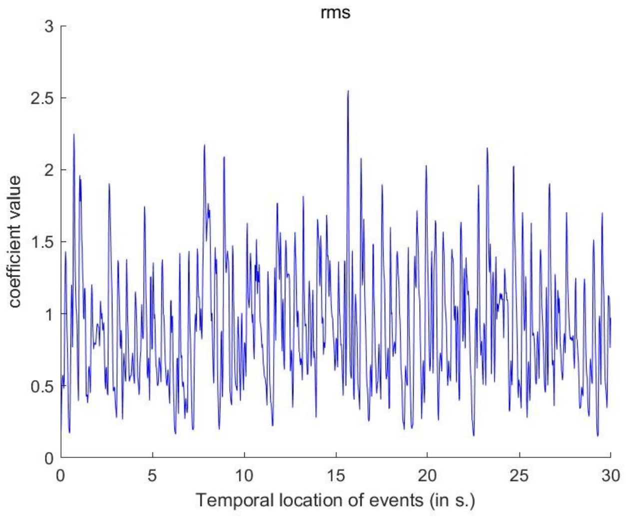




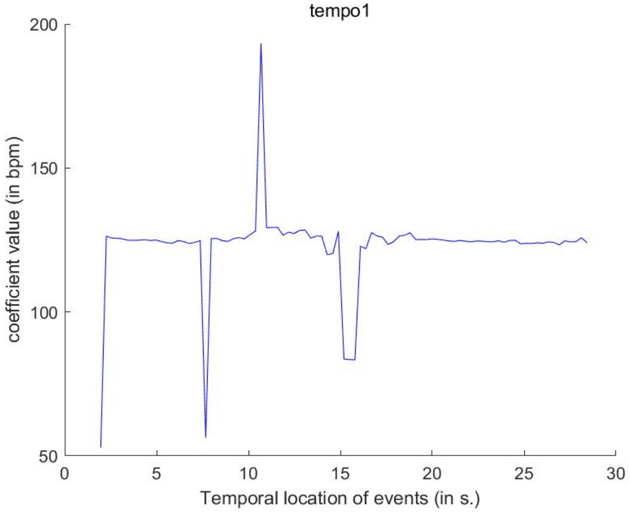
MIRToolbox

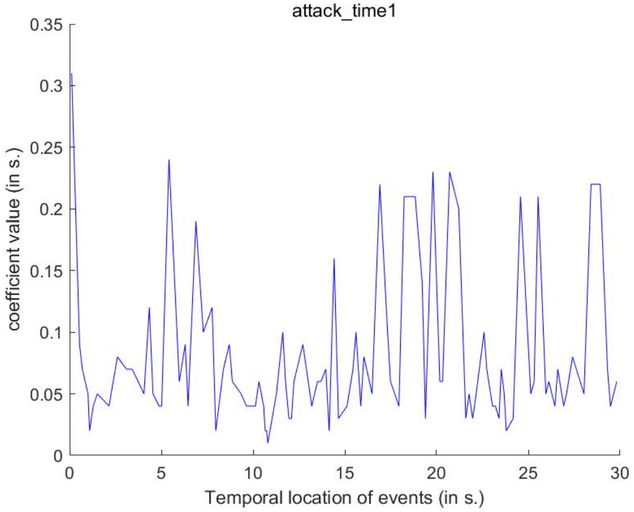
X_axis = Temporal location of events (in s.)

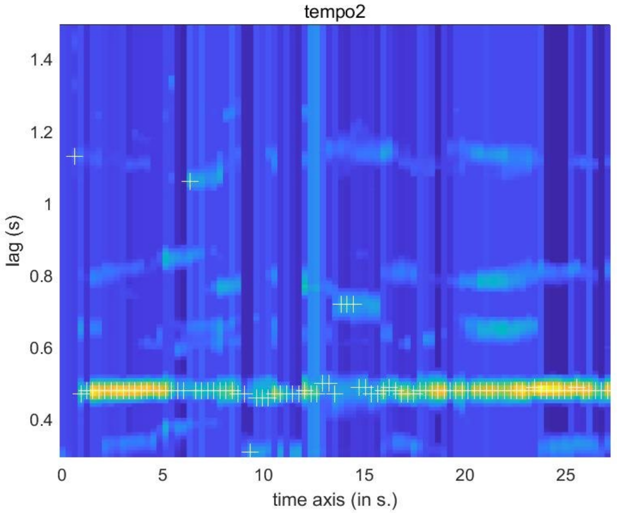
Dynamics

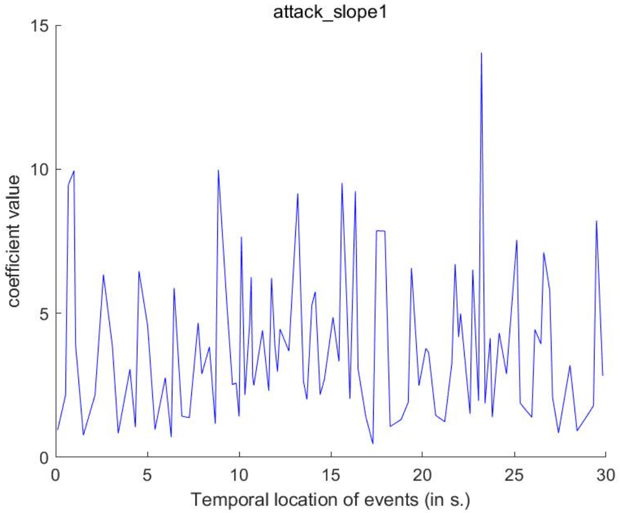


Rhythm





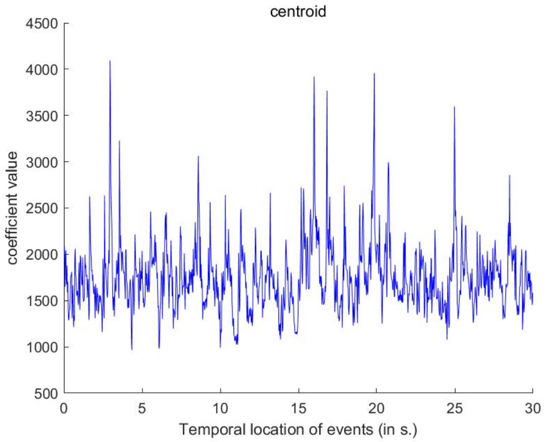


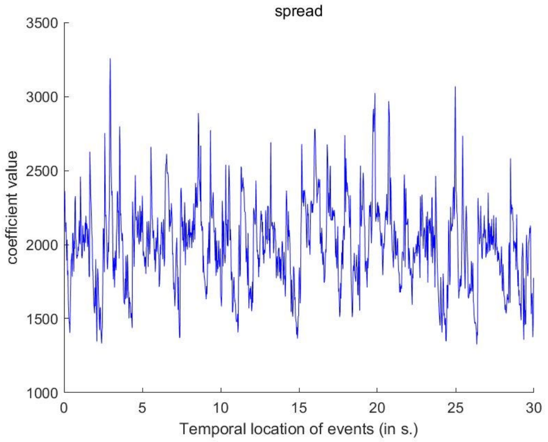


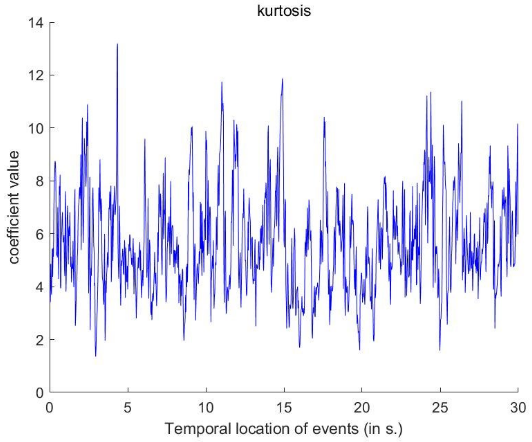
MIRToolbox

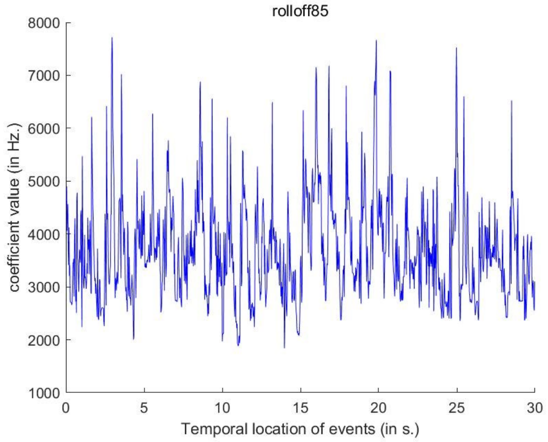
X_axis = Temporal location of events (in s.)

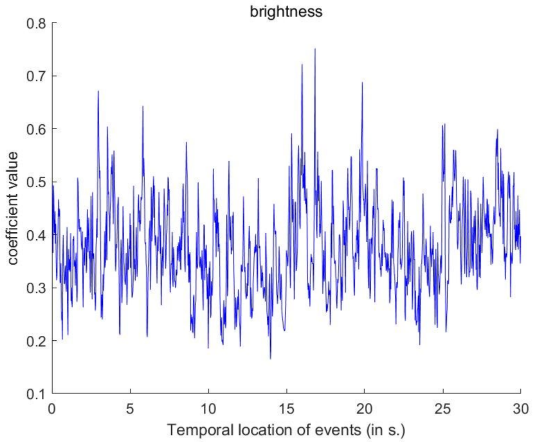
Spectral

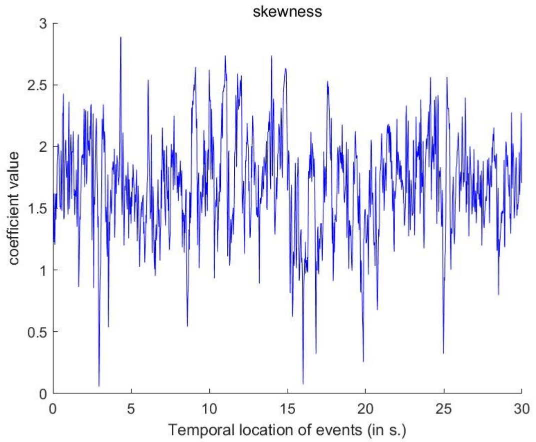


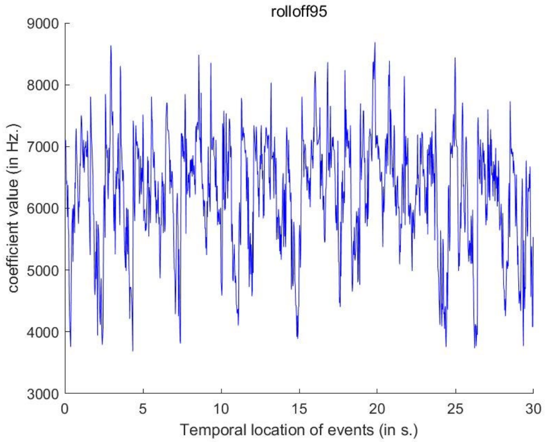








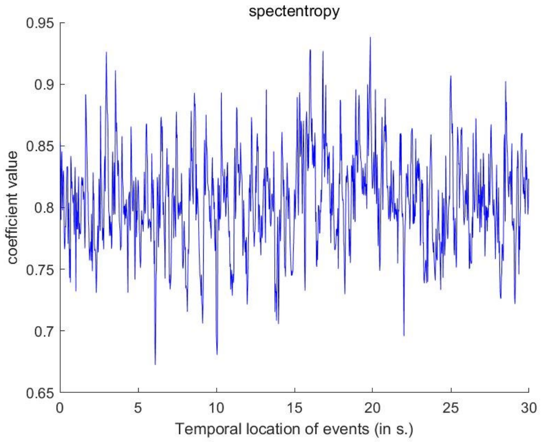


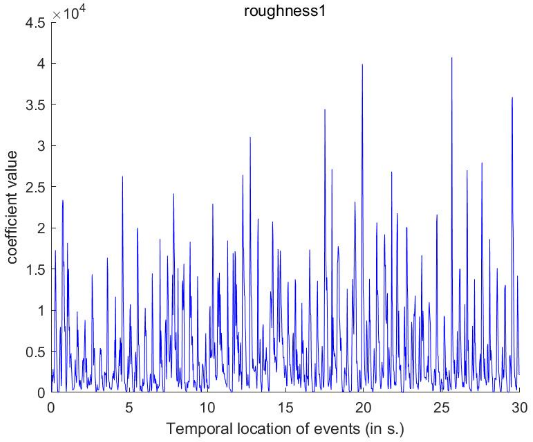


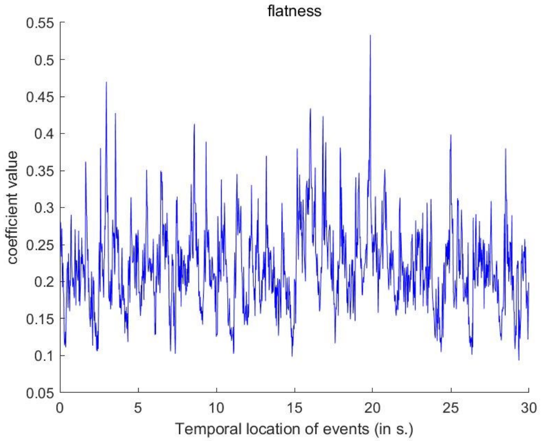
MIRToolbox

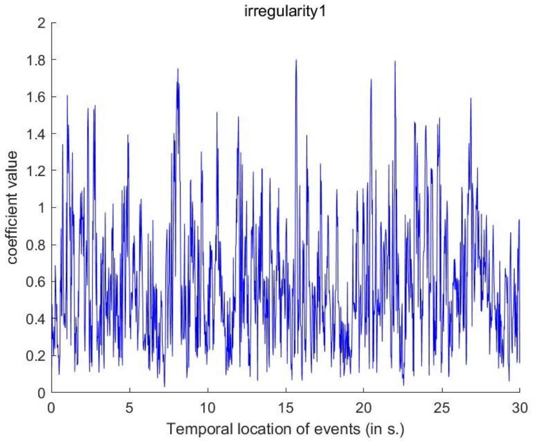
X_axis = Temporal location of events (in s.)

Spectral

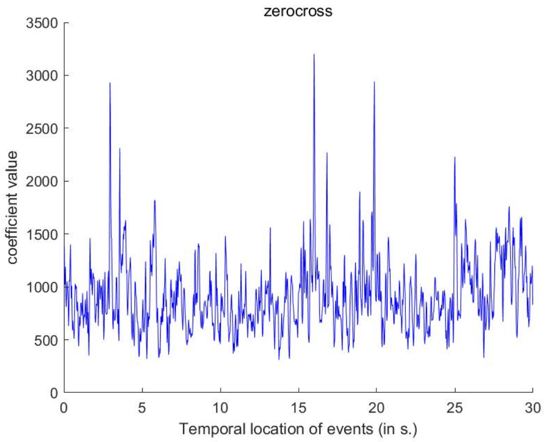


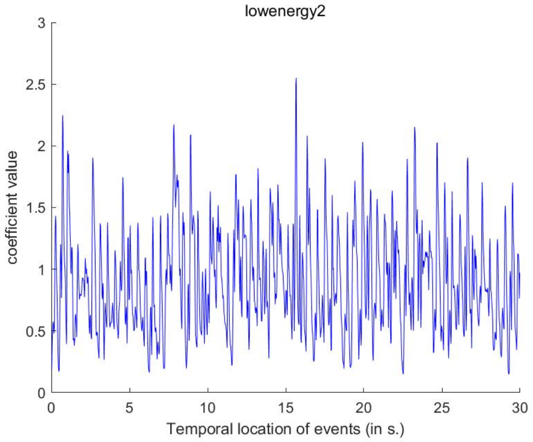


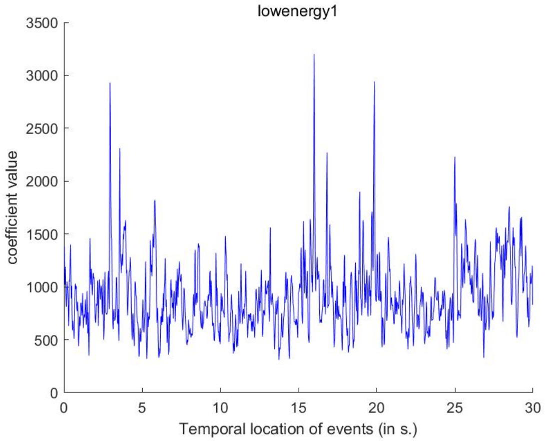


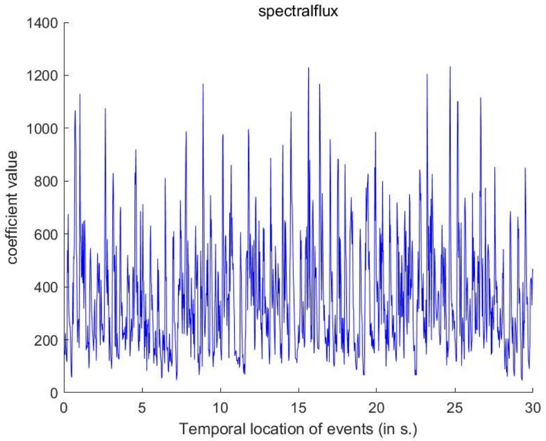


Timbre





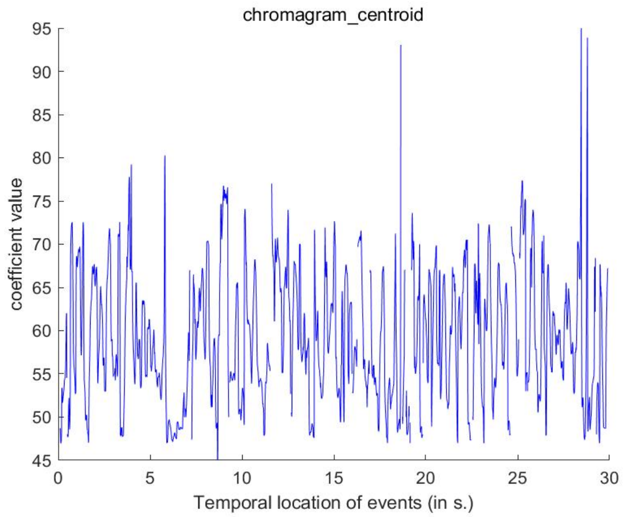


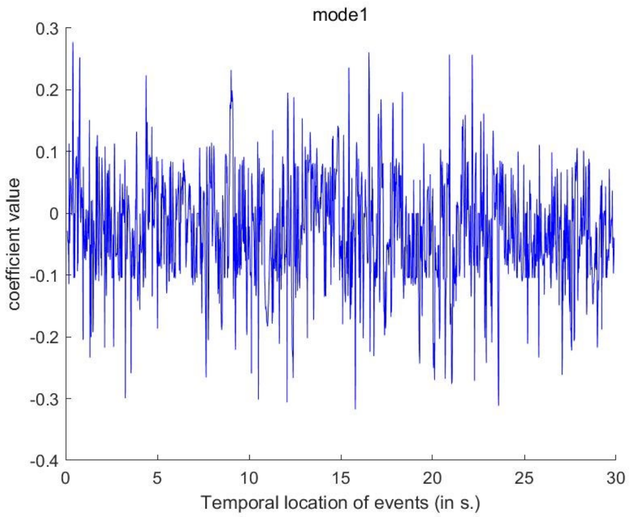


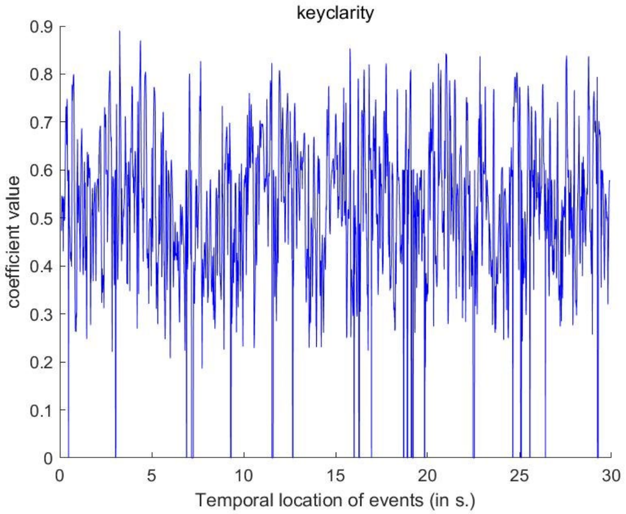
MIRToolbox

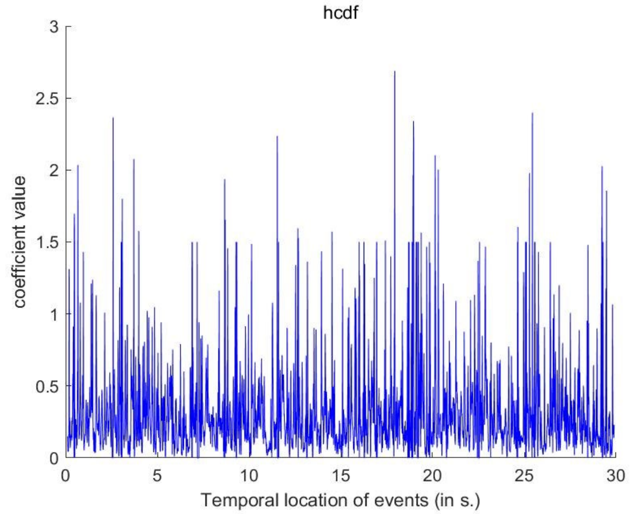
X_axis = Temporal location of events (in s.)

Tonal





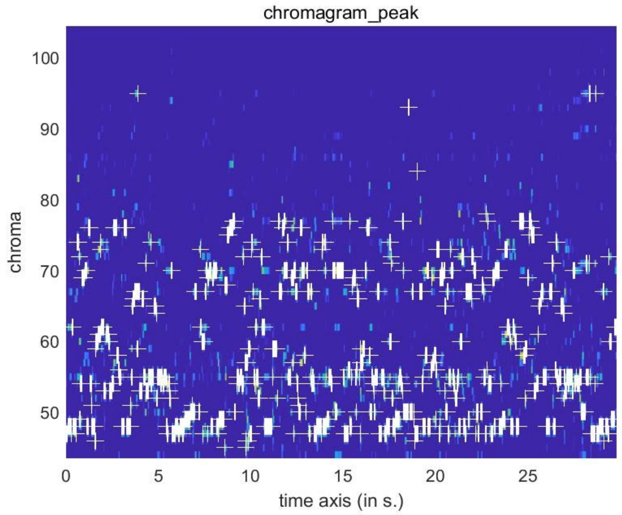


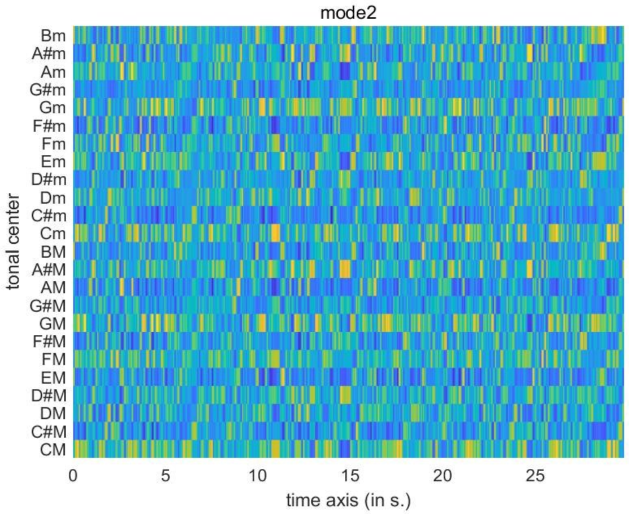


MIRToolbox

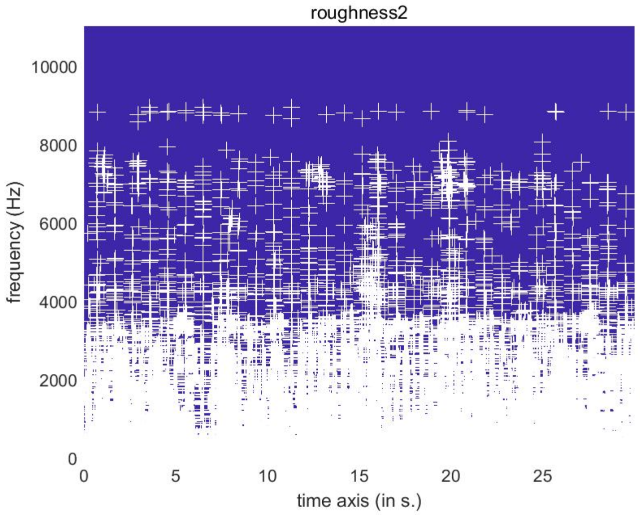
X_axis = time axis (in s.)

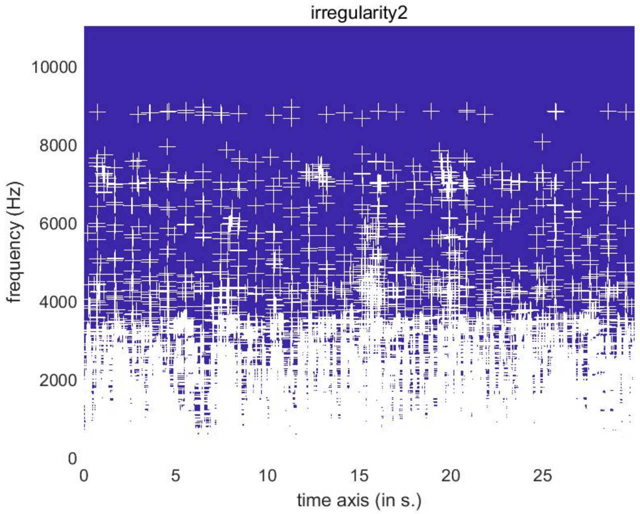
Tonal





Spectral

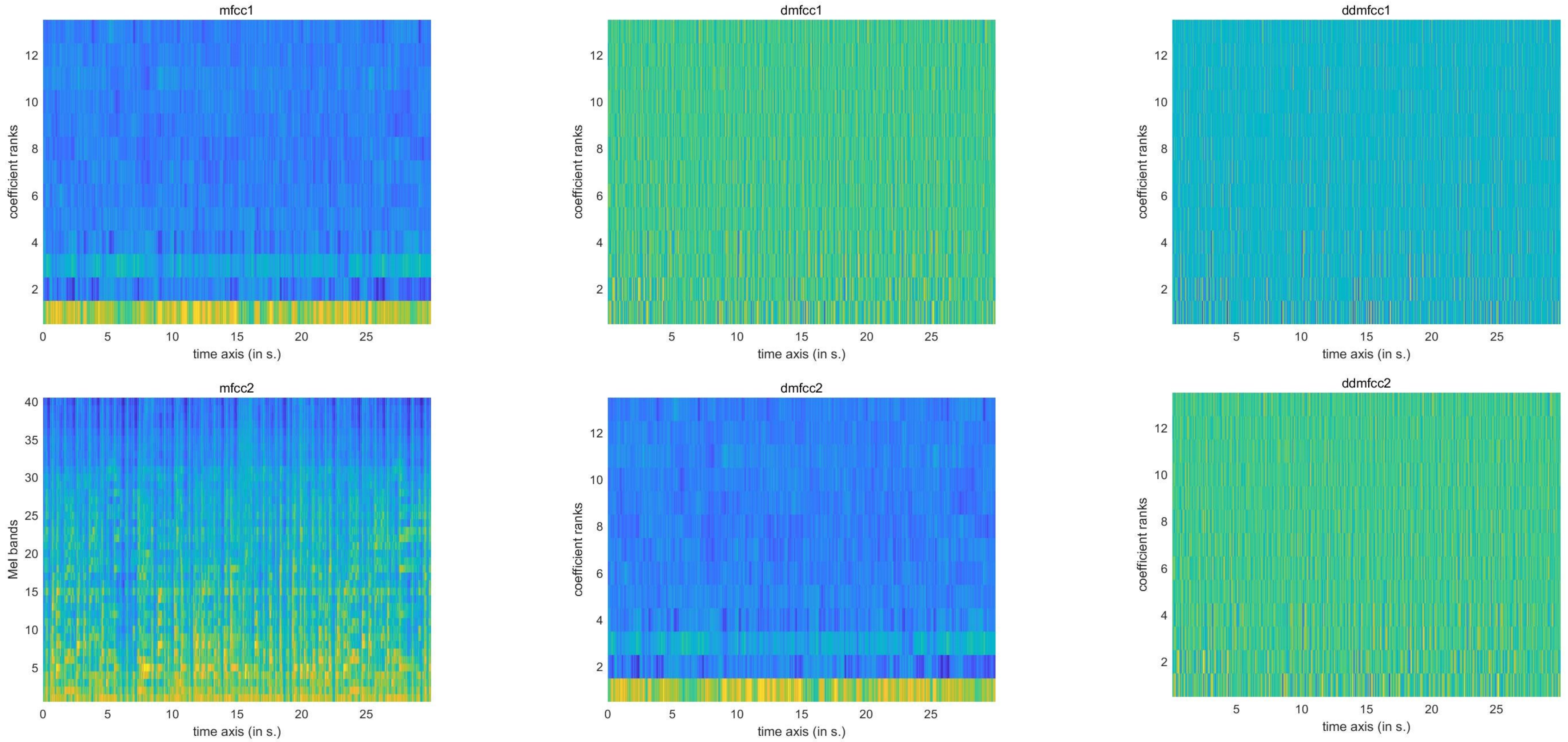




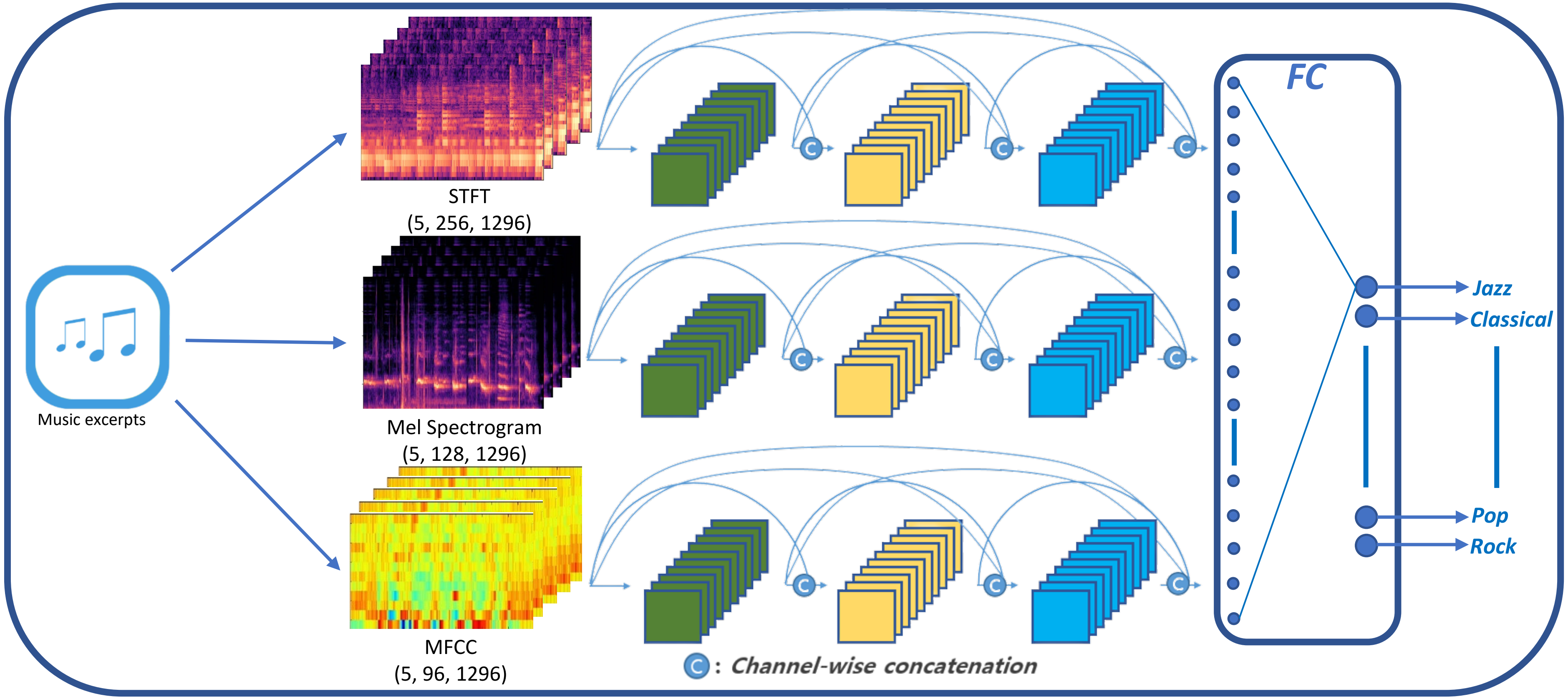
MIRToolbox

X_axis = time axis (in s.)

Spectral



Code Analysis(proposed)

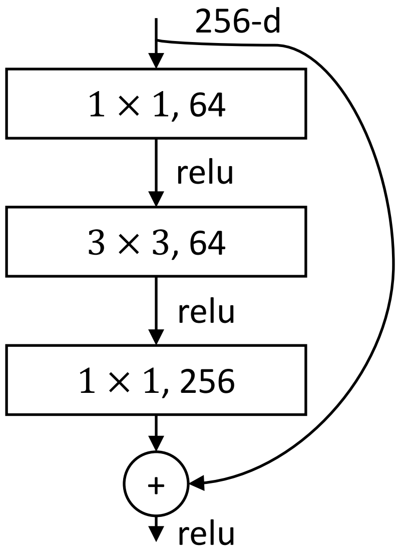


Optimizer : Adam optimization algorithm
Batch size : 32
Learning rate = 0.001
Training set : Test set = 8:2

Code Analysis(ResNet50_trust)

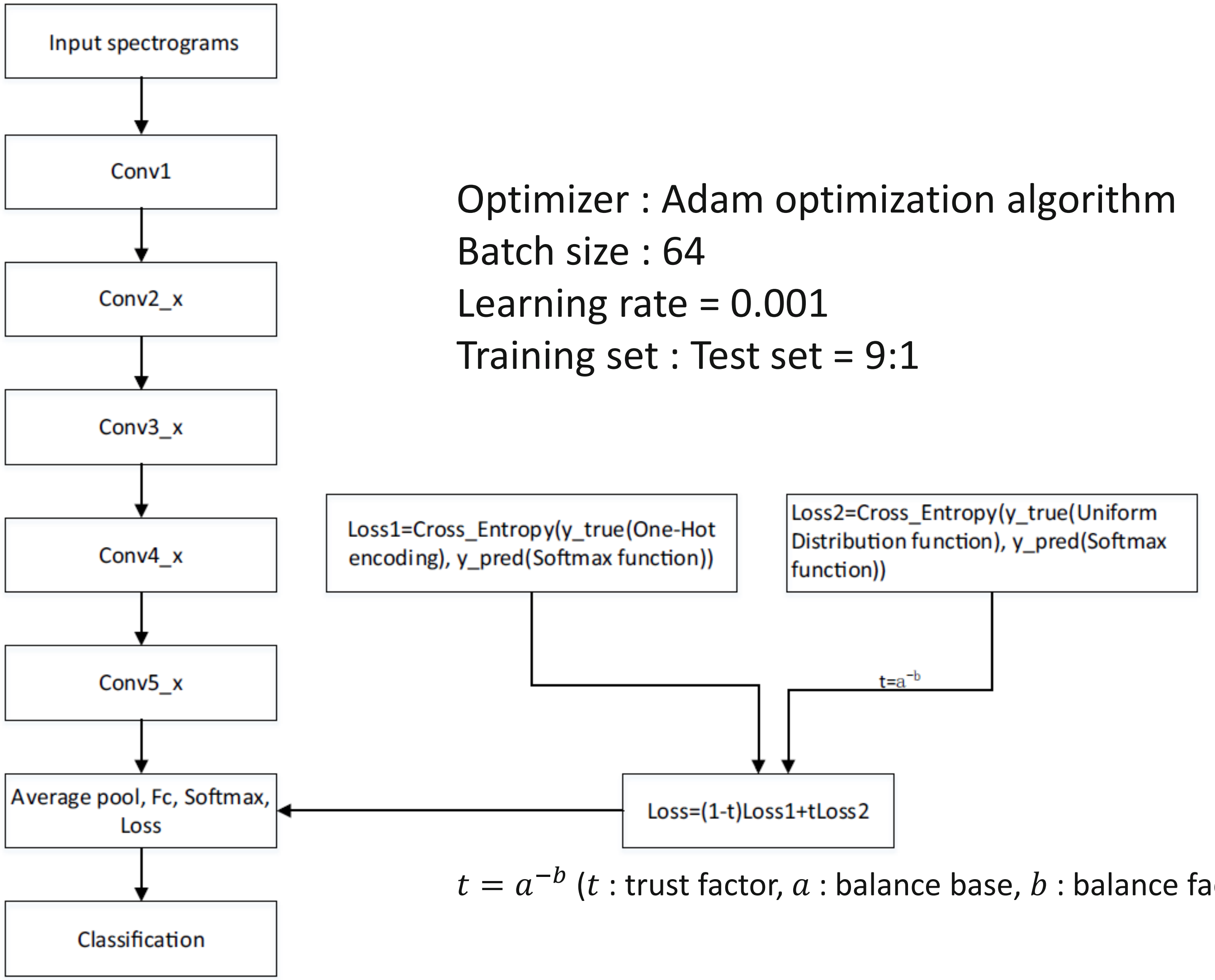
[15] Li, J., Han, L., Li, X., Zhu, J., Yuan, B., & Gou, Z. (2021). An evaluation of deep neural network models for music classification using spectrograms. *Multimedia Tools and Applications*, 1-27.

layer name	ouput size	50-layer
conv1	112×112	7×7, 64, stride 2
conv2 _x	56×56	3×3 max pool, stride 2 $\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
conv4_x	14×14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
conv5_x	7×7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax
FLOPs		3.8×10^9



A “bottleneck” building block for ResNet-50

Code Analysis(ResNet50_trust)



Optimizer : Adam optimization algorithm
Batch size : 64
Learning rate = 0.001
Training set : Test set = 9:1

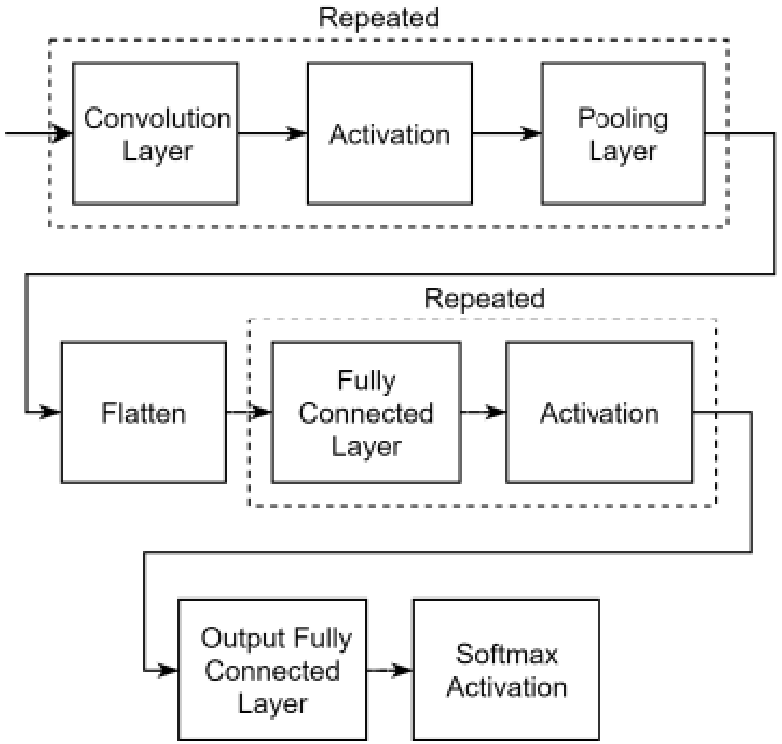
$t = a^{-b}$ (t : trust factor, a : balance base, b : balance factor)

Code Analysis(ResNet50_trust)

Dataset	Types(the number of music audios)	Duration of each audio	Types(the number of music spectrograms)	Total number of music spectrograms
FMA	Classical(300) Electronic(300) Jazz(306) Rock(300)	30s	Classical(1800) Electronic(1800) Jazz(1836) Rock(1800)	7236
GTZAN_4	Classical(100) Disco(100) Jazz(100) Rock(100)	30s	Classical(600) Disco(600) Jazz(600) Rock(600)	2400
GTZAN_10	Blues(100) Classical(100) Country(100) Disco(100) Hiphop(100) Jazz(100) Metal(100) Pop(100) Reggae(100) Rock(100)	30s	Blues(600) Classical(600) Country(600) Disco(600) Hiphop(600) Jazz(600) Metal(600) Pop(600) Reggae(600) Rock(600)	6000
EMA	Angry(160) Happy(188) Quiet(160) Sad(133)	60s	Angry(1574) Happy(1833) Quiet(1536) Sad(1307)	6250

Code Analysis(Elbir et al.)

[16] Elbir, A., Çam, H. B., Iyican, M. E., Öztürk, B., & Aydin, N. (2018, October). Music genre classification and recommendation by using machine learning techniques. In 2018 Innovations in Intelligent Systems and Applications Conference (ASYU) (pp. 1-5). IEEE.



Convolutional Neural Network Configuration

1 - 2D Convolution Layer, 5x5 sized 32 filters, LeakyReLU activation function

2 - Max Pooling Layer

3 - 2D Convolution Layer, 5x5 sized 32 filters, LeakyReLU activation function

4 - Max Pooling Layer

5 - 2D Convolution Layer, 5x5 sized 32 filters, LeakyReLU activation function

6 - Max Pooling Layer

7 - 2D Convolution Layer, 5x5 sized 32 filters, LeakyReLU activation function

8 - Average Pooling Layer

9 - Flatten Layer

10 - Dense Layer, 256 nodes, LeakyReLU activation function

11 - Dense Layer, 128 nodes, LeakyReLU activation function

12 - Dense Layer, 64 nodes, LeakyReLU activation function

13 - Dense Layer, 10 nodes, Softmax activation function, as output layer

Convolutional Neural Network Summarization

Dataset statistics

Dataset (Genre)	Patterns	# of Classes	Average Patterns per Class ± Standard Deviation	Avg. Length(s)
Ballroom	698	8	87.250 ± 17.555	30
Ballroom-Extended	4180	13	321.538 ± 195.702	30
FMA-MEDIUM	24984	16	1561.500 ± 2069.356	30
FMA-SMALL	7996	8	999.500 ± 0.500	30
GTZAN	1000	10	100.000 ± 0.000	30
HOMBURG	1886	9	209.556 ± 134.866	10
ISMIR04	727	6	121.167 ± 95.602	120
MICM	1137	7	162.429 ± 119.717	360
Seyerlehner-Unique	3115	10	311.500 ± 248.349	30

Comparison Results

Dataset (Genre)	Proposed	ResNet50_trust (t =2)	Elbir et al.	-	-	GA	Mel & MFCC
Ballroom	58.79 ± 3.96	48.86 ± 4.41	39.29 ± 3.82	-	-	$\mathbf{83.67 \pm 4.64}$	70.0 ± 3.70
Ballroom-Extended	$\mathbf{85.32 \pm 0.61}$	75.89 ± 3.09	39.07 ± 1.49	-	-	64.49 ± 1.41	81.7 ± 1.55
FMA-MEDIUM	-	-	-	-	-	51.98 ± 0.88	-
FMA-SMALL	51.69	-	37.41 ± 0.74	-	-	20.14 ± 1.56	56.3 ± 0.99
GTZAN	$\mathbf{87.70 \pm 3.01}$	74.50 ± 3.78	58.85 ± 3.79	-	-	75.70 ± 3.03	86.4 ± 2.14
HOMBURG	54.13 ± 1.91	$\mathbf{54.50 \pm 2.93}$	47.14 ± 1.89	-	-	31.17 ± 2.88	-
ISMIR04	80.07 ± 1.64	68.36 ± 1.61	62.28 ± 2.42	-	-	$\mathbf{89.89 \pm 1.97}$	-
MICM	$\mathbf{49.39 \pm 1.79}$	39.56 ± 2.53	39.12 ± 2.78	-	-	7.80 ± 3.20	40.0 ± 2.85
Seyerlehner-Unique	$\mathbf{71.03 \pm 1.49}$	63.88 ± 1.53	55.59 ± 3.45	-	-	65.59 ± 1.92	-
Avg . Rank	-	-	-	-	-	-	-

Individual Studies:

Su-Jeong Han

Progress

- Last Report

- This Report

- 지난 연구까지의 내용
- AX7 segmentation 재진행
- 부피 구하기 자동화 진행 완료

'20.10~ 최신 진행 사항

- **데이터셋 요약 : 예측 모델에 82개 특징 이미지 입력 예정**

	눈	내직근	외직근	상직근	하직근	안와	시신경	상안검	눈물 주머니	지방	합계
AX1	2	2	2			2				2	10
AX2	2	2	2			2	2			2	12
AX3	2	2	2			2	2			2	12
AX4	2	2	2			2	2			2	12
AX5	2	2	2			2					8
AX6									2		2
AX7	2	2	2			2	2		2	2	14
CO1	2										2
CO2		2	2	2	2		2				10
CO3						2					2
RT	1			1	1		1	1			5
LT	1			1	1		1	1			5
합계	**16**	**14**	**14**	**4**	**4**	**14**	**12**	**2**	**4**	**10**	**94**

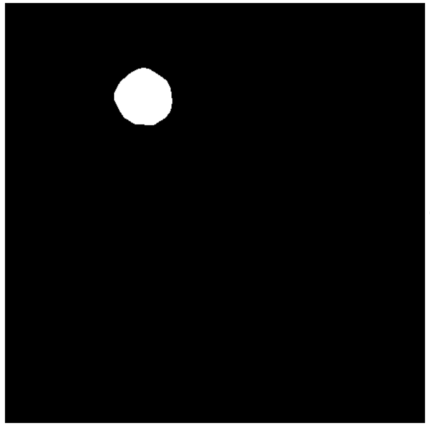
1. Ax7 세그멘테이션

- 입력 : 환자 1237명의 9633개의 original DICOM 데이터
- 출력 : 9633 * 12개의 masking 데이터
- 실험 횟수 : 30회
- 데이터셋 1차 검수 완료 및 1차 수정 작업 완료 -> 검수 진행중
- Predict가 잘 된 data는 다시 train data로 사용하여 다시 test 진행

-> predict 이미지가 삭제되어 실험횟수 3으로 재진행 중

2. 3D 세그멘테이션

<입력 샘플 예시>
(1075 환자의 왼쪽 눈 10040.dcm)



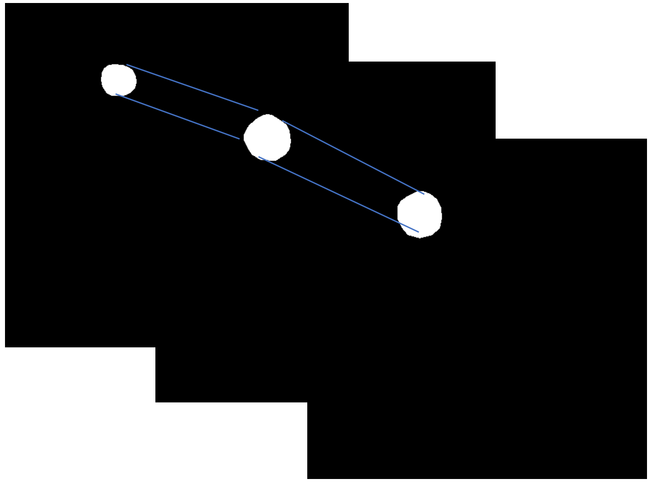
데이터의 사이즈:
512*512

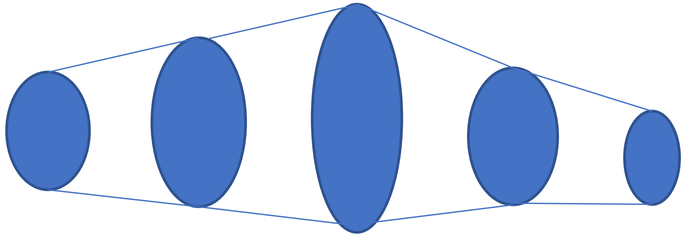
225 =white
0 = black

Segmentation 된 픽셀
Counting

1075 환자 왼쪽 눈
10040.dcm 파일의 넓이
= 가로 * 세로 * selec된 pixel 수

가로 : 0.390625
세로 : 0.390625
높이 : 1.00





각 기둥의 부피를 계산
DICOM 파일에 생긴 간격에
있는 부피 예측

Individual Studies:

Taekyung Song

Introduction

- ImageNet – Image Classification Algorithm

  - AlexNet, VGGNet, GoogLeNet, ResNet

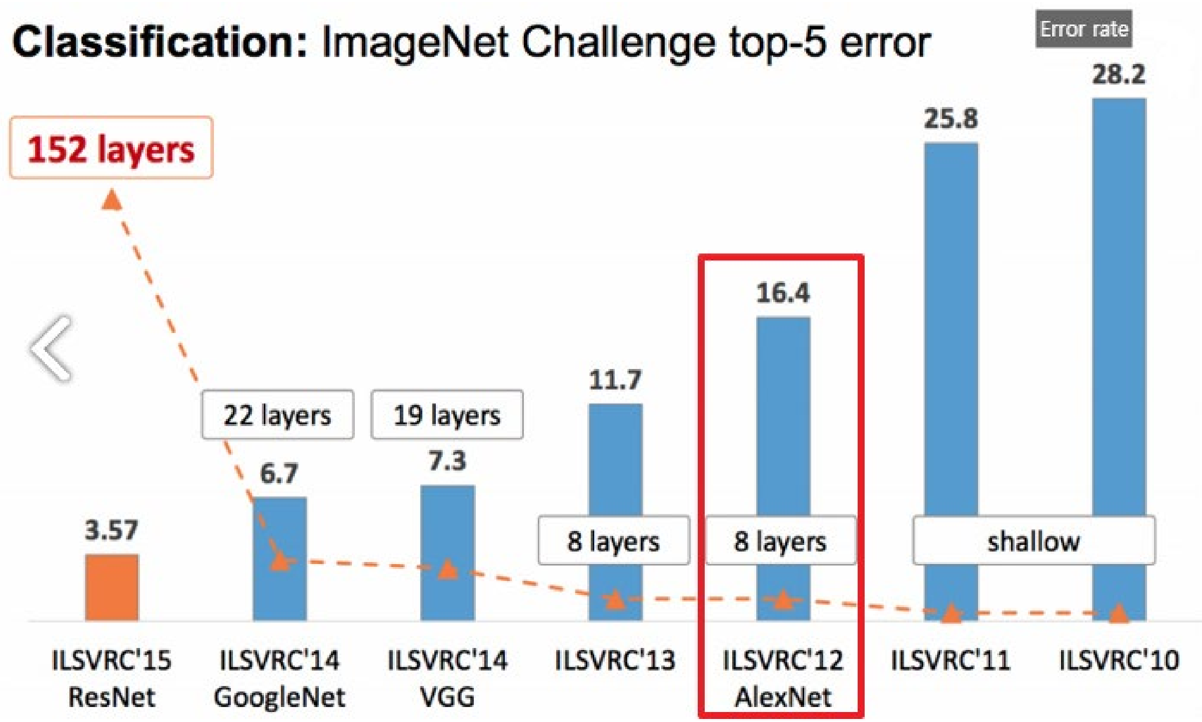
  \+ MobileNet, NASNet, EfficientNet etc.

Paper Review

Paper Title: ImageNet Classification with Deep Convolutional Neural Networks

Authors: Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton

Conference: NIPS 2012

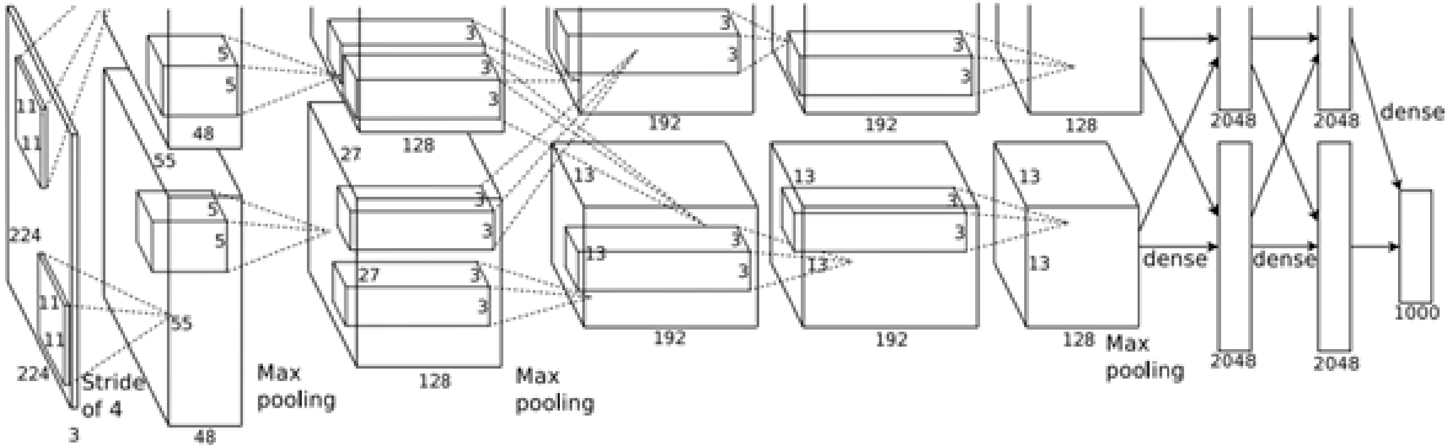


AlexNet

- Data set
  - down-sampled the images to a fixed resolution of 256*256

- Architecture
  - ReLU Nonlinearity
  - Training on Multiple GPUs
  - LRN(Local Response Normalization)
  - Overlapping Pooling

- Prevent Overfiting
  - Data Augmentation
  - Dropout

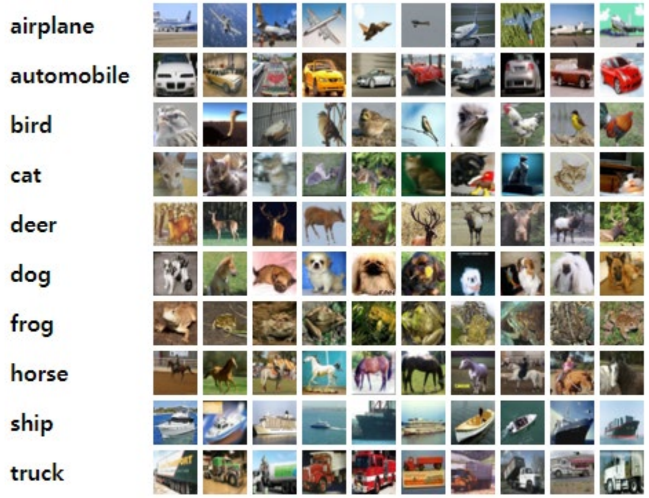


AlexNet

- Model

```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
            Conv2d-1           [-1, 96, 55, 55]          34,944
              ReLU-2           [-1, 96, 55, 55]               0
 LocalResponseNorm-3           [-1, 96, 55, 55]               0
         MaxPool2d-4           [-1, 96, 27, 27]               0
            Conv2d-5          [-1, 256, 27, 27]         614,656
              ReLU-6          [-1, 256, 27, 27]               0
 LocalResponseNorm-7          [-1, 256, 27, 27]               0
         MaxPool2d-8          [-1, 256, 13, 13]               0
            Conv2d-9          [-1, 384, 13, 13]         885,120
             ReLU-10          [-1, 384, 13, 13]               0
           Conv2d-11          [-1, 384, 13, 13]       1,327,488
             ReLU-12          [-1, 384, 13, 13]               0
           Conv2d-13          [-1, 256, 13, 13]         884,992
             ReLU-14          [-1, 256, 13, 13]               0
        MaxPool2d-15            [-1, 256, 6, 6]               0
          Dropout-16                 [-1, 9216]               0
           Linear-17                 [-1, 4096]      37,752,832
             ReLU-18                 [-1, 4096]               0
          Dropout-19                 [-1, 4096]               0
           Linear-20                 [-1, 4096]      16,781,312
             ReLU-21                 [-1, 4096]               0
           Linear-22                   [-1, 10]          40,970
================================================================
Total params: 58,322,314
Trainable params: 58,322,314
Non-trainable params: 0
```

- Data: Cifar-10



50000 32*32 colour images in 10 classes

- Result

```
Accuracy of the network on the 50000 test images: 80 %
Accuracy for class plane is: 81.5 %
Accuracy for class car   is: 85.9 %
Accuracy for class bird  is: 73.1 %
Accuracy for class cat   is: 65.1 %
Accuracy for class deer  is: 78.2 %
Accuracy for class dog   is: 70.3 %
Accuracy for class frog  is: 79.0 %
Accuracy for class horse is: 88.7 %
Accuracy for class ship  is: 89.8 %
```

Next progress

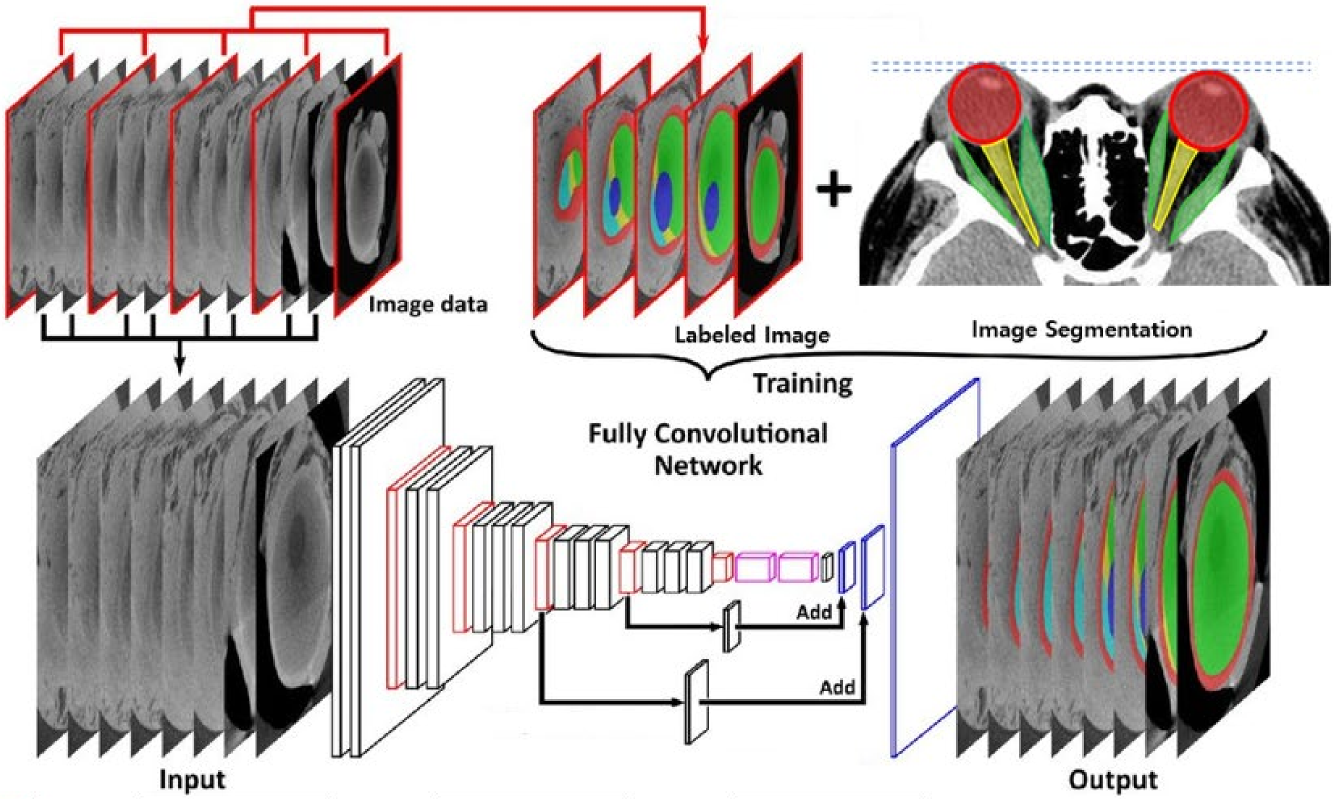
- VGGNet vs GoogLeNet

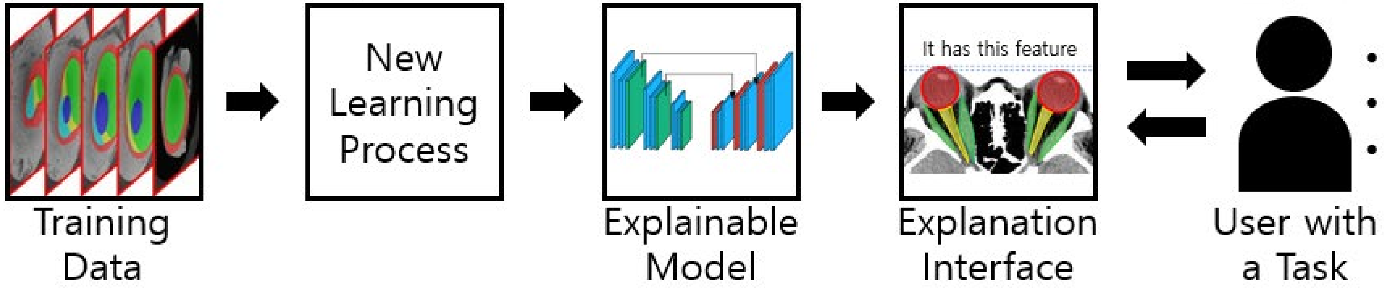
- Simonyan, K., & Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. (2014)

- Christian Szegedy, et al. Going Deeper with Convolutions. CVPR. (2015)

  + Min Lin, Qiang Chen, Shuicheng Yan. Network In Network. arXiv:1312.4400. (2014)

How to pre-process an image and apply it with XAI





- Understand why
- Know when succeed
- Know why error occurred

Individual Studies:

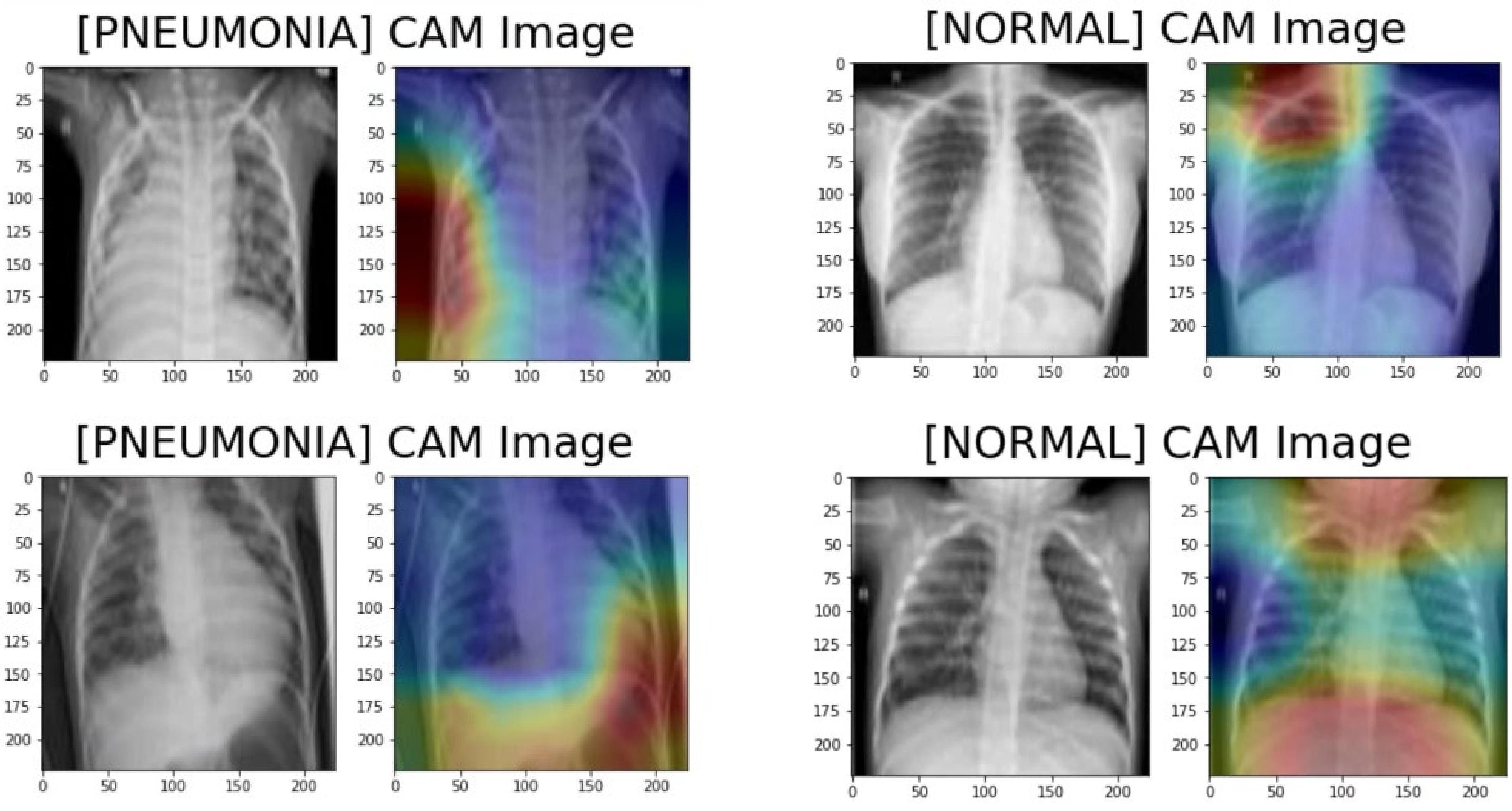
Haesung Oh

Progress

- Last Report
  - Universal Weakly Supervised Segmentation by Pixel-to-segment Contrastive Learning 논문을 직접 구현하는 과정 중 어려움들

- This Report
  - WSSS의 시초가 되는 논문인 CAM과 Grad-CAM의 구현 및 결과 확인

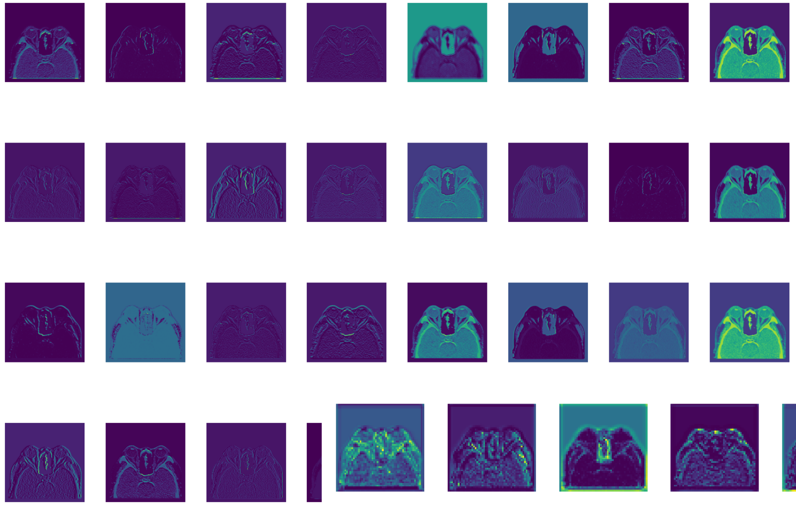
Class Activation Map (CAM)



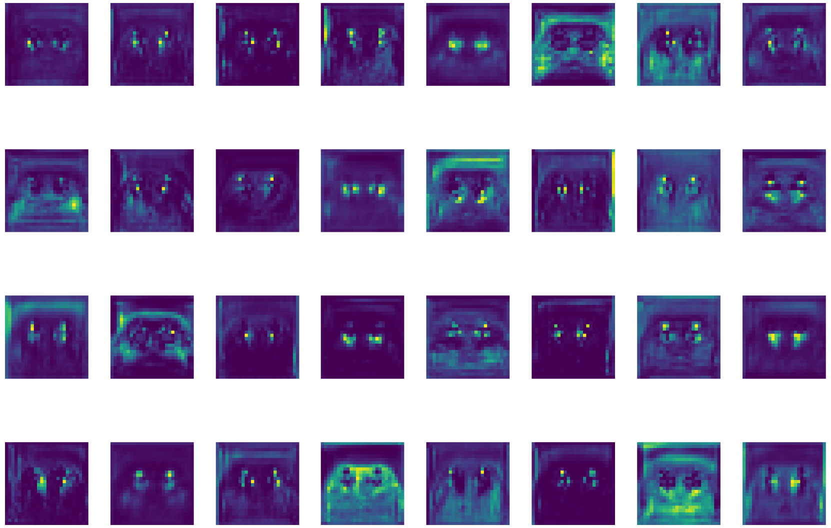
Gradient weighted Class Activation Map (Grad-CAM)

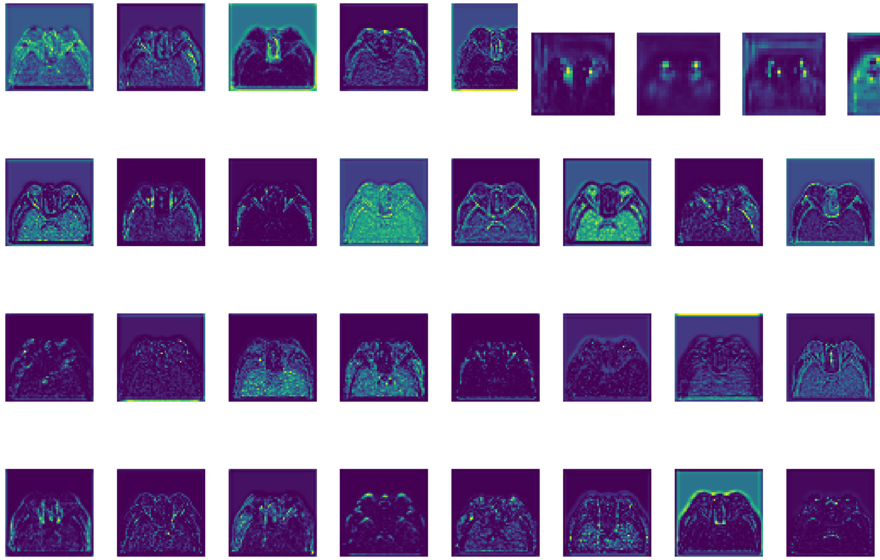
Original

Stage0_Conv_features



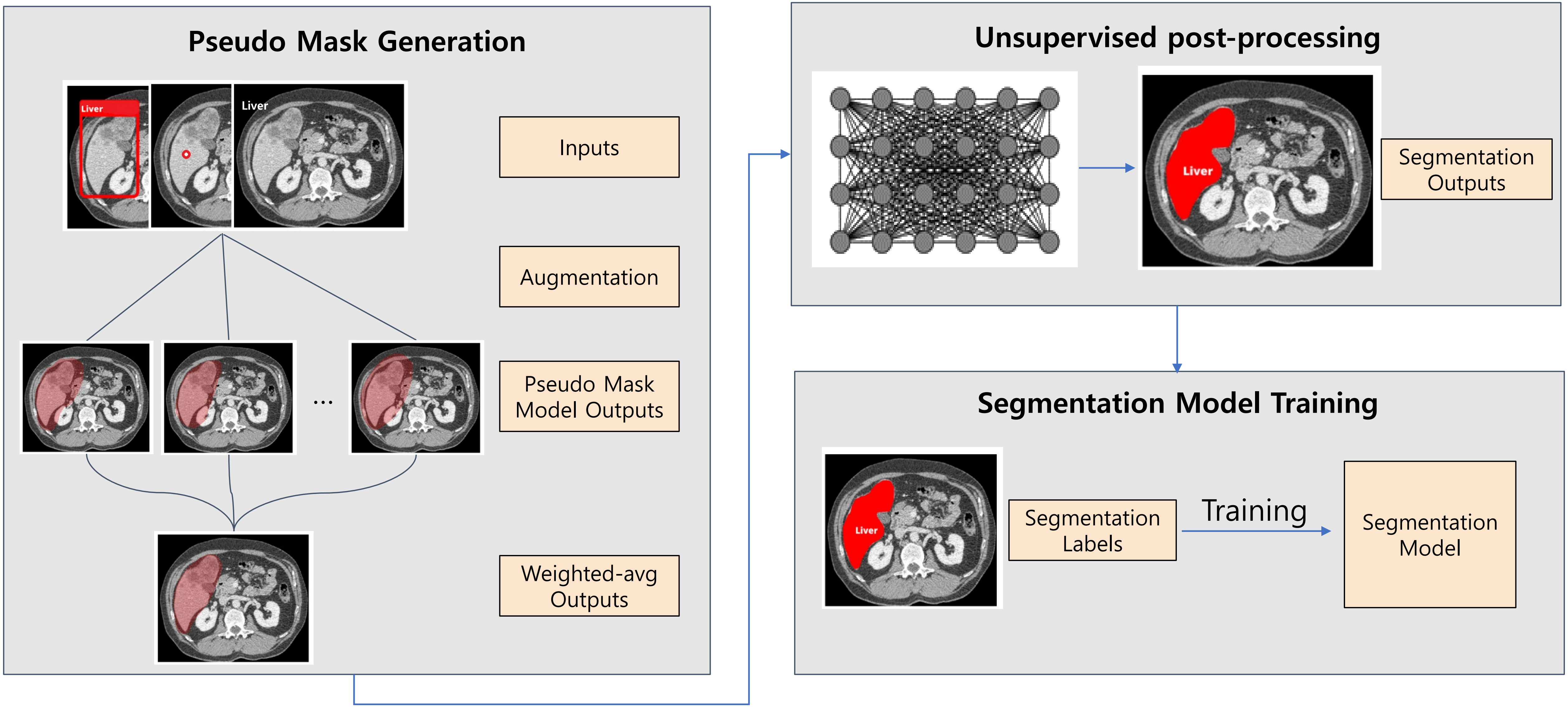
Stage4_Conv_features





Stage18_Conv_features

Medical Weakly Supervised Semantic Segmentation



Government-funded Research Projects:
갑상선 연구
(이상혁, 한수정, 송태경, 오해성)

진행사항

- 지난 보고
  - CT 영상 기반 Activity 추론 연구 실험 최종 보고 및 시각화 보고

- 이번 보고
  - 사진 기반 CAS 점수 추론 연구 Baseline 실험 보고

사진 기반 CAS 추론 연구 내용

<입력 데이터 샘플>

딥러닝 모델

번호	영문 용어	해석
1	Swelling of Eyelid	눈꺼풀 붓기
2	Redness of Eyelid	눈꺼풀 붉기
3	Swelling of Conjuinctiva	결막 붓기
4	Redness of Conjunctiva	결막 붉기
5	Swelling of Caruncle	누구의 붓기

<5개 레이블>

사진 기반 CAS 추론 Baseline 실험 진행 내용

딥러닝 모델 1 → 눈꺼풀 붓기

딥러닝 모델 2 → 눈꺼풀 붉기

딥러닝 모델 3 → 결막 붓기

딥러닝 모델 4 → 결막 붉기

딥러닝 모델 5 → 누구의 붓기

사진 기반 CAS 추론 Baseline 실험 진행 내용

- 실험 모델 : 14종의 딥러닝 모델
  - AlexNet, ResNet, VGGNet, GoogLeNet, DenseNet, InceptionV2, MobileNet, ShuffleNet, EfficientNet (이상 클래식 9종), CSWin, CoAtNet, ECANet, GhostNet, SENet (이상 최신 5종)


딥러닝 모델 1 → 눈꺼풀 붓기 x 14개 모델 실험


딥러닝 모델 2 → 눈꺼풀 붉기 x 14개 모델 실험


딥러닝 모델 3 → 결막 붓기 x 14개 모델 실험


딥러닝 모델 4 → 결막 붉기 x 14개 모델 실험


딥러닝 모델 5 → 누구의 붓기 x 14개 모델 실험

실험 결과(눈꺼풀 붓기, Swelling of Eyelid)

Table 3. Metrics with 95% confidence interval (▼/△ indicates that the corresponding method is **significantly worse** than the best model written in bold based on paired t-test at 5% significance level).

Model	AUROC			AUPRC			Accuracy			F1 Score			Sensitivity TP/(TP+FN)			Specificity TN/(TN+FP)			Precision TP/(TP+FP)		
	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value
GoogLeNet	0.668	0.647 0.689		0.623	0.577 0.669		0.556	0.506 0.607	0.626	0.626	0.603 0.650		0.819	0.685 0.953		0.34	0.144 0.536	0.047	0.534	0.478 0.590	0.776
ECANet	0.599	0.566 0.631	0.001	0.555	0.507 0.602	0.002	0.574	0.532 0.616		0.481	0.364 0.598	0.021	0.477	0.320 0.635	0.011	0.651	0.486 0.816	0.638	0.544	0.490 0.598	
SENet	0.599	0.566 0.631	0.001	0.555	0.507 0.602	0.002	0.574	0.532 0.616		0.481	0.364 0.598	0.021	0.477	0.320 0.635	0.011	0.651	0.486 0.816	0.638	0.544	0.490 0.598	
ShuffleNet	0.595	0.576 0.613	0	0.545	0.520 0.570	0.005	0.56	0.532 0.588	0.403	0.599	0.563 0.635	0.188	0.729	0.616 0.841	0.309	0.406	0.259 0.554	0.069	0.522	0.494 0.550	0.304
ResNet	0.579	0.551 0.608	0	0.535	0.493 0.578	0.002	0.556	0.533 0.579	0.492	0.536	0.421 0.650	0.118	0.631	0.447 0.815	0.145	0.48	0.315 0.645	0.049	0.538	0.482 0.595	0.906
EfficientNet	0.567	0.528 0.606	0	0.523	0.486 0.560	0	0.529	0.507 0.550	0.054	0.539	0.470 0.608	0.037	0.645	0.473 0.818	0.069	0.446	0.283 0.609	0.079	0.508	0.471 0.544	0.112
MobileNetV3	0.54	0.505 0.574	0	0.492	0.457 0.527	0	0.52	0.483 0.557	0.027	0.425	0.283 0.567	0.011	0.483	0.248 0.717	0.033	0.541	0.298 0.784	0.358	0.489	0.443 0.535	0.051
GhostNet	0.523	0.488 0.559	0	0.486	0.451 0.521	0	0.515	0.480 0.550	0.033	0.423	0.247 0.599	0.026	0.538	0.248 0.829	0.053	0.474	0.176 0.771	0.221	0.451	0.370 0.531	0.086
CoAtNet	0.519	0.500 0.539	0	0.481	0.454 0.508	0	0.506	0.478 0.534	0.033	0.352	0.194 0.511	0.002	0.402	0.148 0.655	0.004	0.593	0.340 0.847	0.211	0.438	0.380 0.495	0.009
InceptionV2	0.516	0.484 0.547	0	0.474	0.433 0.515	0	0.51	0.490 0.530	0.016	0.426	0.301 0.552	0.007	0.49	0.255 0.725	0.036	0.532	0.298 0.766	0.287	0.496	0.431 0.561	0.165
AlexNet	0.506	0.473 0.539	0	0.466	0.429 0.502	0	0.499	0.470 0.528	0.002	0.428	0.248 0.608	0.042	0.56	0.276 0.845	0.179	0.424	0.140 0.707	0.106	0.365	0.225 0.505	0.046
CSWin	0.501	0.480 0.521	0	0.477	0.452 0.502	0	0.5	0.464 0.536	0.005	0.316	0.077 0.555	0.02	0.5	0.123 0.877	0.159	0.5	0.123 0.877	0.372	0.231	0.056 0.407	0.008
DenseNet	0.497	0.469 0.525	0	0.465	0.423 0.507	0	0.502	0.474 0.529	0.024	0.475	0.307 0.643	0.072	0.633	0.371 0.896	0.249	0.372	0.117 0.626	0.012	0.448	0.363 0.532	0.033
VGGNet	0.497	0.455 0.538	0	0.471	0.431 0.511	0	0.525	0.499 0.552	0.06	0.242	0.053 0.431	0.001	0.287	0.002 0.572	0.007	0.722	0.451 0.992		0.399	0.284 0.515	0.02

CI, confidence interval; *: p-value < 0.05; **: p-value < 0.01

실험 결과(눈꺼풀 붉기, Redness of Eyelid)

Table 3. Metrics with 95% confidence interval (▼/△ indicates that the corresponding method is **significantly worse** than the best model written in bold based on paired t-test at 5% significance level).

Model	AUROC			AUPRC			Accuracy			F1 Score			Sensitivity TP/(TP+FN)			Specificity TN/(TN+FP)			Precision TP/(TP+FP)		
	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value
GoogLeNet	0.617	0.557 0.678		0.173	0.138 0.209		0.596	0.404 0.788	0.225	0.163	0.113 0.214	0.159	0.467	0.190 0.743	0.289	0.614	0.369 0.859	0.214	0.136	0.104 0.168	0.493
InceptionV2	0.546	0.469 0.623	0.135	0.168	0.099 0.237	0.855	0.532	0.371 0.693	0.032	0.188	0.141 0.235		0.509	0.323 0.694	0.317	0.537	0.335 0.738	0.031	0.183	0.025 0.340	
CoAtNet	0.541	0.479 0.603	0.06	0.123	0.096 0.150	0.002	0.559	0.324 0.793	0.268	0.156	0.119 0.194	0.309	0.452	0.164 0.740	0.257	0.571	0.277 0.865	0.267	0.156	0.094 0.218	0.632
AlexNet	0.538	0.482 0.593	0.079	0.138	0.097 0.180	0.111	0.51	0.247 0.773	0.122	0.115	0.048 0.181	0.087	0.52	0.156 0.883	0.397	0.511	0.175 0.847	0.114	0.095	0.025 0.165	0.31
ResNet	0.537	0.475 0.599	0.025	0.127	0.099 0.155	0.018	0.682	0.524 0.839	0.811	0.141	0.116 0.166	0.075	0.284	0.090 0.478	0.015	0.732	0.534 0.929	0.828	0.117	0.096 0.138	0.37
ShuffleNet	0.532	0.465 0.599	0.011	0.12	0.095 0.145	0	0.643	0.423 0.863	0.592	0.107	0.059 0.156	0.043	0.316	0.051 0.582	0.031	0.69	0.413 0.966	0.639	0.104	0.050 0.157	0.353
VGGNet	0.519	0.467 0.571	0.007	0.135	0.086 0.184	0.06	0.531	0.323 0.739	0.081	0.134	0.065 0.203	0.156	0.479	0.213 0.745	0.295	0.536	0.277 0.795	0.072	0.081	0.037 0.124	0.134
GhostNet	0.517	0.456 0.577	0.032	0.121	0.097 0.146	0.002	0.705	0.607 0.803		0.119	0.061 0.178	0.028	0.24	0.044 0.435	0.043	0.758	0.627 0.889		0.106	0.047 0.164	0.278
MobileNet	0.511	0.478 0.543	0.002	0.12	0.098 0.141	0.009	0.478	0.274 0.682	0.037	0.164	0.135 0.194	0.288	0.533	0.283 0.784	0.458	0.47	0.210 0.730	0.036	0.129	0.088 0.170	0.483
ECANet	0.51	0.464 0.557	0.009	0.124	0.105 0.144	0.04	0.626	0.470 0.782	0.125	0.134	0.082 0.185	0.045	0.352	0.130 0.575	0.083	0.658	0.458 0.857	0.116	0.093	0.061 0.124	0.186
SENet	0.51	0.464 0.557	0.009	0.124	0.105 0.144	0.04	0.626	0.470 0.782	0.125	0.134	0.082 0.185	0.045	0.352	0.130 0.575	0.083	0.658	0.458 0.857	0.116	0.093	0.061 0.124	0.186
DenseNet	0.506	0.457 0.555	0.022	0.121	0.091 0.151	0.003	0.583	0.401 0.765	0.245	0.138	0.090 0.186	0.235	0.391	0.159 0.623	0.034	0.61	0.381 0.838	0.269	0.105	0.060 0.150	0.336
CSWin	0.505	0.450 0.560	0.021	0.146	0.104 0.187	0.219	0.333	0.062 0.603	0.021	0.125	0.061 0.188	0.126	0.7	0.354 1.046		0.3	-0.046 0.646	0.025	0.068	0.034 0.103	0.134
EfficientNet	0.489	0.440 0.537	0.002	0.121	0.087 0.154	0.006	0.662	0.515 0.809	0.637	0.122	0.085 0.159	0.042	0.266	0.112 0.420	0.013	0.71	0.528 0.893	0.676	0.097	0.066 0.128	0.261

CI, confidence interval; *: p-value < 0.05; **: p-value < 0.01

실험 결과(결막 붓기, Swelling of Conjuinctiva)

Table 3. Metrics with 95% confidence interval (▼/△ indicates that the corresponding method is **significantly worse** than the best model written in bold based on paired t-test at 5% significance level).

Model	AUROC			AUPRC			Accuracy			F1 Score			Sensitivity TP/(TP+FN)			Specificity TN/(TN+FP)			Precision TP/(TP+FP)		
	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value
GoogLeNet	0.655	0.623 0.687		0.171	0.123 0.218		0.705	0.507 0.903	0.151	0.181	0.118 0.244		0.409	0.175 0.642	0.051	0.728	0.491 0.965	0.134	0.167	0.083 0.252	
InceptionV2	0.617	0.564 0.669	0.107	0.121	0.097 0.145	0.062	0.493	0.330 0.656	0.002	0.176	0.132 0.221	0.901	0.668	0.473 0.863		0.477	0.287 0.667	0.002	0.107	0.074 0.139	0.165
ShuffleNet	0.606	0.561 0.651	0.062	0.12	0.094 0.146	0.048	0.853	0.798 0.908		0.09	0.011 0.169	0.041	0.134	0.008 0.259	0.001	0.917	0.850 0.984		0.08	0.005 0.154	0.046
ECANet	0.602	0.543 0.661	0.031	0.112	0.089 0.135	0.035	0.669	0.507 0.830	0.047	0.147	0.114 0.180	0.397	0.375	0.169 0.581	0.031	0.694	0.505 0.882	0.041	0.11	0.079 0.140	0.202
SENet	0.602	0.543 0.661	0.031	0.112	0.089 0.135	0.035	0.669	0.507 0.830	0.047	0.147	0.114 0.180	0.397	0.375	0.169 0.581	0.031	0.694	0.505 0.882	0.041	0.11	0.079 0.140	0.202
EfficientNet	0.585	0.523 0.647	0.037	0.138	0.096 0.180	0.077	0.723	0.551 0.895	0.128	0.138	0.079 0.197	0.346	0.32	0.099 0.541	0.021	0.756	0.547 0.965	0.126	0.148	0.050 0.247	0.782
ResNet	0.575	0.512 0.639	0.016	0.123	0.070 0.177	0.172	0.623	0.464 0.781	0.023	0.127	0.081 0.173	0.113	0.428	0.199 0.657	0.061	0.643	0.453 0.833	0.022	0.079	0.052 0.105	0.054
DenseNet	0.563	0.527 0.600	0.002	0.11	0.073 0.148	0.075	0.473	0.251 0.695	0.008	0.159	0.118 0.199	0.522	0.613	0.365 0.861	0.7	0.462	0.197 0.726	0.008	0.103	0.062 0.143	0.17
MobileNet	0.533	0.455 0.611	0.013	0.097	0.071 0.123	0.024	0.731	0.644 0.819	0.005	0.142	0.084 0.200	0.349	0.257	0.147 0.366	0.008	0.772	0.673 0.872	0.003	0.111	0.056 0.165	0.234
AlexNet	0.523	0.469 0.576	0.001	0.083	0.069 0.097	0.006	0.545	0.261 0.829	0.041	0.085	0.028 0.143	0.021	0.457	0.110 0.804	0.192	0.549	0.211 0.887	0.042	0.049	0.017 0.081	0.015
CoAtNet	0.516	0.478 0.553	0	0.108	0.079 0.137	0.031	0.493	0.245 0.741	0.015	0.114	0.064 0.163	0.101	0.529	0.235 0.823	0.368	0.493	0.201 0.785	0.014	0.068	0.037 0.098	0.053
GhostNet	0.509	0.459 0.559	0	0.08	0.067 0.092	0.005	0.66	0.473 0.848	0.06	0.08	0.037 0.122	0.021	0.271	0.034 0.508	0.006	0.694	0.474 0.913	0.064	0.068	0.015 0.120	0.102
VGGNet	0.505	0.432 0.577	0	0.09	0.064 0.116	0.014	0.381	0.196 0.565	0	0.118	0.095 0.142	0.036	0.59	0.353 0.828	0.58	0.366	0.146 0.586	0	0.07	0.057 0.083	0.026
CSWin	0.498	0.448 0.548	0	0.089	0.068 0.111	0.011	0.415	0.104 0.726	0.009	0.087	0.032 0.142	0.078	0.6	0.231 0.969	0.63	0.4	0.031 0.769	0.01	0.047	0.017 0.077	0.027

CI, confidence interval; *: p-value < 0.05; **: p-value < 0.01

실험 결과(결막 붉기, Redness of Conjunctiva)

Table 3. Metrics with 95% confidence interval (▼/△ indicates that the corresponding method is **significantly worse** than the best model written in bold based on paired t-test at 5% significance level).

Model	AUROC			AUPRC			Accuracy			F1 Score			Sensitivity TP/(TP+FN)			Specificity TN/(TN+FP)			Precision TP/(TP+FP)		
	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value
GoogLeNet	0.753	0.729 0.776		0.482	0.428 0.535		0.627	0.514 0.739	0.18	0.456	0.402 0.510		0.703	0.524 0.881		0.6	0.412 0.788	0.046	0.384	0.321 0.448	0.694
InceptionV2	0.652	0.599 0.706	0.001	0.39	0.326 0.454	0.014	0.573	0.452 0.693	0.02	0.337	0.246 0.427	0.035	0.568	0.306 0.830	0.43	0.578	0.352 0.804	0.011	0.309	0.227 0.392	0.235
DenseNet	0.642	0.574 0.711	0.002	0.392	0.302 0.482	0.023	0.587	0.452 0.722	0.142	0.359	0.256 0.463	0.034	0.563	0.346 0.780	0.275	0.605	0.381 0.828	0.066	0.307	0.188 0.425	0.225
ShuffleNet	0.635	0.600 0.669	0	0.368	0.309 0.427	0.001	0.687	0.591 0.783	0.781	0.247	0.163 0.332	0.002	0.303	0.088 0.517	0.009	0.812	0.627 0.996	0.902	0.425	0.238 0.611	
ResNet	0.623	0.568 0.678	0.001	0.338	0.279 0.397	0.004	0.703	0.657 0.750		0.284	0.173 0.395	0.017	0.312	0.141 0.484	0.008	0.823	0.730 0.916		0.322	0.217 0.426	0.385
ECANet	0.615	0.578 0.652	0	0.331	0.301 0.362	0	0.602	0.535 0.670	0.011	0.367	0.313 0.421	0.017	0.538	0.383 0.693	0.042	0.62	0.492 0.747	0.005	0.303	0.258 0.348	0.167
SENet	0.615	0.578 0.652	0	0.331	0.301 0.362	0	0.602	0.535 0.670	0.011	0.367	0.313 0.421	0.017	0.538	0.383 0.693	0.042	0.62	0.492 0.747	0.005	0.303	0.258 0.348	0.167
CoAtNet	0.608	0.538 0.677	0.001	0.349	0.284 0.413	0	0.565	0.397 0.732	0.111	0.308	0.226 0.390	0	0.49	0.253 0.728	0.136	0.593	0.314 0.872	0.11	0.353	0.232 0.474	0.353
EfficientNet	0.602	0.555 0.650	0	0.323	0.275 0.371	0.001	0.687	0.634 0.741	0.555	0.297	0.236 0.358	0.002	0.296	0.212 0.381	0.001	0.803	0.734 0.871	0.633	0.319	0.245 0.393	0.247
AlexNet	0.572	0.531 0.614	0	0.291	0.246 0.337	0	0.544	0.372 0.716	0.066	0.272	0.169 0.375	0.007	0.499	0.211 0.788	0.157	0.562	0.258 0.865	0.07	0.338	0.143 0.533	0.536
GhostNet	0.569	0.515 0.623	0	0.29	0.247 0.334	0.001	0.556	0.437 0.675	0.013	0.28	0.191 0.369	0.013	0.48	0.243 0.716	0.122	0.583	0.361 0.804	0.01	0.288	0.223 0.352	0.189
VGGNet	0.565	0.525 0.605	0	0.281	0.240 0.322	0	0.604	0.478 0.731	0.17	0.253	0.163 0.344	0.001	0.397	0.151 0.642	0.009	0.667	0.432 0.901	0.247	0.379	0.208 0.549	0.707
MobileNetV3	0.563	0.531 0.595	0	0.281	0.254 0.308	0	0.575	0.455 0.695	0.054	0.277	0.181 0.374	0.001	0.461	0.216 0.706	0.004	0.604	0.378 0.831	0.07	0.255	0.233 0.277	0.059
CSWin	0.479	0.451 0.507	0	0.233	0.206 0.259	0	0.565	0.364 0.767	0.154	0.154	0.012 0.297	0.003	0.4	0.031 0.769	0.141	0.6	0.231 0.969	0.238	0.096	0.007 0.184	0.001

CI, confidence interval; *: p-value < 0.05; **: p-value < 0.01

실험 결과(누구 붓기, Swelling of Caruncle)

Table 3. Metrics with 95% confidence interval (▼/△ indicates that the corresponding method is **significantly worse** than the best model written in bold based on paired t-test at 5% significance level).

Model	AUROC			AUPRC			Accuracy			F1 Score			Sensitivity TP/(TP+FN)			Specificity TN/(TN+FP)			Precision TP/(TP+FP)		
	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value
ResNet	0.641	0.547 0.735		0.08	0.018 0.142	0.358	0.65	0.467 0.833	0.288	0.115	0.034 0.196		0.528	0.283 0.773	0.479	0.656	0.458 0.855	0.28	0.078	0.009 0.146	0.733
CoAtNet	0.625	0.539 0.711	0.793	0.08	0.044 0.115	0.413	0.517	0.311 0.723	0.009	0.098	0.061 0.135	0.645	0.644	0.439 0.850		0.511	0.289 0.732	0.008	0.055	0.033 0.078	0.465
ShuffleNet	0.624	0.445 0.804	0.853	0.104	0.027 0.181		0.609	0.363 0.855	0.227	0.076	0.033 0.118	0.222	0.505	0.190 0.820	0.463	0.61	0.344 0.877	0.22	0.043	0.018 0.068	0.31
MobileNetV3	0.623	0.540 0.706	0.728	0.092	0.029 0.156	0.734	0.507	0.279 0.736	0.053	0.069	0.032 0.105	0.184	0.593	0.337 0.848	0.78	0.504	0.256 0.752	0.054	0.037	0.016 0.058	0.251
GoogLeNet	0.615	0.511 0.719	0.729	0.091	0.027 0.154	0.755	0.509	0.276 0.742	0.078	0.106	0.046 0.166	0.675	0.625	0.372 0.878	0.902	0.509	0.260 0.758	0.076	0.082	0.000 0.164	
ECANet	0.611	0.542 0.679	0.569	0.056	0.038 0.075	0.174	0.556	0.333 0.780	0.109	0.087	0.060 0.114	0.474	0.555	0.295 0.816	0.63	0.561	0.319 0.803	0.108	0.056	0.035 0.076	0.488
SENet	0.611	0.542 0.679	0.569	0.056	0.038 0.075	0.174	0.556	0.333 0.780	0.109	0.087	0.060 0.114	0.474	0.555	0.295 0.816	0.63	0.561	0.319 0.803	0.108	0.056	0.035 0.076	0.488
CSWin	0.589	0.509 0.669	0.4	0.061	0.024 0.099	0.327	0.59	0.247 0.933	0.247	0.026	0.001 0.050	0.061	0.4	0.031 0.769	0.228	0.6	0.231 0.969	0.256	0.013	0.000 0.026	0.107
InceptionV2	0.567	0.478 0.656	0.175	0.045	0.029 0.061	0.087	0.572	0.385 0.759	0.112	0.068	0.041 0.094	0.295	0.522	0.255 0.789	0.477	0.577	0.375 0.779	0.115	0.038	0.023 0.052	0.3
DenseNet	0.553	0.413 0.693	0.185	0.052	0.027 0.076	0.073	0.403	0.185 0.620	0.001	0.058	0.031 0.086	0.141	0.546	0.262 0.830	0.481	0.401	0.165 0.637	0.001	0.035	0.015 0.055	0.217
VGGNet	0.534	0.412 0.656	0.149	0.071	0.021 0.121	0.42	0.511	0.323 0.699	0.014	0.074	0.027 0.121	0.296	0.525	0.255 0.795	0.467	0.512	0.309 0.715	0.015	0.041	0.014 0.068	0.3
EfficientNet	0.531	0.448 0.614	0.105	0.05	0.027 0.073	0.103	0.784	0.669 0.899		0.046	0.008 0.084	0.059	0.213	0.031 0.396	0.001	0.805	0.679 0.931		0.027	0.004 0.049	0.1
AlexNet	0.53	0.409 0.650	0.205	0.047	0.027 0.067	0.142	0.45	0.140 0.760	0.077	0.045	0.018 0.071	0.141	0.56	0.222 0.898	0.656	0.451	0.116 0.785	0.081	0.024	0.010 0.039	0.173
GhostNet	0.496	0.388 0.605	0.101	0.043	0.028 0.059	0.089	0.531	0.316 0.746	0.035	0.08	0.040 0.119	0.289	0.499	0.211 0.787	0.257	0.528	0.295 0.760	0.037	0.048	0.021 0.074	0.256

CI, confidence interval; *: p-value < 0.05; **: p-value < 0.01

토의

- 중대 병원 측으로 문의 :

  1) 저희 데이터셋을 보셨을 때, 5개 항목들이 얼마나 판단이 어려운 정도인지. 또한 상대적 난이도 순위가 있을지.

  2) 각 항목에 대해 어떤 부분을 주요하게 보고 판단하시는지.

토의

1) 저희 데이터셋을 보셨을 때, 5개 항목들이 얼마나 판단이 어려운 정도인지. 또한 상대적 난이도 순위가 있을지.

이승현 선생님 답변 발췌 :

우선 **<u>애매한 사진들이 많네</u>**요…

항목들에서 **<u>아주 분명한 경우는 별로 문제가 안 되지만 애매한 경우에는 평가한 의사에 따라서 변경될 수 있는 지표들인 것 같아요.</u>**

게다가 사진으로 찍으니 세세한 변화들이 잘 안 보이거나(conjunctival swelling) 타 부분들과 대조가 적게 되는 것 같기도 합니다.(redness of eyelid)

상대적 난이도는 redness of conjunctiva, swelling of caruncle가 쉬울 것 같고

swelling of eyelid 가 조금 애매하지만 그 다음,

그리고 사진으로는 redness of eyelid나 swelling of conjunctiva는 확실하지 않은 경우 잘 안 보이는 것 같습니다.

- ➢ 애매한 사진 많음. -> 모델이 학습하기 어려울 수 있음.
- ➢ redness of conjunctiva < swelling of caruncle< swelling of eyelid < redness of eyelid = swelling of conjunctiva

토의

- redness of conjunctiva < swelling of caruncle< swelling of eyelid < redness of eyelid = swelling of conjunctiva
- 육안 난이도 순위 : 작을 수록 쉬움.

Target	육안 난이도 순위	Best Model	AUROC Average	AUPRC Average
Swelling of Eyelid	3	GoogLeNet	0.668	0.623
Redness of Eyelid	4	GoogLeNet	0.617	0.173
Swelling of Conjuinctiva	4	GoogLeNet	0.655	0.171
Redness of Conjunctiva	1	GoogLeNet	0.753	0.482
Swelling of Caruncle	2	ResNet	0.641	0.08

- 누구 붓기 클래스 비율 : 1,166 : 44

토의

2) 각 항목에 대해 어떤 부분을 주요하게 보고 판단하시는지. -> 자세한 내용은 생략

- 비교적 분명한 사진들을 몇 개 짚어주셨음.

이승현 선생님 답변 발췌 :

말씀하신 CAS 항목들은 실제 환자 볼 때는 직관적/일반적 으로 생각하는 부분들과 일치하는 부분이 많고

(보통 환자들이 생각하는 것과 저희가 보는 것이 비슷합니다. 또 환자들이 "충혈된다" "눈이 붓는다" "눈꺼풀이 빨갛고 열감이 있다" 등 얘기하기도 합니다.)

대신 **사진으로 분석하자니 어려움이 있는 것 같습니다.**

Government-funded Research Projects:
산학협력프로젝트
(연구실 전체)

진행사항

- 지난 보고
  - 간단한 진행 내용 보고

- 이번 보고
  - 간단한 진행 내용 보고

Q&A

References (Taekyung Song)

[1] Alex Krizhevsky, et al. "ImageNet Classification with Deep Convolutional Neural Networks" Part of Advances in Neural Information Processing Systems 25 (NIPS 2012)

References (Haesung Oh)

[1] Zhou, Bolei, et al. "Learning deep features for discriminative localization." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[2] Selvaraju, Ramprasaath R., et al. "Grad-cam: Visual explanations from deep networks via gradient-based localization." *Proceedings of the IEEE international conference on computer vision*. 2017.