기계학습자동화 연구실
주간 정기 세미나

이상혁, 조성현, 한수정, 송태경, 오해성

2022. 03. 22.

Outline

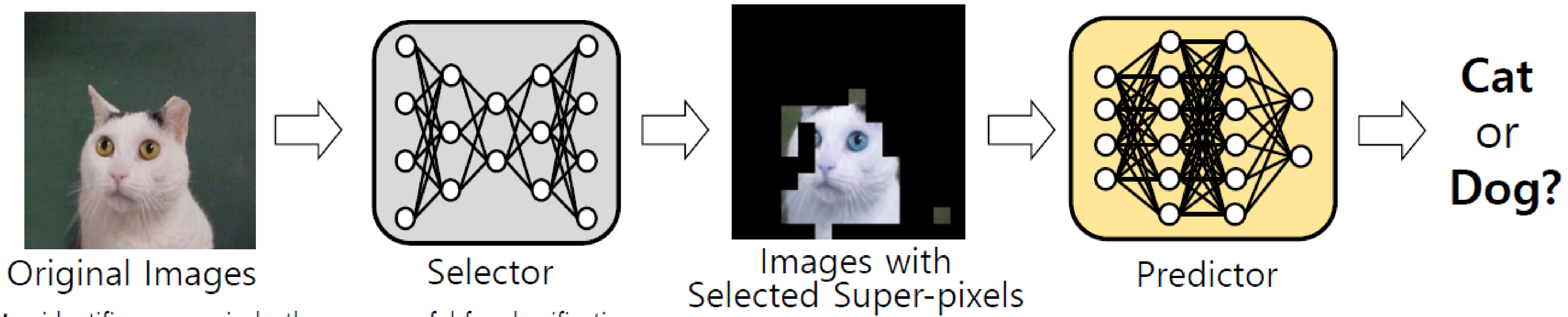
- **Individual Studies**
  - Sanghyuck Lee
  - Sung-Hyun Cho
  - Su-Jeong Han
  - Taekyung Song
  - Haesung Oh
  - Sanghyeong Jin

Outline

- **Government-funded Research Projects**
  - 갑상선 연구(이상혁, 한수정, 송태경, 오해성)
    - ✓ Clinical Activity Score (CAS) 5점 척도 점수 추론 연구 (이상혁)
      - ➢ 지난 보고 : (사진 only) Baseline 실험 보고
      - ➢ 이번 보고 : (사진 only) Baseline (+ Augmentation) 실험 보고
    - ✓ 3D 세그멘테이션 (한수정)
      - ➢ 지난 보고 : AX7 세그멘테이션 진행, 부피 계산 자동화
      - ➢ 이번 보고 : AX7 세그멘테이션 완료, 1명 환자의 1개 타겟 3D 시각화
  - 산학협력프로젝트(연구실 전체)
    - ➢ 간단한 진행 내용 보고

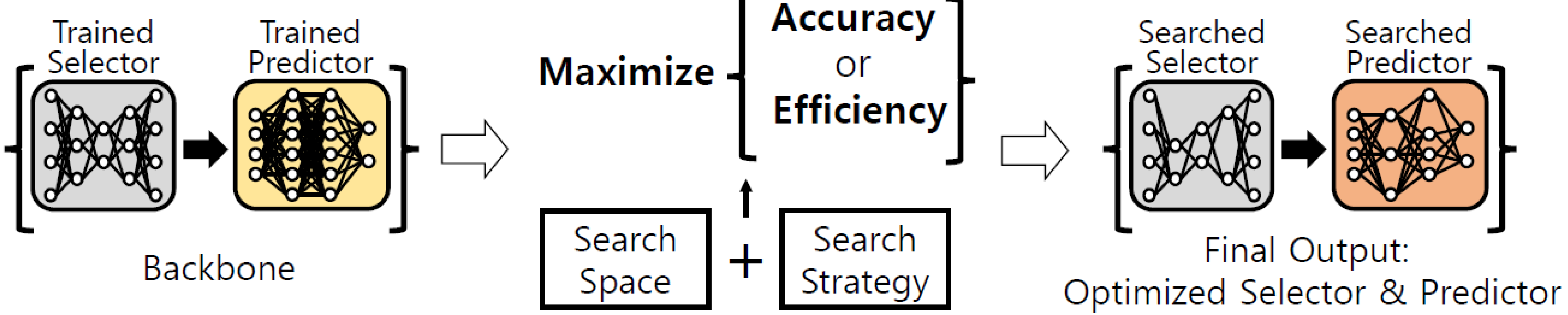
Neural Architecture Search based on Super-pixel Selection

<Selector & Predictor Train Phase>



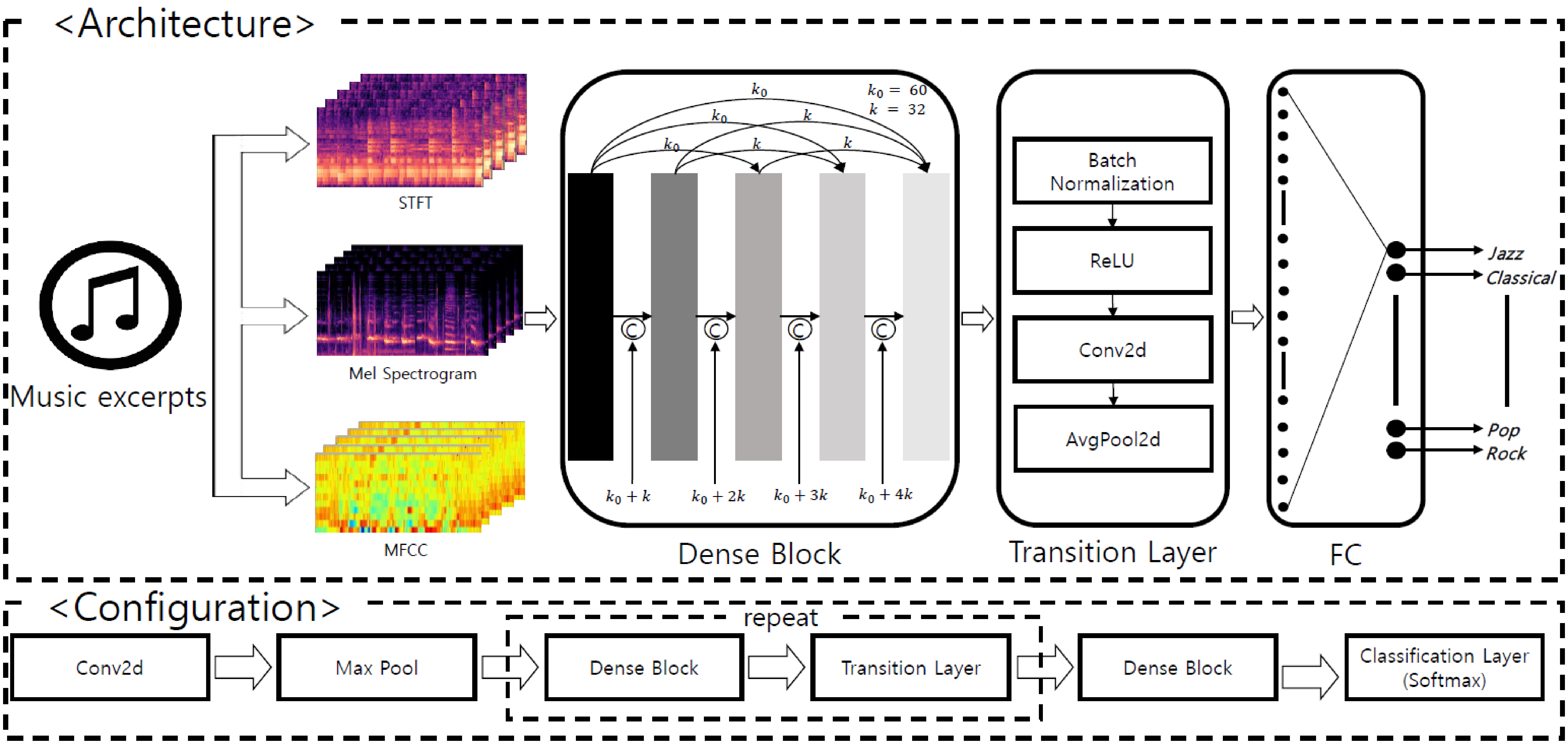
- **Selector** identifies super-pixels, those are useful for classification.
- **Predictor** classifies using selected super pixels.
- **Super-pixel** is a pixel cluster consisting of $n \times n$ pixels.

<Neural Architecture Search Phase>



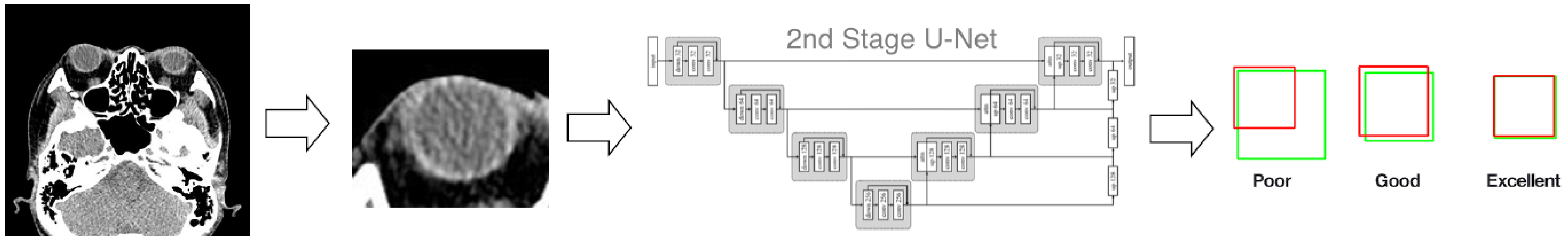
- Through **Neural Architecture Search** Phase using the **trained Selector & Predictor as a Backbone Network**, find **final optimized Selector & Predictor** from **Accuracy or Efficiency** point of view.

Music Genre Classification



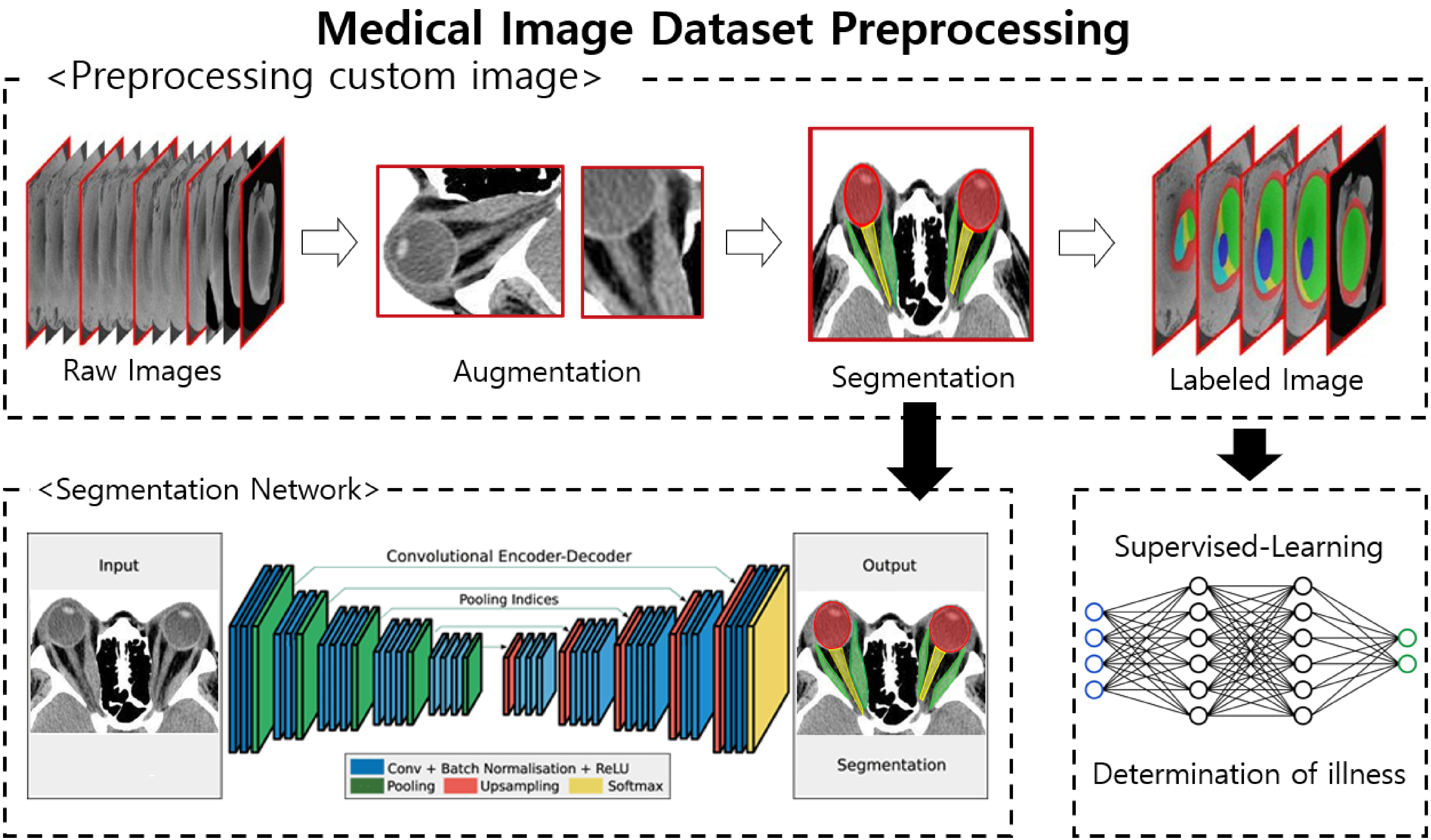
Efficient segmentation using U-Net

< U-Net using OCR >



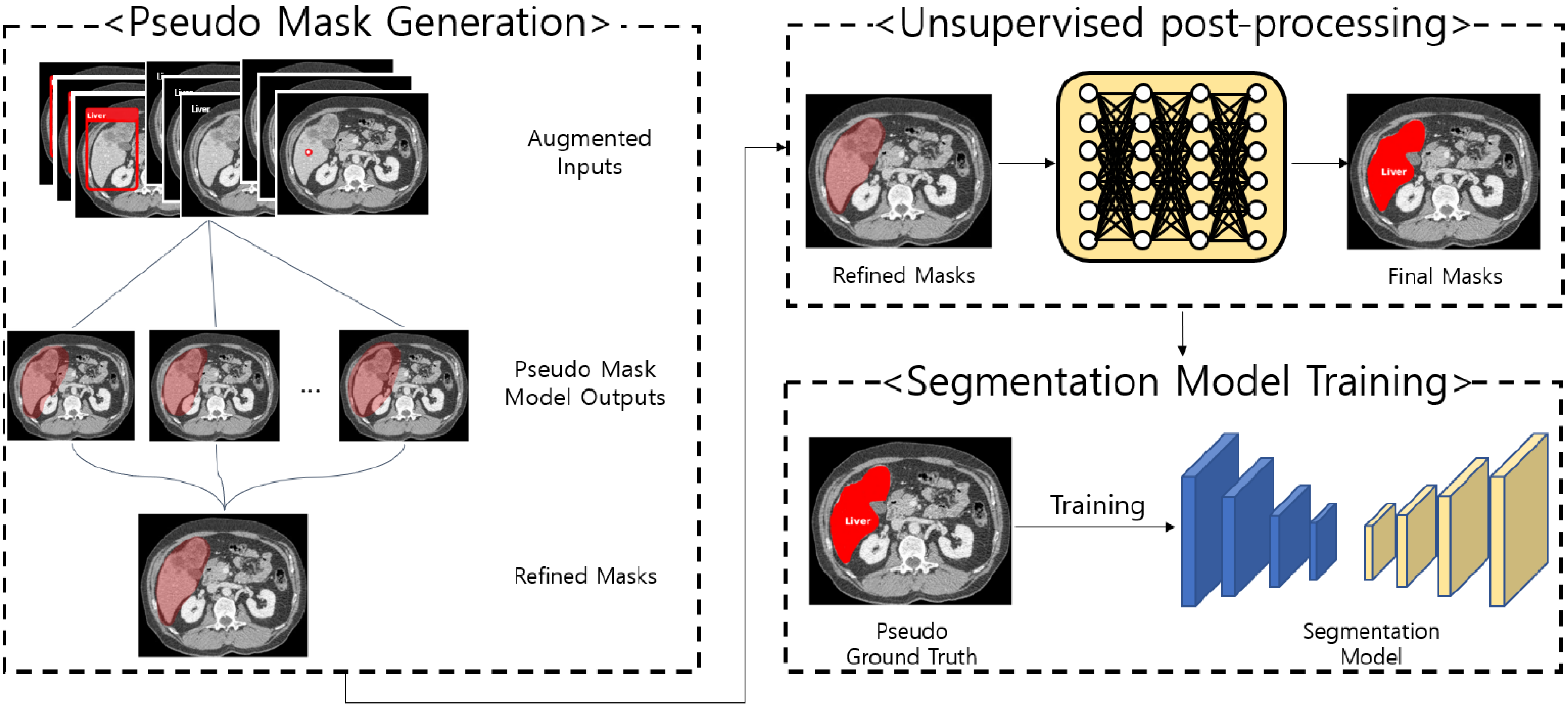
Original Images | Part Detection | 2nd Stage U-Net | Predictor

- **Predictor** : segmentation image.

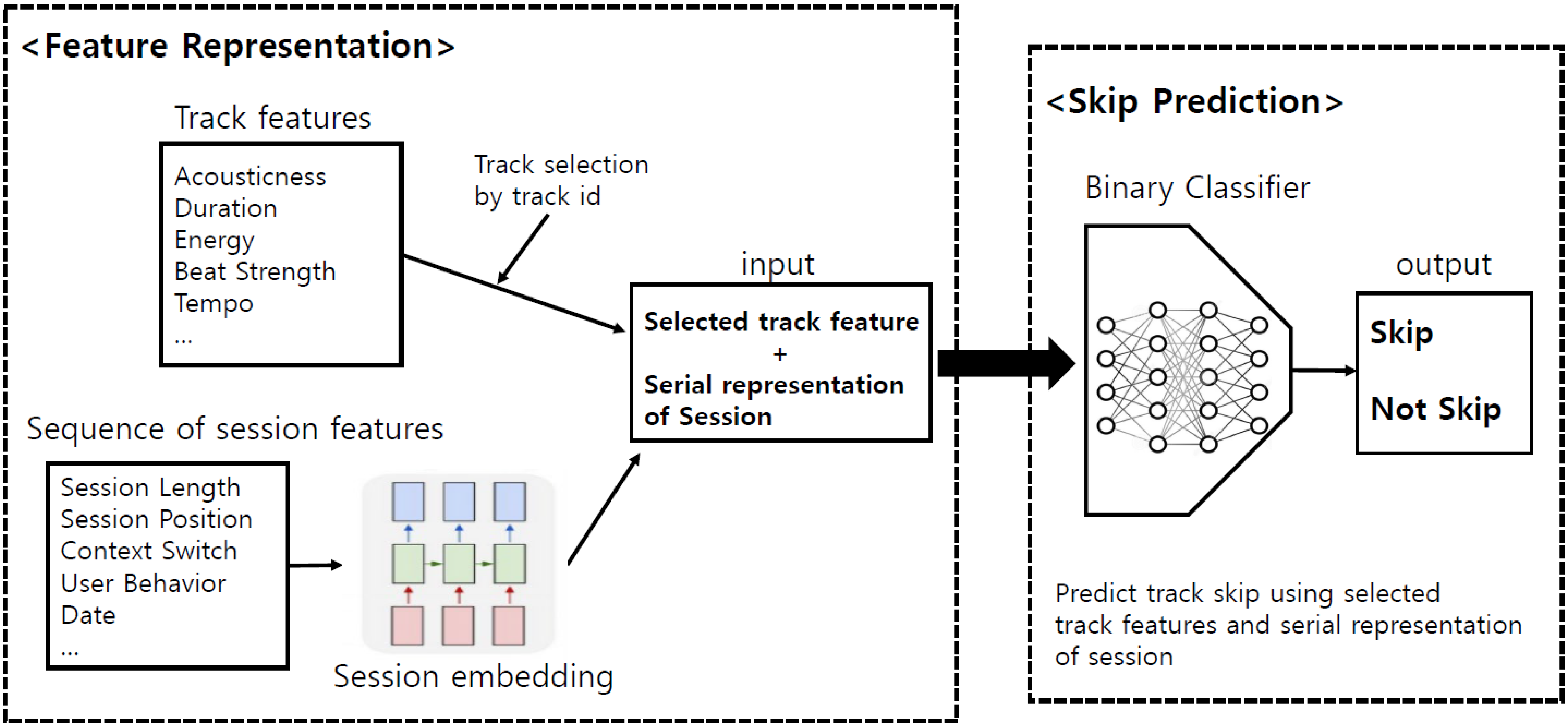


Medical Image Dataset Preprocessing
<Preprocessing custom image>
Raw Images
Augmentation
Segmentation
Labeled Image
<Segmentation Network>
Input
Convolutional Encoder-Decoder
Pooling Indices
Output
Segmentation
Conv + Batch Normalisation + ReLU
Pooling
Upsampling
Softmax
Supervised-Learning
Determination of illness

Medical Weakly Supervised Semantic Segmentation



Track Skip Prediction in Session

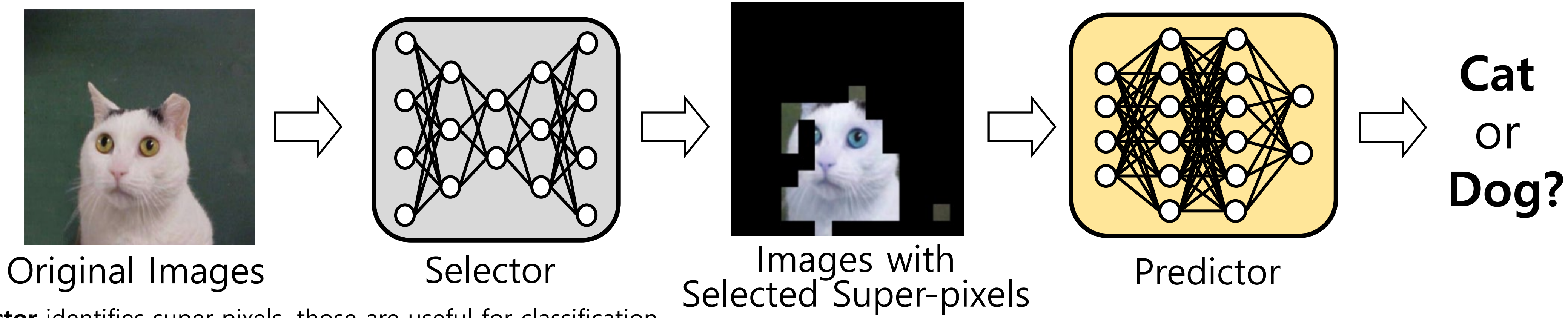


Individual Studies:

Sanghyuck Lee

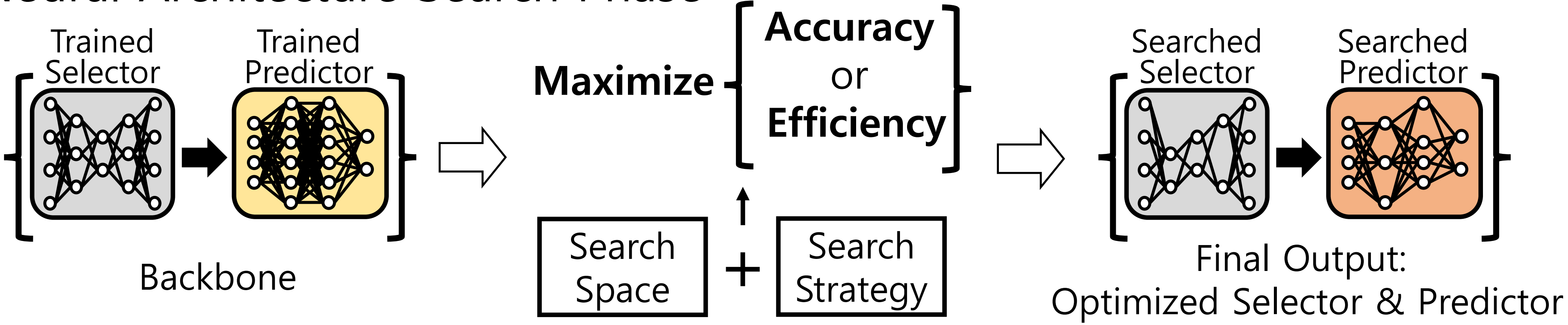
Neural Architecture Search based on Super-pixel Selection

<Selector & Predictor Train Phase>



- **Selector** identifies super-pixels, those are useful for classification.
- **Predictor** classifies using selected super pixels.
- **Super-pixel** is a pixel cluster consisting of $n \times n$ pixels.

<Neural Architecture Search Phase>



- Through **Neural Architecture Search** Phase using the **trained Selector & Predictor as a Backbone Network**, find **final optimized Selector & Predictor** from **Accuracy or Efficiency** point of view.

Progress

- Super-pixel Selection Approach Implementation (Real-X)
  - Jethani et al., 2021

- In the Future:
  - Various Datasets Experiments with Real-X
  - Neural Architecture Search Study and Experiments

Master’s Thesis

- Title:
  - Korean: 정확도 높은 의료 영상 분석을 위한 특징 선택 기반 신경망 구조 탐색
  - English: Neural Architecture Search based on Feature Selection for Accurate Medical Image Analysis

- Abstract

딥러닝 기술에 기반한 의료 영상 분야는 광범위하게 연구되고 있는 분야이다. 최근에는, AutoML이 인공지능 분야에서 상당한 관심을 받으면서, NAS에 기반한 의료용 신경망 기술의 유망함이 널리 알려지고 있다. 하지만, 기존 연구들은 불필요한 입력 값에 대해 효과적으로 대처하지 못하고, 이들로부터 학습에 방해를 받아 낮은 정확도를 보이고, 뿐만 아니라 낮은 효율성까지 보일 수 있다. 본 연구에서는, 특징 선택에 기반하여 신경망을 자동으로 설계하는, NAS 기술을 제안한다. 본 기술은 인간 전문가가 제안한 슬라이스 컷보다 적은 슬라이스를 선택하면서, 더 높은 정확도를 보인다.

Individual Studies:

Sung-Hyun Cho

Progress

- Last Report
  - Comparison result

- This Report
  - Abstract
  - Comparison result

Abstract

Music Genre Classification (MGC) is the most actively studied in the field of Music Information Retrieval (MIR) and reveals its effectiveness in practice. Among many classification algorithms, deep learning models using Convolutional Neural Network (CNN) show the best performance in classification. Existing music exceprts in the form of voice signals are converted to images such as Short-Time Fourier Transform (STFT), mel-spectrogram, and Mel-Frequency Cepstral Coefficient (MFCC) to perform training and testing. However, recent studies focused on classifying music genre using only one input type. In this study, I propose a novel DenseNet model that increases classification accuracy by using all three : STFT, mel-spectrogram, MFCC. Furthermore, to solve existing problem of ????, ???? is implemented in this paper. Extensive experiments were conducted on datasets of a total of 9 music genre domains, and the classification accuracy of proposed model was noteworthy when compared to other models.

Related Work 1

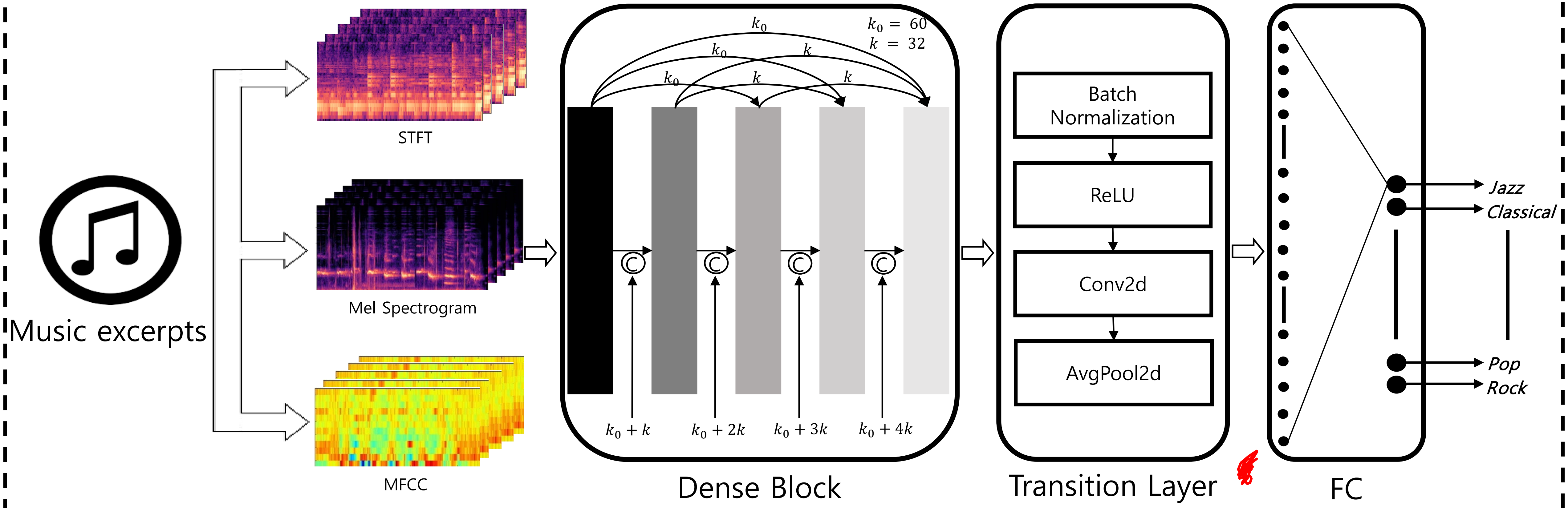
Author	Year	Neural Network	# of employed datasets	Employed Dataset	Input Type	Evaluation Metric	Ref.
Massoudi	2021	CNN	1	Urban Sound Classification Challenge	MFCC	Accuracy	[1]
Ghildiyal	2020	CNN	1	GTZAN	Mel-Spectrogram	Accuracy	[2]
Palanisamy	2020	CNN	3	ESC-50, GTZAN, UrbanSound8K	Log Mel-Spectrogram	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]

Proposed Method

<Architecture>



<Configuration>

Conv2d → Max Pool → [repeat: Dense Block → Transition Layer] → Dense Block → Classification Layer (Softmax)

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	Li et al.	Elbir et al.	-	-
Ballroom	58.79 ± 3.96	48.86 ± 4.41	39.29 ± 3.82	-	-
Ballroom-Extended	$\mathbf{85.32 \pm 0.61}$	75.89 ± 3.09	39.07 ± 1.49	-	-
FMA-MEDIUM	-	-	-	-	-
FMA-SMALL	$\mathbf{54.34 \pm 1.24}$	-	37.41 ± 0.74	-	-
GTZAN	$\mathbf{87.70 \pm 3.01}$	74.50 ± 3.78	58.85 ± 3.79	-	-
HOMBURG	54.13 ± 1.91	$\mathbf{54.50 \pm 2.93}$	47.14 ± 1.89	-	-
ISMIR04	80.07 ± 1.64	68.36 ± 1.61	62.28 ± 2.42	-	-
MICM	$\mathbf{49.39 \pm 1.79}$	39.56 ± 2.53	39.12 ± 2.78	-	-
Seyerlehner-Unique	$\mathbf{71.03 \pm 1.49}$	63.88 ± 1.53	55.59 ± 3.45	-	-
Avg . Rank	-	-	-	-	-

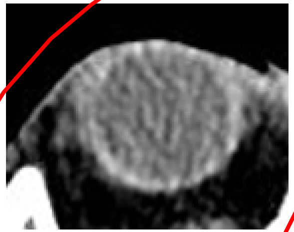
Individual Studies:

Su-Jeong Han

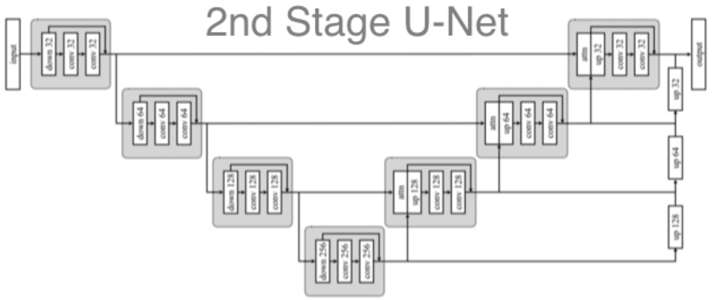
Efficient segmentation using U-Net

< U-Net using OCR >

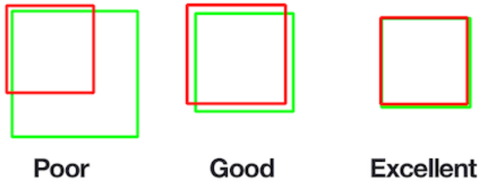
Original Images



Part Detection



2nd Stage U-Net



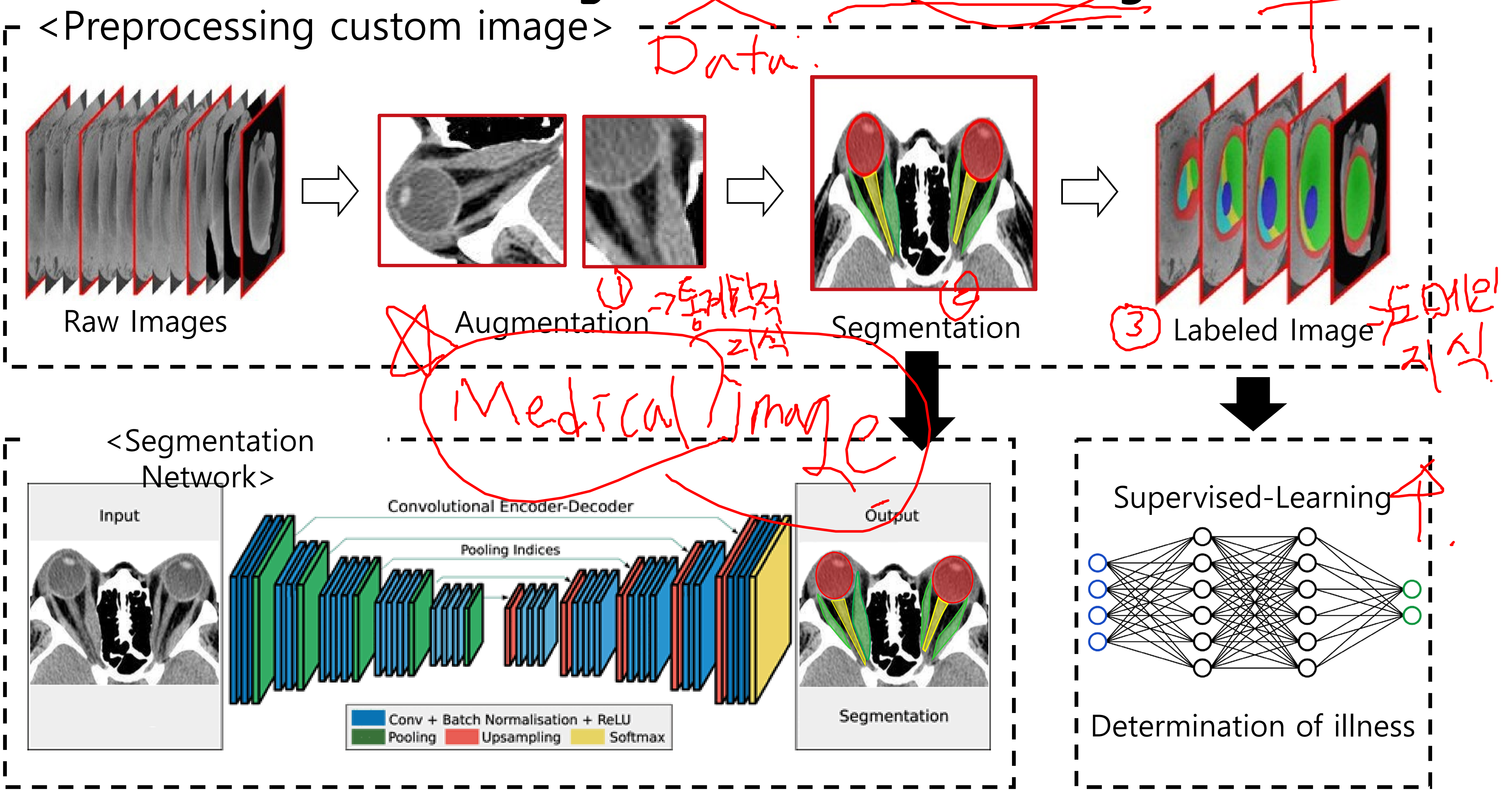
Predictor

- **Predictor** : segmentation image.

Individual Studies:

Taekyung Song

Medical Image Dataset Preprocessing



Progress

Hyper-parameter 조

바꾸는 시도 → 인정X.

찾아낸 HP는 설득 없음.

- Last Report
  - ImageNet Classification with Deep Convolutional Neural Networks
    - Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton / NIPS, 2012
- This Report
  - Very Deep Convolutional Network for Large-Scale Image Recognition
    - Karen Simonyan, Andrew Zisserman / ICLR, 2015
  - Network In Network
    - Min Lin, Qiang Chen, Shuicheng Yan / ICLR, 2014
  - Going deeper with convolutions
    - Christian Szegedy et al / CVPR, 2015

HP 최적화 관련 연구 어려움.

+ On Large-Batch Training for Deep Learning: Generalization Gap and Sharp
  - N. S. Keskar et al / ICLR, 2017

Next progress

- The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset

  - I Kandel, M Castelli / ICT express, 2020

- Bag of Tricks for Image Classification with Convolutional Neural Networks

  - T He, Z Zhang, H Zhang, Z Zhang / CVPR, 2019

- Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift

  - S Ioffe, C Szegedy / ICML, 2015

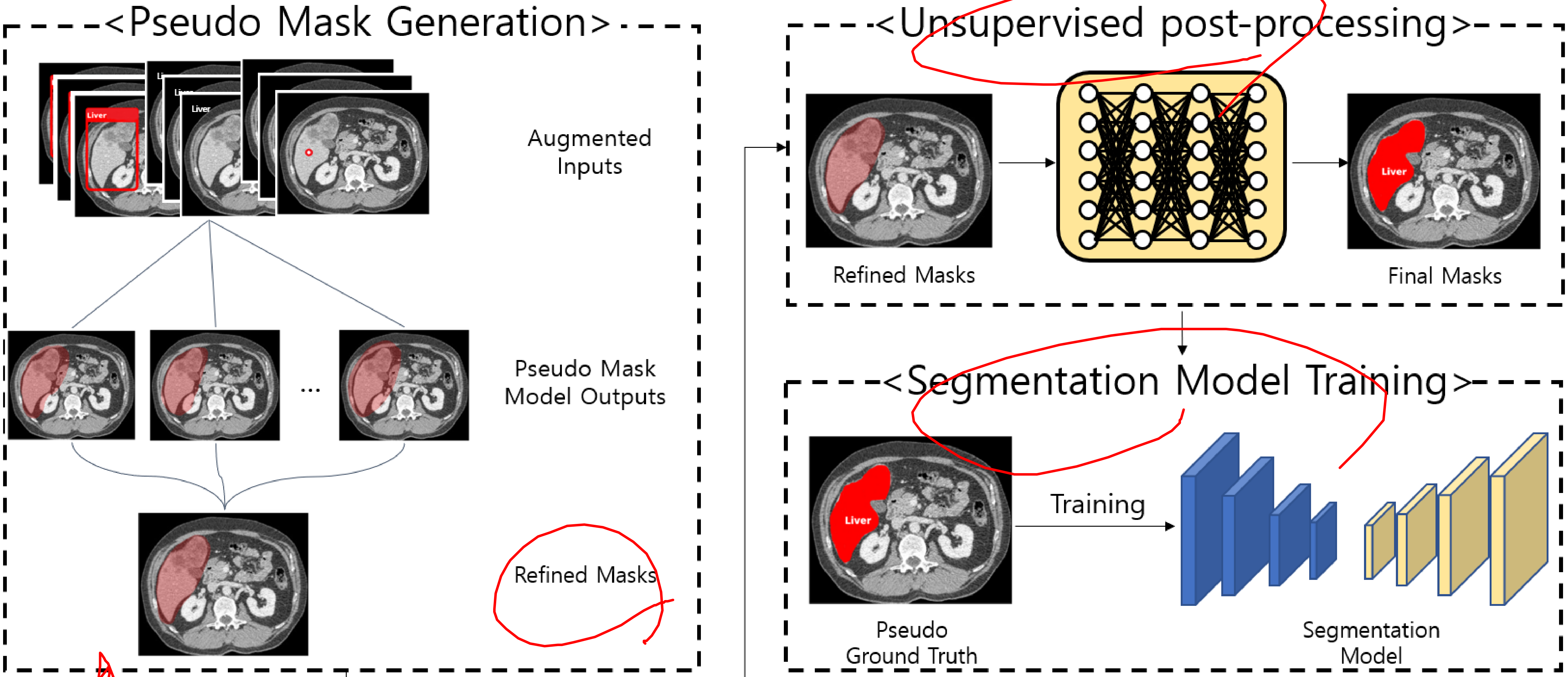
- EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks

  - M Tan, Q Le / ICML, 2019

Individual Studies:

Haesung Oh

Medical Weakly Supervised Semantic Segmentation



Performance measure : IoU

Progress

- Last Report
  - CAM/Grad-CAM 구현 및 결과 확인

- This Report
  - Dense-Conditional Random field (Dense CRF) 구현 및 결과 확인

Dense CRF

Paper Title: **Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials (Dense CRF)**

Authors: Krähenbühl, P., & Koltun, V.

Conference: Neural IPS 2012

Dense CRF

Original Image

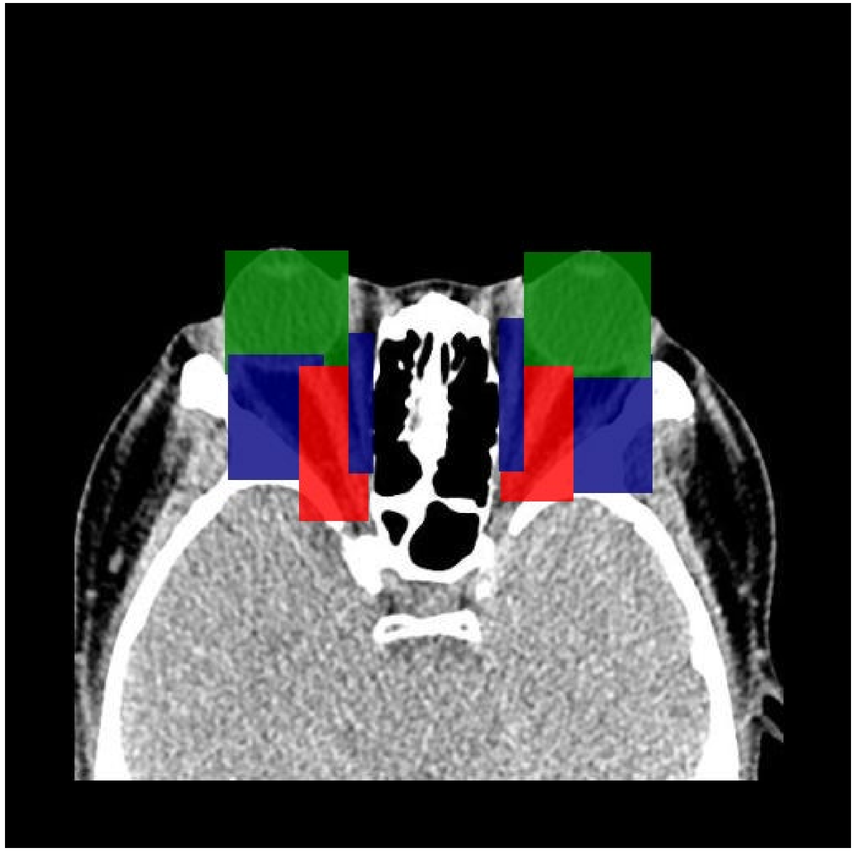
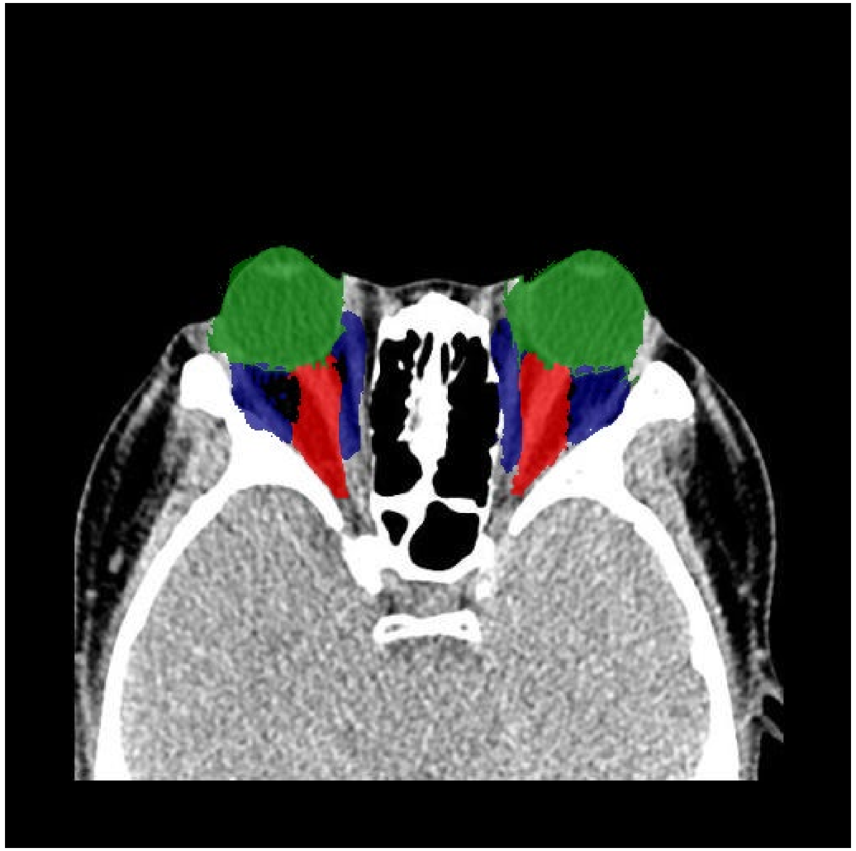


Image + Bounding box

Dense CRF

CRF iter-1



CRF iter-10

Further study

1. CAM / Grad-CAM
2. Dense CRF
3. Deep Seeded Region Growing (DSRG)
4. Fickle Net

Individual Studies:

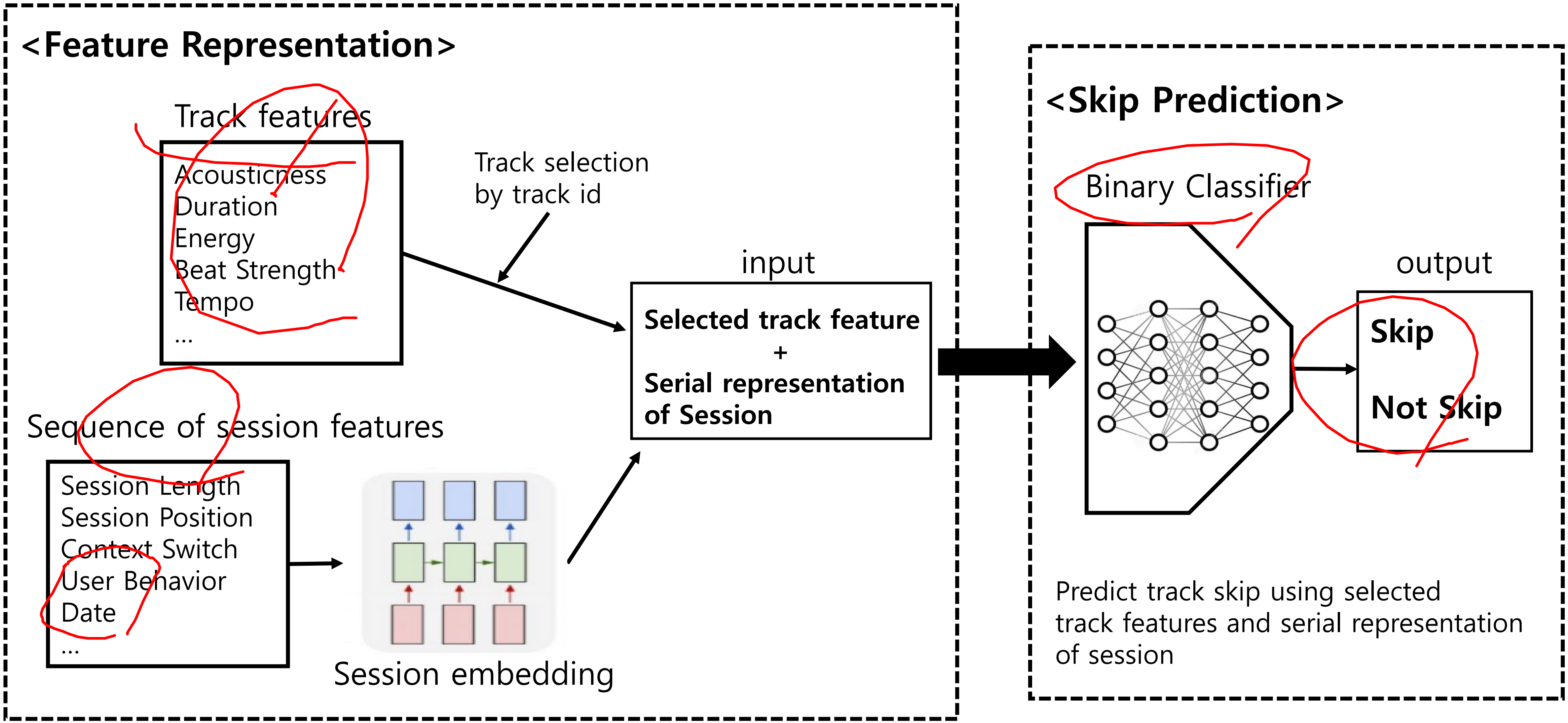
Sanghyeong Jin

Progress

- Last Report
  - Logistic Regression 적용
  - 성능평가 지표 MAA

- This Report
  - Boosting tree 관련 논문(A. Ferraro et al.) 구현
  - Multi-RNN 관련 논문(Christian Hansen et al.) 구현

Track Skip Prediction in Session



Previous - MAA

Ground truth sequence	Predicted sequence	AA
1, 1, 1, 1, 1	0, 1, 1, 1, 1	0.543
1, 1, 1, 1, 1	1, 0, 1, 1, 1	0.643
1, 1, 1, 1, 1	1, 1, 0, 1, 1	0.710
1, 1, 1, 1, 1	1, 1, 1, 0, 1	0.760
1, 1, 1, 1, 1	1, 1, 1, 1, 0	0.800

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

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

T : 예측되는 전체 트랙의 개수
A(i) : i번째 트랙까지의 정확도
L(i) : i번째 정답 여부에 대한 Boolean indicator (0, 1)

-> Skip 문제와 관련하여 제안된 평가 방법
-> Skip 순서에 있어 **첫번째 예측의 정확도가 가장 중요 (First Prediction Accuracy)**
-> 후반부 순서의 Skip 결과는 예측 정확도에 대해 낮은 영향력

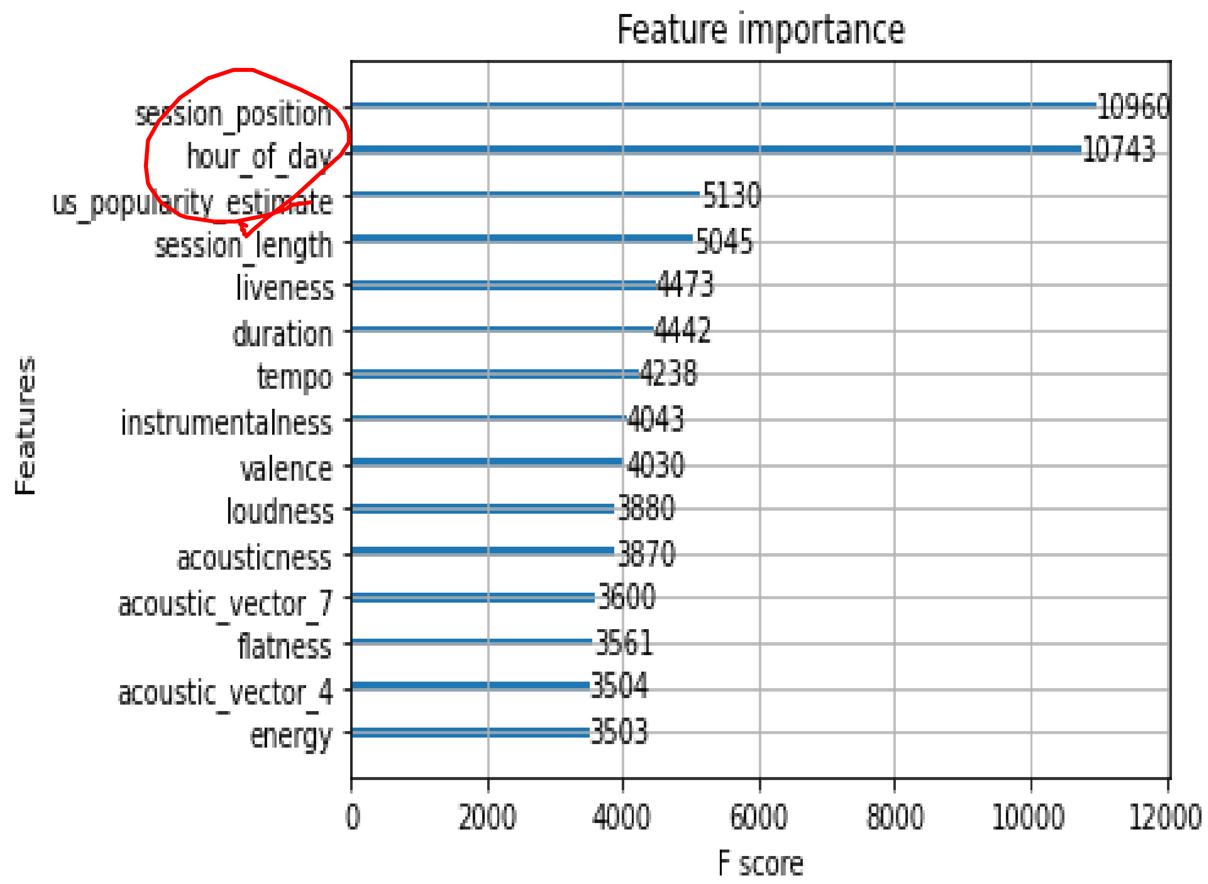
Previous – Logistic Regression

```
                                              Estimate Std. Error z value Pr(>|z|)
(Intercept)                                  0.6421248  2.2020049   0.292 0.770585
duration                                    -0.0021591  0.0001353 -15.961  < 2e-16 ***
release_year                                -0.0021660  0.0011018  -1.966 0.049317 *
acousticness                                -0.2581741  0.1405965  -1.836 0.066317 .
beat_strength                               -1.3457630  0.3174467  -4.239 2.24e-05 ***
bounciness                                   0.7468733  0.2582489   2.892 0.003827 **
energy                                      -0.2819291  0.0778488  -3.621 0.000293 ***
flatness                                    -1.3482182  0.3658634  -3.685 0.000229 ***
instrumentalness                             0.0937035  0.0656026   1.428 0.153191
key                                          0.0034210  0.0022805   1.500 0.133589
liveness                                    -0.1093288  0.0569146  -1.921 0.054741 .
mechanism                                    0.4241828  0.2152769   1.970 0.048792 *
organism                                     0.5422013  0.3345318   1.621 0.105066
tempo                                       -0.0006265  0.0003988  -1.571 0.116183
valence                                      0.1337574  0.0505400   2.647 0.008131 **
acoustic_vector_4                           -0.1317812  0.0427264  -3.084 0.002040 **
acoustic_vector_6                            0.1168963  0.0345982   3.379 0.000728 ***
acoustic_vector_7                           -0.2130206  0.0477388  -4.462 8.11e-06 ***
session_length                               0.0172787  0.0024419   7.076 1.48e-12 ***
context_switch                              -0.4386746  0.0431305 -10.171  < 2e-16 ***
no_pause_before_play                         0.7437025  0.0343424  21.655  < 2e-16 ***
short_pause_before_play                      0.8829114  0.0532501  16.580  < 2e-16 ***
long_pause_before_play                      -0.3168883  0.0589580  -5.375 7.67e-08 ***
hist_user_behavior_n_seekback               -1.3217715  0.0435101 -30.379  < 2e-16 ***
hist_user_behavior_is_shuffle                0.1788305  0.0197577   9.051  < 2e-16 ***
premium                                     -0.1843370  0.0224916  -8.196 2.49e-16 ***
context_type_catalog                        -0.1348362  0.0223796  -6.025 1.69e-09 ***
context_type_charts                         -0.1022746  0.0694048  -1.474 0.140590
context_type_editorial_playlist             -0.0970320  0.0230656  -4.207 2.59e-05 ***
context_type_personalized_playlist          -0.5692851  0.0591190  -9.629  < 2e-16 ***
context_type_radio                           0.0565769  0.0290573   1.947 0.051525 .
hist_user_behavior_reason_start_appload      0.5799226  0.1661346   3.491 0.000482 ***
hist_user_behavior_reason_start_backbtn     -0.3667445  0.1535303  -2.389 0.016906 *
hist_user_behavior_reason_start_clickrow    -0.4512124  0.1500658  -3.007 0.002640 **
hist_user_behavior_reason_start_endplay      3.2354986  0.9492867   3.408 0.000654 ***
hist_user_behavior_reason_start_fwdbtn       0.4047911  0.1494822   2.708 0.006770 **
hist_user_behavior_reason_start_trackdone   -1.7442696  0.1497636 -11.647  < 2e-16 ***
hist_user_behavior_reason_end_backbtn        8.3736786  0.0797148 105.046  < 2e-16 ***
hist_user_behavior_reason_end_clickrow      16.2579912 41.5248300   0.392 0.695410
hist_user_behavior_reason_end_endplay        6.0192884  0.0695008  86.608  < 2e-16 ***
hist_user_behavior_reason_end_fwdbtn         6.6007104  0.0664345  99.357  < 2e-16 ***
hist_user_behavior_reason_end_logout         4.9968260  0.1231531  40.574  < 2e-16 ***
hist_user_behavior_reason_end_remote         5.8277837  0.1528635  38.124  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 232533  on 167879  degrees of freedom
Residual deviance:  93295  on 167837  degrees of freedom
AIC: 93381
```

Stepwise selection으로 Reduce된 Model
- 29개의 track feature 중 17개, Session에서 10개가 선택됨
- 예상보다 결과가 좋았음

Fold	Fold MAA	Mean
1	0.8455	0.8503
2	0.8530	
3	0.8480	
4	0.8503	
5	0.8495	
6	0.8507	
7	0.8467	
8	0.8528	
9	0.8546	
10	0.8515	

Boosting Tree Approach (A. Ferraro et al. 2019.)

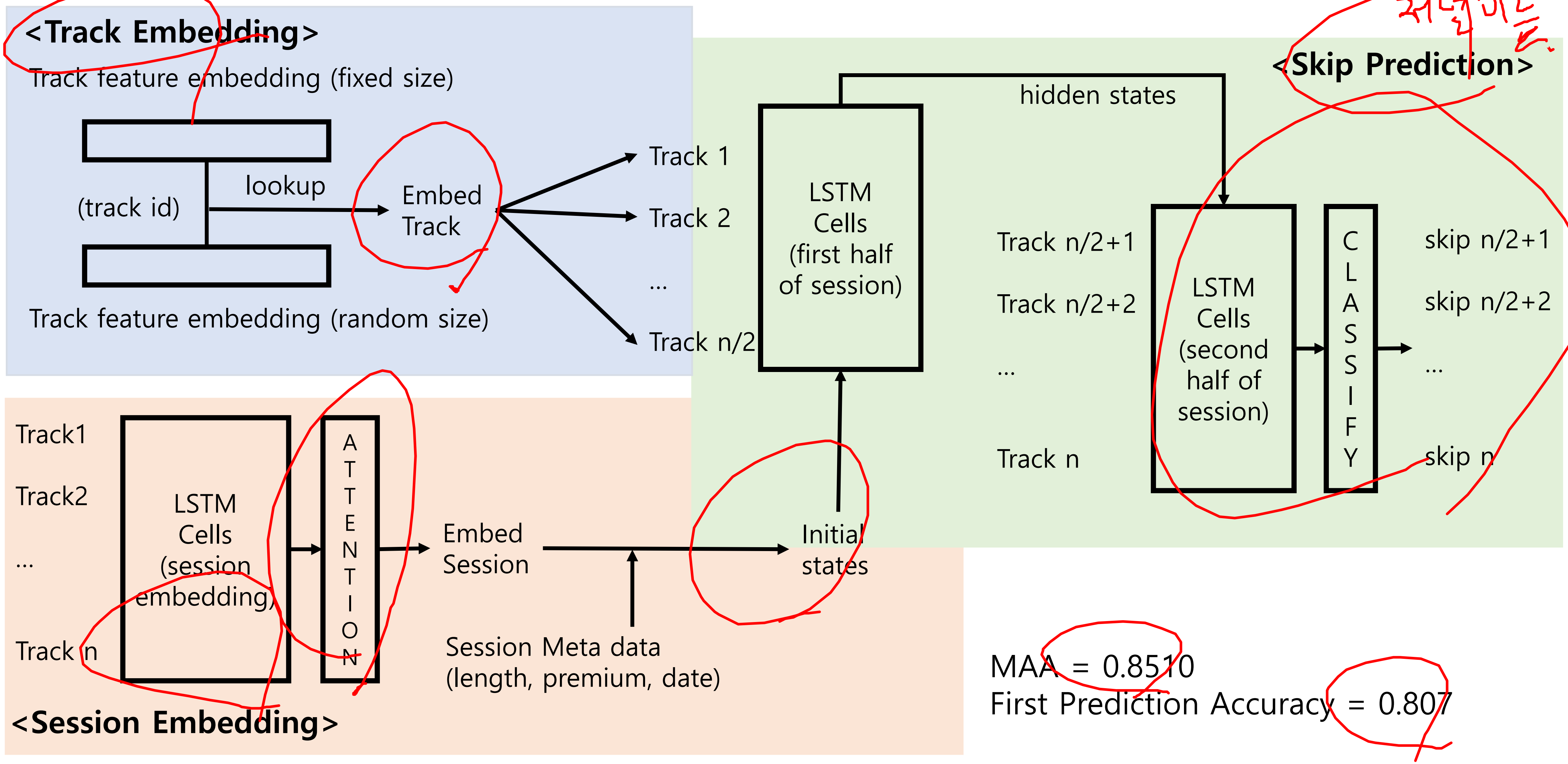


Model = XGBoost, num_boost_round = 200, max_depth = 15, eta = 0.3, colsample_bytree = 1.0, subsample = 1.0

Logistic Regression에 비해 성능은(MAA) 더 높게 측정되었으나 유의미한 차이가 있다고 보기는 힘듦 (First Prediction Accuracy = 0.735)

Fold	Fold MAA	Mean
1	0.8528	0.8565
2	0.8568	
3	0.8567	
4	0.8508	
5	0.8563	
6	0.8592	
7	0.8600	
8	0.8608	
9	0.8587	
10	0.8525	

Multi-RNN Approach (Christian Hansen et al. 2019.) +컨퍼런스 저널이고



MAA = 0.8510

First Prediction Accuracy = 0.807

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

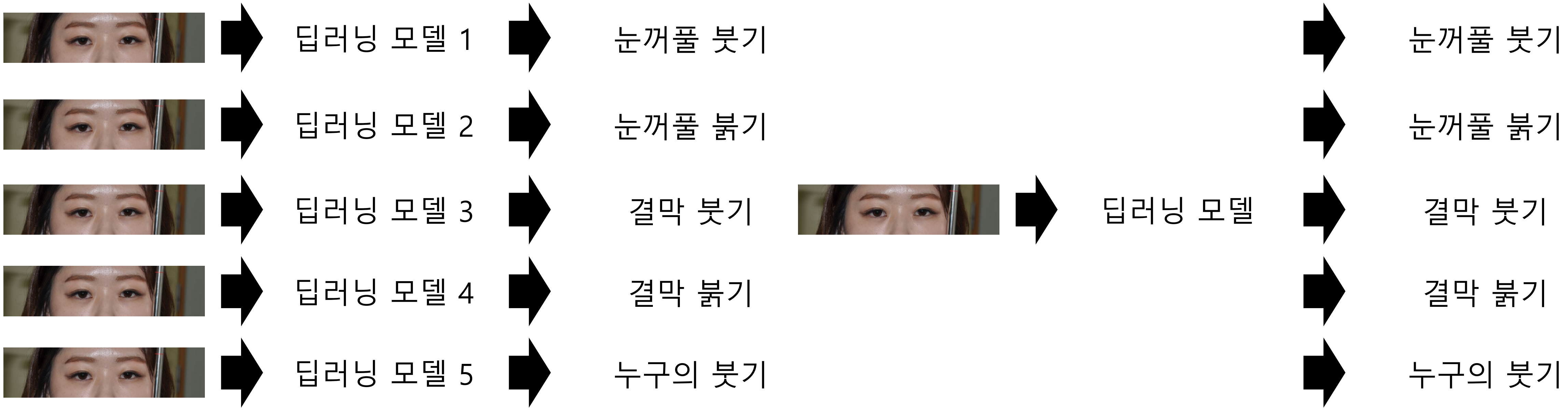
진행사항

- ✓ Clinical Activity Score (CAS) 5점 척도 점수 추론 연구 (이상혁)
  - ➢ 지난 보고 : (사진 only) 5개 모델 사용 Baseline 실험 보고
  - ➢ 이번 보고 : (사진 only) 1개 모델 사용 Baseline (+ Augmentation) 실험 보고

- ✓ 3D 세그멘테이션 (한수정)
  - ➢ 지난 보고 : AX7 세그멘테이션 진행, 부피 계산 자동화
  - ➢ 이번 보고 : AX7 세그멘테이션 완료, 1명 환자의 1개 타겟 3D 시각화

CAS 5점 척도 점수 추론 연구

- 지지난주 : 5개 모델 사용 실험
- 지난주 : 1개 모델 사용 실험



<지지난주 : 5개 모델 사용 실험>

<지난주 : 1개 모델 사용 실험>

CAS 5점 척도 점수 추론 연구

크기 resize x

- 지지난주 : 5개 모델 사용 실험
- 지난주 : 1개 모델 사용 실험
- 이번주 : 1개 모델 Augmentation 실험 => 전체 1,210 환자 중, 증상이 하나라도 있으면 Augmentation (좌우 반전, 밝기, 대비, ~~채도~~, 노이즈 사용), 전체 중, 대략 절반 정도의 환자가 적용됨.

클래스명	클래스 비율 (False vs True)
Swelling of Eyelid	672 : 538
Redness of Eyelid	1,094 : 116
Swelling of Conjuinctiva	1,121 : 89
Redness of Conjunctiva	954 : 256
Swelling of Caruncle	1,166 : 44

Augmentation

딥러닝 모델

1개씩 고려해서

눈꺼풀 붓기

눈꺼풀 붉기

결막 붓기

결막 붉기

누구의 붓기

<이번 주 : 1개 모델 Augmentation 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
ResNet101	0.630	0.579 0.682		0.537	0.485 0.588	0.950	0.564	0.503 0.625		0.565	0.532 0.598	0.105	0.692	0.553 0.830	0.001	0.472	0.268 0.675	0.310	0.512	0.448 0.575	
ECA	0.628	0.596 0.660	0.943	0.530	0.482 0.578	0.673	0.517	0.463 0.571	0.219	0.564	0.499 0.629	0.385	0.792	0.617 0.966	0.024	0.328	0.140 0.515	0.034	0.469	0.428 0.510	0.155
SE	0.628	0.596 0.660	0.943	0.530	0.482 0.578	0.673	0.517	0.463 0.571	0.219	0.564	0.499 0.629	0.385	0.792	0.617 0.966	0.024	0.328	0.140 0.515	0.034	0.469	0.428 0.510	0.155
InceptionV2	0.627	0.603 0.650	0.899	0.519	0.483 0.554	0.166	0.531	0.486 0.576	0.448	0.536	0.451 0.622	0.185	0.708	0.515 0.902	0.008	0.400	0.190 0.609	0.154	0.467	0.432 0.503	0.232
DenseNet264	0.618	0.591 0.646	0.703	0.539	0.509 0.568		0.562	0.521 0.603	0.944	0.538	0.503 0.574	0.030	0.636	0.491 0.781	0.000	0.512	0.340 0.684	0.594	0.504	0.452 0.557	0.841
CoAtNet	0.615	0.584 0.645	0.614	0.527	0.477 0.577	0.630	0.542	0.494 0.590	0.382	0.558	0.483 0.632	0.273	0.738	0.565 0.912	0.008	0.394	0.213 0.576	0.170	0.480	0.448 0.512	0.226
EfficientNet	0.570	0.539 0.602	0.004	0.467	0.430 0.504	0.012	0.551	0.518 0.585	0.642	0.475	0.419 0.531	0.001	0.510	0.380 0.639	0.000	0.579	0.436 0.722		0.477	0.438 0.516	0.191
GhostNet	0.539	0.504 0.574	0.005	0.459	0.428 0.490	0.002	0.529	0.485 0.574	0.240	0.477	0.415 0.538	0.003	0.541	0.394 0.689	0.000	0.525	0.376 0.674	0.543	0.457	0.416 0.498	0.043
MobileNetV3	0.533	0.499 0.568	0.000	0.447	0.419 0.476	0.001	0.491	0.445 0.537	0.032	0.495	0.418 0.573	0.023	0.651	0.459 0.843	0.003	0.378	0.178 0.578	0.159	0.432	0.395 0.470	0.010
GoogLeNet	0.511	0.479 0.543	0.000	0.426	0.394 0.458	0.000	0.419	0.398 0.441	0.001	0.590	0.569 0.611		1.000	1.000 1.000		0.000	0.000 0.000	0.000	0.419	0.398 0.441	0.005
ShuffleNet	0.496	0.468 0.523	0.002	0.449	0.392 0.506	0.017	0.515	0.451 0.579	0.071	0.311	0.088 0.533	0.015	0.503	0.130 0.876	0.015	0.497	0.122 0.872	0.656	0.451	0.209 0.692	0.622
VGGNet19	0.494	0.459 0.530	0.000	0.424	0.389 0.459	0.001	0.452	0.413 0.492	0.006	0.460	0.314 0.606	0.081	0.705	0.436 0.974	0.035	0.276	0.014 0.538	0.040	0.363	0.269 0.458	0.035
CSWin	0.486	0.463 0.509	0.000	0.415	0.379 0.452	0.000	0.419	0.398 0.441	0.001	0.590	0.569 0.611		1.000	1.000 1.000		0.000	0.000 0.000	0.000	0.419	0.398 0.441	0.005
AlexNet	0.484	0.460 0.507	0.000	0.414	0.388 0.441	0.000	0.428	0.396 0.460	0.003	0.574	0.537 0.612	0.343	0.944	0.818 1.070	0.343	0.059	-0.074 0.192	0.000	0.421	0.399 0.444	0.006

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
DenseNet264	0.701	0.636 0.766		0.253	0.180 0.325		0.901	0.890 0.913	0.343	0.000	0.000 0.000	0.000	0.000	0.000 0.000	0.000	0.999	0.998 1.001	0.343	0.000	0.000 0.000	0.115
InceptionV2	0.691	0.638 0.745	0.681	0.205	0.160 0.251	0.102	0.888	0.863 0.913	0.183	0.056	-0.029 0.140	0.010	0.065	-0.036 0.166	0.000	0.978	0.943 1.013	0.187	0.049	-0.025 0.124	0.243
ECA	0.686	0.644 0.728	0.517	0.179	0.148 0.210	0.023	0.891	0.878 0.904	0.205	0.022	-0.014 0.059	0.000	0.018	-0.009 0.046	0.000	0.986	0.965 1.007	0.173	0.039	-0.036 0.114	0.215
SE	0.686	0.644 0.728	0.517	0.179	0.148 0.210	0.023	0.891	0.878 0.904	0.205	0.022	-0.014 0.059	0.000	0.018	-0.009 0.046	0.000	0.986	0.965 1.007	0.173	0.039	-0.036 0.114	0.215
ResNet101	0.685	0.607 0.763	0.604	0.214	0.148 0.279	0.146	0.867	0.792 0.943	0.335	0.029	-0.017 0.076	0.000	0.059	-0.060 0.178	0.000	0.956	0.860 1.052	0.329	0.045	-0.032 0.122	0.203
CoAtNet	0.682	0.623 0.742	0.419	0.224	0.159 0.289	0.319	0.869	0.822 0.916	0.152	0.101	-0.006 0.209	0.154	0.127	-0.016 0.269	0.000	0.951	0.886 1.015	0.118	0.098	-0.004 0.201	0.406
EfficientNet	0.652	0.589 0.715	0.189	0.163	0.130 0.196	0.010	0.882	0.870 0.895	0.005	0.044	-0.021 0.109	0.002	0.042	-0.028 0.112	0.000	0.974	0.956 0.993	0.012	0.055	-0.013 0.124	0.208
GhostNet	0.632	0.564 0.700	0.036	0.174	0.112 0.235	0.030	0.896	0.883 0.909	0.085	0.000	0.000 0.000	0.000	0.000	0.000 0.000	0.000	0.994	0.987 1.001	0.085	0.000	0.000 0.000	0.115
MobileNetV3	0.582	0.526 0.638	0.016	0.143	0.114 0.172	0.013	0.899	0.887 0.911	0.422	0.043	-0.013 0.099	0.001	0.029	-0.014 0.072	0.000	0.995	0.985 1.004	0.233	0.230	-0.068 0.528	
AlexNet	0.573	0.503 0.643	0.002	0.149	0.116 0.183	0.010	0.902	0.890 0.913		0.000	0.000 0.000	0.000	0.000	0.000 0.000	0.000	1.000	1.000 1.000		0.000	0.000 0.000	0.115
VGGNet19	0.554	0.513 0.594	0.001	0.132	0.110 0.155	0.005	0.901	0.889 0.913	0.343	0.000	0.000 0.000	0.000	0.000	0.000 0.000	0.000	0.999	0.998 1.001	0.343	0.000	0.000 0.000	0.115
ShuffleNet	0.548	0.495 0.601	0.001	0.171	0.087 0.255	0.163	0.509	0.207 0.811	0.015	0.097	0.023 0.170	0.017	0.500	0.123 0.877	0.015	0.500	0.124 0.876	0.015	0.054	0.013 0.094	0.192
GoogLeNet	0.540	0.496 0.583	0.000	0.132	0.096 0.169	0.006	0.098	0.087 0.110	0.000	0.179	0.160 0.198		1.000	1.000 1.000		0.000	0.000 0.000	0.000	0.098	0.087 0.110	0.339
CSWin	0.493	0.453 0.533	0.000	0.108	0.080 0.136	0.003	0.902	0.890 0.913		0.000	0.000 0.000	0.000	0.000	0.000 0.000	0.000	1.000	1.000 1.000		0.000	0.000 0.000	0.115

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
InceptionV2	0.715	0.669 0.761		0.190	0.120 0.260	0.742	0.933	0.917 0.948	0.343	0.036	-0.045 0.116	0.055	0.045	-0.057 0.148	0.000	0.993	0.977 1.009	0.343	0.029	-0.037 0.096	0.147
DenseNet264	0.711	0.630 0.793	0.906	0.173	0.115 0.230	0.416	0.935	0.919 0.950	0.341	0.044	-0.010 0.097	0.020	0.027	-0.005 0.059	0.000	0.996	0.991 1.002	0.168	0.150	-0.076 0.376	0.563
ResNet101	0.691	0.616 0.765	0.473	0.204	0.113 0.295		0.921	0.893 0.949	0.251	0.006	-0.007 0.019	0.000	0.011	-0.014 0.036	0.000	0.983	0.952 1.014	0.254	0.004	-0.005 0.013	0.092
ECA	0.683	0.627 0.740	0.273	0.175	0.096 0.253	0.416	0.926	0.910 0.942	0.237	0.034	-0.018 0.087	0.003	0.022	-0.011 0.055	0.000	0.987	0.968 1.006	0.167	0.092	-0.063 0.247	0.306
SE	0.683	0.627 0.740	0.273	0.175	0.096 0.253	0.416	0.926	0.910 0.942	0.237	0.034	-0.018 0.087	0.003	0.022	-0.011 0.055	0.000	0.987	0.968 1.006	0.167	0.092	-0.063 0.247	0.306
CoAtNet	0.672	0.611 0.733	0.195	0.156	0.100 0.212	0.237	0.907	0.863 0.951	0.149	0.068	0.001 0.134	0.134	0.082	-0.038 0.201	0.000	0.963	0.909 1.016	0.151	0.131	-0.004 0.266	0.346
GhostNet	0.599	0.547 0.650	0.007	0.098	0.066 0.131	0.020	0.917	0.891 0.942	0.112	0.068	-0.003 0.139	0.153	0.063	0.001 0.126	0.000	0.975	0.946 1.004	0.083	0.112	-0.039 0.264	0.413
EfficientNet	0.596	0.540 0.652	0.005	0.136	0.087 0.184	0.145	0.902	0.857 0.947	0.078	0.089	0.020 0.159	0.400	0.099	-0.002 0.199	0.000	0.957	0.909 1.004	0.068	0.247	-0.042 0.537	
MobileNetV3	0.544	0.498 0.589	0.000	0.075	0.059 0.092	0.007	0.912	0.889 0.935	0.067	0.015	-0.019 0.050	0.001	0.011	-0.014 0.036	0.000	0.973	0.945 1.002	0.062	0.025	-0.032 0.082	0.082
VGGNet19	0.527	0.471 0.583	0.000	0.088	0.056 0.120	0.013	0.935	0.923 0.948	0.168	0.000	0.000 0.000	0.000	0.000	0.000 0.000	0.000	0.999	0.997 1.001	0.168	0.000	0.000 0.000	0.086
CSWin	0.518	0.459 0.577	0.001	0.079	0.053 0.105	0.009	0.936	0.923 0.950		0.000	0.000 0.000	0.000	0.000	0.000 0.000	0.000	1.000	1.000 1.000		0.000	0.000 0.000	0.086
GoogLeNet	0.513	0.430 0.597	0.000	0.093	0.055 0.131	0.007	0.064	0.050 0.077	0.000	0.119	0.095 0.143		1.000	1.000 1.000		0.000	0.000 0.000	0.000	0.064	0.050 0.077	0.191
AlexNet	0.512	0.454 0.569	0.000	0.080	0.055 0.106	0.014	0.936	0.923 0.950		0.000	0.000 0.000	0.000	0.000	0.000 0.000	0.000	1.000	1.000 1.000		0.000	0.000 0.000	0.086
ShuffleNet	0.494	0.415 0.573	0.000	0.119	0.025 0.213	0.124	0.507	0.178 0.837	0.015	0.066	0.013 0.119	0.015	0.492	0.121 0.864	0.013	0.499	0.123 0.876	0.015	0.035	0.007 0.064	0.150

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
DenseNet264	0.753	0.717 0.789		0.492	0.426 0.558		0.802	0.769 0.836	0.780	0.336	0.220 0.452	0.595	0.306	0.153 0.459	0.000	0.915	0.848 0.983	0.020	0.605	0.452 0.758	
InceptionV2	0.749	0.707 0.791	0.841	0.452	0.395 0.509	0.204	0.729	0.645 0.813	0.088	0.364	0.260 0.468	0.915	0.455	0.266 0.644	0.000	0.799	0.655 0.943	0.012	0.393	0.239 0.548	0.041
CoAtNet	0.740	0.713 0.768	0.213	0.425	0.374 0.477	0.008	0.764	0.720 0.809	0.065	0.305	0.176 0.433	0.419	0.342	0.144 0.539	0.000	0.866	0.768 0.963	0.012	0.356	0.195 0.518	0.019
ECA	0.725	0.690 0.759	0.032	0.412	0.371 0.453	0.013	0.755	0.707 0.802	0.041	0.372	0.296 0.448		0.415	0.266 0.564	0.000	0.838	0.748 0.927	0.003	0.451	0.328 0.575	0.040
SE	0.725	0.690 0.759	0.032	0.412	0.371 0.453	0.013	0.755	0.707 0.802	0.041	0.372	0.296 0.448		0.415	0.266 0.564	0.000	0.838	0.748 0.927	0.003	0.451	0.328 0.575	0.040
ResNet101	0.723	0.690 0.757	0.208	0.400	0.362 0.439	0.016	0.731	0.657 0.805	0.043	0.351	0.279 0.422	0.532	0.408	0.254 0.562	0.000	0.808	0.684 0.931	0.007	0.443	0.290 0.595	0.059
EfficientNet	0.645	0.619 0.671	0.000	0.308	0.270 0.345	0.000	0.763	0.740 0.786	0.014	0.244	0.192 0.295	0.000	0.208	0.147 0.268	0.000	0.898	0.854 0.942	0.001	0.352	0.280 0.423	0.011
GhostNet	0.609	0.579 0.640	0.000	0.303	0.258 0.347	0.001	0.758	0.709 0.807	0.064	0.249	0.163 0.336	0.018	0.234	0.129 0.340	0.000	0.887	0.809 0.965	0.010	0.330	0.210 0.449	0.017
MobileNetV3	0.547	0.503 0.592	0.000	0.256	0.219 0.292	0.000	0.711	0.650 0.772	0.003	0.213	0.150 0.276	0.007	0.232	0.114 0.349	0.000	0.823	0.724 0.922	0.003	0.262	0.194 0.329	0.002
VGGNet19	0.546	0.496 0.595	0.000	0.235	0.193 0.276	0.000	0.729	0.622 0.836	0.141	0.063	-0.024 0.150	0.000	0.117	-0.057 0.291	0.000	0.877	0.704 1.050	0.142	0.053	-0.010 0.115	0.000
GoogLeNet	0.533	0.485 0.582	0.000	0.237	0.201 0.273	0.000	0.194	0.178 0.210	0.000	0.324	0.302 0.347	0.222	1.000	1.000 1.000		0.000	0.000 0.000	0.000	0.194	0.178 0.210	0.000
AlexNet	0.526	0.469 0.584	0.000	0.225	0.177 0.272	0.000	0.806	0.789 0.822		0.000	0.000 0.000	0.000	0.000	0.000 0.000	0.000	0.999	0.998 1.001	0.343	0.000	0.000 0.000	0.000
ShuffleNet	0.503	0.462 0.544	0.000	0.237	0.166 0.308	0.000	0.504	0.274 0.735	0.015	0.166	0.041 0.291	0.031	0.497	0.122 0.872	0.014	0.499	0.123 0.875	0.015	0.100	0.024 0.175	0.000
CSWin	0.502	0.466 0.538	0.000	0.197	0.171 0.222	0.000	0.806	0.790 0.822	0.343	0.000	0.000 0.000	0.000	0.000	0.000 0.000	0.000	1.000	1.000 1.000		0.000	0.000 0.000	0.000

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
InceptionV2	0.755	0.700 0.809		0.164	0.079 0.249		0.965	0.956 0.973		0.000	0.000 0.000	0.000	0.000	0.000 0.000	0.000	1.000	1.000 1.000		0.000	0.000 0.000	0.000
CoAtNet	0.741	0.679 0.802	0.709	0.129	0.074 0.185	0.492	0.960	0.950 0.971	0.280	0.025	-0.032 0.082	0.152	0.040	-0.050 0.130	0.000	0.994	0.983 1.006	0.293	0.018	-0.023 0.059	0.399
ResNet101	0.738	0.659 0.817	0.515	0.141	0.069 0.214	0.599	0.963	0.953 0.973	0.343	0.017	-0.021 0.054	0.017	0.014	-0.018 0.047	0.000	0.998	0.992 1.003	0.343	0.020	-0.025 0.065	0.462
ECA	0.728	0.642 0.814	0.434	0.101	0.076 0.126	0.110	0.963	0.955 0.971	0.193	0.020	-0.025 0.065	0.044	0.014	-0.018 0.047	0.000	0.998	0.994 1.001	0.168	0.033	-0.042 0.109	0.953
SE_ResNet50	0.728	0.642 0.814	0.434	0.101	0.076 0.126	0.110	0.963	0.955 0.971	0.193	0.020	-0.025 0.065	0.044	0.014	-0.018 0.047	0.000	0.998	0.994 1.001	0.168	0.033	-0.042 0.109	0.953
DenseNet264	0.714	0.636 0.792	0.165	0.157	0.091 0.223	0.890	0.964	0.955 0.973	0.343	0.000	0.000 0.000	0.000	0.000	0.000 0.000	0.000	0.999	0.998 1.001	0.343	0.000	0.000 0.000	0.000
EfficientNet	0.668	0.594 0.742	0.084	0.099	0.055 0.142	0.134	0.960	0.949 0.971	0.070	0.017	-0.021 0.054	0.017	0.014	-0.018 0.047	0.000	0.995	0.989 1.001	0.081	0.020	-0.025 0.065	0.462
GhostNet	0.632	0.555 0.708	0.016	0.076	0.042 0.109	0.045	0.964	0.955 0.973	0.343	0.000	0.000 0.000	0.000	0.000	0.000 0.000	0.000	0.999	0.998 1.001	0.343	0.000	0.000 0.000	0.000
MobileNetV3	0.570	0.503 0.637	0.001	0.054	0.037 0.070	0.024	0.958	0.943 0.973	0.081	0.015	-0.019 0.050	0.010	0.014	-0.018 0.047	0.000	0.992	0.984 1.001	0.083	0.017	-0.021 0.054	0.293
GoogLeNet	0.558	0.505 0.610	0.001	0.065	0.023 0.107	0.075	0.035	0.027 0.044	0.000	0.068	0.052 0.084		1.000	1.000 1.000		0.000	0.000 0.000	0.000	0.035	0.027 0.044	
CSWin	0.556	0.524 0.588	0.000	0.053	0.034 0.072	0.026	0.965	0.956 0.973		0.000	0.000 0.000	0.000	0.000	0.000 0.000	0.000	1.000	1.000 1.000		0.000	0.000 0.000	0.000
ShuffleNet	0.551	0.436 0.666	0.003	0.109	0.003 0.216	0.306	0.499	0.149 0.849	0.015	0.033	0.006 0.060	0.025	0.500	0.123 0.877	0.015	0.500	0.124 0.876	0.015	0.017	0.003 0.031	0.026
AlexNet	0.546	0.439 0.653	0.004	0.053	0.025 0.081	0.030	0.965	0.956 0.973		0.000	0.000 0.000	0.000	0.000	0.000 0.000	0.000	1.000	1.000 1.000		0.000	0.000 0.000	0.000
VGGNet19	0.502	0.417 0.586	0.001	0.069	0.010 0.127	0.078	0.964	0.955 0.973	0.343	0.000	0.000 0.000	0.000	0.000	0.000 0.000	0.000	0.999	0.998 1.001	0.343	0.000	0.000 0.000	0.000

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

토의

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
Swelling of Eyelid																					
W/O Aug	**0.627**	**0.582 0.672**		**0.55**	**0.488 0.611**		**0.579**	**0.514 0.644**		**0.503**	**0.391 0.616**	**0.591**	**0.596**	**0.363 0.829**	**0.267**	**0.566**	**0.312 0.821**		**0.547**	**0.446 0.648**	
With Aug	0.627	0.603 0.650	0.977	0.519	0.483 0.554	0.405	0.531	0.486 0.576	0.144	0.536	0.451 0.622		0.708	0.515 0.902		0.4	0.190 0.609	0.124	0.467	0.432 0.503	0.148
Redness of Eyelid																					
W/O Aug	**0.697**	**0.652 0.743**		**0.201**	**0.157 0.246**	**0.899**	**0.901**	**0.889 0.913**		**0**	**0.000 0.000**	**0.172**	**0**	**0.000 0.000**	**0.18**	**0.999**	**0.998 1.001**		**0**	**0.000 0.000**	**0.168**
With Aug	0.691	0.638 0.745	0.869	0.205	0.160 0.251		0.888	0.863 0.913	0.377	0.056	-0.029 0.140		0.065	-0.036 0.166		0.978	0.943 1.013	0.192	0.049	-0.025 0.124	
Swelling of Conjuinctiva																					
W/O Aug	**0.727**	**0.682 0.772**		**0.169**	**0.109 0.230**	**0.613**	**0.92**	**0.872 0.968**	**0.584**	**0.045**	**-0.023 0.113**		**0.058**	**-0.048 0.165**		**0.976**	**0.922 1.030**	**0.529**	**0.117**	**-0.108 0.342**	
With Aug	0.715	0.669 0.761	0.746	0.19	0.120 0.260		0.933	0.917 0.948		0.036	-0.045 0.116	0.858	0.045	-0.057 0.148	0.859	0.993	0.977 1.009		0.029	-0.037 0.096	0.436
Redness of Conjunctiva																					
W/O Aug	0.696	0.647 0.745	0.18	0.407	0.349 0.466	0.315	0.685	0.552 0.817	0.463	0.306	0.221 0.391	0.244	0.43	0.192 0.668	0.815	0.747	0.527 0.967	0.605	0.445	0.294 0.595	
With Aug	**0.749**	**0.707 0.791**		**0.452**	**0.395 0.509**		**0.729**	**0.645 0.813**		**0.364**	**0.260 0.468**		**0.455**	**0.266 0.644**		**0.799**	**0.655 0.943**		**0.393**	**0.239 0.548**	**0.457**
Swelling of Caruncle																					
W/O Aug	0.693	0.613 0.772	0.215	0.105	0.039 0.170	0.128	0.965	0.956 0.973		0	0.000 0.000		0	0.000 0.000		1	1.000 1.000		0	0.000 0.000	
With Aug	**0.755**	**0.700 0.809**		**0.164**	**0.079 0.249**		**0.965**	**0.956 0.973**	**1**	**0**	**0.000 0.000**		**0**	**0.000 0.000**		**1**	**1.000 1.000**		**0**	**0.000 0.000**	

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)	W/O Augmentation		With Augmentation	
	AUROC Average	Sensitivity Average	AUROC Average	Sensitivity Average
Swelling of Eyelid (3)	0.627	0.596	**0.627**	**0.708**
Redness of Eyelid (4)	0.697	0	0.691	0.065
Swelling of Conjuinctiva (4)	0.727	0.058	0.715	0.045
Redness of Conjunctiva (1)	0.698	0.43	**0.749**	**0.455**
Swelling of Caruncle (2)	0.693	0	0.755	0

- 클래스 비율(위 테이블 순서대로) : 672:538 / 1,094:116 / 1,121:89 / 954:256 / 1,166:44

극복 전략

- 지지난주 전략인 5개 모델 사용 전략 -> Sensitivity 확보
- Object Detection 기능을 추가해 사전에 눈을 발견하여, Classification하는 전략
- 증상이 있는 환자 위주로 데이터 추가 요청
- 의료 영상 혹은 사진 등, Augmentation 전략 추가 조사 및 실행

금주 이정규 교수님 미팅을 위한 토의

1) **(사진 only)** 데이터셋 난이도 높은 것으로 보임.

- 1, 4번 클래스를 제외하면, 학습이 어려움.

2) **(CT only)**도 어렵지 않을지 의견 궁금함.

- 각 클래스가 외관을 보고 판단하기 때문에

3) 곧장 **(CT+사진)**으로 가는 것은 어떨지, 의견 궁금함.

- 이 경우, 클래스 분류에 유의미한 CT cut을 선정해야 할듯.

3D 세그멘테이션

- Last Report
- AX7 segmentation 재진행
- 부피 구하기 자동화

- This Report
- AX7 segmentation 완료
- AX7 segmentation 1차 검수 완료
- 1075번 환자 왼쪽 눈 시각화

'20.10~ 최신 진행 사항

- 데이터셋 요약 : 예측 모델에 94개 특징 이미지 입력 완료

	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	12
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**

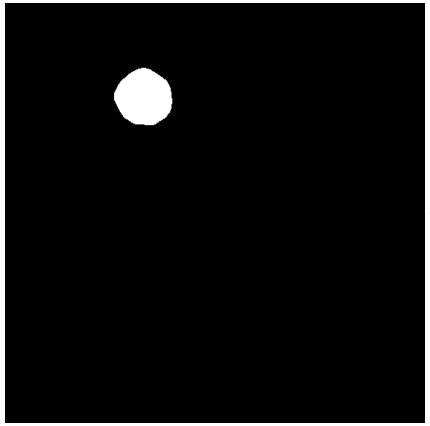
Ax7 세그멘테이션

- 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 진행 완료**

- Segmentation이 잘된 데이터를 모델에 넣어 predict 재진행

3D 세그멘테이션

<입력 샘플 예시>
(1075 환자의 왼쪽 눈 10040.dcm)



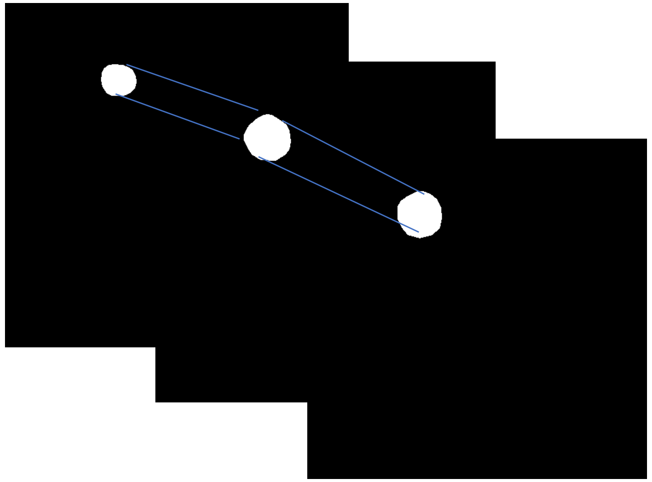
데이터의 사이즈:
512*512

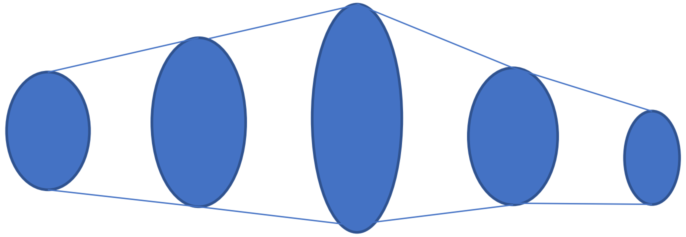
225 =white
0 = black

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



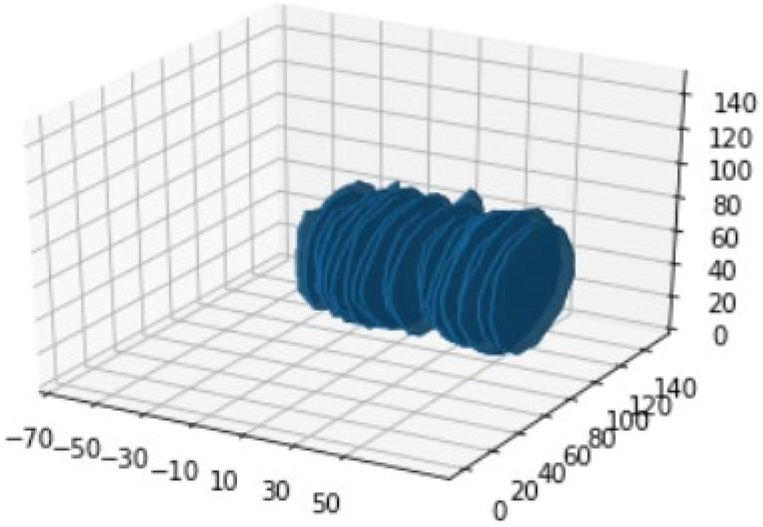


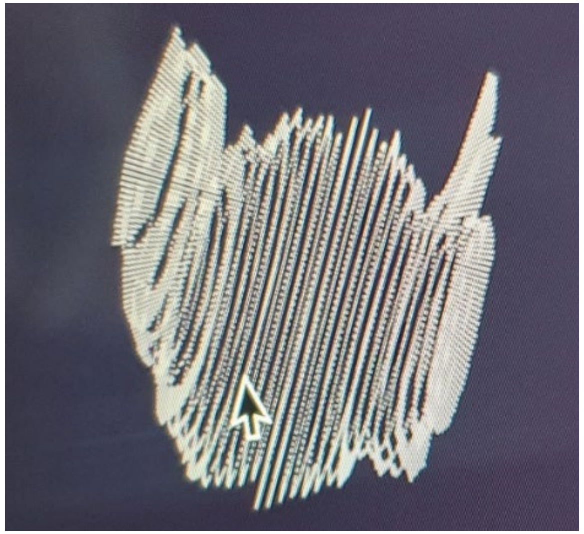
각 기둥의 부피를 계산
DICOM 파일에 생긴 간격에
있는 부피 예측

3D 시각화

1075번 환자

왼쪽 눈 : 89개의 데이터 시각화





분업으로 수동 segmentation 을 진행하였는데 각자가 생각한 눈의 기준이 달라서 좌측 결과를 초래

-> 다시 imageJ를 활용하여 수동 segmentation을 진행 할 예정

수동 segmentation 자체가 잘못되어 있음을 파악

Government-funded Research Projects:
산학협력프로젝트
(연구실 전체)

진행사항

- 지난 보고
  - Software 3건 등록 진행 중(~2월)

- 이번 보고
  - Software 3건 등록 진행 중(~4월)

Q&A

References

References (Sung-Hyun Cho)

[1] Massoudi, M., Verma, S., & Jain, R. (2021, January). Urban Sound Classification using CNN. In 2021 6th International Conference on Inventive Computation Technologies (ICICT) (pp. 583-589). IEEE.

[2]Ghildiyal, A., Singh, K., & Sharma, S. (2020, November). Music genre classification using machine learning. In 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 1368-1372). IEEE.

[3] Palanisamy, K., Singhania, D., & Yao, A. (2020). Rethinking cnn models for audio classification. arXiv preprint arXiv:2007.11154.

[4] Yu, Y., Luo, S., Liu, S., Qiao, H., Liu, Y., & Feng, L. (2020). Deep attention based music genre classification. Neurocomputing, 372, 84-91.

[5] Won, M., Ferraro, A., Bogdanov, D., & Serra, X. (2020). Evaluation of cnn-based automatic music tagging models. arXiv preprint arXiv:2006.00751.

[6] Bisharad, D., & Laskar, R. H. (2019, October). Music Genre Recognition Using Residual Neural Networks. In TENCON 2019-2019 IEEE Region 10 Conference (TENCON) (pp. 2063-2068). IEEE.

[7] Zhang, S., Gu, H., & Li, R. (2019). MUSIC GENRE CLASSIFICATION: NEAR-REALTIME VS SEQUENTIAL APPROACH.

[8] Chillara, S., Kavitha, A. S., Neginhal, S. A., Haldia, S., & Vidyullatha, K. S. (2019). Music genre classification using machine learning algorithms: a comparison. Int Res J Eng Technol, 6(5), 851-858.

References (Sung-Hyun Cho)

[9] Zhang, W., Lei, W., Xu, X., & Xing, X. (2016, September). Improved music genre classification with convolutional neural networks. In Interspeech (pp. 3304-3308).

[10] Choi, K., Fazekas, G., Sandler, M., & Cho, K. (2017, March). Convolutional recurrent neural networks for music classification. In 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 2392-2396). IEEE.

[11] Su, Y., Zhang, K., Wang, J., & Madani, K. (2019). Environment sound classification using a two-stream CNN based on decision-level fusion. Sensors, 19(7), 1733.

[12] Choi, K., Fazekas, G., & Sandler, M. (2016). Automatic tagging using deep convolutional neural networks. arXiv preprint arXiv:1606.00298.

[13] Li, S., Yao, Y., Hu, J., Liu, G., Yao, X., & Hu, J. (2018). An ensemble stacked convolutional neural network model for environmental event sound recognition. Applied Sciences, 8(7), 1152.

[14] Ashraf, M., Geng, G., Wang, X., Ahmad, F., & Abid, F. (2020). A Globally Regularized Joint Neural Architecture for Music Classification. IEEE Access, 8, 220980-220989.

[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.

[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.

References (Haesung Oh)

[1] Krähenbühl, Philipp, and Vladlen Koltun. "Efficient inference in fully connected crfs with gaussian edge potentials." *Advances in neural information processing systems* 24 (2011).

References (Sanghyeong Jin)

[1] A. Ferraro, D. Bogdanov, X. Serra. “Skip prediction using boosting trees based on acoustic features of tracks in sessions”, 12th ACM International Conference on Web Search and Data Mining. 2019.

[2] 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.

Appendix

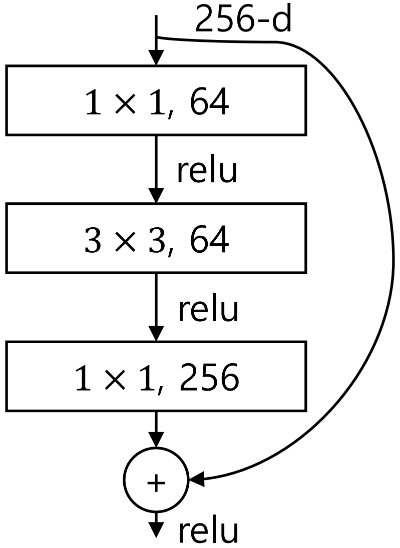
Individual Studies:

Sung-Hyun Cho

Code Analysis(ResNet50_trust)

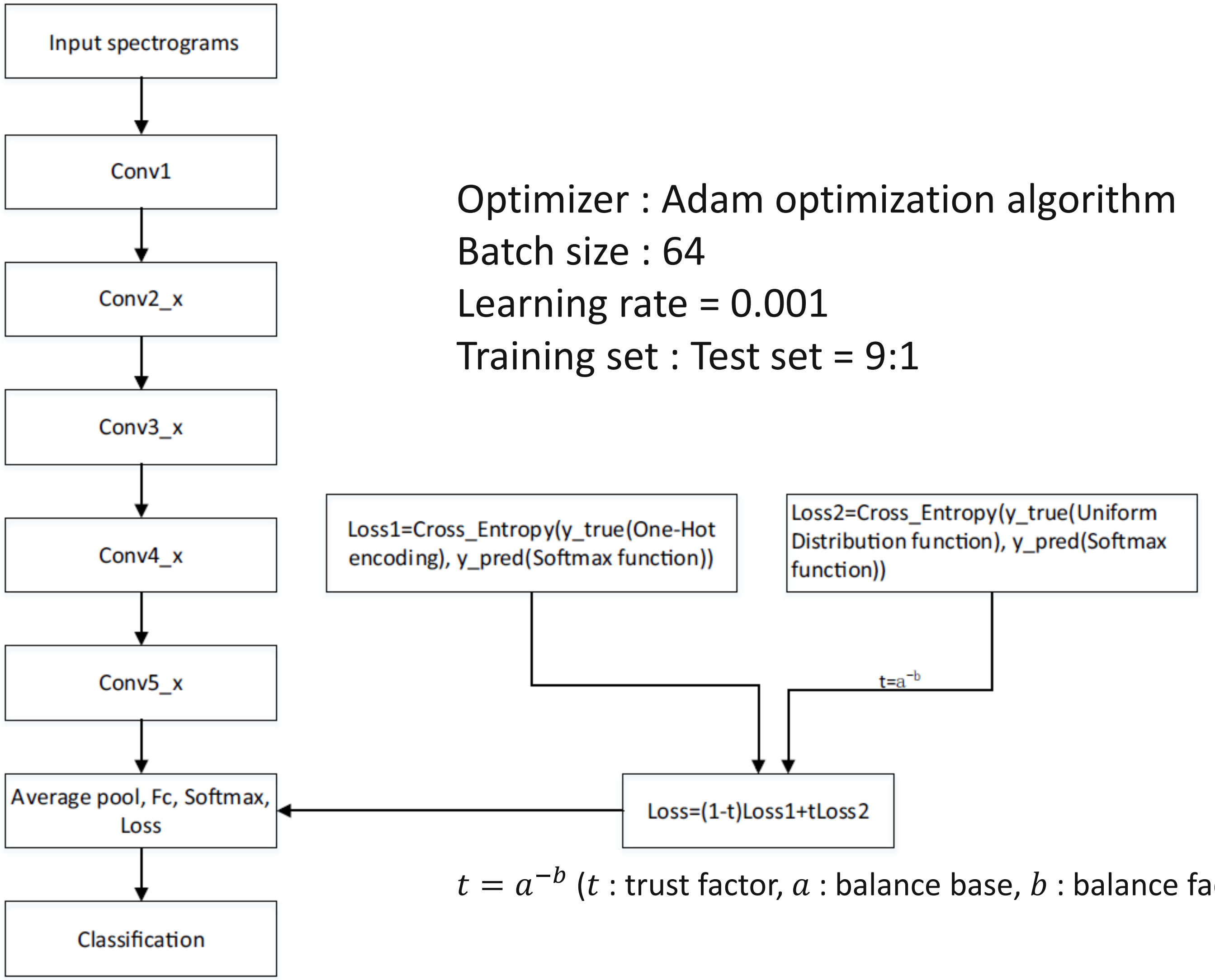
[15] Li, J., Han, L., Li, X., Zhu, J., Yuan, B., & Gou, Z. (2021). An evaluation of deep neural network models for music classification using spectrograms. *Multimedia Tools and Applications*, 1-27.

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

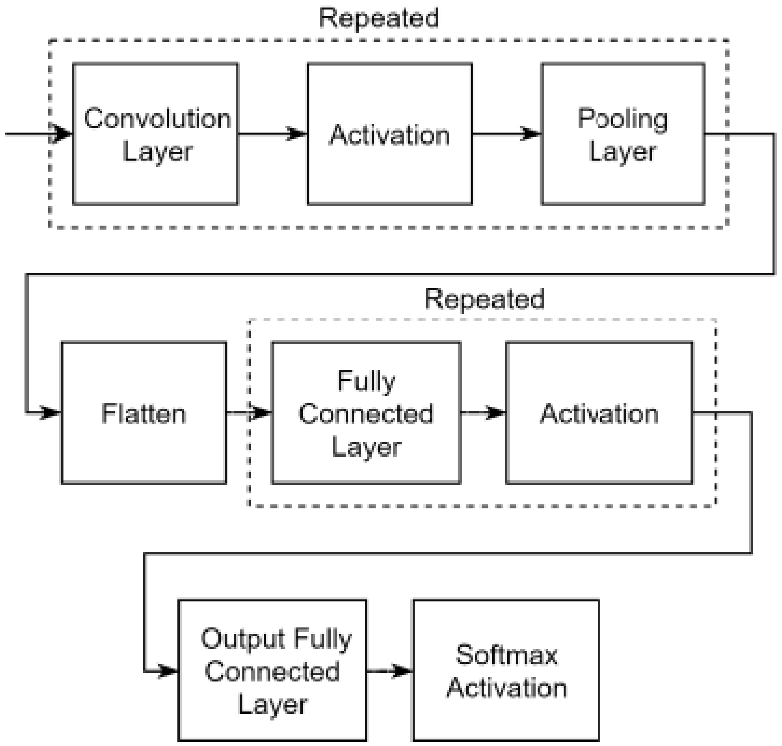
$t = a^{-b}$ (t : trust factor, a : balance base, b : balance factor)

Code Analysis(ResNet50_trust)

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

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

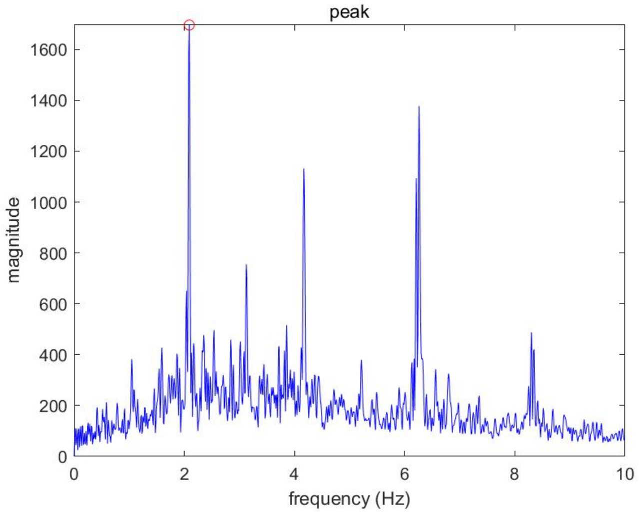
Convolutional Neural Network Summarization

Appendix A

MIRToolbox

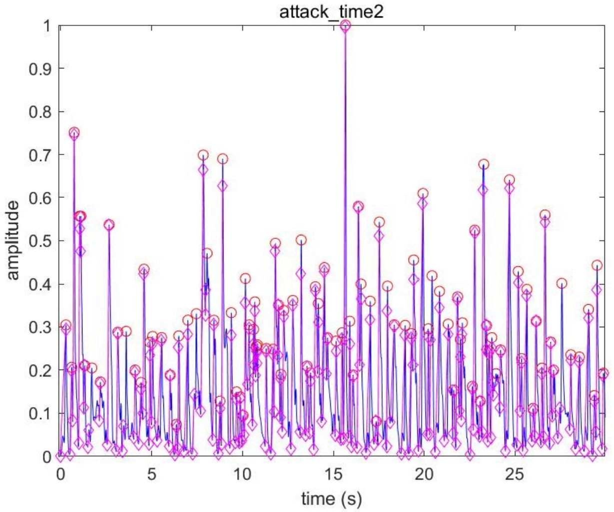
X_axis = frequency

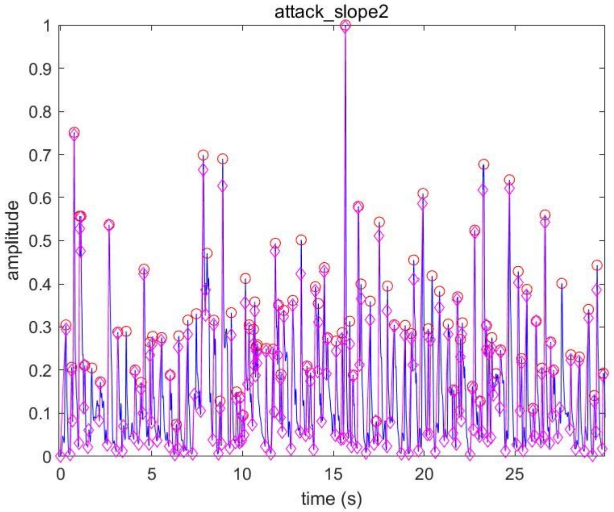
Fluctuation



X_axis = time

Rhythm

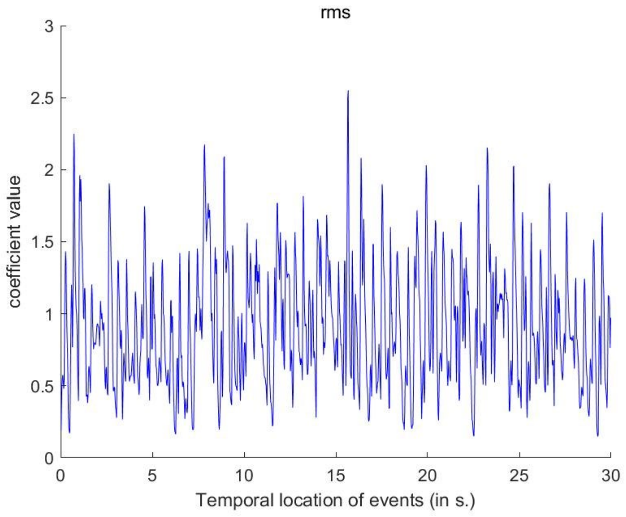




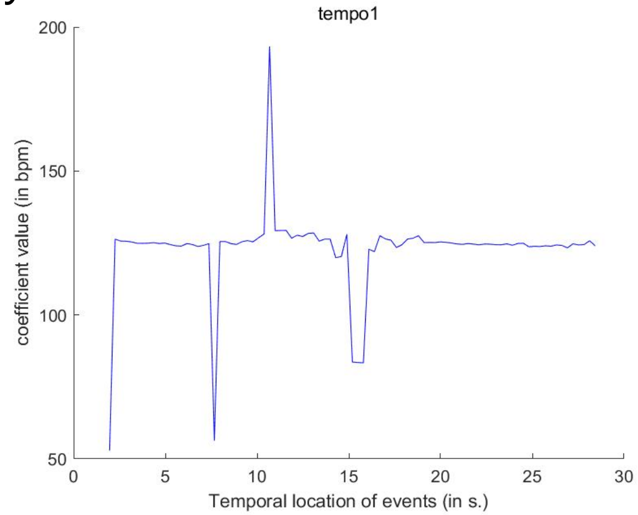
MIRToolbox

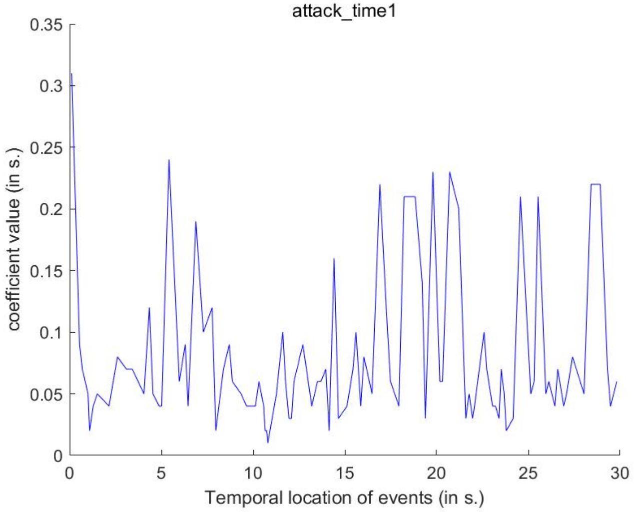
X_axis = Temporal location of events (in s.)

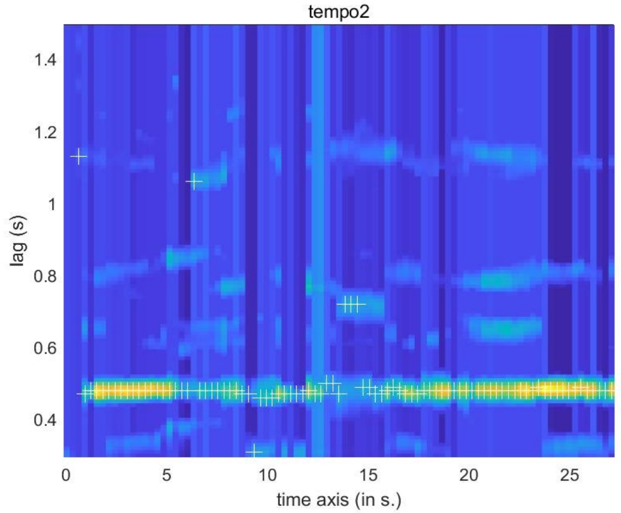
Dynamics

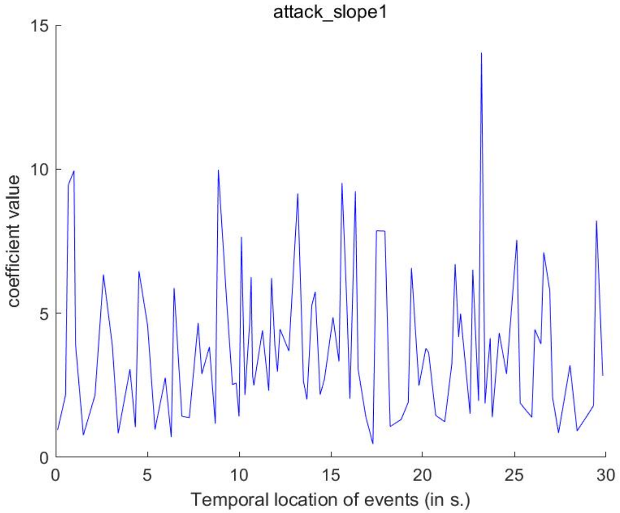


Rhythm





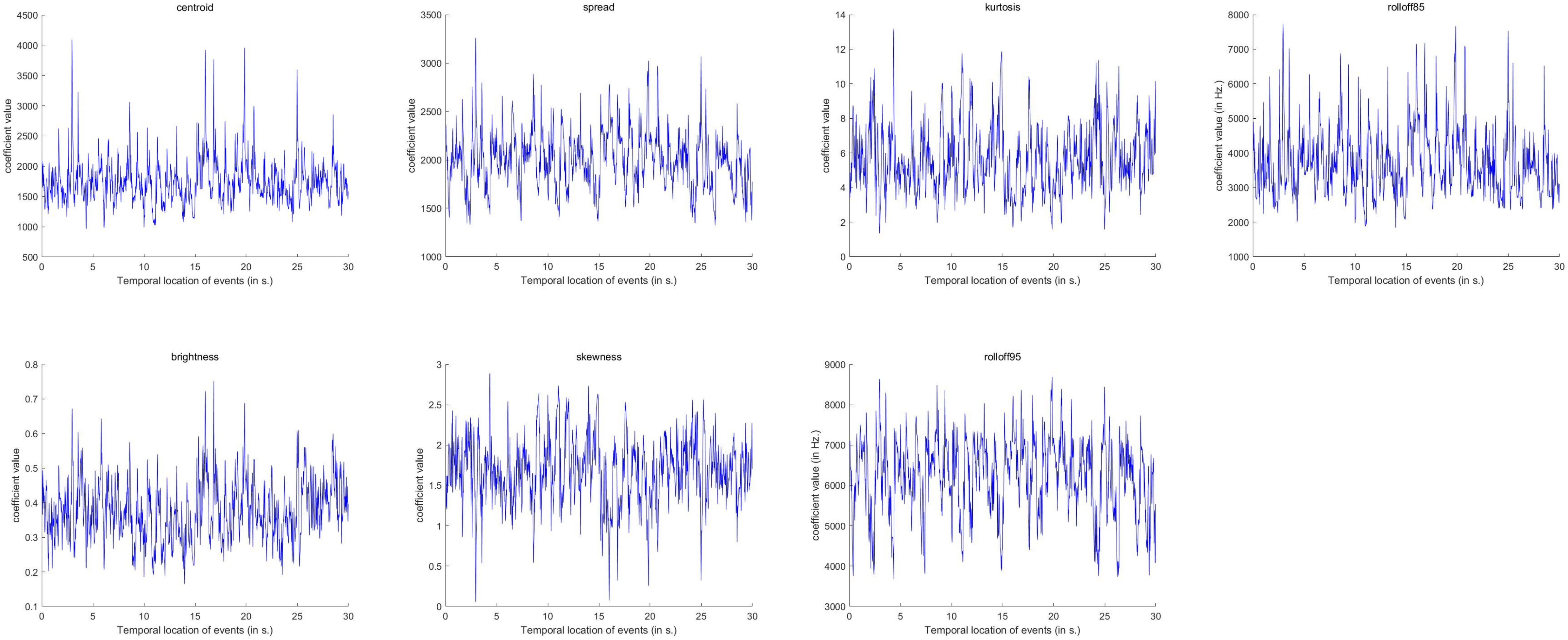




MIRToolbox

X_axis = Temporal location of events (in s.)

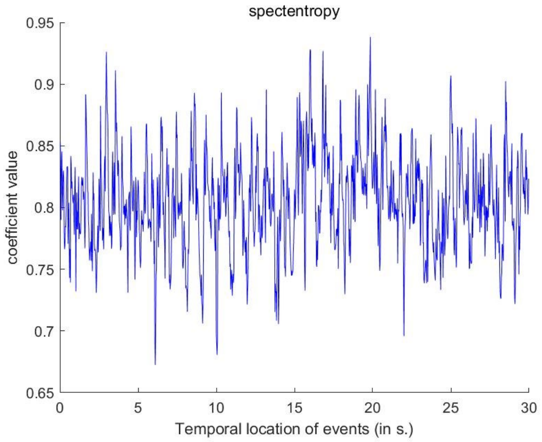
Spectral

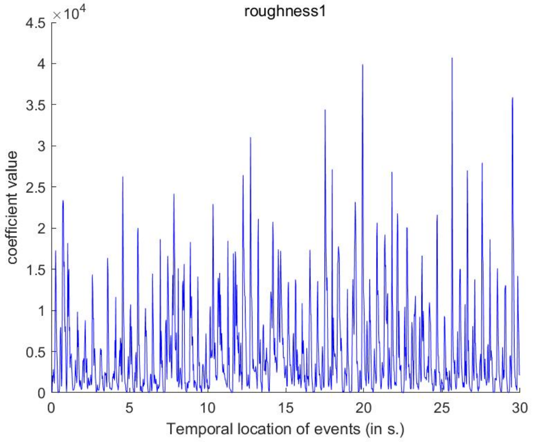


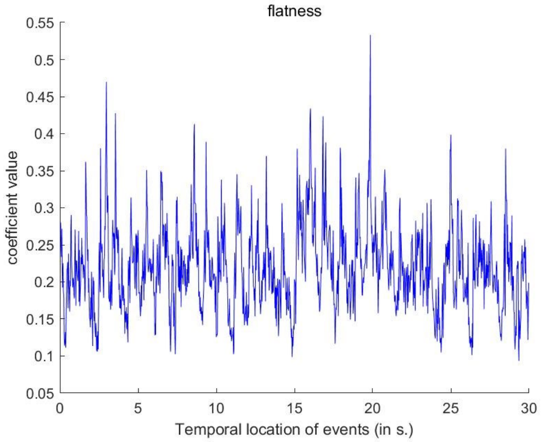
MIRToolbox

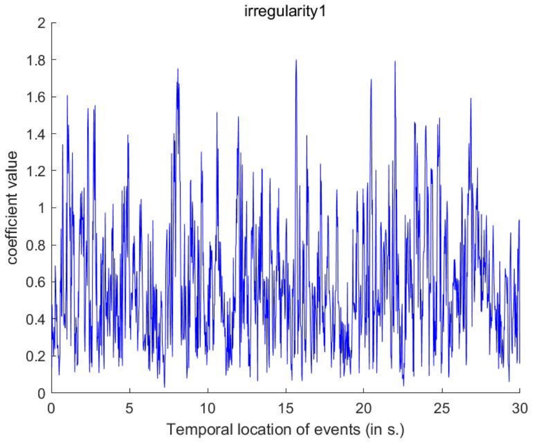
X_axis = Temporal location of events (in s.)

Spectral

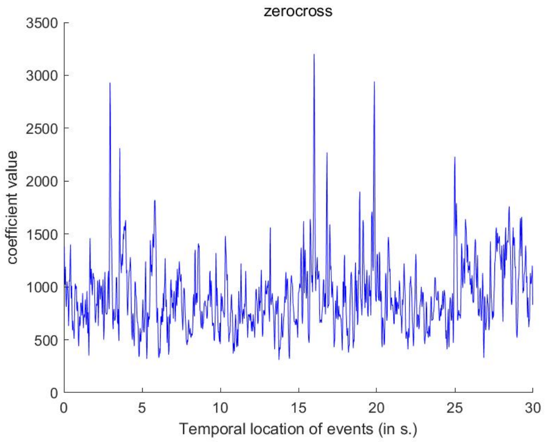


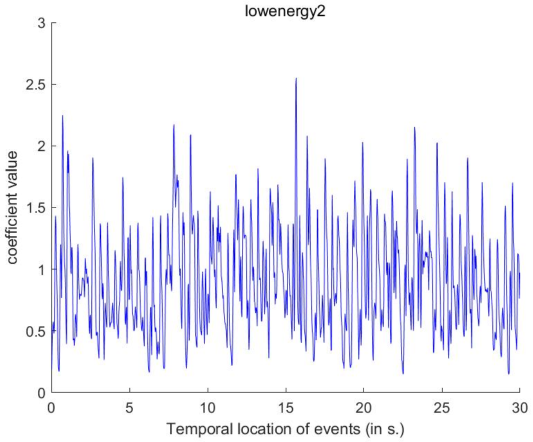


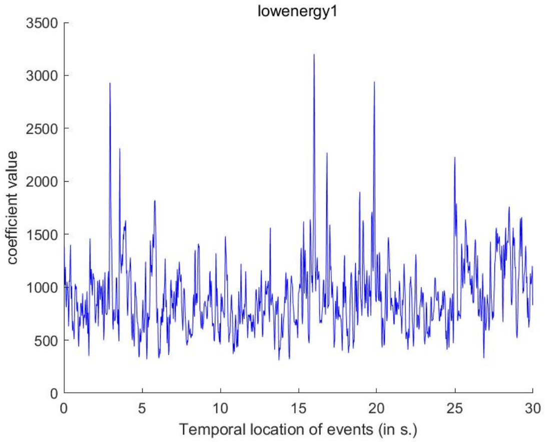


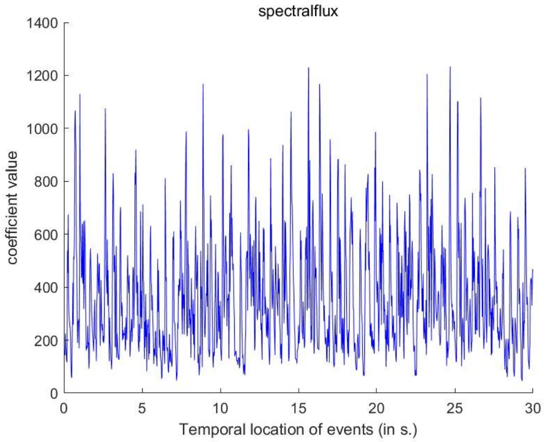


Timbre





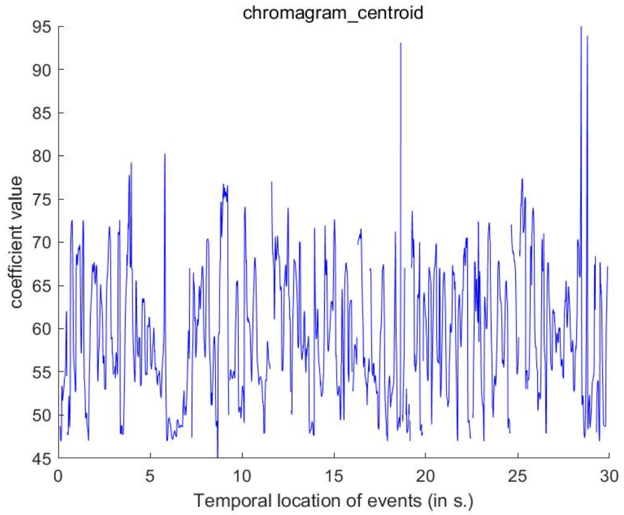


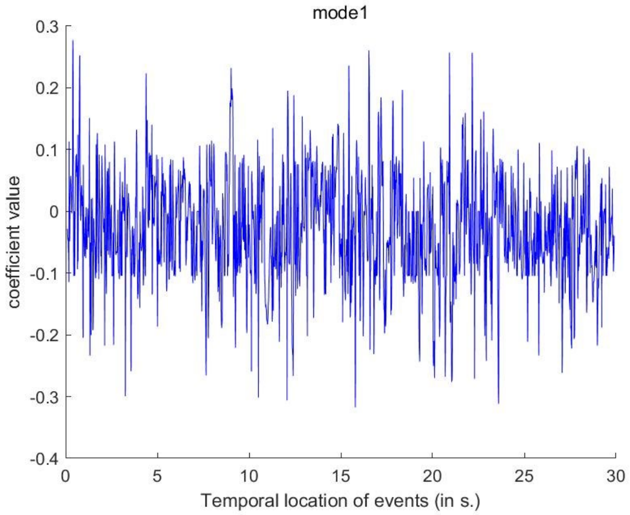


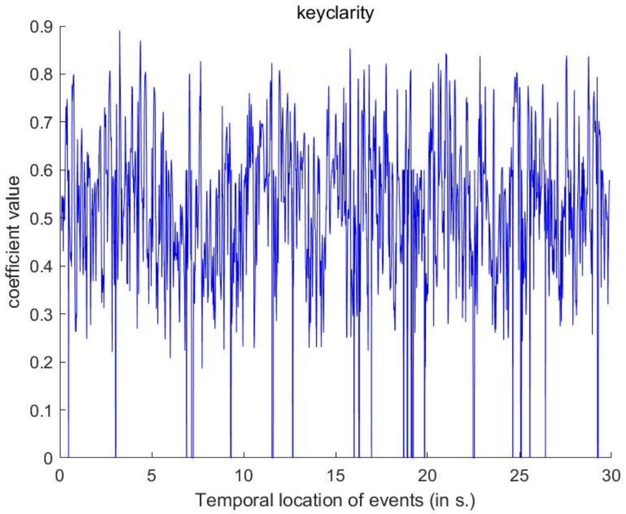
MIRToolbox

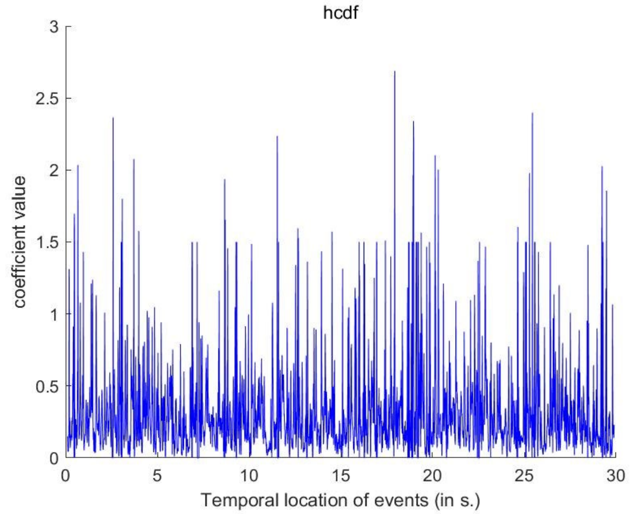
X_axis = Temporal location of events (in s.)

Tonal





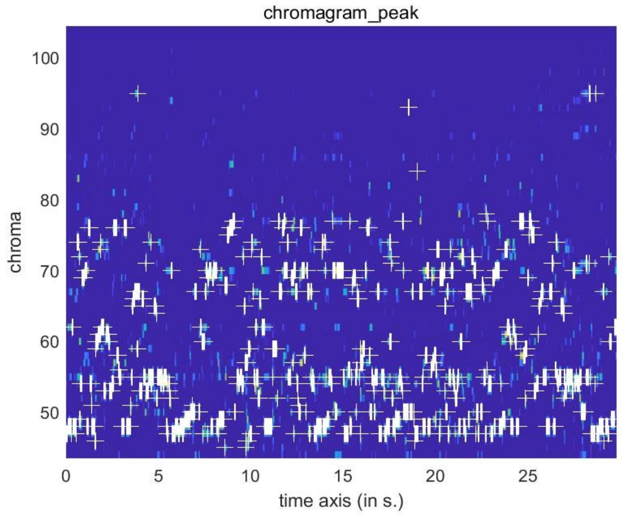


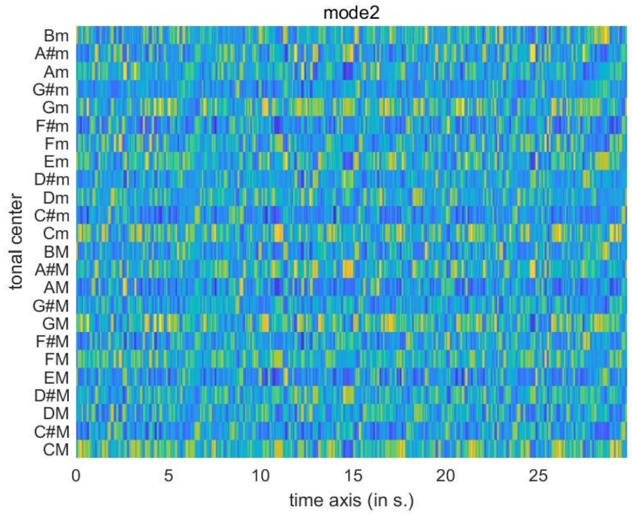


MIRToolbox

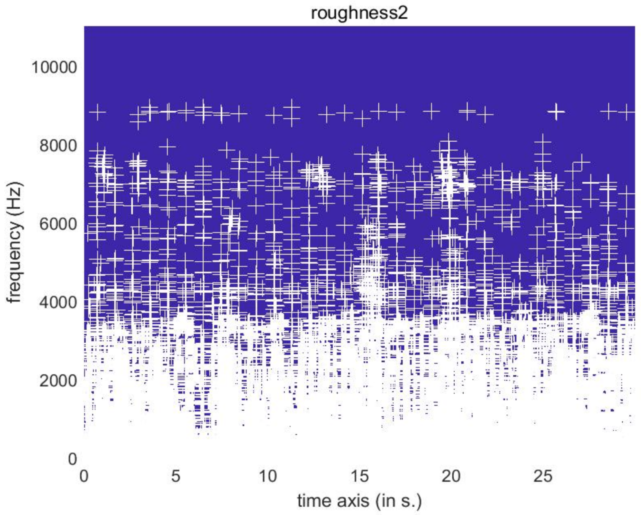
X_axis = time axis (in s.)

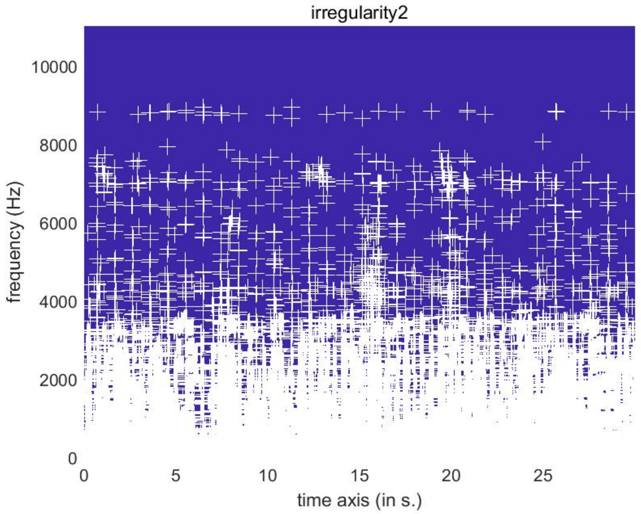
Tonal





Spectral





MIRToolbox

X_axis = time axis (in s.)

Spectral