

제 13회 담낭 용종 연구 보고

2021.05.11



실험 세팅

➤ 데이터 구성

- 177명의 환자 데이터 (서울대 전체 데이터)
- Neoplastic : 100명, Non-neoplastic : 77명

➤ Train & Test Set

- 2,888장의 데이터셋
- Train Set : 141명(2256장)
- Test Set : 36명(632장)

➤ Input Image

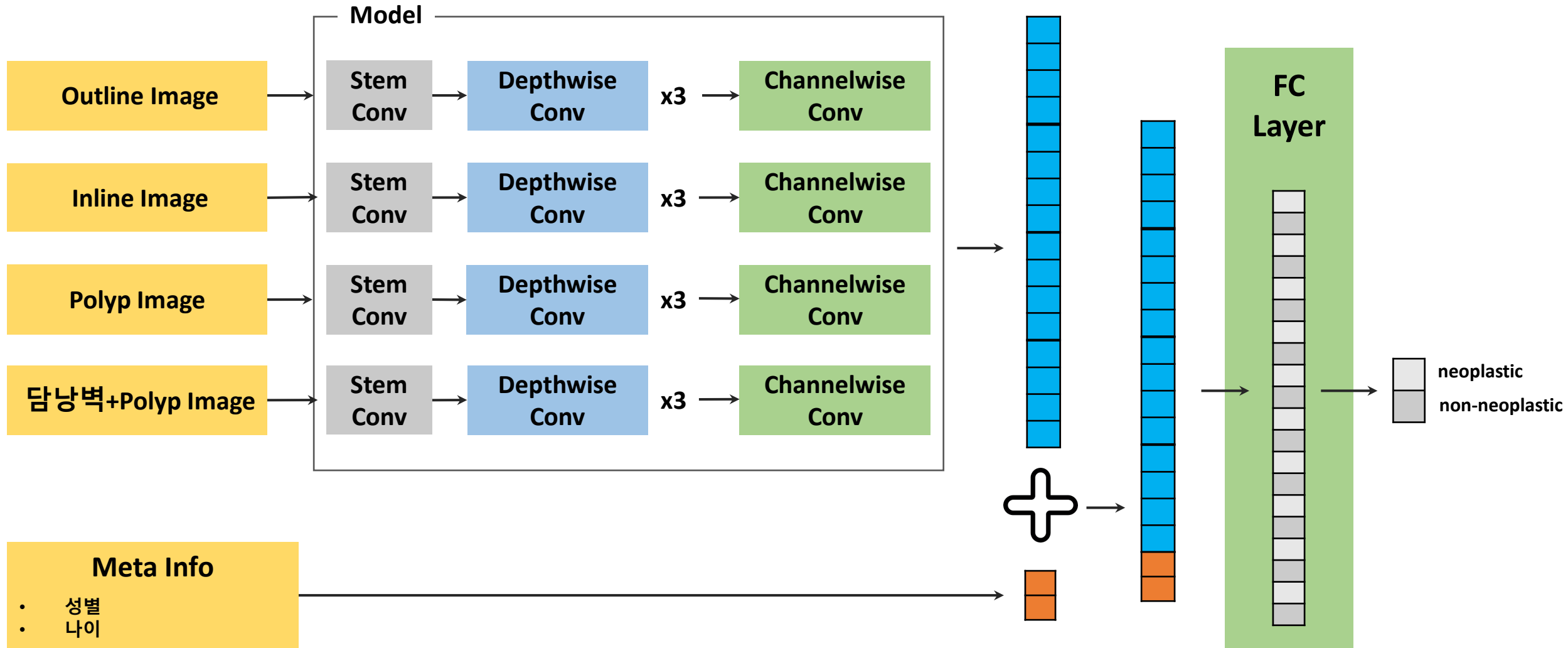
- Outline, Inline, Polyp, 담낭벽+Polyp 이미지

➤ 모델 구성

- 갑상선에서 사용한 모델을 담낭 용종 연구에 맞게 조정한 모델
- 데이터 불균형 문제를 해소하기 위한 Focal loss function 적용

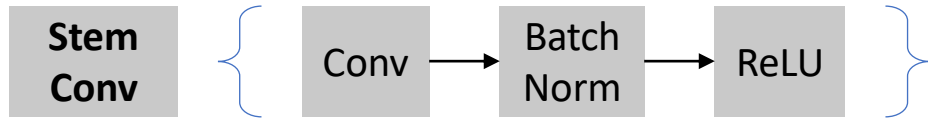
➤ 30번 반복 실험

모형 파이프 라인



모형 파이프 라인 설명

Stem Block



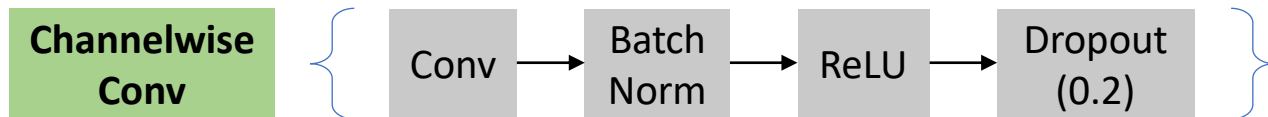
- Conv [kernel size = 3x3, stride = 2, padding = 1]

Depthwise Block



- 3x3 Conv [kernel size = 3, stride = 1, padding = 1]
- 1x1 Conv [kernel size = 1, stride = 1, padding = 0]

Channelwise Block



- Conv [kernel size & stride = Input image size, padding = 0]

Experimental Method

1. Baseline Model

- 기존 모델, 실험 세팅 및 데이터를 이용한 방법

2. Noise Removal in Image

- Baseline model에서 노이즈 이미지를 제거한 데이터를 이용한 방법
- 전체 입력 데이터의 약 25%가 노이즈 이미지가 포함되어 있음

* 전체 입력 데이터(2,888장의 데이터셋 x 4종류의 입력 이미지) 11,552장 중 2,936장이 노이즈 이미지

3. Band Normalization

- Baseline model에서 유사하게 인식하고 있는 부분을 구별하기 위해 각 영역을 구분

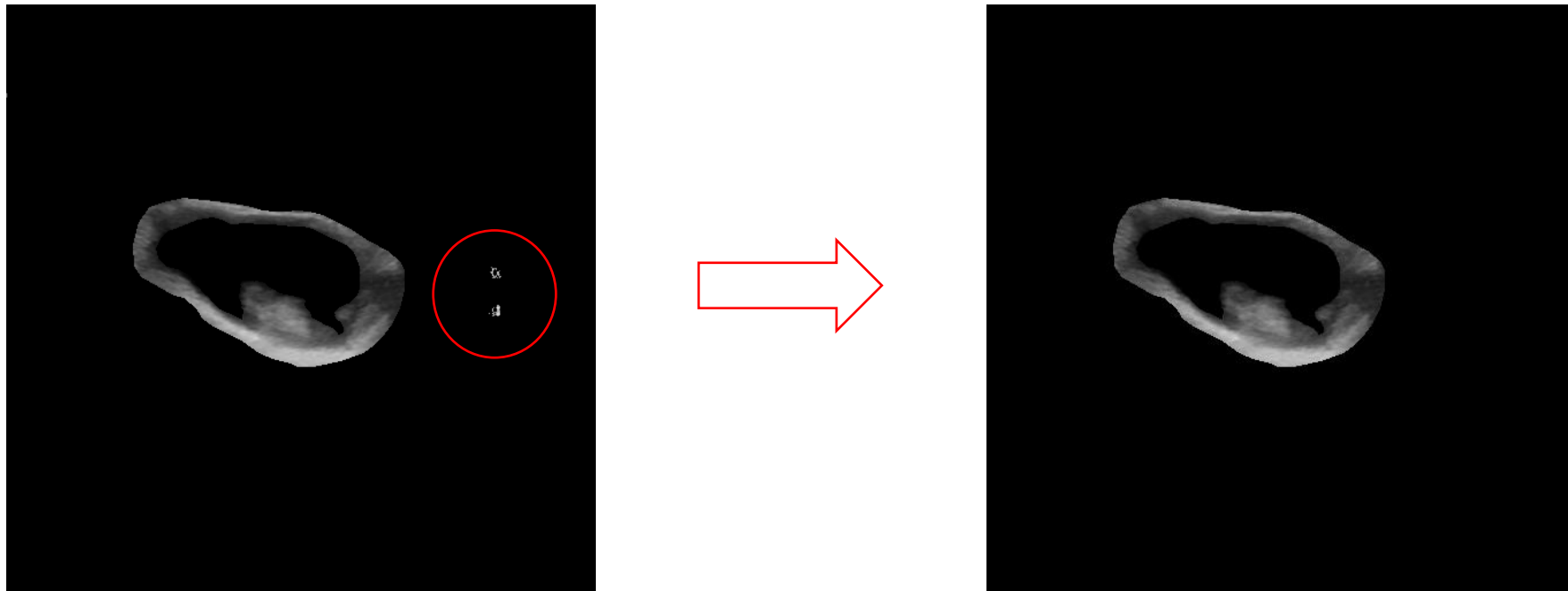
4. Fractional Max-Pooling

- 학습 과정 중 Max-Pooling 영역이 매번 임의의 크기로 조합
- 다양한 크기의 영역들의 예측 결과 값의 평균을 이용하여 정확도 향상

Noise Removal in Image

➤ 입력 이미지에 대한 noise 제거

- 기존 데이터 전처리 과정에서 제거되지 않은 노이즈 이미지를 제거
- 노이즈 이미지가 제거된 데이터를 이용한 Baseline model 성능 확인

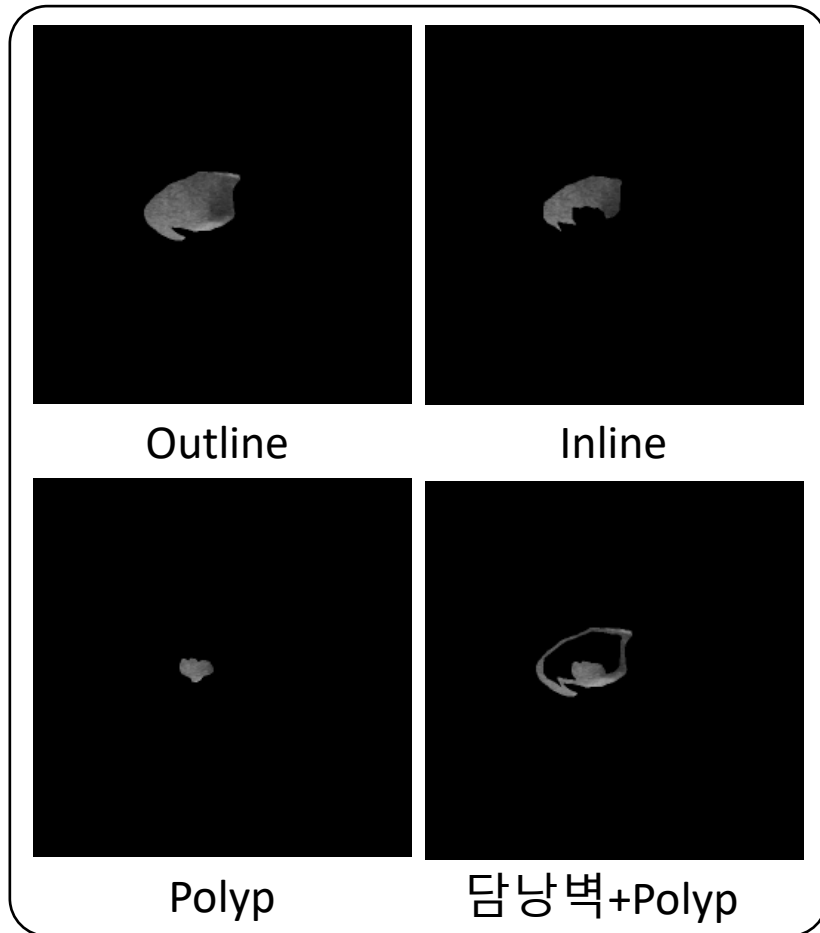


Band Normalization

➤ Band Normalization

- 각 입력 이미지에 대해 0에서 1사이 값으로 Normalization 범위 지정

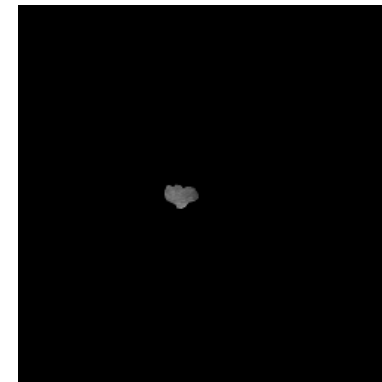
Input Image



Inline



(0.0 ~ 0.3)



Polyp



(0.3 ~ 0.6)



담낭벽

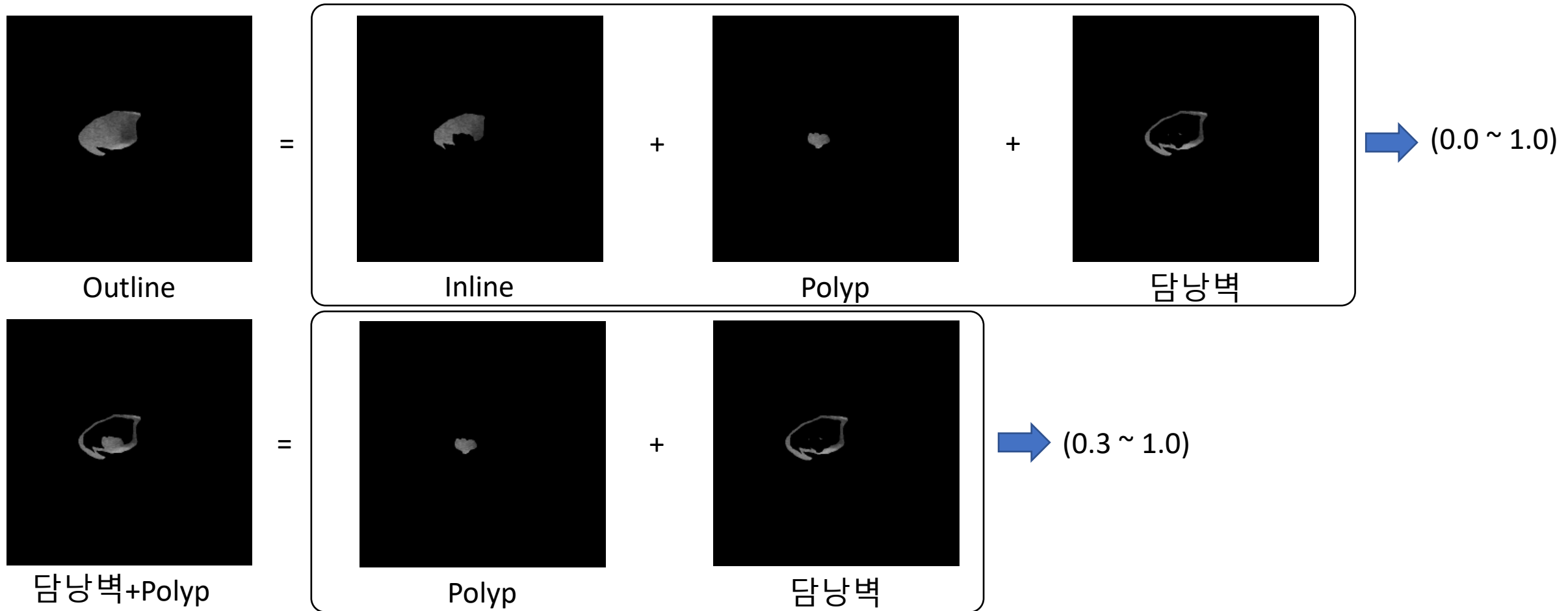


(0.6 ~ 1.0)

Band Normalization

➤ Band Normalization

- Outline 이미지와 담낭벽+Polyp의 경우 Inline, polyp, 담낭벽으로 구성되어 있음

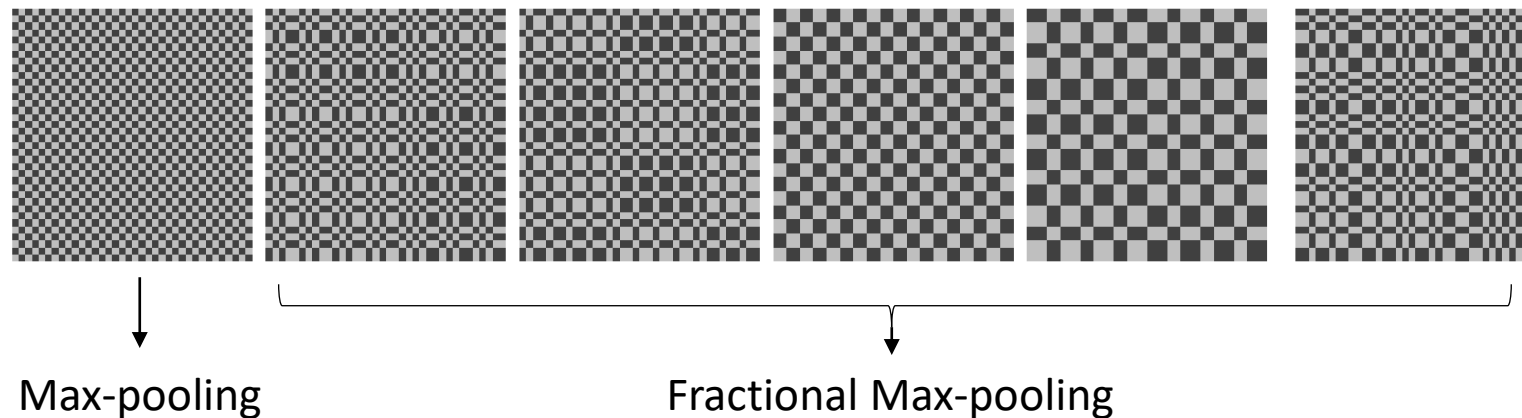
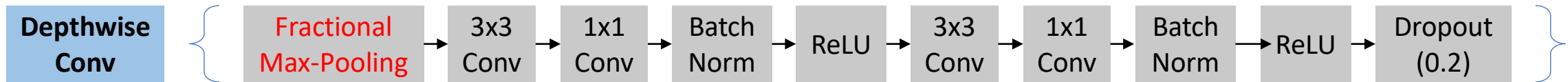


Fractional Max-Pooling

➤ Fractional Max-Pooling

- Baseline model의 depthwise block에서 Max-Pooling 부분을 Fractional Max-Pooling으로 대체
- Max-Pooling 영역이 다양한 크기로 조합
- 조합된 영역에 대한 예측 결과의 평균 값을 이용

Depthwise Block



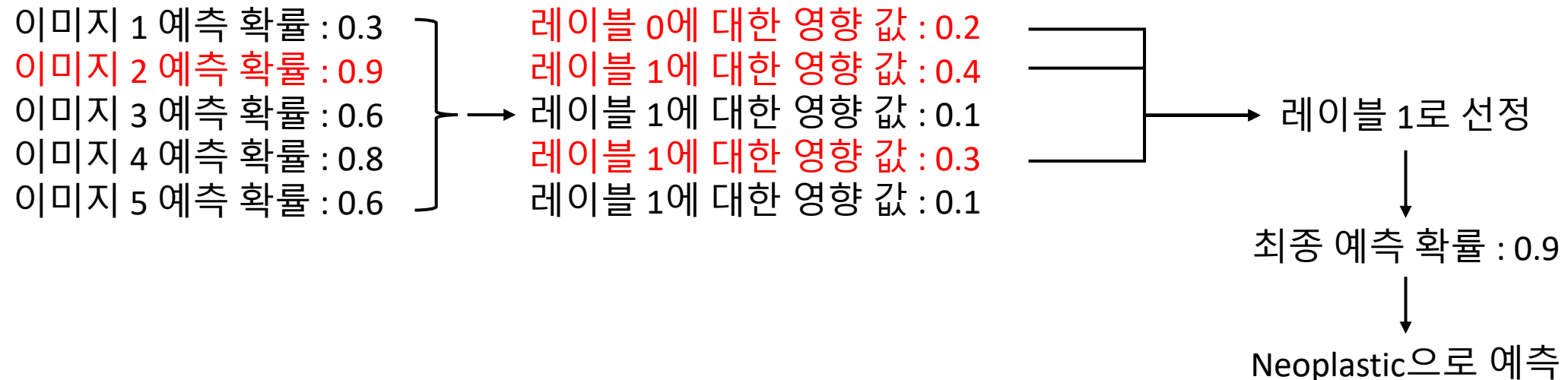
Top-k Voting Method

➤ Top-k Voting Method

1. 모든 이미지 중 레이블 결정에 대한 영향 값 계산 $f = \begin{cases} |\hat{y} - 0.5| & \text{if } \hat{y} \geq 0.5 \\ |0.5 - \hat{y}| & \text{otherwise.} \end{cases}$
2. 큰 영향을 미치는 이미지 k개를 선택
3. k개의 이미지에서 다수의 레이블에 대한 확률 값 중에서 가장 큰 값을 최종 결과값으로 선정

Ex) Top-3 - 환자 1명에 대해 5장의 이미지로 예측하는 경우

- 0 ~ 1의 확률 값을 가지고 0.5를 기준으로 0에 가까우면 Non-neoplastic, 1에 가까우면 Neoplastic
- 선정 레이블의 확률 값 중 가장 큰 값을 최종 결과값으로 선정



➤ Soft Voting Method

➤ Soft Voting Method Process

- 모든 이미지에 대해 예측한 레이블 값의 예측 확률 평균을 구한 뒤 가장 확률이 높은 레이블 값을 최종 결과값으로 선정

Ex) 환자 1명에 대해 5장의 이미지를 예측하는 상황

- 0 ~ 1의 확률 값을 가지고 0.5 이상이면 1(Neoplastic)로 예측

이미지 1 예측 확률 : 0.3	}	→	평균 예측 확률 : 0.62	→	Neoplastic으로 판별
이미지 2 예측 확률 : 0.9					
이미지 3 예측 확률 : 0.6					
이미지 4 예측 확률 : 0.7					
이미지 5 예측 확률 : 0.6					

Top-3 Vote Method Classification Performance Table

Table . Metrics with 95% confidence interval of four NNs (▼/△ indicates that the corresponding method is significantly worse/better than Top-3 vote method based on paired t-test at 5% significance level).

Model	Metrics	AUC Mean	95% CI		p-value
			Lower	Upper	
Baseline	AUC	0.925	0.909	0.940	
	Accuracy	0.844	0.821	0.867	
	Sensitivity	0.817	0.781	0.853	
	Specificity	0.877	0.830	0.924	
Noise Removal	AUC	0.911	0.891	0.930	0.385
	Accuracy	0.839	0.816	0.862	0.803
	Sensitivity	0.858	0.821	0.896	0.240
	Specificity	0.815	0.767	0.862	0.181
Band Normalization	AUC	0.899	0.883	0.916	0.124
	Accuracy	0.824	0.801	0.847	0.346
	Sensitivity	0.835	0.810	0.860	0.516
	Specificity	0.810	0.760	0.860	0.166
Fractional Max-pooling	AUC	0.910	0.896	0.925	0.307
	Accuracy	0.819	0.799	0.838	0.219
	Sensitivity	0.827	0.792	0.861	0.753
	Specificity	0.808	0.758	0.859	0.159

AUC, area under the curve; CI, confidence interval; NN, neural network. *: p-value < 0.01, **: p-value < 0.001

Top-10 Vote Method Classification Performance Table

Table . Metrics with 95% confidence interval of four NNs (▼/△ indicates that the corresponding method is significantly worse/better than Top-10 vote method based on paired t-test at 5% significance level).

Model	Metrics	AUC Mean	95% CI		p-value
			Lower	Upper	
Baseline	AUC	0.923	0.908	0.938	
	Accuracy	0.833	0.812	0.854	
	Sensitivity	0.770	0.723	0.817	
	Specificity	0.913	0.876	0.950	
Noise Removal	AUC	0.916	0.899	0.933	0.621
	Accuracy	0.833	0.811	0.855	1.000
	Sensitivity	0.880	0.845	0.915	0.031
	Specificity	0.775	0.723	0.827	0.020
Band Normalization	AUC	0.906	0.890	0.923	0.265
	Accuracy	0.821	0.795	0.847	0.567
	Sensitivity	0.848	0.820	0.877	0.069
	Specificity	0.788	0.733	0.842	0.030
Fractional Max-pooling	AUC	0.917	0.905	0.929	0.617
	Accuracy	0.820	0.799	0.842	0.505
	Sensitivity	0.847	0.813	0.881	0.083
	Specificity	0.788	0.731	0.844	0.032

AUC, area under the curve; CI, confidence interval; NN, neural network. *: p-value < 0.01, **: p-value < 0.001

Soft Vote Method Classification Performance Table

Table . Metrics with 95% confidence interval of four NNs (▼/△ indicates that the corresponding method is significantly worse/better than Soft vote method based on paired t-test at 5% significance level).

Model	Metrics	AUC Mean	95% CI		p-value
			Lower	Upper	
Baseline	AUC	0.919	0.904	0.934	
	Accuracy	0.825	0.802	0.848	
	Sensitivity	0.733	0.683	0.783	
	Specificity	0.940	0.912	0.968	
Noise Removal	AUC	0.908	0.891	0.925	0.447
	Accuracy	0.849	0.829	0.869	1.000
	Sensitivity	0.803	0.765	0.842	0.125
	Specificity	0.906	0.867	0.946	0.294
Band Normalization	AUC	0.897	0.882	0.911	0.137
	Accuracy	0.833	0.812	0.854	0.567
	Sensitivity	0.752	0.719	0.784	0.616
	Specificity	0.935	0.891	0.980	0.885
Fractional Max-pooling	AUC	0.904	0.886	0.921	0.313
	Accuracy	0.804	0.778	0.830	0.207
	Sensitivity	0.682	0.634	0.729	0.266
	Specificity	0.956	0.934	0.978	0.479

AUC, area under the curve; CI, confidence interval; NN, neural network. *: p-value < 0.01, **: p-value < 0.001