

제 07회 담낭 용종 연구 보고

2021.01.07



실험 세팅

➤ 데이터 구성

- 178명의 환자 데이터 (서울대 전체 데이터)
- Neoplastic : 100명, Non-neoplastic : 78명

➤ Train & Test Set

- 2,916장의 데이터셋
- Train Set : 142명(2243장)
- Test Set : 36명(673장)

➤ Input Image

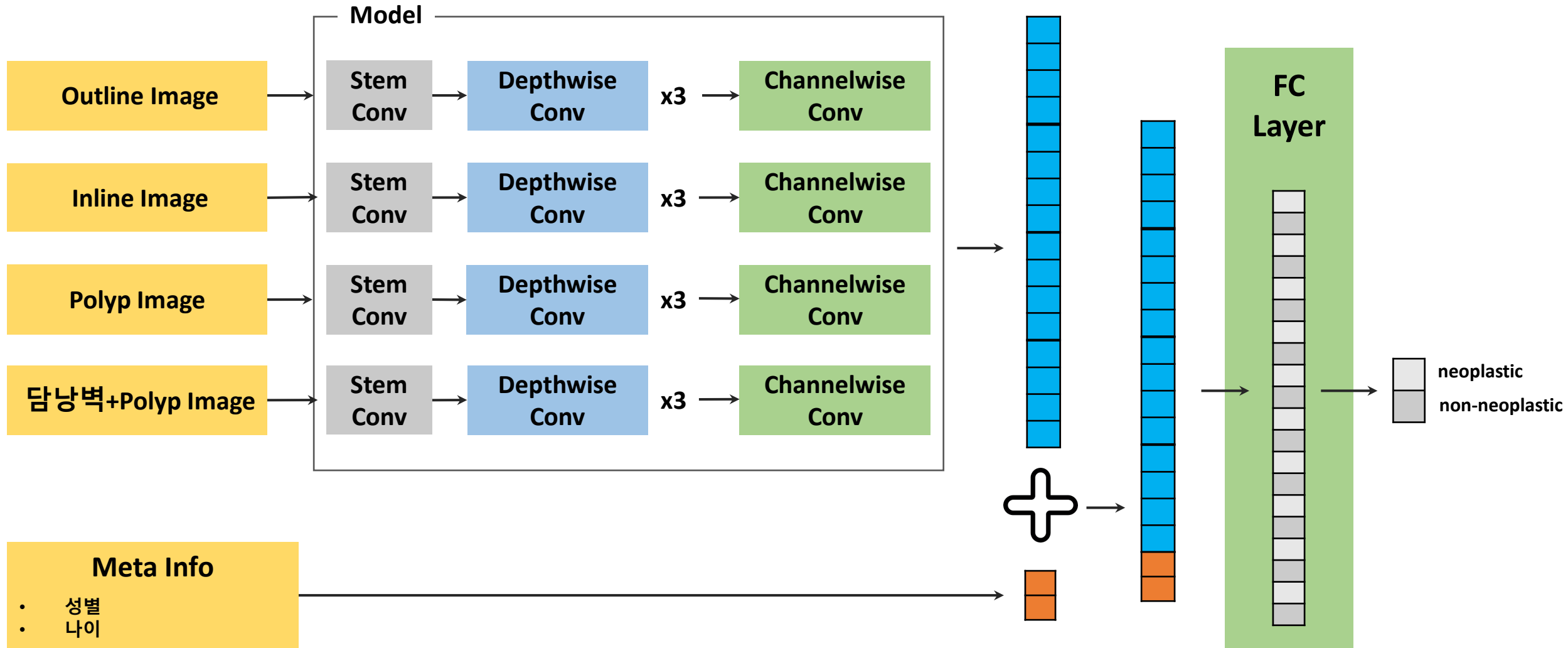
- Outline, Inline, Polyp, 담낭벽+Polyp 이미지

➤ 모델 구성

- 갑상선에서 사용한 모델을 담낭 용종 연구에 맞게 조정한 모델

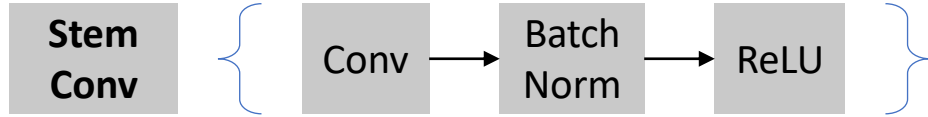
➤ 30번 반복 실험

모형 파이프 라인



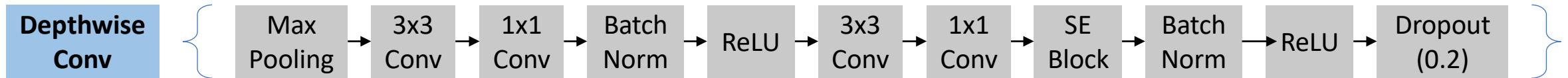
모형 파이프 라인 설명

Stem Block



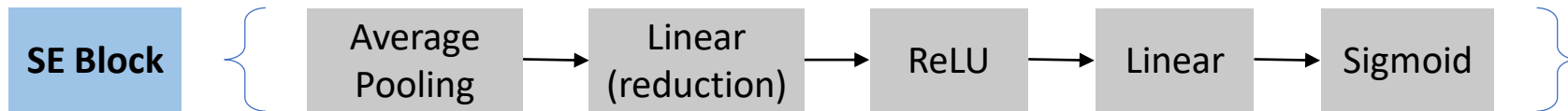
- Conv [kernel size = 3x3, stride = 2, padding = 1]

Depthwise Block



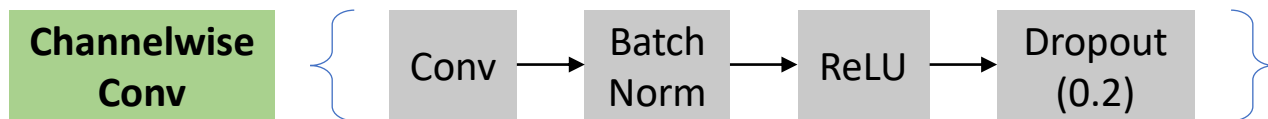
- 3x3 Conv [kernel size = 3, stride = 1, padding = 1]
- 1x1 Conv [kernel size = 1, stride = 1, padding = 0]

SE Block



- Linear [reduction = 2] $\longrightarrow$ out channel = (channel / reduction)

Channelwise Block



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

Classification Performance Table

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

Model	Metrics	AUC Mean	95% CI		p-value
			Lower	Upper	
Module Network	AUC	0.870	0.853	0.888	
	Accuracy	0.813	0.795	0.831	
	Sensitivity	0.917	0.904	0.930	
	Specificity	0.565	0.525	0.605	

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

- Test set에서 환자의 모든 이미지를 사용한 결과

Voting Method 설명

➤ Hard Voting Method

- 모든 이미지에 대해 예측한 결과값을 최종 결과값으로 선정

➤ Soft Voting Method

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

➤ Positive Instance Voting Method

- 모든 이미지에서 최소 한 장의 양성 판별 이미지가 있다면 양성으로 선정하고 그에 대한 예측값을 최종 결과값으로 선정

➤ Soft Voting Method Process

➤ 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 확률값을 가지고 0.5 이상이면 1(Neoplastic)으로 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			

Vote Method Classification Performance Table

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

Vote Method	Metrics	AUC Mean	95% CI		p-value	
			Lower	Upper		
Hard	AUC	0.870	0.853	0.888		
	Accuracy	0.813	0.795	0.831		
	Sensitivity	0.917	0.904	0.930		
	Specificity	0.565	0.525	0.605		
Soft	AUC	0.897	0.873	0.920	0.561	
	Accuracy	0.809	0.780	0.837	0.183	
	Sensitivity	0.912	0.884	0.939	0.622	
	Specificity	0.680	0.623	0.737	0.072	
Positive	AUC	0.883	0.859	0.908	0.873	
	Accuracy	0.673	0.650	0.696	8.01e-13	**▼
	Sensitivity	0.977	0.964	0.989	1.33e-05	**△
	Specificity	0.293	0.245	0.341	5.46e-18	**▼

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