

# 환자 정보를 이용한 자동 Colposcopy 정확도 향상 방법

정리본

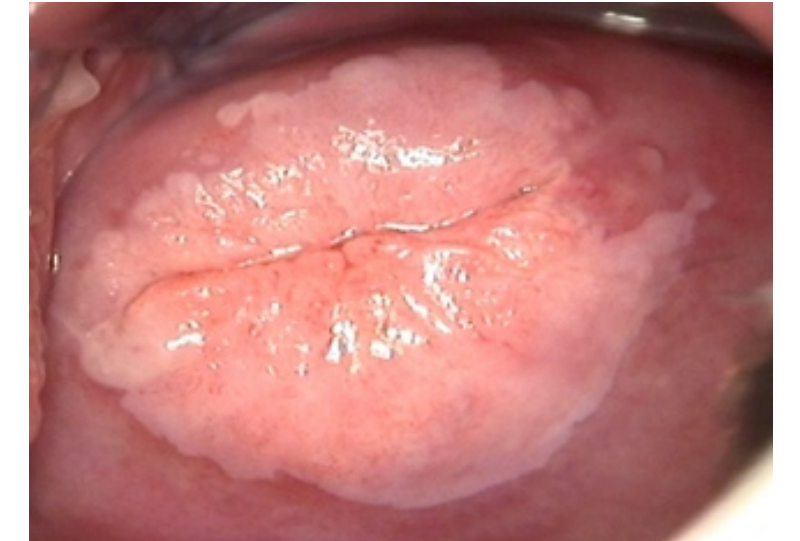
# Colposcopy

- 자궁 경부를 확대하여 이상 부위를 관찰하는 검사
- 자궁 경부암(cervical cancer) 진단
- 생리식염수(saline)<sup>1</sup>, 아세트산(acetic)<sup>2</sup>, 필터(green filter)<sup>3</sup>, 아이오딘(iodine)<sup>4</sup>
- 자궁 경부암 진단에 **추가적인 정보** 사용
  - HPV Status, Transformation Type 등..

1. saline



2 acetic



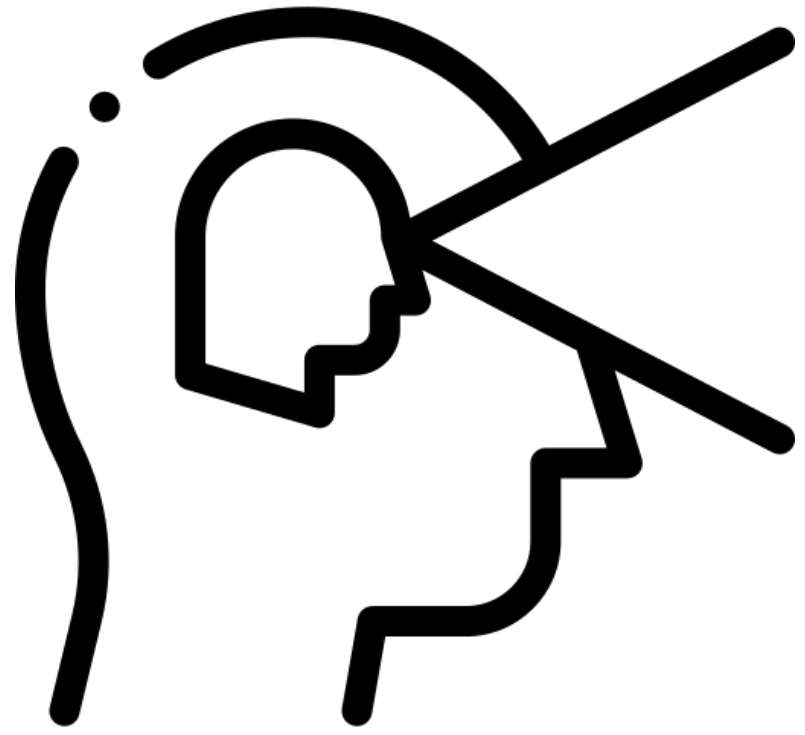
3. acetic green



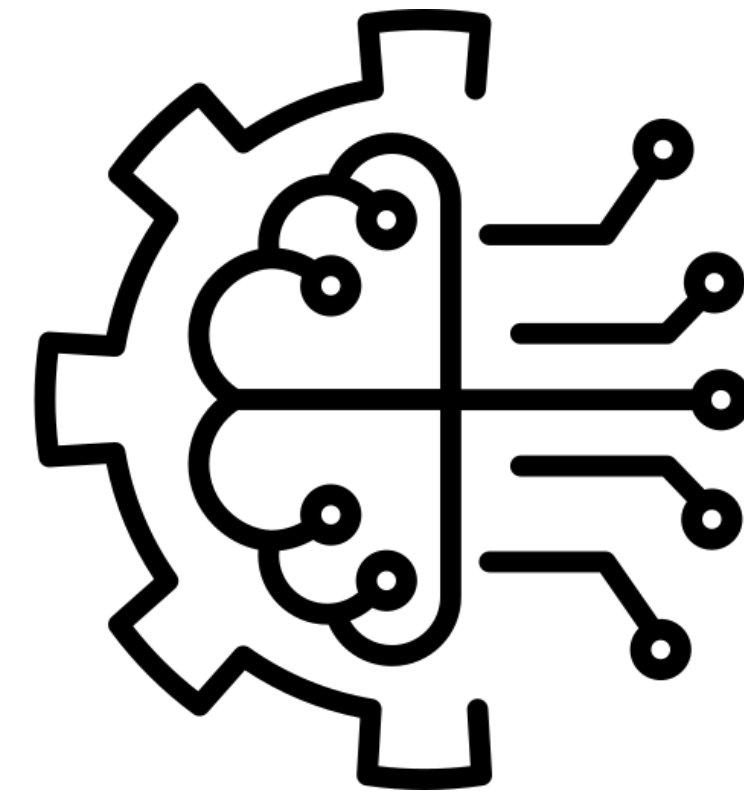
4. iodine



# Colposcopy Problem



**Colposcopy**



**Computer Aided Diagnosis**

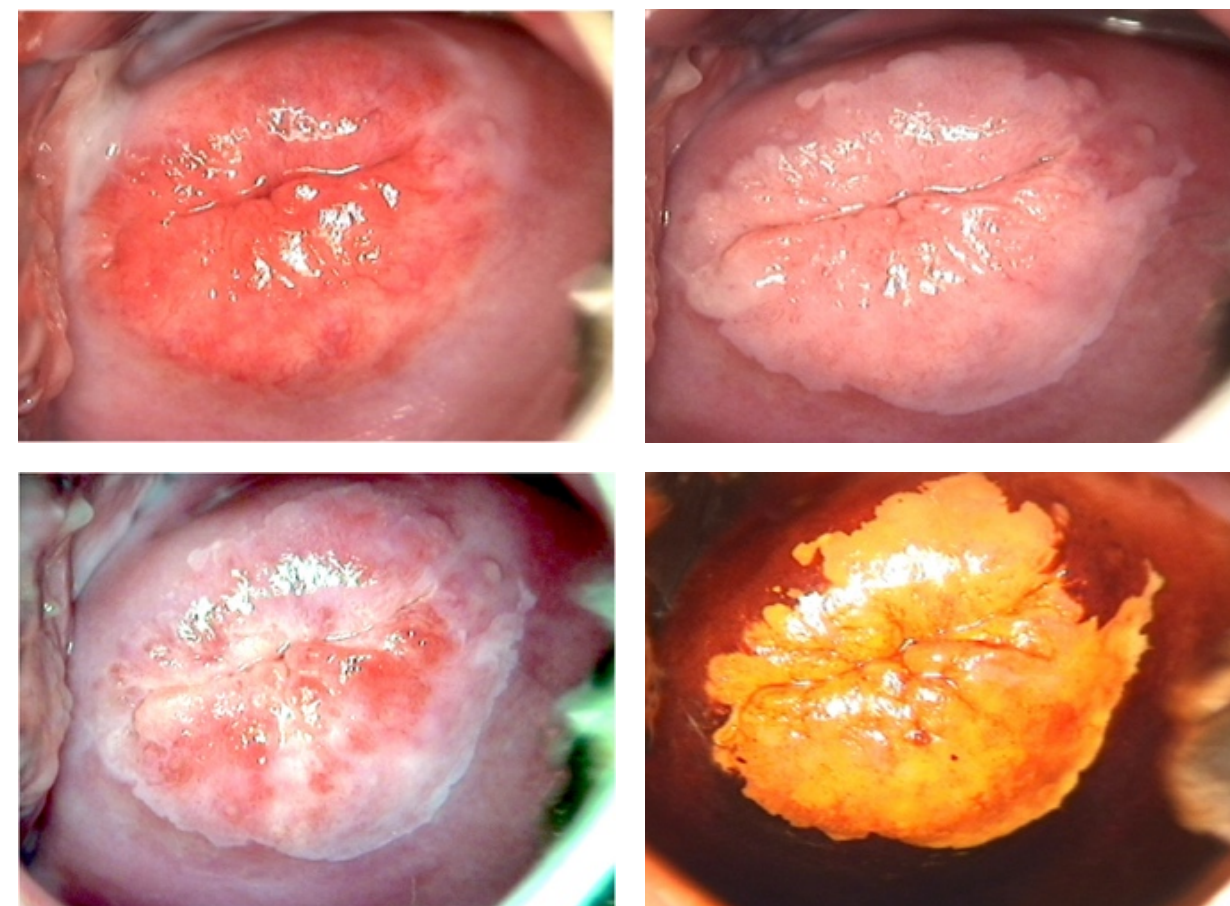
# Prior Work

Paper	Data	Model
Application of deep learning to the classification of images from colposcopy <sup>[1]</sup>	<b>Images Only</b> - acetic, green filter	<b>CNN</b>
Classification of cervical neoplasms on <u>colposcopic</u> photography using deep learning <sup>[2]</sup>	<b>Images Only</b> - saline, acetic	<b>Inception-resnet-v2</b>

- 다양한 용액이나 추가적인 정보를 충분히 반영 X



# Topic



+

**auxiliary  
information**

[HPV Status, TZ Type, Age]

# Dataset

- <https://screening.iarc.fr/atlascolpo.php>
- image, Category, Case, Age, PHV Status, TZ type, Solution, Diagnosis
- 총 198명 환자 데이터 - 약 870개 이미지 데이터

# Dataset

## pre-processing

- **Age** (min:25, max:60) : 3가지 Group  
25~34 / 35~44 / 45~60
- **TZ Type** (Transformation Zone) : 자궁경부 내외부의 접합부 근처에서 세포의 형태 변화가 나타나는 곳  
Type 1, Type 2, Type 3
- **HPV Status** : HPV 보균자 여부  
0 : 음성  
1 : 양성

# Dataset

## pre-processing

- **Solution** : 용액&필터 사용 정보  
acetic  
acetic green = acetic & green filter  
saline  
saline green = saline & green filter  
iodine
- **Label (0/1)** : 진단 결과  
0 = normal  
1 = abnormal



- 모델은 크게 **Image encoder, Auxiliary encoder**, 그리고 **Prediction Method**로 나뉘어 있다.

## Model



# Model architecture

## 이미지 단위 진단

### Image encoder

- 환자의 image를 사전학습된 *resnet50*을 사용하여 vector space  $s$  로 치환  
 $s = resnet50(img)$

### Auxiliary encoder

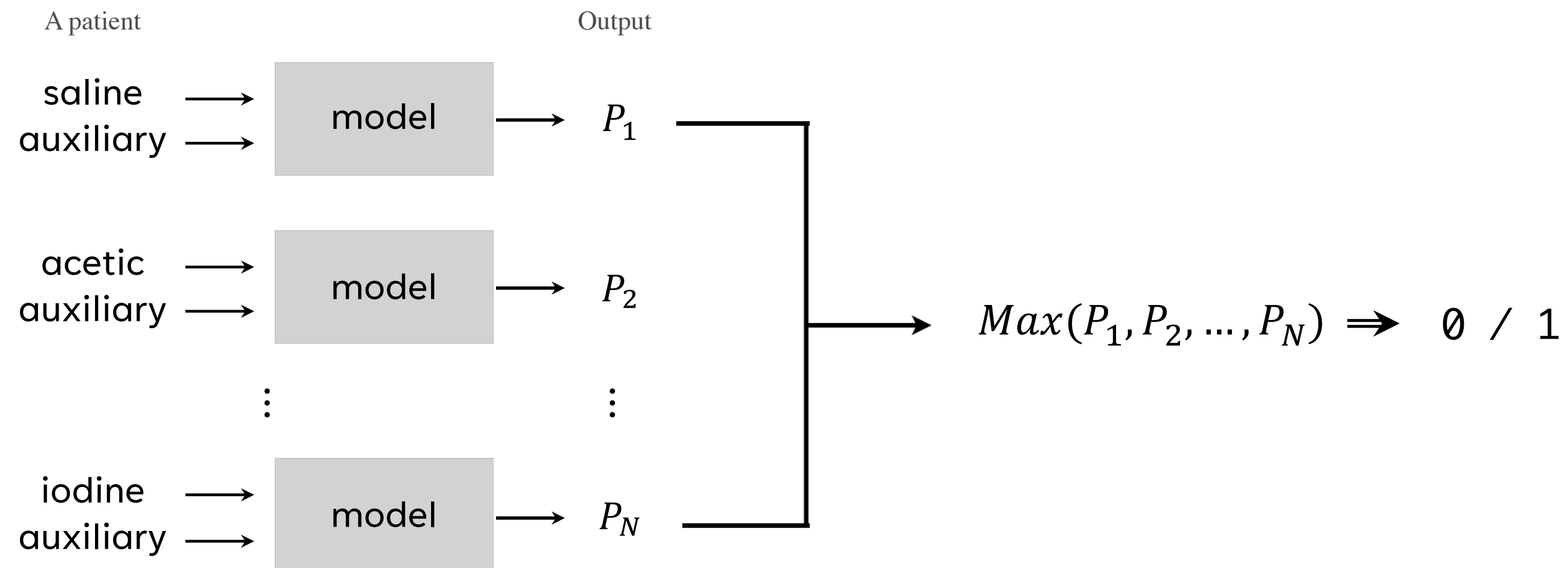
- 3개의 부가정보(Age, HPV Status, TZ Type)를 **concatenation**한 후 vector space  $z_2$  로 치환  
 $\mathbf{x}_i = \text{Concat}(x_{i1}, x_{i2}, x_{i3})$   
 $\mathbf{z}_1 = \text{ReLU}(W_1 \mathbf{x}_i + b_1)$   
 $\mathbf{z}_2 = W_2 \mathbf{z}_1 + b_2$

## Prediction Method

- $s$ 와  $z_2$ 를 **concatenation**한 후 최종 확률  $P_i$  예측  
 $\mathbf{z}_3 = \text{ReLU}(\text{Concat}(z_2, s))$   
 $P_i = W_3 \mathbf{z}_3 + b_3$

# Model architecture

## 환자 단위 진단



- 한 환자의 여러 이미지가 각각 모델을 통해 생성된 확률값을 비교
- 확률값 중 최대값을 구해 그 값이 **threshold(0.5)**를 넘는지 비교 후 최종 진단 결과 도출  
 $P_i = \max(P_1, \dots, P_N) > TH$

# Experiment Evaluation

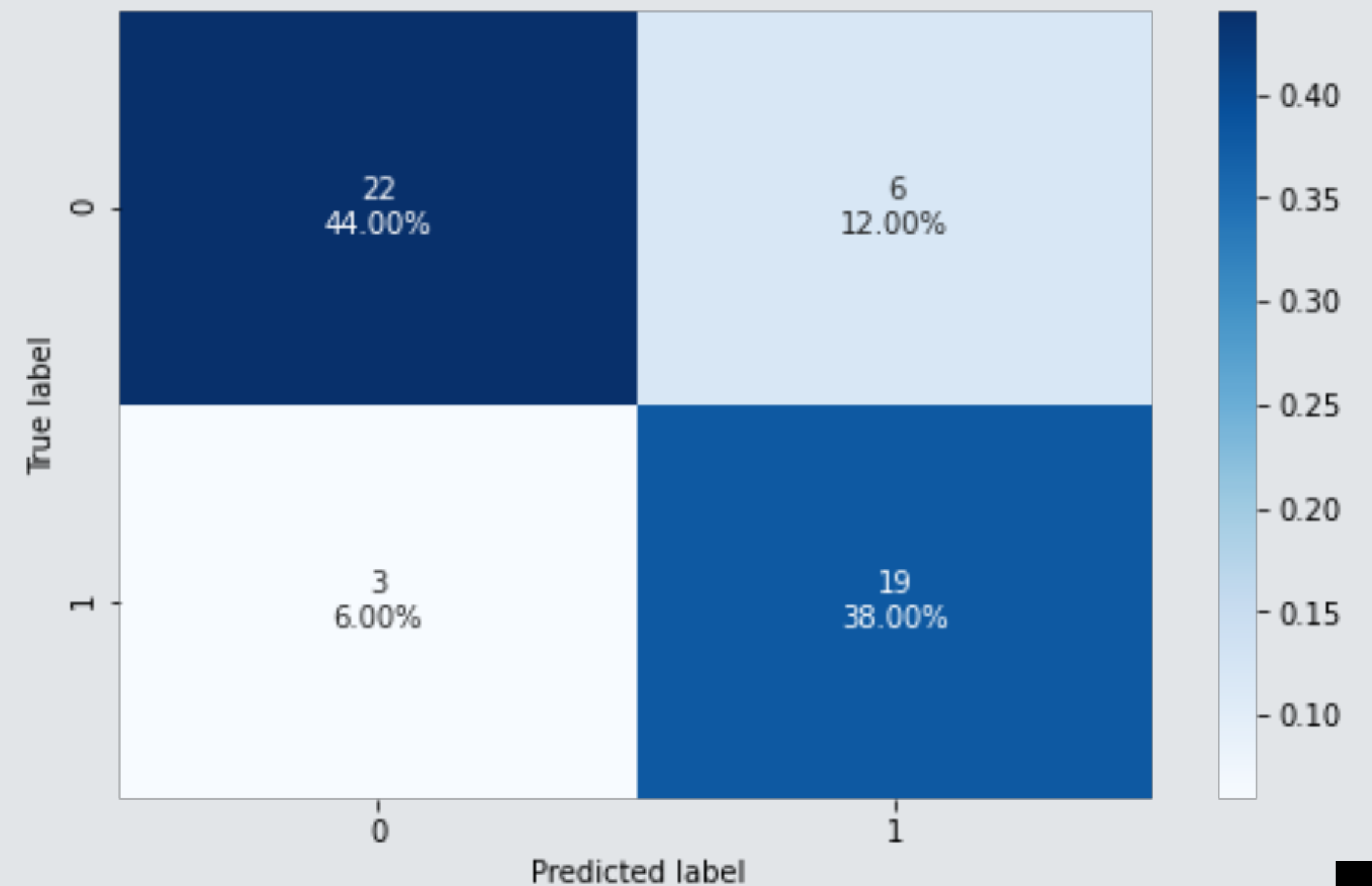
## – 기존 연구와 성능 비교

	Precision (%)	Recall (%)	Accuracy (%)	F1-Score (%)
CNN	44	100	44	61.11
Inception-resnet-v2	50	4.54	56	8.32
Proposed Model	76	86.36	82	80.85

Precision : 모델이 Positive라 분류한 것 중 실제 값이 Positive인 비율

Recall : 실제 값이 Positive인 것 중 모델이 Positive라고 분류한 비율

F1 Score : Precision과 Recall의 조화평균  $F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall}$



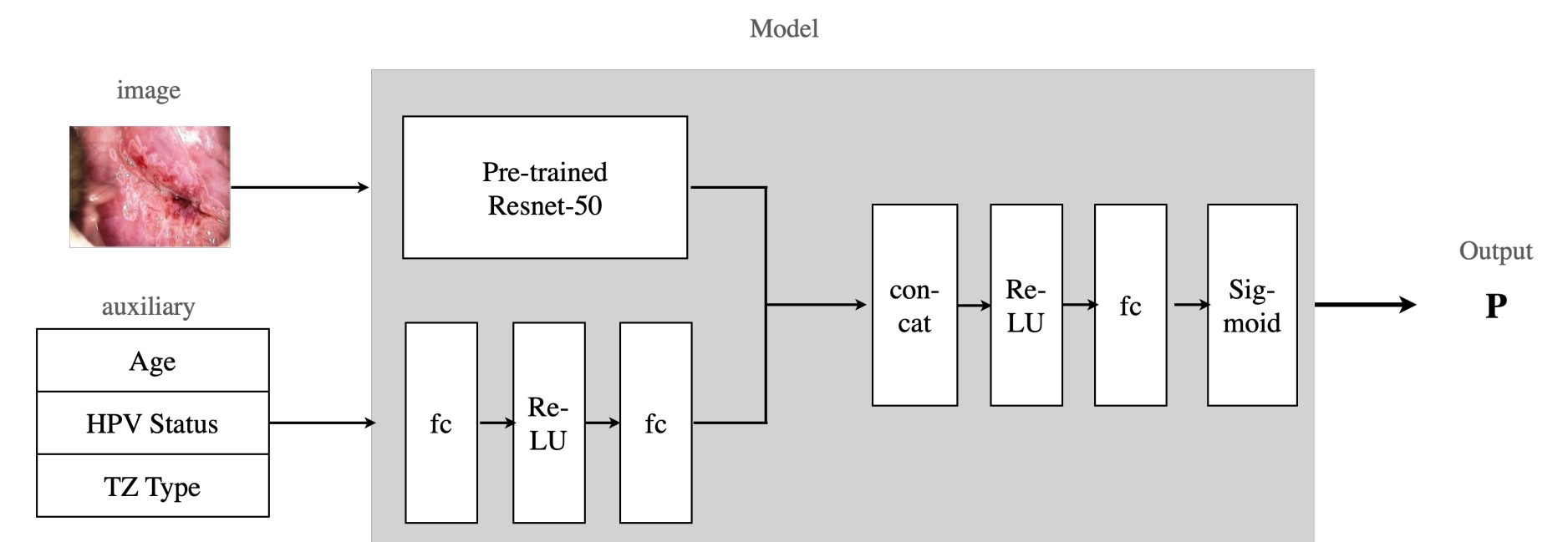
# Experiment Evaluation

## – 컴포넌트 별 효과 분석

	Precision (%)	Recall (%)	Accuracy (%)	F1 Score (%)
<b>Base Model</b>	0	0	56	0
<b>+ Pre-trained</b>	92.31	57.14	75	70.59
<b>+ auxiliary information</b>	85.71	85.71	85	85.71

**Base Model** : 사전학습된 모델을 사용하지 않았을 때 결과  
**+Pre-trained** : 사전학습된 resnet50 모델을 사용했을 때 결과  
**+auxiliary information** : 사전학습된 모델과 부가정보를 이용했을 때 결과

단순히 이미지를 사용하는 것 보다 부가정보를 사용했을 때 결과가 훨씬 좋은 것을 보여줌



**Precision** : 모델이 Positive라 분류한 것 중 실제 값이 Positive인 비율

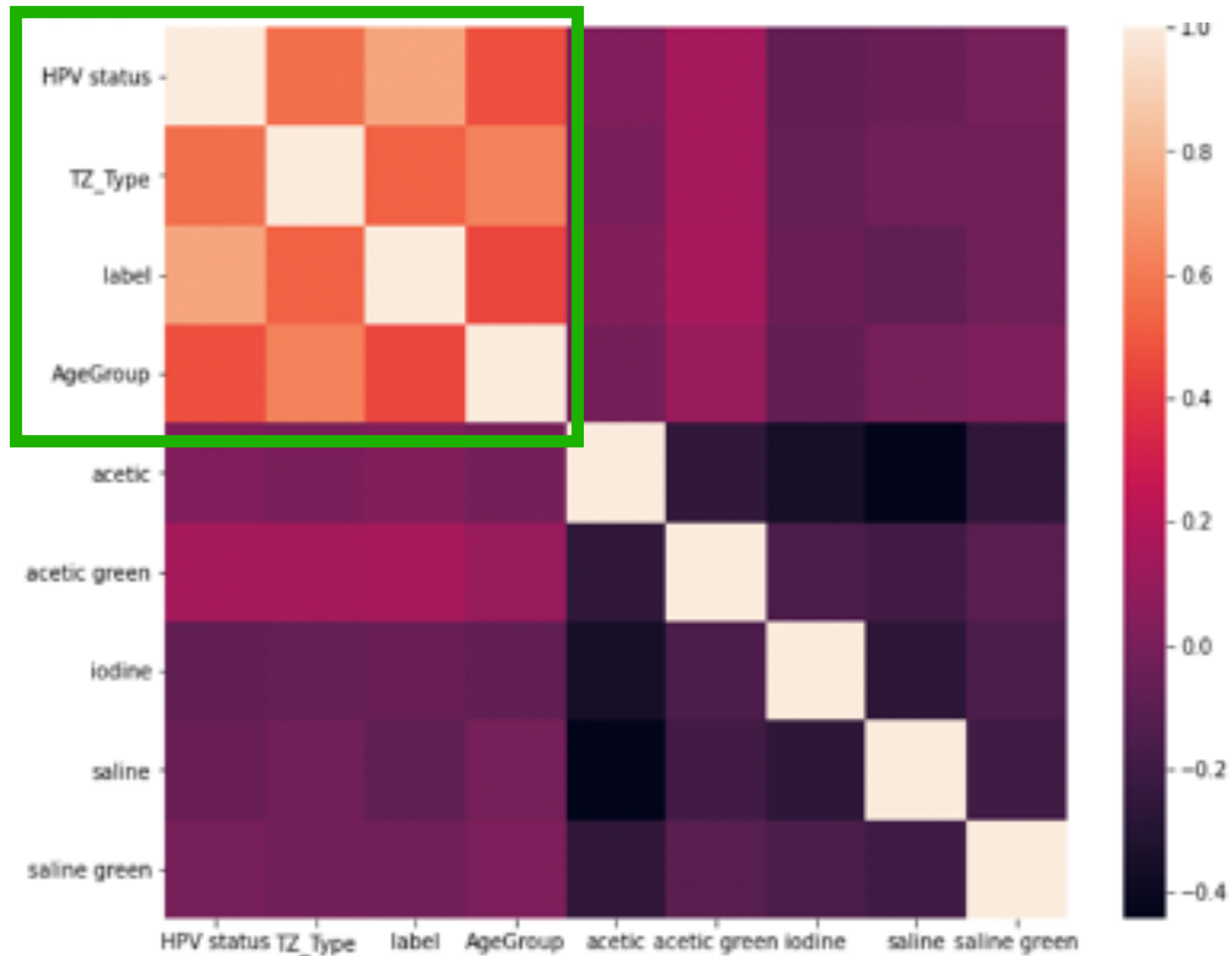
**Recall** : 실제 값이 Positive인 것 중 모델이 Positive라고 분류한 비율

**F1 Score** : Precision과 Recall의 조화평균

$$F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall}$$

# 부가정보가 정확도에 미치는 영향

## 부가정보와 결과의 상관관계



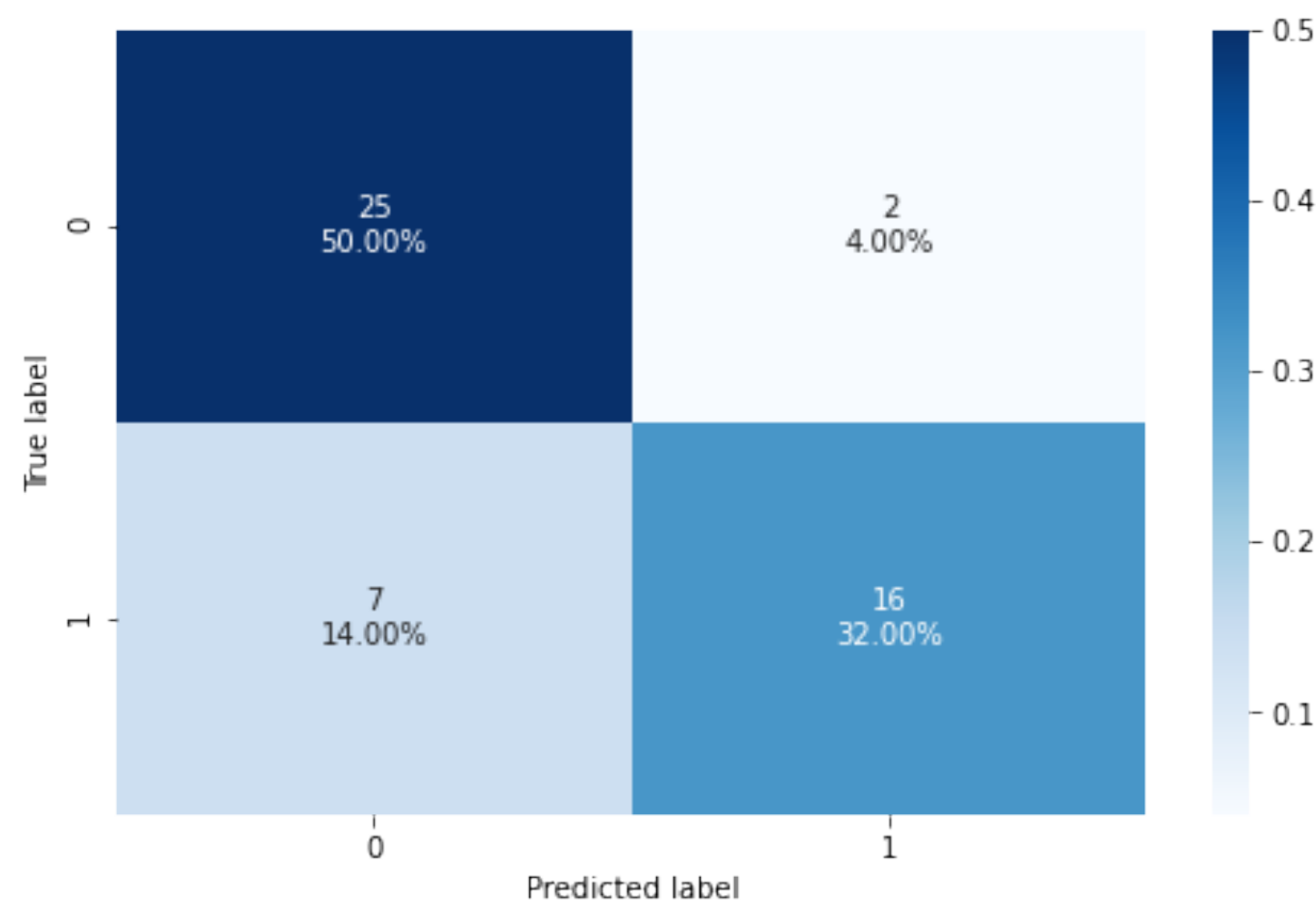
- 1 또는 -1에 가까울 수록 높은 상관관계를, 0에 가까울수록 낮은 상관관계를 나타냄
- 부가정보(HPV Status, TZ Type, AgeGroup)가 결과(label)와 **높은 상관관계**를 보여주고 있음
- 따라서 이미지 외에도 **부가정보를 함께 사용** 하여 자궁경부암 진단을 하는 것이 **훨씬 좋은 결과**를 예측할 수 있음을 나타낸다.



# 부가정보가 정확도에 미치는 영향

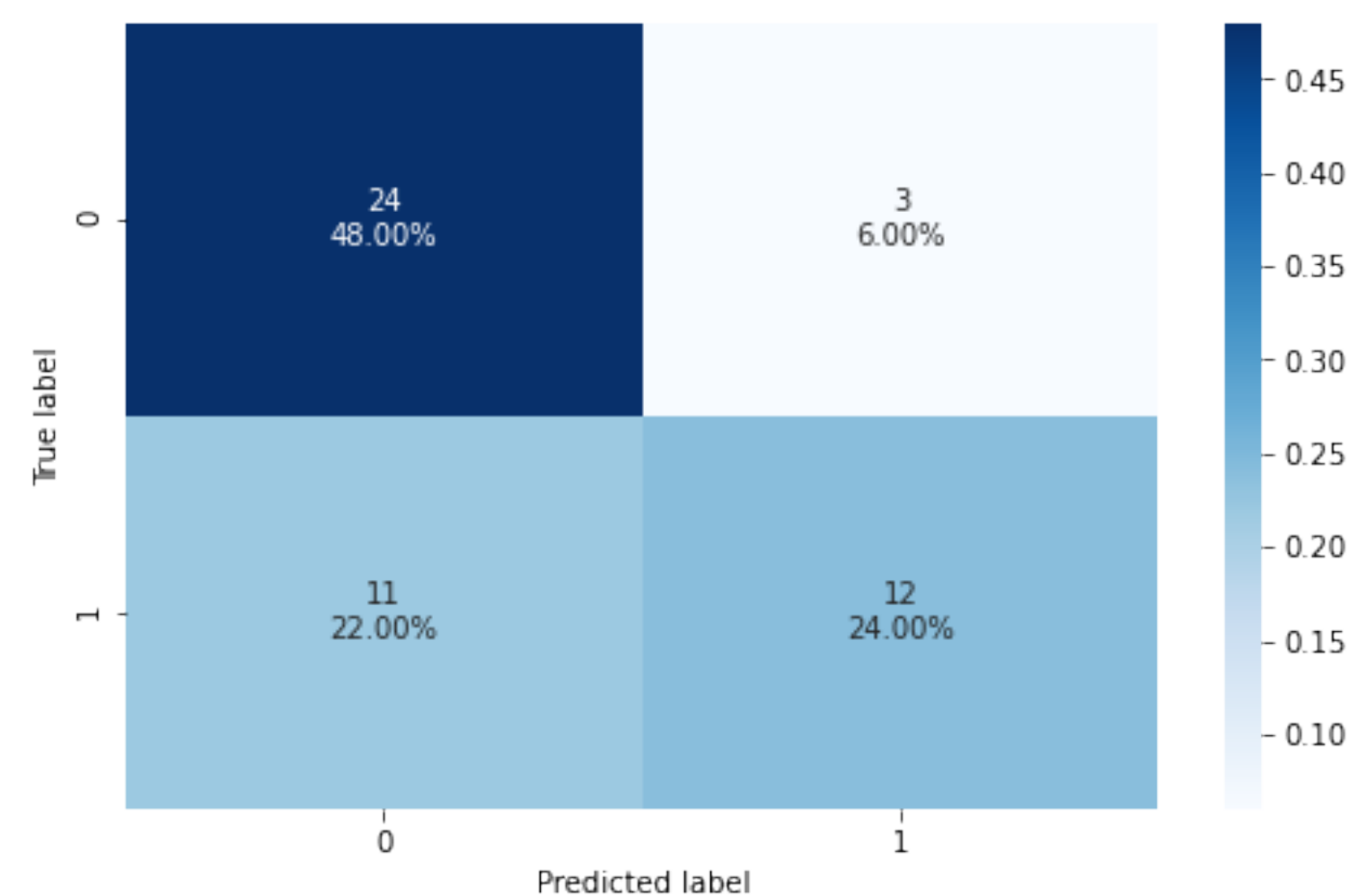
## 각 부가정보를 사용했을 때 정확도

### HPV Status만 사용



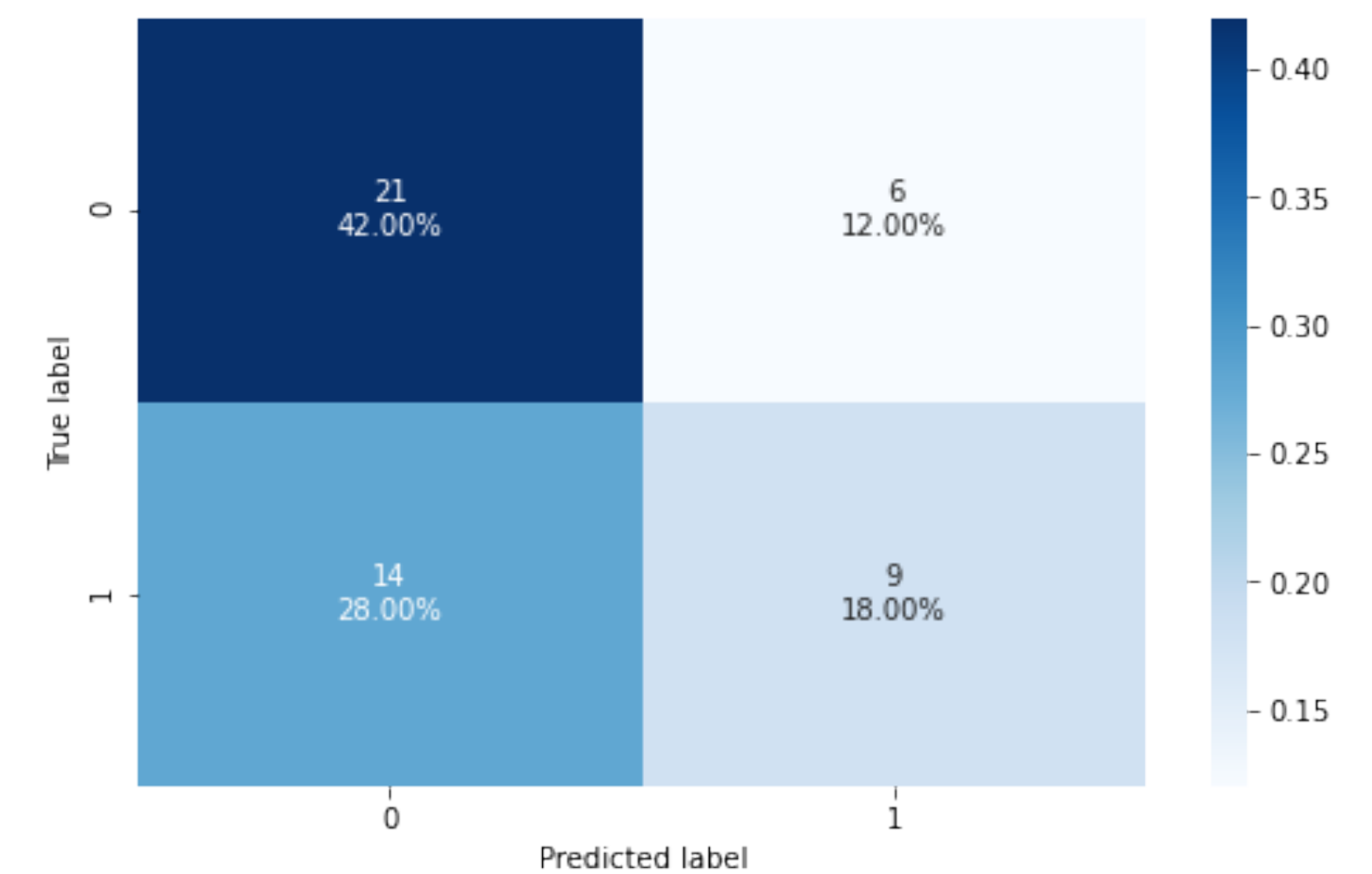
**Accuracy : 0.82**  
Recall : 0.696  
Precision : 0.889  
F1 : 0.78

### TZ Type만 사용



**Accuracy : 0.72**  
Recall : 0.522  
Precision : 0.8  
F1 : 0.632

### Age만 사용



**Accuracy : 0.6**  
Recall : 0.391  
Precision : 0.6  
F1 : 0.474

- 모든 부가정보를 사용하는 것이 각 부가정보를 사용하는 것보다 높은 성능을 보여줌
- 부가정보를 따로 적용했을 때 **HPV Status**가 결과와 가장 높은 연관성을 보여줌

# Conclusion

- 환자의 이미지만을 사용해 진단하던 기존 Colposcopy 검사에 추가적인 **부가정보**를 이용하여 더 안정성과 정확도가 높은 진단 방법 고안
- 모델 구조를 크게 Image encoder, Auxiliary encoder, Prediction Method로 나누어 제안
- Colposcopy 검사와 상관관계가 높은 세가지 **부가정보(HPV Status, TZ Type, Age)**를 추가적으로 사용
- 각 부가정보를 사용했을 때의 정확도를 비교하여 HPV Status, TZ Type, Age 순으로 연관성이 높음을 확인