



YOLO 학습 성능 향상을 위한 1채널 데이터의 3채널 합성 연구

고려대학교 세종캠퍼스 병렬알고리즘 연구실

이승우, 유승현, 서종웅, 백화평, 정용화

이승우 rainup4632@korea.ac.kr

유승현 tidlsld44@korea.ac.kr

서종웅 qornrs156@korea.ac.kr

백화평 seojongwoong@korea.ac.kr

정용화 ychungy@korea.ac.kr

Introduction

본 연구는 돼지 농장에서 스마트팜 도입 시 발생하는 데이터 부족 문제를 해결하기 위해 단일 채널(Grayscale) 영상 활용 방안을 제안한다. 기존 CNN 모델은 RGB 이미지 기반으로 설계되어 있어 단일 채널 데이터 사용 시 구조적 제약과 성능 저하가 발생할 수 있다. 이를 개선하기 위해 연구에서는 외곽선(edge) 강조 및 텍스처 변형 기법을 적용하여 모델의 일반화 성능을 향상시키고자 하였다. 실험을 통해 이러한 방법이 학습 성능 및 정확도 개선에 효과적임을 분석하였다.

Proposal method

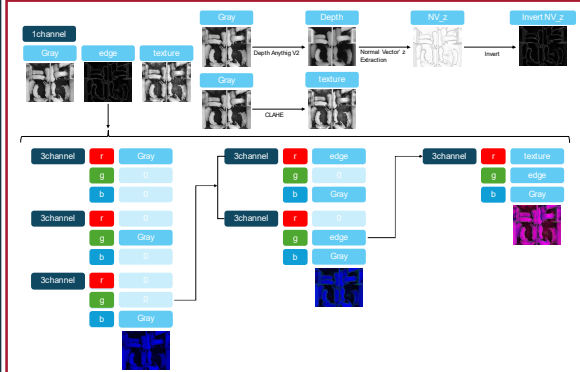


Figure 1. 제안된 방법 개요.

학습데이터 구성 방식	Box Precision↑ (%)	Box Recall↑ (%)	Box AP50↑ (%)	F1-score↑ (%)
1 channel				
Gray				
0 (Baseline)	Gray	Gray	Gray	84.6 76.7 84.5 80.4
	Gray	0	0	85.3 79.7 87.7 82.4
	0	Gray	0	91.3 78.1 85.0 79.7
	0	0	Gray	82.5 80.1 87.1 81.3
		0	0	85.3 79.7 87.7 82.4
X	Gray	Edge	Texture	88.5 78.7 88.7 83.3
				89.7 83.8 90.8 86.6

Figure 2. 모델의 성능 지표 비교.

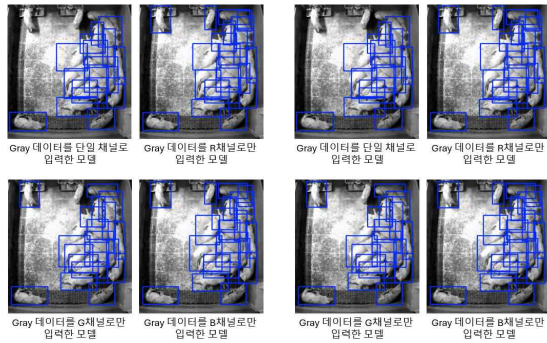


Figure 2. 모델의 predict 비교.

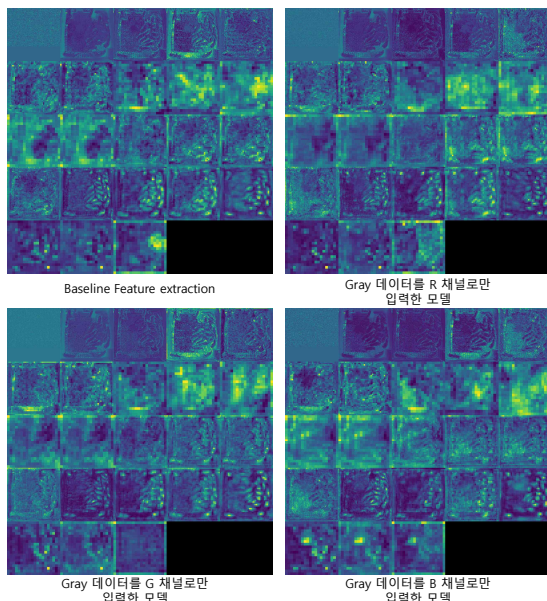
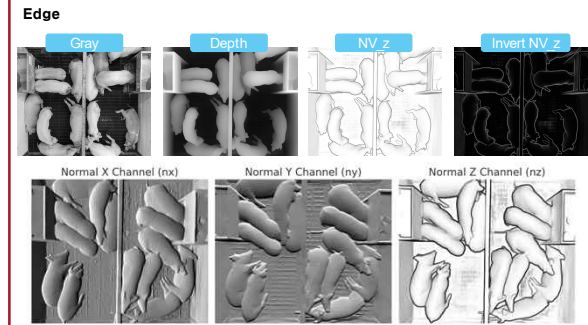
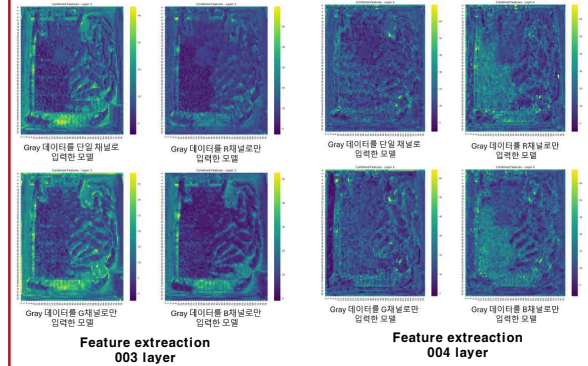
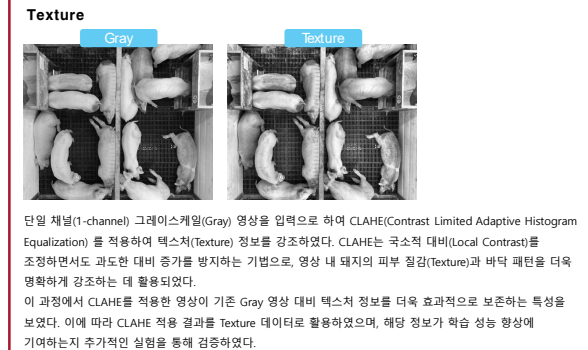


Figure 2. 모델의 Feature extraction 비교.



단일 채널(1-channel) 그레이스케일(Gray) 영상을 입력으로 하여 Depth-Anything2 모델을 활용해 1채널(1-channel) Depth 영상을 생성하였다. 이후, 해당 Depth Map에 대해 중앙 차분법(Central Difference Method)을 적용하여 법선 벡터(Normal Vector)를 계산하였으며, 이를 N_x, N_y, N_z 3개의 축으로 분리하여 분석하였다. 이 과정에서 N_z 채널이 돼지의 윤곽(edge) 정보를 가장 효과적으로 보존하는 특성을 보였다. 이에 따라 N_z 법선 벡터를 후처리하여 Edge 데이터로 활용하였으며, 이 정보가 학습에 기여하는지 추가적인 실험을 통해 검증하였다.



단일 채널(1-channel) 그레이스케일(Gray) 영상을 입력으로 하여 CLAHE(Contrast Limited Adaptive Histogram Equalization)를 적용하여 텍스처(Texture) 정보를 강조하였다. CLAHE는 국소적 대비(Local Contrast)를 조정하면서도 과도한 대비 증가를 방지하는 기법으로, 영상 내 돼지의 피부 질감(Texture)과 바닥 패턴을 더욱 명확하게 강조하는 데 활용되었다.

이 과정에서 CLAHE를 적용한 영상이 기존 Gray 영상 대비 텍스처 정보를 더욱 효과적으로 보존하는 특성을 보였다. 이에 따라 CLAHE 적용 결과를 Texture 데이터로 활용하였으며, 해당 정보가 학습 성능 향상에 기여하는지 추가적인 실험을 통해 검증하였다.

Data	R/G/B Count	R/G/B Ratio (%)	Data	R/G/B Count	R/G/B Ratio (%)
German Color Train	17880163200	33.57	German Color Train	105358609700	33.14
	181068451072	34		106209891981	33.4
	172741542862	32.43	Tilted	106378399016	33.46
German Color Val	47286251179	33.47	German Color Train	7344302600	34.21
	47811465362	33.84		74858559091	34.87
	46200349907	32.7	Top	66363143846	30.91
Jochiwon Color Val	2884671971	35.49	German Color_val	80741219875	34.02
	2834390697	34.87		78765384279	33.18
	2409193407	29.64	German Color_segment Train	77860668354	32.8



Conclusion

본 연구에서는 단일 Gray 스케일을 입력 데이터로 활용할 때 최적의 채널 배치를 탐색하였다. 분석 결과, 돼지와 배경의 RGB 픽셀값 비율이 유사함에도 불구하고, 채널별로 데이터가 다뤄지는 방식이 다르다는 점을 확인하였다. 이를 기반으로 각 채널에 적절한 데이터를 배치하는 것이 모델 성능 향상에 중요한 요소를 검증하였다.

첫 번째 실험에서는 Gray 스케일을 R, G, B 채널 중 하나에 독립적으로 할당한 결과, B 채널에 데이터를 입력했을 때 AP50 지표가 87.7%, F1-score가 82.4%로 가장 우수한 성능을 기록하며, 특정 채널에 데이터를 적절히 배치하는 것이 효과적임을 확인하였다.

두 번째 실험에서는 B 채널에 Gray 스케일을 고정한 상태에서 G 채널에 Edge 데이터, R 채널에 Texture 데이터를 추가한 최적의 조합을 통해 AP50 지표는 90.8%, F1-score는 86.6%로 향상시키며, 단순 Gray 스케일 입력 방식 대비 각각 6.3%와 6.2%의 성능 향상을 기록하였다.

이 결과는 단일 Gray 스케일을 활용할 때 채널 배치와 추가 특성 조합이 모델 학습 성능을 극대화할 수 있음을 시사하며, 향후 단일 채널 기반 데이터 활용과 모델 설계에 새로운 가능성을 제공할 것이다.