

MobileNetV2 기반의 딥러닝 모델을 활용한 넙치 다중 질병 분류 연구

이채령, 이명철*

건양대학교 의공학부, *한국전자통신연구원

chaeryounginji@gmail.com, *mclee@etri.re.kr

Multi-Disease Classification of Flatfish Using a MobileNetV2-Based Deep Learning Model

Chaeryoung Lee, Myungcheol Lee*

Konyang Univ., *Electronics and Telecommunications Research Institute

요 약

본 논문은 국내 양식 산업에 심각한 경제적 손실을 초래하는 넙치 질병을 신속하고 정확하게 진단하기 위하여 딥러닝 기술을 활용한다. 기존 연구에서 사용된 VGG16과 같은 모델은 과도한 연산량과 상대적으로 낮은 정확도로 인해 모바일 환경에서의 실시간 진단에 한계가 있다. 이에 따라 MobileNetV2 기반의 경량 딥러닝 모델을 제안한다. 넙치 질병 이미지 데이터셋에 MobileNetV2를 미세 조정(fine-tuning)하여 다중 질병 분류 성능을 최적화하였다. 또한, 기존 VGG16을 경량화한 TFLite 모델과 비교 분석을 통해 MobileNetV2 모델의 성능과 모바일 적용 효율성을 검증하였다. 제안 모델은 분류 성능 측면에서 기존 모델 대비 약 10% 증가한 87.53%의 F1 score를 달성했으며, 모델 크기 약 4.8배 감소 및 추론 속도 6.56ms로 크게 향상하였다. 이는 넙치 질병 진단에 있어 정확성과 실시간성을 동시에 확보하여, 양식 산업의 효율적인 질병 관리에 기여할 수 있다.

I. 서 론

국내 양식 산업에서 넙치는 가장 중요한 양식어종 중 하나로, 국내 생산량에서 높은 비중을 차지하며 산업 전반에서 핵심적인 위치를 확보하고 있다[1]. 그러나 넙치 양식에서는 다양한 감염성 질병으로 인한 대규모 폐사가 발생해 양식 산업의 경제적 손실로 이어질 수 있으며[2], 이러한 질병은 다중 감염 형태로 빈번하게 발생하여 신속하고 정확한 다중 질병 진단이 중요하다.

기존 연구에서는 VGG16 기반의 딥러닝 모델을 활용하여 넙치 질병 이미지를 분류하였으나[3], VGG16은 구조적으로 연산량과 모델 크기가 커서 모바일 환경에서의 실시간 적용에는 적합하지 않다. 이에 본 연구에서는 MobileNetV2 기반 다중 질병 분류 모델을 설계하고, TensorFlow Lite(TFLite)로 변환하여 모바일 환경에서도 적용할 수 있는 모델을 제안하며, 기존 VGG16 기반 모델과의 성능을 비교하여 경량 모델의 현장 적용 가능성을 검증하였다.

II. 본 론

가. 데이터셋 구성

본 연구는 국내 양식 넙치의 주요 질병 8종을 대상으로 경량화된 딥러닝 기반 다중 질병 분류 모델을 설계하였다. 데이터는 AI-Hub에서 제공하는 넙치 질병 이미지로, 그림 1과 같이 바이러스성출혈성패혈증, 림포시스티스병, 여위병, 스쿠티카병, 연쇄구균증, 비브리오병, 에드워드병, 기타로 총 8개 질병 클래스로 구성된다. 이때 비브리오병은 5개의 세부 질병(h, a, ang, s, pd)을 하나로 통합하여 단일 클래스로 정의하였다.

원본 이미지는 약 6,000×4,000 크기로 제공되었으며, 학습을 위해 224×224로 리사이즈하였다. 전체 데이터는 총 31,482장으로 train, valid,

test 데이터를 8:1:1 비율로 나누어 사용하였으며, 데이터 불균형 문제를 완화하기 위해 소수 클래스는 오버샘플링과 데이터 증강을 적용하고, 일부 클래스에 대해서만 언더샘플링하였다.

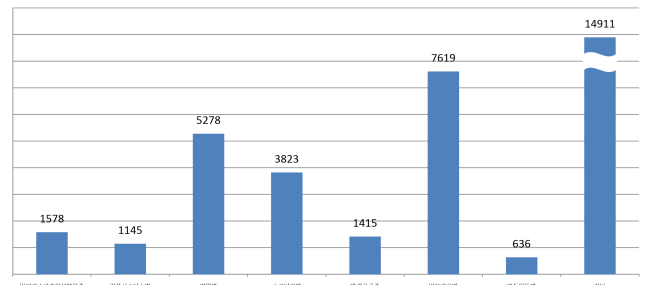


그림 1. 넙치 질병 클래스별 이미지 분포

나. 모델 구성

VGG16 기반 모델[4]과 MobileNetV2 기반 모델[5]을 TFLite 형식으로 사용하여 모바일 환경에서의 적용 가능성을 검증하였다.

먼저 사전에 학습된 VGG16 기반 모델은 Conv Block 1-5로 구성된 특징 추출부와 Flatten 이후 fully connected layers(1024-512-7)로 이어지는 분류기로 이루어진다.

반면 MobileNetV2 기반 모델은 그림 2와 같이 특징 추출부의 block 14~16만 학습하고, 나머지 블록은 동결(freeze)하였다. 이는 초기 블록의 사전 학습 가중치를 유지하여 기본적인 특징 표현을 보존하면서, 상위 블록만을 미세 조정하여 질병 분류 성능을 향상하고 과적합을 방지하기 위함이다. Global Average Pooling을 통해 특성을 벡터화한 뒤, fully connected layers(1024-512-256-8)를 거쳐 최종 출력을 도출하였으며, 각 층에는 Dropout과 L2 정규화를 적용하여 경량화된 구조에서도 안정적인 학습과 과적합 방지를 도모하였다.

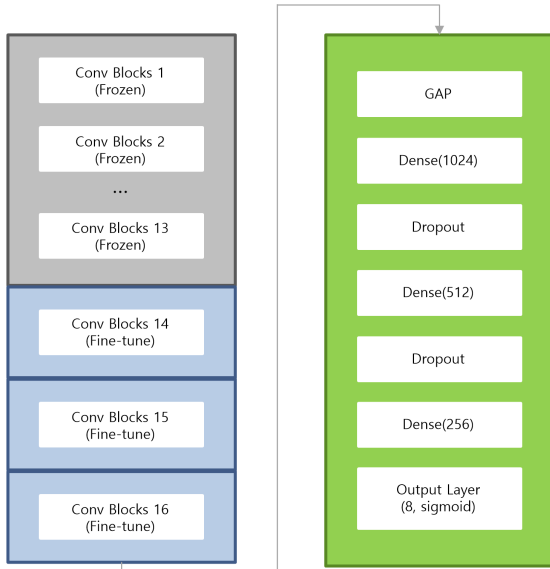


그림 2. MobileNetV2 기반 넙치 질병 분류모델 구조

다. 손실함수 및 평가지표

본 연구의 넙치 질병 분류 문제는 한 이미지에 여러 질병이 동시에 나타날 수 있는 다중 레이블 분류 문제이다. 이에 따라 출력층에는 클래스별 확률을 독립적으로 산출하기 위해 시그모이드(sigmoid) 활성화 함수를 적용하였다. 학습 단계에서 손실함수는 Asymmetric Loss(ASL)를 사용하였다. ASL은 희소 클래스의 학습 신호를 강화하고 hard negative의 영향을 완화하여 초기 수렴을 촉진하는 손실함수로, 본 연구에서는 구현 시 프레임워크 기본 설정에 따라 클래스별 손실을 평균하여 적용하였다.

$$ASL_{(y,p)} = - \sum_{c=1}^C [y_c(1-p_c)^{\gamma^+} \log(p_c) + (1-y_c)p_c^{\gamma^-} \log(1-p_c)]$$

여기서 y_c 는 c 번째 클래스의 실제 레이블이며, p_c 는 예측 확률, γ^+ 와 γ^- 는 각각 양성 및 음성 표본에 대한 조정 파라미터를 의미한다.

검증 단계에서는 F1 score를 최대화하기 위하여 validation 데이터에 기반한 클래스별 임계값(threshold) 탐색을 수행하였다. 추가적으로 Top-K 방식을 적용하여 각 입력의 예측 확률 벡터에서 상위 K개 클래스를 우선적으로 선택함으로써, 확률 분포가 여러 클래스에 분산될 때 발생할 수 있는 과도한 양성 판정을 완화하고 보다 안정적인 평가를 가능하게 하였다.

모델 성능 평가는 F1-score, Precision, Recall, Specificity, AUPRC를 활용하였다[6]. 이때 F1-score는 모든 클래스에 대한 참-거짓 예측을 표본 단위로 합산하는 micro 방식을 사용하였다. 또한, 모델의 실용성을 평가하기 위해 모델 크기(MB)와 이미지 한 장당 처리 시간(ms)을 측정하였다.

라. 실험 결과

본 비교는 TFLite 변환 모델을 기준으로 수행하였다. VGG16은 과거 학습된 모델로 기타 질병 클래스가 포함되지 않아, 본 데이터셋에서도 이를 제외한 7개 클래스(입력 112×112)로 평가하였고, MobileNetV2는 기타 질병을 포함한 8개 클래스(입력 224×224)로 학습 평가하였다. 두 모델 모두 비교의 일관성을 확보하기 위하여 임계값 0.8을 적용하였으며, 사전 검토에서도 해당 설정이 가장 안정적이고 우수한 성능을 보였다. 최종 결과는 표 1에 제시하였다. VGG16 TFLite의 F1-score는 77.49%였던 반면, MobileNetV2 TFLite는 87.53%로 VGG16 대비 약 10% 향상되었으며, 모델 크기도 76.16MB에서 15.98MB로 약 4.8배 감소하였다. 또한 TFLite 모델 기준 CPU 추론시간도 크게 단축되었다. 이러한 결과는 MobileNetV2

가 정확도와 효율성 모두에서 우수하여 모바일 환경에서의 실시간 진단에 적합함을 보여준다.

모든 학습은 NVIDIA A40 GPU 4장(각 48GB 메모리)을 탑재한 서버에서 TensorFlow 2.13.1 기반, Python 3.8.10 환경에서 수행하였으며, CPU는 Intel Xeon Gold 6448Y(64코어, 128스레드)를 사용하였다.

표1. 실험 결과 비교

	VGG16	VGG16 TFLite	MobileNetV2	MobileNetV2 TFLite
F1-score	75.75	77.49	87.54	87.53
Precision	81.13	83.74	86.73	86.71
Recall	71.04	72.10	88.36	88.32
Specificity	95.28	95.90	97.29	97.18
AUPRC	70.55	72.77	94.82	94.80
Model Size	228.98	76.16	19.45	15.98
Time(GPU)	1.72	-	2.13	-
Time(CPU)	54.92	103.23	12.60	6.56

III. 결 론

본 연구에서는 넙치 질병 8종을 분류할 수 있는 MobileNetV2 기반 경량 모델을 제안하고, TFLite 변환을 통해 모바일 환경에서의 실시간 적용 가능성을 검증하였다. 제안 모델은 기존 VGG16 대비 모델 크기와 연산량을 크게 줄이면서도 F1-score와 AUPRC 등 정확도 지표가 향상되었으며, 추론 시간 역시 6.56ms로 단축되었다. 향후 실제 양식장 환경에서 모바일 앱과 연계한 실시간 검증을 통해 제안 모델의 실용성을 확인하고 국내 양식 산업의 질병 관리 효율성 향상에 기여할 수 있을 것으로 기대된다.

ACKNOWLEDGMENT

- 이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. 2021-0-00136, 다양한 산업 분야 활용성 증대를 위한 대규모/대용량 블록체인 데이터 고확장성 분산 저장 기술 개발).
- 이 연구는 과학기술정보통신부의 재원으로 한국지능정보사회진흥원의 지원을 받아 구축된 "넙치 질병 데이터"를 활용하여 수행된 연구입니다. 본 연구에 활용된 데이터는 AI 허브(aihub.or.kr)에서 다운로드 받으실 수 있습니다.

참 고 문 헌

- [1] 한국통계정보서비스(KOSIS), "시도·시군구별 양식방법별 어종별 양식현황(입식, 생산, 사육, 먹이[사료] 투입량)," 자료갱신일 2025-05-30.
- [2] Shim, J.D., Hwang, S.D., Jang, S.Y., Kim, T.W., and Jeong, J.M., "Monitoring of the mortalities in olive flounder (*Paralichthys olivaceus*) farms of Korea," Journal of Fish Pathology, vol. 32, no. 1, pp. 29-35, 2019.
- [3] 한국지능정보사회진흥원(NIA), AI-Hub, "넙치 질병 이미지 데이터셋," 2020.
- [4] Simonyan, K., and Zisserman, A., "Very Deep Convolutional Networks for Large-Scale Image Recognition," ICLR, 2015.
- [5] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.C., "MobileNetV2: Inverted Residuals and Linear Bottlenecks," CVPR, pp. 4510-4520, 2018.
- [6] Zhang, M.L., and Zhou, Z.H., "A Review on Multi-Label Learning Algorithms," IEEE TKDE, vol. 26, no. 8, pp. 1819-1837, 2014.