

넙치의 외부 및 내부 질병 증상을 이용한 넙치 질병 예측 알고리즘 비교 연구

(Comparative Study on Prediction Algorithms for Paralichthys olivaceus Diseases Using External and Internal Symptoms)

신유식*, 임창현**, 신영학***

(Yusik Shin, Changhyeon Im, Younghak Shin)

요약

양식장에서 서식하는 어류들은 질병에 걸리면 전파 속도가 매우 빨라 이를 사전에 방지하지 않으면 양식장에 크나큰 손실을 주게 된다. 현재 다양한 딥러닝 기술을 통하여 어류의 외부와 내부 질병 증상을 검출하고 예측하는 연구가 활발히 진행되고 있다. 하지만 특정 질병이라도 질병별로 발현되는 증상이 명확하지 않고, 증상이 여러 질병에 걸쳐 발현되는 경우가 빈번하여 증상을 바탕으로 한 질병 분류 성능은 아직 낮은 단계이다. 본 연구에서는 딥러닝 기술을 이용하여 어류의 외부와 내부 질병 증상을 찾고, 이를 사용해 어류의 질병을 예측하는 다양한 방법론을 제안하고 성능 비교를 통해 가장 우수한 질병 예측 알고리즘을 선정한다. 질병 예측 알고리즘으로는 어류 질병 전문가가 분석한 질병 증상 표를 기반으로 한 방법, 데이터 통계를 기반으로 하는 질병 예측 방법, 앙상블 모델 및 딥러닝 모델 학습을 기반으로 한 예측 방법으로 총 세 가지 방법론을 제안한다. 실험 결과, 질병 증상 표와 데이터 통계 기반 예측 방법은 각각 48.45%와 50.91%의 예측 성능을 보인 반면, 앙상블 모델 및 딥러닝 모델 학습을 활용한 질병 예측 방법은 70.10%의 성능을 기록하며 가장 높은 예측 정확도를 보였다.

■ 중심어 : 어류 질병 예측 ; 데이터 통계 ; 앙상블 ; 딥러닝 ;

Abstract

Fish that live in fish farms spread very quickly when they get sick, so if they do not prevent it in advance, they will cause significant losses to fish farms. Currently, research is actively being conducted to detect and predict external and internal disease symptoms of fish through various deep learning technologies. However, even for a specific disease, the symptoms that are expressed for each disease are not clear, and symptoms are frequently expressed across multiple diseases, so the disease classification performance based on symptoms is still low. In this study, we use deep learning techniques to find external and internal disease symptoms in fish, propose various methodologies to predict disease in fish, and select the best disease prediction algorithm through performance comparison. We propose a total of three methodologies as disease prediction algorithms: a method based on a disease symptom table analyzed by a fish disease expert, a disease prediction method based on data statistics, an ensemble model, and a prediction method based on deep learning model training. Experiments show that disease symptom tables and data statistics-based prediction methods performed 48.45% and 50.91%, respectively, while disease prediction methods using ensemble model and deep learning model training performed 70.10%, showing the highest prediction accuracy.

■ keywords : Fish Disease Prediction ; Data Statistics ; Ensemble ; Deep Learning ;

I. 서론

최근 몇 년간 국내에서 넙치 양식에 관한 관심과 투자가 급증함에 따라 넙치를 주요 양식어종

* 준회원, 목포대학교 컴퓨터공학과

** 준회원, 목포대학교 컴퓨터공학과

*** 정회원, 목포대학교 컴퓨터공학과

이 논문은 2024년도 해양수산부 재원으로 해양수산과학기술진흥원의 지원을 받아 수행된 연구입니다(RS-2022-KS221673, 빅데이터 기반 양식생산성 향상기술).

접수일자 : 2024년 10월 29일

수정일자 : 2024년 11월 14일

게재확정일 : 2024년 11월 18일

교신저자 : 신영학 e-mail : younghak@mokpo.ac.kr

으로 하는 양식산업이 증가하고 있다. 그에 따라 생태계의 보존과 환경 보호를 모두 고려하여 넙치의 지속적인 생산성을 위한 양식 방안을 모색하고 있다[1].

양식장에서 서식하는 넙치어류는 에드워드병, 비브리오팀, 연쇄구균병, 스쿠티카병 등의 세균성 질환과 바이러스성 질환에 매우 취약하다. 이러한 질병을 일으키는 바이러스와 박테리아는 물속에서 빠르게 퍼져 나가 양식장 내 어류들에게 전염되며, 심할 경우 대규모 폐사를 초래할 수 있다. 이를 방지하기 위해서는 질병을 초기에 발견하고 치료하는 것이 매우 중요하다[2].

어류의 질병은 전통적으로 질병 전문가의 지식을 바탕으로 진단해 왔지만 진단 및 대처에 많은 시간과 비용이 발생하며 늦어지는 경우 양식장에 심각한 손실을 일으킬 수 있다[3].

현재 딥러닝 기반 모델을 활용하여 어류의 질병을 예측하는 시도들이 점차 증가하고 있다. 논문[4]은 넙치의 질병 증상을 효과적으로 인식하기 위해 넙치 이미지에서 장기를 추출하고 장기에서 발생하는 질병 증상을 최신 딥러닝 기반 분류 모델을 이용하여 분류한다. 넙치의 장기를 추출 후 분류 모델 학습을 통해 장기 증상을 인식하는 방법이 해부 이미지에서 장기의 증상을 검출하는 방법보다 높은 정확도를 보이는 것을 확인하였다.

논문[5]은 인도네시아 재래식 양식업체에서 주로 사용되는 시각적 식별 방식을 보완하기 위해 이미지 인식 기술을 적용하여 질병 식별의 효율성을 높였다. 이 연구는 EUS 초음파내시경 질병 감지를 위해 MobileNetV2 모델과 HSV 임계값 기반 객체 분할 기법을 사용하여, 감염된 부위를 84% 정확도로 식별했다.

딥러닝을 활용한 질병 증상 예측 성능이 높게 나오는 연구들이 존재하지만, 증상을 바탕으로 한 질병 분류 성능은 아직 낮은 단계이다. 이는 특정 질병이라도 질병별로 발현되는 증상이 명확하지 않고, 증상이 여러 질병에 걸쳐 발현되는

경우가 빈번하기 때문이다.

본 논문에서는 질병 증상을 바탕으로 어류의 질병을 정확하게 예측하기 위해 다양한 방법론을 제시하고, 질병 예측의 정확도를 비교 분석하였다. 먼저 넙치의 외부 증상과 내부 증상을 딥러닝 모델 학습을 통해 예측한다. 그 후, 예측된 증상들을 종합하여 넙치 어류의 일곱 가지 질병을 예측한다. 세 가지 질병 예측 방법을 사용하여 비교 분석 후 평가하였다. 첫 번째는 어류 질병 전문가가 분석해서 제작한 증상의 중요도표를 바탕으로 질병을 예측하는 방법이며, 두 번째는 현재 어류 데이터의 질병과 증상 개체 수로 통계를 내어 예측하는 방법이다. 세 번째 방법은 어체의 질병별 증상들을 학습하여 기계학습 앙상블 모델 및 딥러닝 모델을 사용하여 질병을 예측하는 방법이다.

본 논문의 구성은 2장에서는 넙치의 외부 증상 검출과 내부 증상 분류, 질병 예측에 관해서 기술한다. 3장에서는 질병을 예측하기 위해 설계한 알고리즘의 학습 및 성능 비교에 관해서 기술한다. 4장에서는 본 논문의 결론과 향후 연구 방향에 관하여 기술한다.

II. 어류 증상 검출 및 질병 예측

본 연구에서는 넙치 어류의 외부 증상 검출과 내부 증상 분류를 통해 증상을 찾고, 이를 이용해 질병을 예측하는 다양한 방법론을 제안한다. 그림 1은 어류의 외부 및 내부 장기 이미지를 입력하여 최종적으로 어류의 질병을 예측하는 알고리즘을 나타낸 도식도이다.

먼저 넙치 어류의 개체별 실험 데이터에서 넙치의 외부 증상 객체 검출과 내부 장기를 세분화 후 증상 분류를 진행하였다[4][6]. 넙치의 외부 증상 검출과 내부 증상 분류를 학습한 모델을 사용하여 외부 및 내부 증상을 예측하고, 최종 질병 분류를 진행한다.

질병 분류 방법은 어류 질병 전문가가 직접 만든 증상 가중치표를 이용하는 방법, 전체 어류

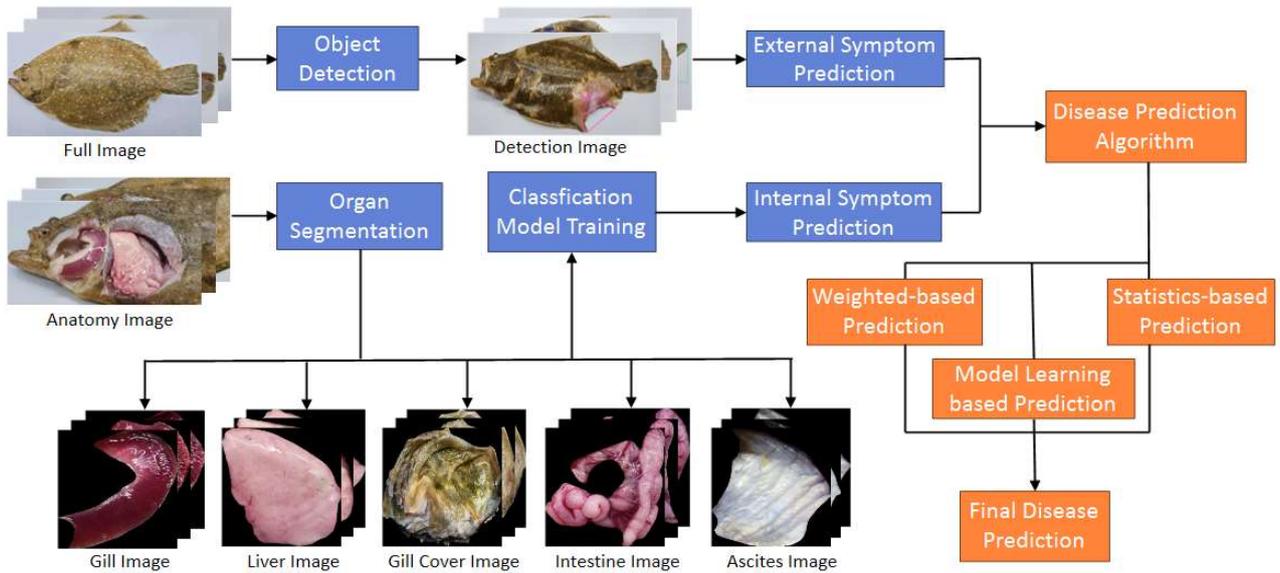


그림 1. 외부 증상 검출과 내부 증상 분류 후 여러 질병 예측 방법을 통한 어류 질병 예측 알고리즘

데이터의 질병별 증상 개체 수 통계를 바탕으로 질병을 예측하는 방법, 기계학습 앙상블 모델 및 딥러닝 모델을 사용하여 질병을 예측하는 방법을 사용하였다. 최종적으로 세 가지 방법의 성능을 비교하여 어떤 질병 예측 방법이 더 효과적인지 분석하였다.

1. 외부 증상 검출

(1). 외부 증상 검출 실험 데이터

본 연구에서는 외부 증상 검출 실험 데이터로 넙치의 전체적인 외형이 나타난 이미지를 사용한다. 그림 2는 넙치의 외부 증상을 검출하기 위해 촬영한 이미지로 (a)는 0도, (b)는 45도, (c)는 90도에서 촬영하였다. 그림 3은 넙치의 외형에 나타난 외부 증상 예시 이미지와 해당 질병 증상의 마스크 레이블 예시이다. (a)는 항문탈장으로 넙치의 장기가 튀어나와 외부로 나타나 있는 증상이다. (b)는 체표궤양으로 넙치의 체표 표면에 발생하는 병변으로 붉은 반점이나 상처가 나타나 있는 증상이다. (c)는 주둥이궤양으로 넙치의 입 주변에 나타나는 병변으로 주둥이 부분이 부풀어 오르거나 붉게 변하거나 상처가 나타나 있는 증상이다. (d)는 꼬리지느러미 결손으로 넙치의 꼬리지느러미 일부분이 훼손되어 있는 증상이다. (e)는 지느러미 기부 궤양으로 지느러미 결손과

궤양이 같이 나타나는 증상이다.



(a) 넙치의 0도 이미지



(b) 넙치의 45도 이미지



(c) 넙치의 90도 이미지

그림 2. 넙치의 각도별 외형 이미지



(a) 항문탈장



(b) 체표궤양



(c) 주둥이케양



(d) 꼬리지느러미 결손



(e) 지느러미 기부 케양

그림 3. 넙치 외부 증상 이미지

(2). 외부 증상 검출 모델

본 연구에서는 넙치의 외부 질병 증상을 검출하기 위해 YOLO 버전 중 하나인 YOLOv8 모델 [7]을 사용하였다.

YOLOv8은 객체 검출, 세분화 등 다양한 컴퓨팅 비전 작업을 지원하며, 이전 버전에 비해 새로운 기능을 도입하여 효율성, 성능 등을 향상한 모델이다[7].

표 1은 클래스별 외부 증상 데이터의 구성으로 각 클래스마다 8 : 2의 비율로 나누어 학습 및 테스트 데이터를 구성하였다. 표 2는 YOLOv8을 사용하여 외부 증상을 검출한 뒤 클래스별로 성능을 나타낸 것이다. 다른 클래스들에 비해 꼬리지느러미 결손의 성능이 낮는데, 데이터 수가 적은 것과 데이터의 특징이 일정하지 않고 다양한 경우가 많기 때문으로 판단된다.

표 1. 클래스별 외부 증상 데이터 구성

	train	test	Total
항문탈장	232	14	246
체표케양	436	21	457
주둥이케양	844	58	902
꼬리지느러미결손	160	12	172
지느러미기부케양	132	14	146

표 2. 클래스별 외부 증상 검출 성능

클래스 \ 성능	Precision	Recall	mAP50
항문탈장	0.922	1.0	0.995
체표케양	0.876	0.81	0.898
주둥이케양	0.873	0.85	0.843
꼬리지느러미결손	0.9	0.75	0.837
지느러미기부케양	0.93	0.89	0.936

2. 내부 증상 분류

(1). 내부 장기 세분화

넙치의 내부 장기 증상 분류를 위해서 먼저 넙치의 해부 이미지에서 장기를 추출해야 한다. 따라서 본 연구에서는 YOLOv8을 사용하여 넙치의 내부 장기 세분화를 진행하였다 [6]. 그림 4는 내부 장기를 세분화 후 추출할 때 사용하는 넙치의 장기 해부 예시 이미지이다.



그림 4. 넙치 해부 이미지

(2). 내부 증상 분류 실험 데이터

내부 데이터의 경우 넙치의 간, 복수, 아가미, 아가미뚜껑, 장 데이터를 활용한다. 그림 5는 넙치의 장기 증상 데이터의 예시 이미지이다. (a)는 정상 간으로 분홍빛을 띠고, (b)는 울혈 간으로 정상 간보다 진한 붉은빛을 띠고 있다. (c)는 정상 아가미로 진한 검붉은빛을 띠고 있으며, (d)는 아가미빈혈로 연한 분홍빛을 띠고 있다. (e)는 정상 아가미뚜껑이고, (f)는 아가미뚜껑 출혈로 아가미뚜껑에서 출혈이 발생하여 일부 붉은색이 나타난다. (g)는 정상 장으로 분홍빛을 띠고 있고, (h)는 염증 장으로 연노란빛이며 정상 장에 비해 조금 부은 것을 확인할 수 있다. (i)는 정상 복수로 흰색을 띠고 있고, (j)는 맑은 복수로 정상 복수에 비해 더 탁해보이는 것을 확인할 수 있다. (k)는 출혈성 맑은 복수로 맑은 복수에서 부분적으로 출혈이 나타나 있으며, (l)은 출

혈성복수로 복수 전체에서 출혈이 발생하여 복수 색이 바뀐 것을 확인할 수 있다.



그림 5. 넙치 장기별 증상 이미지

(3). 내부 증상 분류 모델

넙치의 내부 장기 증상 분류를 위해 최신 이미지 분류 모델인 Coatnet, Nfnet, EfficientnetV2, ConvnextV2의 성능 비교 후 가장 높은 성능이 나온 모델을 사용하였다.

Coatnet은 기존 CNN과 Transformer의 장점을 결합하여 만든 모델로 일반적으로 사용되는 Depthwise Convolution과 Self-Attention을 단순 Relative Attention을 통해 결합이 가능하고, 수직으로 Convolutional Layer와 Attention Layer를 쌓으면 일반화와 용량의 성능이 높아지는 모델이다[8]. Nfnet은 배치 정규화를 사용하지 않고 AGC 기법을 사용하여 모델의 성능을 높이는 모델이다[9].

EfficientNetV2는 EfficientNet 모델을 개선하여 더 적은 파라미터를 사용해도 높은 성능이 나오도록 하였으며 빠른 학습에 초점을 맞춘 모델이다[10]. ConvnextV2는 기존 Convnext의 아키텍처에 완전한 컨볼루션 마스크 오토인코더 프레임워크를 추가하여 순수 컨볼루션 신경망의 성능을 크게 향상시킨 모델이다[11].

표 3은 클래스별 내부 증상 데이터의 구성으로 (a)는 간, (b)는 아가미, (c)는 아가미뚜껑, (d)는

표 3. 클래스별 내부 증상 데이터의 구성

		train	test	Total
간	울혈간	270	67	337
	정상간	1156	289	1445
아가미	아가미빈혈	606	150	756
	정상아가미	944	236	1180
아가미뚜껑	아가미뚜껑 출혈	34	9	43
	정상 아가미뚜껑	1078	269	1347
장	염증장	84	20	104
	정상장	1710	428	2138
복수	맑은복수	78	20	98
	정상복수	1400	351	1751
	출혈성 맑은 복수	50	14	64
	출혈성복수	100	24	124

장, (e)는 복수 데이터이다. 재라벨링을 거친 증상이 명확한 내부 데이터를 각 클래스마다 8 : 2의 비율로 나누어 학습 및 테스트 데이터를 구성하였다[12]. 표 4는 이미지 분류 모델을 사용하여 클래스별 장기 분류 성능을 나타낸 것이다. 실험 결과, 네 가지 이미지 분류 모델의 성능이 모든 클래스에 대해서 높게 나오며, ConvnextV2가 성능이 가장 높은 것이 확인되었으므로 연구에서 내부 증상을 분류하기 위한 모델로 사용하였다.

표 4. 클래스별 장기 분류 성능

클래스 모델	간	복수	아가미	아가미 뚜껑	장
Coatnet	0.98	0.97	0.98	1	0.99
Nfinet	0.97	0.96	0.97	0.99	0.97
EfficientNetV2	0.98	0.97	0.98	0.99	0.97
ConvNextV2	1	0.97	1	1	0.99

3. 질병 예측

(1). 질병 예측 실험 데이터

질병 실험 데이터의 경우, 질병은 에드워드, 비브리오, 스쿠티카, 연쇄구균, 활주세균, 바이러스성출혈성패혈증이며, 정상 상태를 포함해 총 7가지로 분류된다. 에드워드는 4장, 비브리오는 14장, 스쿠티카는 14장, 연쇄구균은 14장, 활주세균은 3장, 바이러스성출혈성패혈증은 6장, 정상은 55장을 실험에 사용하였다.

(2). 질병 예측 방법론

본 연구에서는 어류에 나타난 증상을 토대로 질병을 예측하기 위해서 세 가지 방법론을 제안한다. 첫 번째 방법론은 질병 증상의 가중치를 기반으로 질병을 예측하는 방법이다. 어류에 질병 증상이 나타났을 때 어류 질병 전문가가 이 증상이 어떤 질병에서 가장 큰 영향을 미치는지에 대한 가중치를 점수로 표현한 어류의 질병 증상 표를 바탕으로 질병을 예측한다. 표 5는 실제 어류 전문가가 제작한 질병 증상 가중치 표이며, (a)는 외부 증상, (b)는 내부 증상의 가중치를 나타낸 것이다. 어류 개체 외부 증상과 내부 장기

증상을 예측하고 가중치 표를 이용하여 그 질병별 증상 점수를 모두 더하여 가장 높은 점수를 가진 질병을 그 어류 개체의 질병으로 예측하는 방법이다.

구분	병원체	체표	복부	입	눈	지느러미	항문
1	에드워드병	발적(1), 근육출혈(2)	평만(5)	-	돌출(1)	출혈(1)	염증(3), 탈장(5)
2	비브리오병	발적(3), 궤양(4), 출혈(2)	평만(1)	주둥이 궤양(2)	돌출(1)	결손(2) 출혈(1), 문드러짐(2)	-
3	연쇄구균병	출혈(2), 근육출혈(3) 미방부 출혈(4)	평만(5)	-	돌출(2), 출혈(3) (적색원형)	출혈(1)	탈장(1)
4	활주세균병	변진궤양(5), 궤양(3)	-	주둥이 궤양(3)	-	문드러짐(2), 결손(3)	-
5	어림증	어림(5), 두부양물(5)	-	-	-	-	-
6	스쿠티카병	문드러짐(2) 궤양(4)	-	주둥이 궤양(3)	출혈(1) 돌출(1)	결손(3) 출혈(1), 기부궤양(4)	-
7	바이러스성 출혈성패혈증	발적(1), 근육출혈(3)	평만(5)	-	-	출혈(1)	탈장(5)

(a) 외부 증상 가중치 표

구분	병원체	아가미	복수	간	신장	비장	장	생식소
1	에드워드병	빈혈(1), 문드러짐(1) 아가미뚜껑 안쪽 출혈(1)	맑은 복수(1), 출혈성 복수(5)	섬유화(5), 간부종(3), 농양(5), 염증(1), 출혈(1)	비대(4), 농양(5), 결절(5)	비대(2)	염증(2), 출혈(2), 얇은장벽 (2)	출혈(1)
2	비브리오병	빈혈(2), 문드러짐(2)	맑은 복수(1), 출혈성 복수(1)	출혈(3), 염증(2)	비대(1)	비대(2)	출혈(2)	-
3	연쇄구균병	아가미뚜껑 안쪽 출혈(3), 빈혈(2), 문드러짐(3)	맑은 복수(1) 출혈성 복수(4)	출혈(4)	비대(1)	비대(2)	염증(2), 출혈(2)	출혈(2), 국소적 후화(4)
4	활주세균병	빈혈(2), 문드러짐(3) 결손(4)	-	-	-	-	-	-
5	어림증	-	-	흑적색(5) 출혈(2)	-	-	출혈(3) 염증(2)	-
6	스쿠티카병	빈혈(1), 문드러짐(3), 결손(3)	-	-	-	-	-	-
7	바이러스성 출혈성패혈증	빈혈(5)	맑은 복수(5) 출혈성 맑은복수(4)	출혈(4)	비대(4)	비대(2)	출혈(2)	-

(b) 내부 증상 가중치 표

표 5. 낚치의 외부 및 내부 증상 가중치 표

두 번째 방법론은 전체 어류 데이터의 질병과 증상 개수로 통계를 내어서 질병별로 가장 많이 발견되는 증상 순으로 높은 점수를 준다. 표 6은 전체 데이터에서 질병과 증상 개수를 통계로 낸 후 점수를 준 표를 나타낸 것이다. 실제 데이터의 통계에 기반해 점수를 매기는 방식으로, 질병 전문가가 만든 증상 가중치 표와 차이가 있다.

세 번째 방법론은 XGBoost, CatBoost 등의 앙상블 모델 및 딥러닝 모델을 이용해 증상 집합을 학습하여 질병을 예측하는 방법이다. 이 방법은 어류의 개체별로 증상 집합을 기록한 엑셀 파일로 학습한 후 정답 질병을 맞추도록 기계학습 모델을 사용하여 질병을 예측한다.

24년 개제 수 정리						
	스쿠티카	연쇄구균	바이러스	비브리오	에드워드	활주세균
개제수	167	179	72	179	56	39
체표궤양	8	2	1	6		2
항문탈장		2	3	6	6	
주둥이궤양	9	3		5	1	3
지느러미기부궤양	7			4		3
꼬리지느러미결손	4	1	1	5		1
아가미빈혈	10	9	9	9	7	7
아가미뚜껑출혈	2	5	2	1	3	
출혈성복수	5	7	5	6	8	4
맑은복수	1	3	6	2	4	3
출혈성맑은복수	1	4	4	3	2	
출혈간	3	8	7	8	5	5
염증장	6	6	8	7	6	6

표 6. 어류의 질병 및 증상 통계 후 증상별 점수

XGBoost는 Gradient Boosting 알고리즘을 분산 환경에서도 실행할 수 있도록 구현한 라이브러리로 Regression, Classification 문제를 모두 지원하며 성능과 자원 효율이 뛰어나다[13]. CatBoost는 범주형 데이터를 처리하는 새로운 방법을 제시하고 기존 Gradient Boosting 알고리즘을 조작하여 예측 시점에서 사용할 수 없는 데이터가 데이터셋에 포함되는 오류를 개선한다[14].

딥러닝 모델은 질병 예측을 위해 다층 퍼셉트론 기반 모델을 사용하였다. 모델은 입력층에서 시작하여 6개의 은닉층과 출력층으로 구성되며, 각 은닉층은 입력 차원이 1024에서 시작하여 층별로 512, 256, 128, 64로 점진적으로 줄어들어, 최종적으로 출력 클래스 수에 도달하도록 설계되었다. 모든 은닉층에는 비선형성을 추가하기 위해 ReLU 활성화 함수를 적용하였으며, 학습 안정성 및 일반화 성능을 높이기 위해 배치 정규화를 도입하였다. 과적합을 방지하기 위해 각 은닉층에 드롭아웃을 추가하였으며, 마지막 출력층은 소프트맥스 함수를 통해 각 클래스에 대한 확률 분포를 계산한다.

세 가지 방법론을 사용하여 질병을 예측한 결과를 Top-1, Top-2, Top-3 Accuracy로 나눠 비교 분석하여 전체적으로 가장 높은 Accuracy가 나온 방법론을 선정하였다. 증상 가중치 표와 데이터 통계 기반의 질병 예측 방법에서 Top-N

Accuracy는 개체별로 질병별 증상 점수를 계산한 뒤, 이 점수가 높은 순서대로 N개의 질병을 선택하여 전체 데이터에서 실제 질병이 Top-N 질병 내에 포함되는 비율로 정확도를 평가한다. 앙상블 및 딥러닝 모델 학습 기반의 질병 예측 방법에서는 Top-N Accuracy가 모델이 예측한 상위 N개의 클래스 중 실제 클래스가 포함되는지를 확인하는 방식으로 계산된다. 수식 (1)은 Top-N Accuracy를 산출도로 나타낸 것이다.

$$\text{Top-N Accuracy} = \frac{\text{Top-N 내 정확하게 예측한 데이터 수}}{\text{전체 데이터 수}} \quad (1)$$

Top-1은 첫 번째 결과와 실제 어류 질병이 일치하는 비율, Top-2는 상위 두 가지 결과 중 하나가 실제 질병과 일치하는 비율, Top-3는 상위 세 가지 결과 중 하나가 실제 질병과 일치하는 비율이다.

III. 실험 결과

1. 질병 예측 방법론 성능 결과

어류의 질병을 예측하는 가장 적합한 방법론을 선정하기 위해 앞서 소개한 세 가지 질병 예측 방법에 대해, 질병 예측에 사용되는 증상을 라벨링된 정답을 그대로 사용하여 질병을 예측한 경우와 외부 증상 검출 모델과 내부 분류 모델을 사용하여 증상을 예측하고 질병을 예측한 경우의 Top-1, Top-2, Top-3 Accuracy 결과를 확인하였다.

표 7은 어류 질병 전문가가 만든 질병 증상 가중치 표를 바탕으로 질병을 예측한 후 예측 결과를 Top-1, Top-2, Top-3 Accuracy로 나누어서 나타낸 것이다. 내부 및 외부 증상 예측을 통해 질병을 예측한 결과, Top-1 Accuracy는 42.73%의 성능이 나온 것을 확인할 수 있다. 증상이 라벨링된 정답을 그대로 사용하여 질병을 예측했다 해도 72.73%의 성능이 나왔다. 라벨링된 정답을 사용해도 성능이 높게 나오지 않은 이유는 앞서 언급한 것처럼 질병별 증상이 다르고 서로 다른 질병에서 같은 증상들이 발현되며, 어류 질병에 따른 명확한 증상이 정해져 있지 않기 때문으로 판단된다.

표 7. 증상 가중치 표 기반 질병 예측 성능

	Accuracy	가중치 기반
증상 모델예측	Top-1 Accuracy	42.73
	Top-2 Accuracy	54.55
	Top-3 Accuracy	56.94
증상 정답처리	Top-1 Accuracy	72.73
	Top-2 Accuracy	86.36
	Top-3 Accuracy	88.18

표 8은 전체 어류 데이터의 질병과 증상 개체수의 통계를 바탕으로 질병을 예측하고 예측 결과를 나타낸 것이다. 증상을 예측해서 질병을 예측한 결과 Top-1 Accuracy는 50.91%의 성능이 나온 것을 확인할 수 있다.

표 8. 데이터 통계 기반 질병 예측 성능

Accuracy	통계 기반
Top-1 Accuracy	50.91
Top-2 Accuracy	60.00
Top-3 Accuracy	65.45

표 9는 앙상블과 딥러닝 모델로 학습한 후 가장 높은 정확도가 나온 모델로 질병을 예측하고 예측 결과를 나타낸 것이다. 앙상블과 딥러닝 모델을 학습한 결과 가장 높은 성능이 나온 경우는 딥러닝 모델로 학습했을 때이며, 증상을 예측한 결과 Top-1 Accuracy는 58.18%의 성능이 나온 것을 확인할 수 있다.

표 9. 앙상블 및 딥러닝 모델 기반 질병 예측 성능

	Top1	Top2	Top3
XGBoost	49.09	60.00	64.55
CatBoost	45.45	61.82	77.27
딥러닝	58.18	73.64	83.64

그림 6은 질병 증상 표, 통계, 딥러닝 모델 학습을 기반으로 각각 수행된 질병 예측 결과를 그래프로 나타낸 것이다. 그래프를 보면 질병 증상 표와 통계를 기반으로 질병을 예측했을 때는 비슷한 성능이지만 딥러닝 학습을 통해 질병을 예측할 때 가장 높은 성능이 확인되었다.

어류 질병은 모든 증상이 나타나지 않을 수 있

질병 예측 방법론 성능 비교(증상 모델예측)

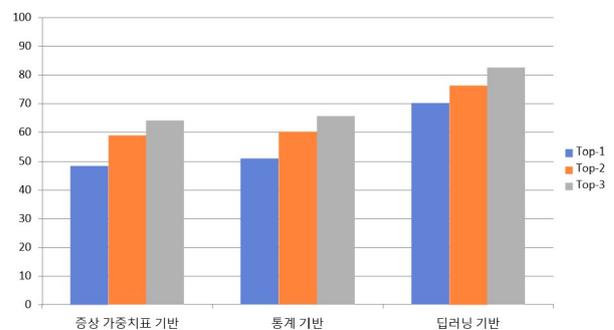


그림 6. 세 가지 방법론 결과 비교 그래프

으며, 서로 다른 질병이 동일한 증상 집합이 발현될 수 있기 때문에 질병 증상 표나 통계를 기반으로 하는 것보다는 질병 증상 집합의 딥러닝 학습을 통해서 질병을 예측하는 것이 가장 높은 결과를 보여주는 것으로 판단된다.

IV. 결 론

본 연구에서는 어류의 질병 증상으로 어떻게 어류의 질병을 정확하게 예측할 것인지에 대한 방법론을 제안하였다. 어류의 외부 증상 검출 및 내부 증상 분류 결과를 바탕으로 하여 어류 질병 전문가가 제작한 증상 가중치 표, 증상 데이터 통계, 증상 집합을 학습한 앙상블 모델 및 딥러닝 모델 기반 방법으로 각각 질병을 예측한 뒤 정확도를 비교 분석하였다. 실험 결과, 증상 가중치 표 기반과 데이터 통계 기반 방법은 각각

48.45%와 50.91%의 예측 성능을 보인 반면, 질병별 증상 집합을 활용한 딥러닝 모델 학습 기반 방법으로 예측하는 방식은 70.10%의 성능으로 가장 높은 정확도를 보였다. 이는 어류 질병은 모든 질병 증상이 발현되는 것이 아니라 몇 가지만 발현될 가능성이 있으며, 서로 다른 질병에서 같은 증상들이 발현될 수 있기 때문인 것으로 판단된다.

향후 연구에서는 현재 딥러닝 모델을 사용하여 나온 질병 예측 정확도를 기준으로 어떻게 하면 질병을 더 잘 예측할 수 있을 지에 대해 연구를 진행할 예정이다.

REFERENCES

- [1] Korea Institute of Science and Technology Information. (2018). Select POR Search Report. Korea Institute of Science and Technology Information. <https://scienceon.kisti.re.kr/srch/selectPORSrchReport.do?cn=TRKO201800014913> (accessed Feb., 1, 2024).
- [2] National Institute of Fisheries Science, "Fish Disease Incidence Statistics," Available: <https://www.nifs.go.kr/fishguard/disease01.do>, 2019, (accessed Feb., 1, 2024).
- [3] National Institute of Fisheries Science, "Fish Disease Control," Available: <https://www.nifs.go.kr/fishguard/disease02>, 2019, (accessed Feb., 1, 2024).
- [4] 임창현, 신영학. "넙치의 내부 장기 이미지를 이용한 질병 증상 분류 및 검출 비교 연구," *한국지능시스템학회 논문지*, 제34권, 제1호, 51-57쪽, 2024년 2월
- [5] Rachman F., M. N. S. Akbar, and E. Putera. "Fish disease detection of epizootic ulcerative syndrome using deep learning image processing technique." *Proceedings International Conference on Fisheries and Aquaculture*. vol. 8. no. 1, pp. 23-34, 2023.
- [6] 손현승, 임한규, 최한석, "딥러닝 기술을 이용한 넙치의 질병 예측 연구," *스마트미디어저널*, 제11권, 제4호, 62-68쪽, 2022년 5월
- [7] Ultralytics, "Ultralytics YOLOv8 Documentation," Available: <https://docs.ultralytics.com/ko>, 2023, (accessed Feb., 1, 2024).
- [8] Zihang Dai, Hanxiao Liu, Quoc V Le, Mingxing Tan. "Coatnet: Marrying convolution and attention for all data sizes," *Advances in neural information processing systems* 34, pp. 3965-3977, 2021.
- [9] Andy Brock, Soham De, Samuel L Smith, Karen Simonyan. "High-performance large-scale image recognition without normalization," *International conference on machine learning*, PMLR 139, pp 1059-1071, 2021.
- [10] Tan, Mingxing, and Quoc Le. "Efficientnetv2: Smaller models and faster training," *International conference on machine learning*, PMLR, pp. 16133-16142, 2021.
- [11] Sanghyun Woo, Shoubhik Debnath, Ronghang Hu, Xinlei Chen, Zhuang Liu, In So Kweon, Saining Xie. "Convnextv2: Co-designing and scaling convnets with masked autoencoders," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 16133-16142, 2023.
- [12] 임창현, 신영학, "머신러닝 특징 공간 분석 기반의 효율적인 리 라벨링을 통한 넙치의 장기 증상 분류 성능 향상 연구," *한국지능시스템학회 논문지*, 제34권, 제3호, 267-276쪽, 2024년 6월
- [13] Chen, Tianqi, and Carlos Guestrin. "Xgboost: A scalable tree boosting system," *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785-794, 2016.
- [14] Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, Andrey Gulin. "CatBoost: unbiased boosting with categorical features," *Advances in neural information processing systems* 31, 2018.

저 자 소 개



신유식(준회원)

2025년 국립목포대학교 컴퓨터공학과
공학사 예정

<주관심분야 : Artificial Intelligence, Deep Learning>
ORCID Number : 0009-0005-2493-1656
E-mail : dbtldrld@gmail.com



임창현(준회원)

2025년 국립목포대학교 컴퓨터공학과
공학사 예정

<주관심분야 : Artificial Intelligence, Deep Learning>
ORCID Number : 0009-0001-9893-6207
E-mail : uckdekf@gmail.com



신영학(정회원)

2009년:광운대학교 전자통신공학과
공학사

2011년:광주과학기술원 정보통신공학부 공
학석사

2016년:광주과학기술원 정보통신공학부 공
학박사

2018년~2020년:LG CNS AI빅데이터연구소
책임연구원

2020년~현재:목포대학교 컴퓨터공학과 교수

<주관심분야 : Artificial Intelligence, Deep Learning,
Medical Imaging>
ORCID Number : 0000-0003-2213-5232
E-mail : younghak@mnu.ac.kr