

넙치 양식장 수질 데이터를 활용한 딥러닝 기반 폐사량 예측 연구

(Deep Learning-Based Prediction of Mortality in Olive Flounder (*Paralichthys olivaceus*)
Aquaculture Using Water Quality Data)

손명준*, 신창선**

(MyeongJun Son, ChangSun Shin)

요약

국내 수산업에서 양식 산업은 안정적인 수산물 공급을 가능하게 한다는 점에서 매우 중요한 역할을 맡고 있다. 하지만 최근 기후 변화에 따른 해수 고수온 현상과 수질 악화 등 환경적 요인으로 인한 집단 폐사 문제가 빈번히 발생하고 있으며, 이러한 집단 폐사 문제는 양식 어가의 생산성과 수익성 저하로 이어지고 있다. 특히 국내에서 가장 널리 양식되는 넙치는 수온과 수질 변화에 민감한 특성을 지녀, 여름철 반복되는 대규모 폐사로 인한 피해가 심각한 실정이다. 피해 저감을 위해 기존에는 영상 기반의 예측 기술이 일부 도입되었으나, 해당 방식은 질병 발생 이후의 상태를 인식하기에 사전 대응이 어렵고 장비 설치 및 환경 제약 등의 한계가 존재한다. 이에 본 연구에서는 수온, 용존산소(DO), pH, 아질산성질소 등 수질 데이터를 기반으로, 시계열 딥러닝 모델(LSTM, GRU, TCN)을 적용하여 넙치의 폐사량을 예측하고, 각 모델의 성능을 비교 분석하였다. 모델 학습은 7일 단위의 입력과 출력을 기반으로 구성하였으며, 성능 지표로는 RMSE, MmRMSE, R^2 , adj- R^2 등을 활용하여 정량적 평가를 수행하였다. 분석 결과, TCN 모델이 RMSE 35.868, MmRMSE 0.106, R^2 0.753, adj- R^2 0.750의 예측 성능을 보여 다른 모델과 대비하여 가장 우수한 결과를 나타냈다. 특히 고폐사 구간에서도 실제 폐사량의 추세를 효과적으로 반영하는 특성을 보였다.

■ 중심어 : 넙치 ; 수질 데이터 ; 딥러닝 ; 시계열 예측 ; 폐사량 예측

Abstract

In Korea's aquaculture industry, fish farming plays a key role in ensuring a stable supply of aquatic products. However, environmental changes such as rising sea temperatures and water quality deterioration have led to frequent mass mortality events, reducing productivity and profitability. Olive flounder, the most widely farmed species in Korea, is highly sensitive to temperature and water quality changes, making it vulnerable to large-scale mortality during summer. Although vision-based prediction technologies have been introduced to mitigate such damage, they are limited to post-disease detection and constrained by equipment requirements and environmental factors. This study applied time-series deep learning models (LSTM, GRU, TCN) using water quality data—temperature, dissolved oxygen, pH, and nitrite nitrogen—to predict mortality. The models were trained using a 7-day input/output structure and evaluated with RMSE, MmRMSE, R^2 , and adjusted R^2 . The TCN model achieved the best performance (RMSE 35.868, MmRMSE 0.106, R^2 0.753, adj- R^2 0.750), effectively capturing trends during high-mortality periods.

■ keywords : Olive flounder ; Smart aquaculture ; Deep learning ; Mortality prediction ; Water quality data

I. 서론

국내 수산업에서 양식 산업은 안정적인 수산물

공급을 가능하게 한다는 점에서 매우 중요한 역할을 맡고 있다. 하지만 최근 기후 변화에 따른 여름철 해수 고수온 현상과 수질 악화와 같은 환경적인 요인으로 인한 집단 폐사가 자주 발생하

* 정회원, 국립순천대학교 스마트농업전공 석사과정

** 종신회원, 국립순천대학교 인공지능공학부 교수

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 지역지능화혁신인재양성사업임(IITP-2025-FS-2020-II201489)

접수일자 : 2025년 07월 14일

수정일자 : 2025년 08월 06일

게재확정일 : 2025년 08월 19일

교신저자 : 신창선 e-mail : csshin@sncu.ac.kr

고 있으며, 이러한 현상은 양식 어가의 생산성과 수익성 저하로 직결되고 있다[1,3,4].

그중에서도 넙치는 국내에서 가장 널리 양식되는 어종 중 하나로, 2024년 기준 생산량은 전체 82천 톤 중 40천 톤으로 48.78%의 높은 비중을 차지하고 있으며, 생산금액은 전체 12,112억 원 중 6,842억 원으로 56.49%의 비중을 차지하고 있다[2]. 이렇게 국내 양식산업에 있어 넙치 양식은 매우 중요한 역할을 담당한다. 하지만 넙치는 수온과 수질 변화에 민감한 어종으로 환경 변화에 따른 폐사 발생률이 높아, 매년 여름철마다 반복되는 양식장 대규모 폐사는 양식 어가뿐만 아니라 양식 산업 전반에 걸친 생산성 저하 및 수익성 저하로 직결되고 있어 반드시 해결해야 하는 과제로 부각되고 있다[3,5].

수산업 분야에서는 이러한 환경 요인에 의한 폐사 문제를 해결하기 위해 영상 기반의 모니터링 및 질병 예측 기술에 대한 연구가 주로 이루어지고 있다. 이러한 기술은 주로 질병 발생 이후의 어류 상태를 인식하는 방법으로, 사후 대응의 성격이 강하고, 분석 영상을 촬영하기 위해 별도의 장비 설치가 필요하다는 점에서 환경·경제적 제약이 크다는 단점이 존재한다[3,5].

반면, 수질 데이터는 센서를 통해 비교적 정량적이고 영상 장비에 비해 저비용으로 데이터 수집이 가능하며, 어류의 생존에 직접적인 영향을 미치는 중요한 환경 변수로 폐사량과의 밀접한 관계를 가진 지표로 활용될 수 있다[4,6,7]. 그럼에도 불구하고 현재까지는 수질 데이터를 기반으로 폐사량을 예측하고자 하는 연구는 제한적이며, 특히 넙치를 대상으로 수질 데이터를 활용한 딥러닝 기반 예측 모델을 비교 분석한 사례는 드문 현실이다[3,4,7,8,15].

넙치의 폐사량 예측을 위해 활용될 수 있는 딥러닝 모델로, LSTM(Long Short-Term Memory) 모델은 순환 신경망(RNN) 모델의 장기 의존성 문제를 개선한 구조이다. 순환 신경망 모델과 같이 체인 구조를 가지지만 하나의 셀 안

에 여러 개의 게이트(입력, 망각, 출력 게이트)를 포함하고 있다. 저장된 정보에 현재의 정보를 업데이트한 후 셀 상태에 따라 활성 함수에 의해 출력을 결정한다[9]. GRU(Gated Recurrent Unit) 모델은 LSTM 모델과 유사한 구조의 시계열 모델이면서 LSTM보다 구조가 간단하고 학습 효율이 높다는 특징이 있다. 업데이트 및 리셋 게이트를 통해 정보의 흐름을 조절하며 상대적으로 적은 학습 파라미터를 통해 효과적인 결과를 도출할 수 있다[10]. TCN(Temporal Convolutional Network) 모델은 합성곱 기반의 시계열 모델로, CNN 모델의 일종이다. 순환 구조 없이도 긴 시계열 기반의 정보를 학습할 수 있으며 과거 시점의 정보만을 기반으로 예측을 수행한다. 병렬 연산이 가능해 보다 빠르고 안정적인 학습이 가능하다[11].

본 논문은 양식 환경에서 접근성이 높은 수질 데이터를 기반으로, 시계열 예측 모델(LSTM, GRU, TCN)을 적용하여 넙치의 폐사량을 예측하고 모델의 성능을 비교함으로써 가장 적합한 모델을 도출하고자 한다. 이를 통해 향후 양식 어가에서 실시간 수질 데이터를 기반으로 한 선제적 대응 체계를 구축하고, 스마트 양식 시스템의 지능화 및 생산성 향상에 기여할 수 있는 기초 자료를 제공하는데 목적이 있다.

II. 본 론

1. 데이터 수집 및 분석 환경

가. 연구 대상 및 데이터 구성

(1) 연구 대상

본 연구는 전라남도 완도군에 위치한 양식장에서 사육 중인 넙치를 대상으로 수행되었다. 해당 양식장은 유수식 환경으로 운영되고 있으며, 수

조별로 수질 데이터와 폐사량을 포함한 생육 데이터를 정기적으로 수집하고 있다.

데이터 수집 기간은 2023년 03월부터 2024년 07월까지로, 총 2개 수조에서 수집한 데이터를 활용하였다.

(2) 데이터 구성 및 기준

표 1. 완도 넙치 양식장 수집 데이터 구성

구분	항목명	단위	범위
사육 정보	수집일	yyyy-mm-dd	2023.01 ~ 2024.07
	입식량	마리	1,800 ~ 7,200
	폐사량	마리	0 ~ 31
	현재 미수	마리	1,800 ~ 3,500
수질 정보	수온	℃	20.9 ~ 29.3
	용존산소(DO)	mg/L	-
	pH	pH	-
	ORP	mV	-
	암모니아성 질소	mg/L	0.02 ~ 0.20
	아질산성 질소	mg/L	0.004 ~ 0.018
성장 정보	BQ_VALUE	-	934 ~ 5,216
	무게	g	70 ~ 530
	전장	cm	10 ~ 33
기타	체고	cm	9 ~ 15
	급이 유무	여부	0 or 1

* 수집 기간 확보된 총 데이터 건수 : 556일(건)

표 1과 같이 수집된 데이터는 총 556일(건)의 시계열 데이터를 수집하였으며, 이중 결측된 10일(건)간의 데이터를 선형보간의 방법을 통해 사용하였다. 해당 데이터는 사육 정보, 수질 정보, 성장 정보로 구분된다.

사육 정보에는 수집일, 입식량, 폐사량, 생존 개체 수 등이 포함되어 있으며, 이는 양식장 환경에서 사육하는 개체의 생존 변화를 반영한다.

수질 정보는 수온, 용존산소(DO), pH, ORP, 암모니아성 질소, 아질산성 질소, BQ_VALUE등을 포함하며, 양식장 수조 내 환경 요인을 정량적으로 파악하기 위한 지표들로 구성되어 있다.

성장 정보에는 넙치 개체의 무게, 전장, 체고가

포함되어 있으며, 이는 양식 개체의 생육 정보 등을 파악하기 위해 사용된다.

이 외에도 급이 유무, 유입수등 여러 항목이 수집되었으나, 본 연구에서는 예측 모델 성능에 주요한 영향을 미치는 변수 중심으로 정리하였다.

2. 수질 지표와 폐사량 간 상관관계 분석

가. 분석 목적

(1) 분석 개요

넙치 양식을 포함한 대부분의 양식 환경에서 수질은 어류의 생존과 관련하여 매우 중요한 요인으로 작용한다. 따라서 이러한 넙치 양식장 내 폐사량과 수질·생육 지표 간의 관계를 정량적으로 분석하기 위해, 양식장에서 수집된 데이터를 기반으로 피어슨 상관관계수(Pearson Correlation Coefficient)를 활용한 상관관계 분석을 수행하였다.

피어슨 상관관계수는 모집단 내에서 두 변수 간 선형적 연관성의 정도와 방향을 나타내는 지표로써, 수치에 따른 선형 관계 정도를 나타낸다. 상관관계수가 +1에 가까울수록 강한 양의 선형 관계, -1에 가까울수록 강한 음의 선형 관계를 의미하며, 0에 가까울수록 두 지표간 연관성이 없음을 의미한다[12].

이러한 상관관계수를 표현하는 공식 중에서 일반적으로 가장 널리 사용되는 표현 방법은 공분산(covariance)을 기반으로 하며 해당 피어슨 상관관계수 식은 수식(1)과 같이 정의할 수 있다[12].

$$p_{xy} = \frac{E(XY) - E(X)E(Y)}{\sigma_X \sigma_Y} \quad (1)$$

$E(XY)$ 는 두 데이터 세트의 해당 위치 원소 곱의 평균, $E(X)$, $E(Y)$ 는 각각 X , Y 변수의 평균값을 의미한다. σ_X , σ_Y 는 X , Y 의 표준편차를 나

타낸다.

본 연구에서는 이러한 피어슨 상관관계수 분석 방법을 통해 수질 데이터와 폐사량 간의 상관관계를 분석하고, 유의미한 변수를 선별하여 딥러닝 예측 모델의 입력 변수로 활용하였다.

(2) 해석 기준

표 2. 피어슨 상관관계수 해석 기준

상관계수	선형 관계 해석
0.7 ~ 1.0	매우 강한 양의 상관관계
0.3 ~ 0.7	강한 양의 상관관계
0.1 ~ 0.3	약한 양의 상관관계
-0.1 ~ 0.1	거의 무시될 수 있는 상관관계
-0.3 ~ -0.1	약한 음의 상관관계
-0.7 ~ -0.3	강한 음의 상관관계
-1.0 ~ -0.7	매우 강한 음의 상관관계

표 2와 같이 피어슨 상관관계수의 값을 기준으로 변수 간 상관관계의 정도를 해석하였다.

본 연구에서는 폐사량 예측 모델 설계에 영향을 미치는 요인만을 선별하기 위해 상관계수의 절댓값이 $|r| \geq 0.2$ 이상인 경우에만 유의미한 상관관계를 가진다고 판단하였다[7].

나. 분석 방법 및 결과

(1) 분석 방법

본 연구에서는 넙치 양식장 내 폐사량과 수질 지표 간의 상관관계를 분석하기 위해 다음과 같은 절차로 진행하였다.

첫째, 분석 대상 넙치 양식장에서 수집된 수질·생육 데이터를 일자 기준으로 통합하여 분석 가능한 형태로 재구성하였다. 폐사량은 일일 기준으로 재구성하였으며, 수질 항목으로는 수온, 용존산소(DO), pH, ORP, 암모니아성 질소, 아질산성 질소, BQ_VALUE등을 포함하였다.

둘째, 분석 대상 기간 동안 각 변수의 결측치를 확인하고, 일부 결측 항목에 대해 평균 대체 혹은 제거 등의 정제를 통해 상관관계 분석이 가능

한 수준으로 데이터를 정비하였다.

셋째, 정제된 데이터를 기반으로 넙치 폐사량과 각 수질 지표 간의 상관관계를 파악하기 위해 피어슨 상관관계수를 산출하였다. 분석 결과는 수질 지표별로 폐사량과의 상관 정도와 방향성을 나타내며, 상관계수의 절댓값이 0.2 이상인 경우를 유의미한 상관관계로 보고 주요 예측 변수로 선정하였다.

(2) 분석 결과

수집 수질 데이터와 생육 데이터를 기반으로, 폐사량과 수질 지표 간의 피어슨 상관관계수를 분석한 결과, 폐사량과 유의미한 상관관계를 가지는 주요 수질 지표들이 다수 확인되었다. 상관계수 분석은 각 변수와 폐사량간의 Pearson 계수를 산출한 뒤, 절댓값이 $|r| \geq 0.2$ 이상인 경우를 유의미한 상관관계로 간주하여 시각화하였다.

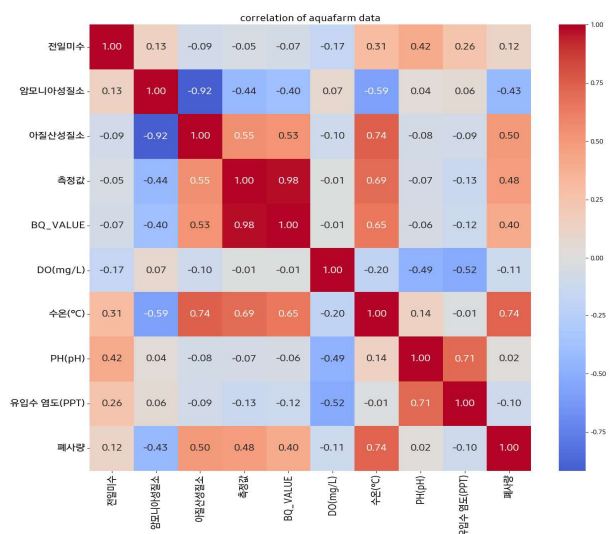


그림 1. 폐사량과 주요 수질 지표간 상관계수 히트맵

그림 1은 넙치 양식장에서 수집된 수질 데이터와 폐사량간의 상관관계를 히트맵 형태로 시각화한 것이다. 폐사량과 수온은 0.74로 높은 양의 상관관계를 보인다. 이는 수온이 상승할수록 넙치의 폐사량이 증가하는 경향이 있음을 보여준다. 또한 아질산성질소는 폐사량과 0.50의 상관

계수를 보이고 있으며, 이를 통해 직접적인 연관이 있는 지표로 볼 수 있다. 측정값(세균값)과 BQ_VALUE도 각각 0.48, 0.40의 높은 양의 상관관계를 보여 납치의 폐사량에 영향을 미치는 요소로 볼 수 있다.

암모니아성질소와 같은 경우 폐사량과 -0.43의 상관관계를 보이고 있어 직접적인 연관을 미치는 지표로 볼 수 있지만, 일반적으로 암모니아성질소의 농도가 높을수록 납치의 폐사량의 증가한다는 일반적인 상식과 상반되는 결과를 보이고 있다. 이는 해당 수집 데이터 결측치 등의 이유로 데이터가 정상적으로 수집되지 않았다고 볼 수 있어 폐사량 예측 모델의 주요 지표로 선정하지 않았다[13].

위 상관관계 분석을 토대로 본 연구에서는 수온, 아질산성질소, 측정값(세균값), BQ_VALUE 등의 수질 데이터를 기반으로 폐사량 예측 모델의 주요 입력 변수로 활용하였다.

3. 딥러닝 기반 예측 모델 학습

가. 모델 입력 구성 및 전처리 과정

(1) 모델 입력 변수 선정 기준

앞선 상관관계 분석 결과를 바탕으로 납치 폐사량 예측을 위한 딥러닝 모델 학습에 사용하기 위한 주요 입력 변수로, 수온, 아질산성질소, 측정값(세균값), BQ_VALUE 등을 선정하였다.

해당 변수들은 상관관계수가 0.4 이상인 변수들로 폐사량과 직접적인 영향이 있다고 판단된 변수들을 모델 학습에 사용하였다.

(2) 정규화 및 시계열 데이터 구성 방식

모델 입력 변수로 선정된 수질 데이터를 기반으로, 모델 학습을 진행하기 위한 데이터셋을 구축하기에 앞서 데이터셋에 대한 정규화 과정을

수행하였다. 모든 입력 변수는 MinMaxScaler를 이용하여 0~1의 범위로 스케일링 하였으며, 정규화 기준은 학습용 데이터셋을 기반으로 설정하여 이후 검증 및 테스트 데이터셋에도 동일한 기준을 적용하였다. 이를 통해 전체 데이터에 걸쳐 스케일링의 일관성을 유지하고, 모델의 수렴 안정성을 확보하였다.

데이터셋은 전체 시계열 데이터를 기준으로 학습(Train), 검증(Validation), 테스트(Test) 세 구간으로 분할하였다. 분할 비율과 같은 경우 전체 데이터의 약 60%는 학습용, 20%는 검증용, 20%는 테스트용으로 구성하였다. 이러한 분할 방식은 데이터 누수(Leakage)를 방지하고, 실제 운영 환경에서의 예측 성능을 검증할 수 있도록 시계열 특성을 반영하여 설계하였다.

(3) 전처리 및 학습 데이터셋 구축

본 연구에서는 시계열 기반 딥러닝 모델의 학습을 위해서 슬라이딩 윈도우(Sliding Window) 방식을 활용하여 입/출력 구조의 시계열 학습 데이터를 구성하였다. 슬라이딩 윈도우는 일정 길이의 시계열 구간을 순차 이동시키며 입력 시점과 예측 대상 구간을 정의하는 방식으로, 시계열 예측에서 많이 쓰이는 전처리 기법이다[14,15].

본 모델에서는 납치 양식 현장에서의 어류 생육 관리 주기인 주 단위에 맞춰 7일간의 수질 데이터를 입력 시퀀스로 설정하였고, 그 후 7일간의 폐사량 데이터를 출력 시퀀스로 구성하였다. 이처럼 입/출력 시퀀스 길이를 각각 7일로 설정함으로써, 과거 7일간의 수질 데이터를 기반으로 향후 7일간의 폐사량을 예측할 수 있도록 설계하였다. 이러한 방식을 통해 데이터를 한 시점(7일)씩 이동시키며 학습용 시계열 샘플을 생성함으로써, 시계열 정보의 시간 흐름을 보존하고 미래 예측이라는 딥러닝 모델 목적에 부합하는 데이터셋을 구축하였다.

위와 같은 전처리 방식을 통해 총 334건의 시

계열 학습 데이터가 생성되었으며, 본 연구의 LSTM, GRU, TCN 모델은 모두 동일한 구조와 동일한 양을 가진 입력 데이터를 기반으로 학습되었다.

나. 모델 구성 및 학습 환경

본 연구에서는 시계열 수질 데이터를 기반으로 담체의 폐사량을 예측하기 위해 다음 세가지 딥러닝 모델을 설계하고 비교하였다. 사용된 딥러닝 모델은 LSTM(Long Short-Term Memory), GRU(Gated Recurrent Unit), TCN(Temporal Convolutional Network)이며, 해당 모델들은 모두 시간의 흐름에 따른 데이터의 패턴을 학습하고 단기 및 장기 의존성을 포착하는데 강점을 가진 구조를 가지고 있다.

각 모델은 동일한 입력 및 출력 조건에서 학습을 진행하였으며, 7일간의 수질 데이터를 입력받아 이후 7일간의 폐사량을 예측하도록 구성하였다. 모델 학습은 Python 기반의 TensorFlow 환경에서 수행하였으며, 학습 데이터셋은 정규화 및 시계열 분할을 통해 전처리된 데이터를 기반으로 구성되었다.

표 3. 모델 학습을 위한 공통 하이퍼파라미터 설정

항목	설정 값
Optimizer	Adam
초기학습률 (Learning Rate)	0.001
손실 함수 (Loss Function)	Mean Squard Error(MSE)
배치 크기(Batch Size)	32
최대 학습 Epoch	512
조기 종료 조건	Validation Loss 기준, 128 Epoch 무개선 시 종료
Best Model 저장 기준	최소 Validation Loss
정규화 방식	MinMaxScaler(0~1)
시계열 입력 구간 (입력 길이)	7일 (7개 시점의 수질 데이터)
시계열 출력 구간 (예측 길이)	7일 (7개 시점의 폐사량 데이터)

모든 딥러닝 모델은 동일한 학습 조건 하에서

성능 비교를 위해 표 3과 같은 공통 하이퍼파라미터 수치를 적용하였다.

Optimizer는 널리 사용되는 Adam을 사용하였으며, 학습률, 손실 함수등은 Adma 사용시 일반적으로 안정적인 성능을 보이는 설정값을 적용하였다. 배치 크기와 Epoch와 같은 경우 학습 환경(GPU 자원, 학습 시간 등)을 고려하여 구성하였다.

다. 모델 예측 결과 및 성능 비교

본 연구에서는 LSTM, GRU, TCN 모델간의 성능 비교를 위해 Hidden Dim, Layer 수, Bidirectional 여부를 조합하여 총 8가지 모델을 구성하였으며, 실험은 동일 입/출력 시계열 구간(7일)과 동일 하이퍼파라미터 설정에 따라 진행하였다. 평가 지표는 RMSE, MmRMSE, R^2 (결정계수), adj- R^2 (수정 결정계수)를 사용하였다.

(1) LSTM 모델 예측 성능 분석

표 4. LSTM 모델 하이퍼파라미터 조합별 성능 비교

Hidden Dim	Layer	Bidrectional	RMSE	Mm RMSE	R^2	adj- R^2
16	1	No	50.317	0.149	0.514	0.509
16	1	YES	56.653	0.168	0.384	0.377
16	2	No	62.944	0.187	0.239	0.231
16	2	YES	63.429	0.188	0.228	0.220
8	1	No	60.731	0.180	0.292	0.284
8	1	YES	64.214	0.190	0.208	0.200
8	2	No	67.844	0.201	0.116	0.107
8	2	YES	61.758	0.183	0.268	0.260

LSTM 모델의 하이퍼파라미터 조합별 성능 비교 결과는 표 4와 같으며 하이퍼파라미터 모델 중 Hidden Dim = 16, Layer = 1, Bidirectional = False의 조건일 때, 가장 낮은 RMSE(50.317)을 기록하였으며, R^2 또한 0.514로 가장 우수한 예측 성능을 나타냈다. 반면 Layer 수가 2이거나 Bidirectional 구조를 적용한 경우에는 오히려 성능이 저하되는 현상이 발생했다.

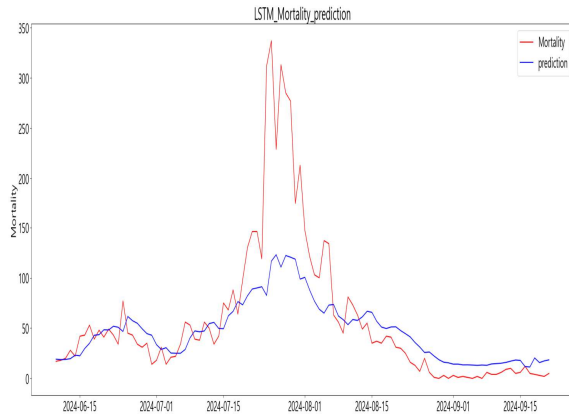


그림 2. LSTM 모델 최적 하이퍼파라미터 조합 시각화

그림 2는 LSTM 모델의 예측 결과를 하이퍼파라미터별로 시각화한 그래프이다. 예측 결과 LSTM 모델은 전반적으로 폐사량이 낮은 구간에서는 일정한 예측력을 보이고 있으나, 폐사량이 급격하게 증가할수록 예측의 정확도가 상대적으로 낮은 특징이 나타난다.

(2) GRU 모델 예측 성능 분석

표 5 GRU 모델 하이퍼파라미터 조합별 성능 비교

Hidden Dim	Layer	Bidirectional	RMSE	Mm RMSE	R ²	adj-R ²
16	1	No	51.711	0.153	0.487	0.481
16	1	YES	47.427	0.141	0.569	0.564
16	2	No	55.274	0.164	0.413	0.407
16	2	YES	70.775	0.210	0.038	0.282
8	1	No	43.094	0.128	0.643	0.640
8	1	YES	50.716	0.150	0.506	0.501
8	2	No	61.073	0.181	0.284	0.277
8	2	YES	58.082	0.172	0.152	0.346

GRU 모델의 하이퍼파라미터 조합별 성능 비교 결과는 표 5와 같으며 하이퍼파라미터별 모델 중 Hidden Dim = 8, Layer = 1, Bidirectional = False일 때 가장 낮은 RMSE(43.09)를 기록하였으며, R² 또한 0.643로 가장 우수한 예측 성능을 나타냈다. 특히 GRU 모델은 LSTM 모델에 비해 폐사량이 급격하게 증가하는 구간에서도 일정 수준의 추세를 예측하는걸 확인할 수 있다.

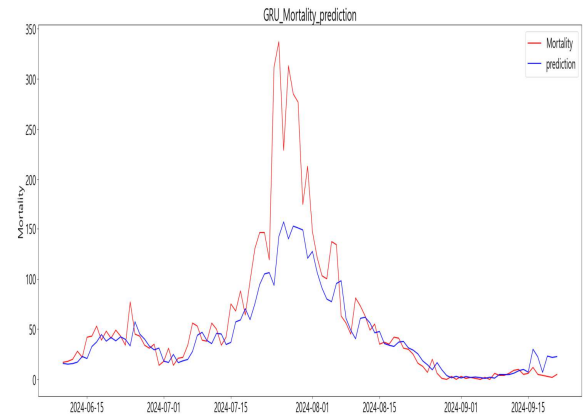


그림 3 GRU 모델 최적 하이퍼파라미터 조합 시각화

그림 3은 GRU 모델의 예측 결과를 하이퍼파라미터별로 시각화한 그래프이다. 예측 결과 GRU 모델은 전체적으로 실제 폐사량의 추세를 일정 부분 반영하고 있으며, 고폐사 구간에서의 예측 또한 LSTM 모델에 비해 더 나은 예측 정확도와 추세 반영을 보이는걸 알 수 있다.

(3) TCN 모델 예측 성능 분석

표 6 TCN 모델 하이퍼파라미터 조합별 성능 비교

Hidden Dim	Layer	Bidirectional	RMSE	Mm RMSE	R ²	adj-R ²
16	1	No	38.881	0.115	0.710	0.707
16	1	YES	49.564	0.147	0.528	0.523
16	2	No	56.671	0.168	0.383	0.377
16	2	YES	37.264	0.110	0.733	0.731
8	1	No	48.392	0.144	0.550	0.546
8	1	YES	35.868	0.106	0.753	0.750
8	2	No	49.613	0.147	0.527	0.522
8	2	YES	58.295	0.173	0.348	0.340

TCN 모델의 하이퍼파라미터 조합별 성능 비교 결과는 표 6과 같으며 하이퍼파라미터별 모델 중 Hidden Dim = 8, Layer = 1, Bidirectional = True일 때 가장 낮은 RMSE(35.868)를 기록하였으며, R² 또한 0.753로 실제 폐사량의 추세를 가장 잘 반영하고 있으며, 과적합 없이 안정적인 일반화 성능을 나타내는 것을 확인할 수 있었다.

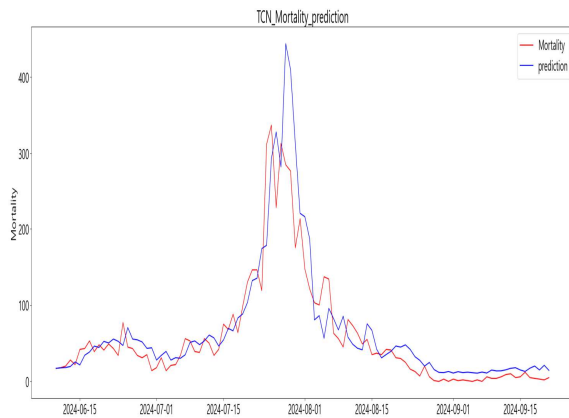


그림 4 TCN 모델 최적 하이퍼파라미터 조합 시각화

그림 4는 TCN 모델의 예측 결과를 하이퍼파라미터별로 시각화한 그래프이다. 예측 결과 TCN 모델은 하이퍼파라미터 값에 따라 일부 과대/과소적합 문제가 발생하긴 하는 경향이 있으나, 전반적으로 실제 폐사량의 추세를 안정적으로 예측하는걸 확인할 수 있다. 또한 고폐사 구간에서의 예측 또한 기존 LSTM, TCN 모델보다 높은 정확도를 보이는데 알 수 있다.

(4) 시계열 예측 모델 성능 비교

표 7 폐사량 예측 모델 성능 비교

모델	Hidden Dim	Layer	Bidirectional	RMSE	MmRMSE	R ²	adj-R ²
LSTM	16	1	No	50.317	0.149	0.514	0.509
GRU	8	1	No	43.094	0.128	0.643	0.640
TCN	16	2	Yes	35.868	0.106	0.753	0.750

표 7은 LSTM, GRU, TCN 모델 각각에서 가장 우수한 예측 성능을 보인 하이퍼파라미터 조합을 기준으로 성능 지표를 정리한 표이다.

LSTM 모델과 같은 경우 RMSE는 50.317, R²는 0.514로 비교적 낮은 성능을 보였으며, 고폐사량 발생 구간에서는 과소 예측되는 경향이 나타났다.

GRU 모델은 RMSE는 43.094, R²는 0.643로 LSTM 모델보다 향상된 성능을 보였으며, 고폐사량 발생 구간에서도 일정 수준의 추세를 예측

하는 성능을 나타냈다.

TCN 모델과 같은 경우 RMSE는 35.868, R²는 0.753으로 세 모델 중 가장 뛰어난 예측 성능을 나타냈다. 이는 TCN 모델의 병렬 처리 및 인접 시점 간 의존성 처리 구조가 넙치 양식장 수질 기반 폐사량 예측에 적합한 모델임을 의미한다. 특히 TCN 모델은 고폐사량 구간에서도 추세와 진폭 모두 실제 값을 가장 잘 예측하는 경향을 보였으며, 이는 실제 양식장 운영에 있어 의사결정 지원 도구로 활용 가능성이 높다.

III. 결 론

본 연구에서는 전라남도 완도군 유수식 넙치 양식장에서 수집한 수질 데이터를 기반으로, 폐사량 예측이 가능한 딥러닝 기반 시계열 예측 모델을 구성하고 성능을 비교·분석하였다. 폐사량 예측에 활용된 수질 지표로는 수온, 아질산성질소, BQ_VALUE, 측정값(세균값)등으로 구성되었으며, 해당 지표들은 상관관계 분석을 통해 폐사량과의 연관성이 높은 지표를 선정하였다.

해당 지표를 기반으로, LSTM(Long Short-Term Memory), GRU(Gated Recurrent Unit), TCN(Temporal Convolutional Network)의 3가지 딥러닝 모델을 대상으로 각기 8개의 하이퍼파라미터 조합을 통해 총 24개의 예측 실험을 수행하였다. 모든 모델은 동일한 조건(입력 및 출력 구간 7일, Optimizer: Adam, learning rate: 0.001, batch size: 32, epoch: 512, loss: MSE, EarlyStopping(patience=128)) 하에 학습되었으며, 성능 평가 지표로는 RMSE, MmRMSE, R²(결정계수), adj-R²(수정 결정계수)를 사용하여 모델 간 성능을 비교하였다.

실험 결과, TCN 모델이 전체적으로 가장 우수한 성능을 보였으며, 특히 Hidden Dim = 8, Layer = 2, Bidirectional = False 조합에서 RMSE 35.866, R² 0.753으로 가장 뛰어난 결과를 기록하였다. 해당 모델은 폐사량이 급격히 증가하는 구간에서도 예측 추세를 비교적 정확하게

반영하였다. 반면, LSTM 및 GRU 모델은 고폐사 구간에서 과소 예측 경향을 보이는 등 예측의 정확도가 부정확하며, 전체적으로 평탄한 추세를 따르는 경향을 보였다.

결과적으로 TCN 모델이 실제 폐사량의 변동성과 경향성을 상대적으로 잘 반영하는 것으로 나타났으며, 이는 TCN의 병렬 처리 및 dilated convolution 기반의 구조적 특성 덕분에, 수질 기반 시계열 예측에서 효과적인 성능을 보인 것으로 판단된다.

본 연구를 통해 수질 기반 딥러닝 예측 모델의 가능성과 효과성을 실증적으로 확인하였으며, 이는 향후 양식장 운영에 있어 의사결정 지원 도구 및 스마트양식 통합 관리 시스템 기술 구현에 기여할 수 있을 것으로 기대된다. 특히 본 연구는 수질 데이터를 기반으로 폐사 발생 조기 징후를 포착하고 선제적 대응이 가능한 최적 딥러닝 모델을 도출하고자 한다는 점에서 의의가 있다.

다만, 본 연구는 특정 지역의 제한된 수조 데이터를 기반으로 하여 다양한 외부 변수들을 충분히 반영하지 못한 한계가 있다. 또한, 시계열 모델 중 3가지 모델(LSTM, GRU, TCN)에 한정하여 수행하였기에 다른 모델과의 비교가 더 필요하다는 점도 제약사항이 될 수 있다.

향후 연구에서는 예측 정확도와 적용 범위를 더욱 확장시키기 위해 다양한 지역의 데이터를 수집하여 모델의 일반화 성능을 검증하고, 최신 시계열 예측 모델을 적용해 보다 높은 성능의 모델을 도출하는 방향의 접근이 필요하다.

REFERENCES

- [1] T. Li, J. Lu, J. Wu, Z. Zhang, and L. Chen, "Predicting Aquaculture Water Quality Using Machine Learning Approaches," *Water*, vol. 14, no. 18, Article 2836, Sep. 2022.
- [2] 통계청, "2024년 어류양식동향조사 결과(잠정)", 보도자료, 2025년
- [3] 조준모, 박상수, 김철수, 장승우, 박성은, "Multi step 딥러닝 모델을 활용한 진동만 저층 용존산소 예측 연구," *한국해양환경·에너지학회지*, 제28권, 제2호, 99 - 107쪽, 2025년
- [4] 손현승, 임한규, 최한석, "딥러닝 기술을 이용한 넙치의 질병 예측 연구", *스마트미디어저널*, 제11권, 제4호, 62-68쪽, 2022년
- [5] 조경원, 김지윤, 홍지훈, "넙치 질병 증상 분류를 위한 객체 탐지 딥러닝 모델 성능 평가," *스마트미디어저널*, 제12권, 제10호, 71 - 84쪽, 2023년
- [6] 정종호, 김유진, 김진성, "인공지능 기술 기반 넙치 질병 증상 감지 및 분류 기법," *스마트미디어저널*, 제13권, 제11호, 36 - 48쪽, 2024년
- [7] Hu, Z., Zhang, Y., Zhao, Y., Xie, M., Zhong, J., Tu, Z., & Liu, J., "A water quality prediction method based on the deep LSTM network considering correlation in smart mariculture", *Sensors*, Vol. 19, No. 6, 1420, 2019.
- [8] 김종민, "생육환경 데이터의 딥러닝을 활용한 지능형 양식 시스템 자율제어에 관한 연구", *동신대학교 일반대학원 박사학위논문*, 2023년
- [9] Sepp Hochreiter and Jurgen Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735 - 1780, 1997.
- [10] Kyunghyun Cho, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio, "Learning Phrase Representations using RNN Encoder - Decoder for Statistical Machine Translation," *Proc. of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1724 - 1734, 2014.
- [11] Colin Lea, Michael D. Flynn, Rene Vidal, Austin Reiter, and Gregory D. Hager, "Temporal Convolutional Networks: A Unified Approach to Action Segmentation," **European Conference on Computer Vision (ECCV)**, pp. 47 - 54, 2016.
- [12] Xia, Zhi, Song, Y., Ma, J., Zhou, L., & Dong, Z., "Research on the Pearson Correlation Coefficient Evaluation Method of Analog Signal in the Process of Unit Peak Load Regulation", *13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI)*, 2017.
- [13] M. Naeem, S. Raza, M. Asif, et al., "Toxic effects of ammonia exposure on growth and hematological response of **Clarias batrachus**," **Applied Ecology and Environmental Research**, vol. 21, no. 6, pp. 5055 - 5067, 2023.
- [14] Majid Vafaeipour et al., "Application of sliding window technique for prediction of wind velocity time series," *International Journal of Energy and Environmental Engineering*, vol. 5, Article 105, 2014.
- [15] Zarándy, Á., Rekeczky, C. & Szolgay, P., "Simulation of Future Dissolved Oxygen Distribution in Aquaculture Waters Based on Sliding Window Temporal Convolutional

Network and Trend Surface Analysis,” **AgTech Digital Twins Aquaculture**, 2021.

저 자 소 개



손명준(정회원)

2022년 국립순천대학교 컴퓨터교육과
학사 졸업.

2025년 국립순천대학교 스마트농업전공
석사 과정.

<주관심분야 : 딥러닝, 머신러닝, IoT, 데이터 분석 >



신창선(증신회원)

1996년 우석대학교 전산학과 학사 졸업.

1999년 한양대학교 컴퓨터교육과 석사
졸업.

2004년 원광대학교 컴퓨터교육과 공학
박사 졸업.

2005년 국립순천대학교 인공지능공학부
교수.

<주관심분야 : 분산컴퓨팅, 실시간 객체모델, 농업 IoT>