

특집논문 (Special Paper)

방송공학회논문지 제31권 제4호, 2026년 7월 (JBE Vol.31, No.4, July 2026)

<https://doi.org/10.5909/JBE.2026.31.4.614>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

생성형 AI 기반 수산물 신선도 설명 생성 및 관능검사 지원 시스템

백 채 윤^{a)}, 윤 중 수^{b)}, 배 재 형^{b)}, 황 준 엽^{a)}, 최 중 원^{a)b)†}

Generative AI-Based Freshness Explanation Generation and Sensory Inspection Support System for Seafood

Chaeyoon Baek^{a)}, JongSu Youn^{b)}, Jaehyung Bae^{b)}, Junyeup Hwang^{a)}, and Jongwon Choi^{a)b)†}

요 약

본 논문은 수입 수산물 관능검사 지원을 위해 비전 기반 신선도 분석과 생성형 AI 기반 설명 생성을 결합한 시스템을 제안한다. 제안 시스템은 색상 보정판 기반 전처리와 VFNet 기반 관심 부위 검출을 통해 어체 전체와 주요 부위를 분석하고, 신선도 등급과 화학 지표인 pH, VBN을 동시에 예측한다. 또한 관능검사 채점표, 다중 평가자 결과, 모델 예측값, 입력 생선 이미지를 생성형 언어모델에 결합하여 최종 신선도 판정을 보조하고 자연어 설명문을 생성한다. 이를 통해 정량 지표와 관능 기준을 함께 반영한 설명 가능한 관능 검사 지원 결과를 제공한다. 어류 5종 실험에서 제안 시스템은 82.98% - 87.77%의 최종 정확도를 보였으며, 생성형 AI 기반 판정 보조와 설명 출력의 활용 가능성을 확인하였다.

Abstract

We propose a system that combines vision-based freshness analysis with generative AI-based explanation generation to support sensory inspection of imported seafood. The proposed system analyzes the whole body and key local regions through color calibration chart-based preprocessing and VFNet-based region of interest detection, and simultaneously predicts freshness grades and chemical indicators, including pH and VBN. In addition, a generative language model integrates the sensory inspection score sheet, multi-rater results, model predictions, and input fish images to support final freshness judgment and generate natural language explanations. Through this framework, the system provides explainable sensory inspection support results that reflect both quantitative indicators and sensory evaluation criteria. Experiments on five fish species achieved final accuracies ranging from 82.98% to 87.77%, demonstrating the feasibility of generative AI-based judgment support and explanation generation.

Keyword : Generative AI, Seafood Freshness Evaluation, Sensory Inspection, Explainability, Multitask Learning

I. 서론

수산물 신선도 평가는 형태, 윤택, 눈, 아가미, 절단면, 냄새 등 다양한 관능 정보를 종합하여 수행되는 핵심 품질 판정 과정이다. 그러나 실제 검사 현장에서는 검사자의 숙련도, 피로도, 촬영 환경 차이 등에 따라 동일 시료에 대해서도 판정 결과가 달라질 수 있어 평가의 일관성 확보가 어렵다. 기존의 인공지능 기반 수산물 신선도 평가 연구는 주로 최종 등급 분류나 화학지표 예측 정확도 향상에 초점을 두어 왔으나, 판정 결과를 검사자가 이해할 수 있는 자연어 설명 형태로 함께 제공하는 기능은 상대적으로 부족하다.

본 연구에서는 이를 보완하기 위해 비전 기반 신선도 분석과 생성형 AI 기반 설명문 생성을 결합한 관능검사 지원 시스템을 제안한다. 제안 시스템은 색상 보정판 기반 전처리로 데이터의 색상 일관성을 확보하고, VFNet^[1] 기반 관심영역 탐지를 통해 어체 전체와 눈, 아가미, 절단면 등 주요 부위를 검출한다. 검출된 부위 특징으로부터 신선도 등급과 pH, VBN을 동시에 예측한 뒤, 예측 결과와 입력 생선 이미지, 관능검사 채점표를 생성형 언어모델에 함께 입력하여 최종 설명문을 생성한다. 이를 통해 제안 시스템은 단순한 등급 출력에 그치지 않고, 생성형 AI를 활용해 판정 결과를 자연어로 해석 가능한 형태로 제공함으로써 관능검사자의 의사결정을 지원한다.

II. 관련 연구

수산물의 신선도 평가는 전통적으로 검사자의 경험에 기

반한 관능검사에 의존해 왔다. 관능검사는 사람의 감각을 이용하여 식품의 품질을 판단하는 평가 방법으로, 수산물의 경우 시각, 후각, 촉각 등을 통해 외관, 색상, 눈의 혼탁도, 아가미 색상 및 점액, 표피 윤택, 육질 탄력, 냄새 등의 항목을 종합적으로 확인하는 절차를 의미한다. FAO는 생선 품질 평가 방법을 관능적 방법과 기기적 방법으로 구분하며, 화학적·기기적 지표 역시 실제 품질 판단에 활용되기 위해서는 관능평가와의 상관성이 중요하다고 설명한다^[2]. 또한 Codex의 어류 및 패류 관능평가 지침은 관능검사자가 제품의 감각적 속성을 기준으로 품질을 평가하기 위한 절차와 기준을 제시한다^[3]. 국내 수입 수산물 검사 기준에서도 수산물의 관능검사 기준은 식품의 기준 및 규격의 수산물 성상 기준을 따르며, 제품별로 외관, 색깔, 선별, 활력도, 풍미, 선도, 건조 및 유소 등 항목을 활용하도록 규정하고 있다. 그러나 실제 현장에서는 검사자의 숙련도와 피로도, 조명 환경 등에 따라 결과 편차가 발생할 수 있다^[4]. 또한 관능검사는 정성적 판단에 기반하므로 pH나 VBN과 같은 정량 지표를 직접 반영하기 어렵다는 한계가 있다.

이를 보완하기 위해 최근에는 스마트폰과 센서를 결합한 비파괴 신선도 평가 기술이 활발히 연구되고 있다. 스마트폰 기반 독소 검출 시스템^[5], TVB-N 색 변화 분석^[6], 형광 sensing tag 기반 비접촉 모니터링^[7], 색 변화 기반 food spoilage 분석 플랫폼^[8] 등이 대표적이다. 그러나 이러한 방법은 별도의 센서나 특정 촬영 환경에 대한 의존성이 높다는 한계를 가진다.

한편, 영상 자체를 이용한 인공지능 기반 신선도 판별 연구도 지속적으로 이루어지고 있다. 눈과 아가미 색 특징을 이용한 ANN, SVM 기반 분류^[9], common carp 영상을 이용한 딥러닝 기반 진단^[10], 생선 눈 영상을 활용한 모바일 신선도 판별 시스템^[11] 등이 주로 보고되었다. 그러나 이러한 연구들은 대부분 최종 등급 분류에 초점을 두고 있어, 부위별 판단 근거를 구조적으로 제시하거나 정량 지표를 함께 활용하는 데에는 한계가 있다. 영상 기반 인공지능 분야에서는 객체 검출 결과나 전경 정보를 활용하여 복수의 태스크를 동시에 수행하는 멀티태스크 분석 구조도 연구되어 왔다^[12]. 이러한 접근은 영상 내 관심 영역을 분리하고, 해당 영역의 특징을 기반으로 다양한 예측 결과를 산출할 수 있다는 점에서 부위 기반 신선도 분석과 관련된다. 그러

a) 중앙대학교 메타버스융합학과(Department of Metaverse Convergence, Chung-Ang University)

b) 중앙대학교 첨단영상대학원 영상학과(Department of Advanced Imaging, GSAIM, Chung-Ang University)

‡ Corresponding Author : 최종원(Jongwon Choi)

E-mail: choijw@cau.ac.kr

Tel: +82-2-820-5870

ORCID: <https://orcid.org/0000-0001-9753-8760>

※ 본 연구는 식품의약품안전처(25192선위안007, 수입식품 현장 관능검사 지원 체계 개선 연구)의 지원을 받아 수행된 연구이다.

· Manuscript received June 25, 2026; Revised July 7, 2026; Accepted July 7, 2026.

나 기존 연구는 일반적인 영상 분석 태스크를 대상으로 하며, 수산물 관능검사에서 요구되는 어체 전체, 눈, 아가미, 절단면 등 부위별 시각 단서와 pH, VBN 같은 화학지표를 함께 고려하는 구조와는 차이가 있다. 따라서 실제 현장에서 활용 가능한 시스템을 위해서는 일반 촬영 환경의 단일 이미지로부터 주요 부위를 자동 탐지하고, 부위별 시각 정보와 pH, VBN 등의 정량 지표를 함께 예측하며, 그 결과를 설명 가능한 형태로 제공하는 구조가 필요하다.

III. 제안 방법

그림 1은 제안하는 설명 가능 수산물 신선도 평가 시스템의 전체 파이프라인을 나타낸다. 입력 영상은 색상 보정판 기반 전처리, 관심 부위 검출, 멀티태스크 예측, 설명 결과 생성 단계를 거쳐 최종 신선도 평가 결과로 출력된다.

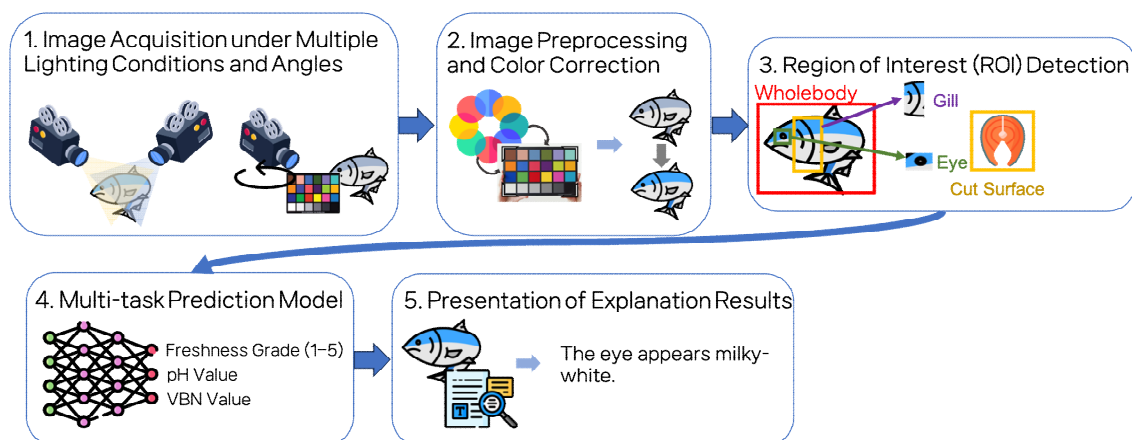


그림 1. 제안 시스템의 전체 파이프라인

Fig. 1. Overall Pipeline of the Proposed System

표 1. 어종별 촬영 부위에 따른 데이터 수

Table 1. Number of Images by Fish Species and Captured Region

Fish Species	Number of Fish	Whole Body, Eye, and Gill Images	Cut Surface Images	Total
Salmon	154	22,176	2,228	24,464
Monkfish	255	29,340	2,280	31,620
Mackerel	276	41,952	2,208	44,160
Flounder	219	17,520	1,752	19,272
Pollock	392	59,664	3,136	62,800
Total	1,296	170,652	11,664	182,316

1. 데이터 취득 및 전처리

데이터 취득 단계에서는 복수의 조명 조건과 촬영 각도에서 어체 전체 이미지와 주요 부위 이미지, 절단면 이미지를 확보하고, 동일 프레임 내에 색상 보정판을 함께 촬영하였다. 개발 대상은 어류 5종(고등어, 연어, 아귀, 명태, 각시 가자미)이며, 각 어종에 대해 홀바디, 눈, 아가미, 절단면 등 관능검사에 활용되는 부위를 중심으로 데이터를 수집하였다. 본 연구에서는 총 1,296마리의 시료로부터 182,316장의 이미지를 확보하였다.

표 1은 어종별 촬영 부위에 따른 데이터 수를 요약한 것이다. 또한 촬영 환경의 다양성을 반영하기 위해 조명과 각도 조건을 달리하여 이미지를 취득하였고, 색상 정보의 일관성을 확보하기 위해 모든 이미지에 색상 보정판을 함께 포함하도록 구성하였다.

데이터 라벨은 관능평가지 채점 기준과 화학지표를 바탕

으로 동일 시료에 대해 단계적으로 구축하였다. 먼저 냉동 어류에 대해 1차 관능평가를 실시한 뒤, 밀봉 상태로 흐르는 물에 단기간 해동하고 시료를 1마리씩 지퍼백에 소분하여 보관하였다. 이후 일정 기간 냉장 보관하면서 어종별 부패 진행 속도를 반영한 조건에서 반복 평가를 수행하였다. 예를 들어 고등어는 24시간 간격으로 4일간, 명태, 가자미, 아귀는 24시간 간격으로 5일간 검증하여 다양한 신선도 상태의 시료를 확보하였다. 각 시점마다 숙련된 평가자 3인 이상이 관능평가를 수행하였으며, 평가자별 점수는 각각 독립적인 학습 target으로 사용하였다. 최종 신선도 등급은 추론 단계에서 세 평가자 head가 예측한 등급을 중앙값 보팅하여 산출하였다. 동시에 동일 시료에 대해 화학적 분석을 수행하여 VBN 값을 실측하였으며, 해당 실측값은 관능평가 결과와 함께 데이터베이스에 저장하였다. 본 연구에서 pH와 VBN은 데이터 구축 단계에서는 실측 화학

지표이며, 멀티태스크 모델 학습 단계에서는 각각의 회귀 target으로 사용된다. 예를 들어 5 - 10 mg%는 신선, 15 - 25 mg%는 보통, 30 - 40 mg%는 부패 초기, 50 mg% 이상은 부패로 해석하였다. 또한 부위별 신선도 라벨은 식약처 기준을 준용하여 눈의 색 및 탁도, 아가미의 색 및 탁도, 절단 후 육색 및 내장 상태, 전체 바디의 탄력도 및 색, 냄새 항목을 중심으로 구축하였다. 관능평가 항목은 공통적으로 신선도 판단에 필요한 시각적 요소를 포함하되, 세부 항목 구성과 채점 기준은 어종별 특성을 반영하여 일부 다르게 적용하였다.

딥러닝 기반 객체 인식 모델의 성능은 학습 데이터의 annotation 품질에 크게 영향을 받으며, 이를 위해 영상 데이터에 대한 체계적인 라벨링 도구와 데이터 구축 절차의 중요성이 강조되고 있다^[13]. 본 연구에서도 각 이미지에 대한 annotation은 GUI 기반 입력 환경에서 수집하였다. 그림 2

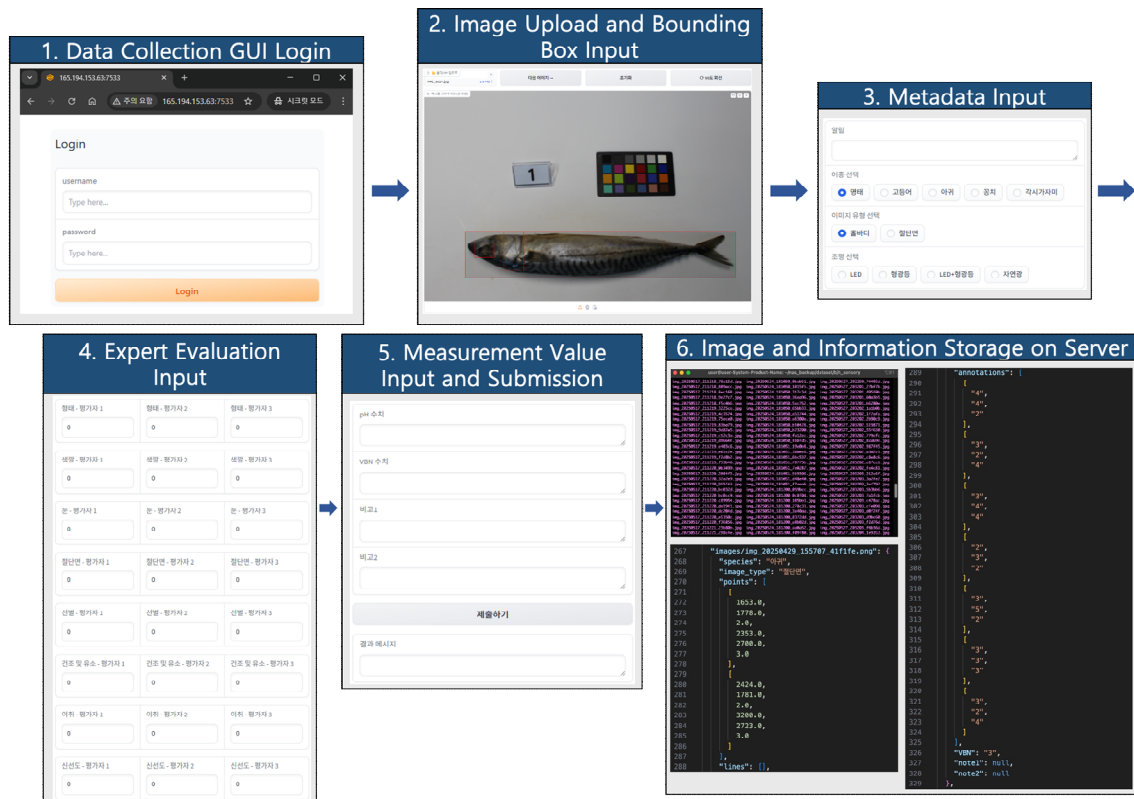


그림 2. GUI 기반 데이터 수집 및 저장 절차
Fig. 2. GUI-based Data Collection and Storage Process

는 데이터 수집 GUI의 전체 흐름을 보여준다. 사용자는 먼저 GUI에 로그인한 뒤 이미지를 업로드하고 bbox를 입력하며, 이어서 어종, 촬영 부위, 조명, 각도와 같은 메타데이터를 입력한다. 이후 평가자 3인은 형태, 색깔, 눈, 절단면, 선별, 아가미, 건조, 이취, 신선도 항목의 점수를 입력하고, pH 및 VBN과 같은 측정 수치를 함께 제출한다. 최종적으로 입력된 이미지, 부위 annotation, 평가자별 관능평가 점수, 메타데이터 및 화학지표는 서버 내에 통합 저장되도록 구성하였다.

표 2는 데이터 수집 GUI를 통해 저장되는 주요 항목을 정리한 것으로, 이미지 파일과 메타데이터뿐 아니라 부위 annotation, 평가자별 관능평가 점수, pH 및 VBN과 같은 화학지표가 함께 저장됨을 보여준다. 따라서 본 데이터셋은 단순한 이미지와 등급 쌍이 아니라 이미지, 부위 annotation, 평가자별 관능평가 점수, 화학지표가 연결된 다중 정보 데이터베이스로 구축되었으며, 이는 시각 정보와 정량 지표를 동시에 활용하는 멀티태스크 학습의 기반이 된다.

또한 동일 개체에서 다수 촬영된 이미지가 학습 및 검증 데이터에 동시에 포함될 경우 발생할 수 있는 data leakage를 방지하기 위해, 본 연구에서는 개별 이미지 단위가 아닌 촬영 폴더 단위로 train, validation 데이터를 분할하였다.

각 촬영 폴더는 동일 개체 또는 동일 촬영 세트에 해당하므로, 동일 폴더 내 이미지는 모두 하나의 데이터 분할에만 배정하였다.

표 3은 어종별 train, validation 분할 결과를 나타낸다. 대부분의 어종은 약 8:2 비율을 목표로 분할하였으나, 실제 이미지 수 기준 비율은 촬영 폴더별 이미지 수 차이로 인해 어종별로 차이를 보였으며, 최종적으로 약 78.5:21.5에서 81.9:18.1 사이의 분포를 보였다. 따라서 본 데이터셋은 이미지 단위 무작위 분할에서 발생할 수 있는 동일 개체 누수 가능성을 줄인 상태에서 학습 및 검증에 사용되었으며, 이를 통해 동일 개체의 유사 이미지로 인해 성능이 과대평가되는 문제를 완화하고자 하였다.

전처리 단계에서는 조명 및 촬영 환경에 따른 색상 편차를 줄이기 위해 색상 보정판 기반 보정을 수행하였다. 어류 신선도 평가는 색깔, 눈, 아가미, 절단면 등 색상 정보에 크게 의존하므로, 촬영 조건 변화에 따른 색상 분포 차이는 학습 및 평가 안정성을 저하시킬 수 있다. 이에 본 연구에서는 입력 영상 내 색상 보정판을 기준 참조체로 활용하여, 색 보정 행렬을 추정하고 이를 영상 전체에 적용하는 전처리 과정을 설계하였다.

표 2. 데이터베이스 저장 항목

Table 2. Database Storage Items

Category	Stored Items
Image Information	Image filename, storage path, sample ID
Metadata	Fish species, captured region, lighting condition, shooting angle
Region Annotation	Bbox coordinates
Sensory Evaluation Scores	Shape, color, eye, cut surface, selection, gill, dryness, off-odor, freshness (detailed items may vary depending on fish species)
Evaluator Information	Evaluator 1, Evaluator 2, Evaluator 3
Chemical Indicators	pH, VBN
Final Label	Aggregated final freshness value

표 3. 어종별 train, validation 데이터 분할 결과

Table 3. Train, validation split results for each fish species

Fish Species	Train images	Validation images	Image-level split ratio	Train folders	Validation folders
Salmon	4,811	1,064	81.9:18.1	134	30
Monkfish	7,330	1,836	80:20	204	51
Mackerel	7,920	1,977	80:20	221	55
Flounder	8,381	2,268	78.7:21.3	251	63
Pollock	10,958	2,997	78.5:21.5	244	82

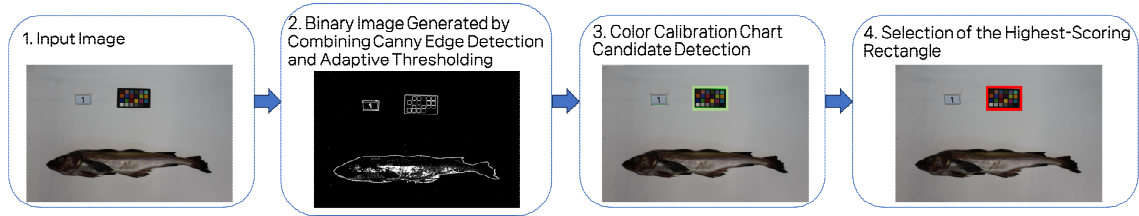


그림 3. 색상 보정판 검출 및 최종 선택 과정

Fig. 3. Color Calibration Chart Detection and Final Selection Process

그림 3은 입력 이미지에서 보정판 후보를 생성하고 최종 보정판을 선택하는 전체 과정을 보여준다. 보정판 검출은 개별 색상 패치를 직접 탐지하기보다, 전체 보정판의 외곽 형상을 먼저 안정적으로 검출하는 방식으로 수행하였다. 이는 패치 단위 탐지가 조명 변화, 블러, 반사, 그림자, 저해상도 촬영 등으로 인해 일부 경계가 약해질 경우 오검출이나 잘못된 대응으로 이어질 수 있는 반면, 전체 색상판은 비교적 일정한 직사각형 형상과 고정된 중횡비를 가지므로 보다 강건하게 검출될 수 있기 때문이다. 이를 위해 영상에 대해 Canny edge detection^[14]과 adaptive thresholding^[15]을 적용하여 후보 사각형을 생성하고, 각 후보에 대해 영역의 넓이와 기준 중횡비 1.5:1에 대한 적합도를 계산하였다. 최종적으로는 넓이가 충분히 크고 기준 보정판의 형상과 가장 유사한 후보를 최종 보정판으로 선택하였다.

선택된 보정판 영역은 perspective warping^[16]을 통해 기준 크기로 정렬한 뒤, 6×4 격자로 분할하여 총 24개의 색상 패치 평균값을 추출하였다. 이후 추출된 패치 색상과 기준 보정판의 24개 색상을 일대일로 대응시켜 색 보정 행렬을

계산하고, 이를 동일 프레임의 원본 영상 전체에 적용하였다. 이러한 과정은 어종 고유의 색차는 유지하면서도 조명 환경, 촬영 시점, 카메라 설정 차이에 의해 발생하는 전역적인 색상 분포 변동을 줄여, 색상 기반 관능 항목에 대한 입력 일관성을 높이는 역할을 한다. 특히 눈의 혼탁도, 아가미 색상, 절단면 색 변화와 같이 신선도 판정에 직접적으로 활용되는 시각 단서의 상대적 안정성을 확보하는 데 유리하다.

또한 보정판 오탐이나 부분 가림이 존재할 경우, 잘못 추정된 색 보정 행렬이 오히려 영상의 색 편향을 증가시킬 수 있다. 이를 방지하기 위해 본 연구에서는 색 보정 이후 자동 신뢰도 점검 단계를 추가로 수행하였다. 구체적으로, 보정된 sRGB 영상을 Lab 색공간으로 변환한 뒤 충분히 밝은 배경 영역만을 추출하고, 해당 영역의 비율과 chroma 중앙값, RGB spread 중앙값을 계산하였다. 일반적으로 촬영 배경은 중립적인 밝은 색에 가까워야 하므로, 보정 이후에도 배경 영역에서 높은 chroma 값이나 채널 간 큰 차이가 나타나는 경우 전역적인 색 치우침이 남아 있는 것으로 판

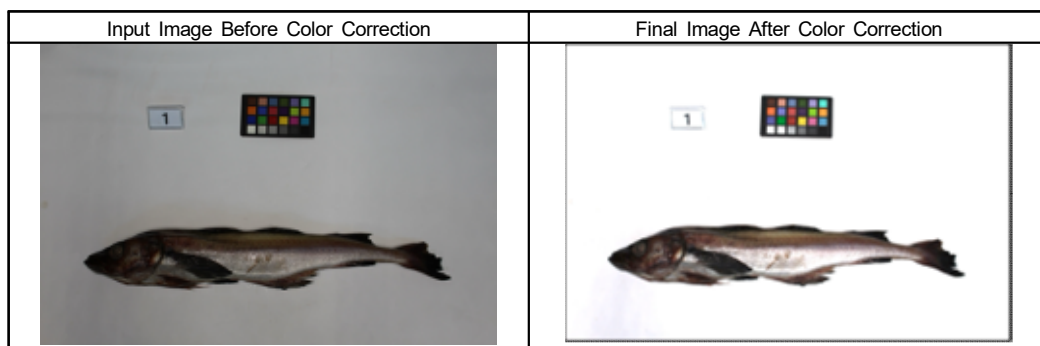


그림 4. 색 보정 전후 영상 비교 예시

Fig. 4. Example Comparison of Images Before and After Color Correction

단하였다. 이러한 기준에 따라 오보정 후보 이미지를 자동으로 검출하고, 품질이 낮은 샘플은 후속 학습 및 평가 단계에서 제외하였다.

그림 4는 색 보정 적용 전후의 영상 비교 예시를 나타내며, 본 전처리 과정이 조명 조건 변화에 따른 색 편차를 완화하는 데 기여함을 시각적으로 보여준다.

2. 관심 부위 검출 및 멀티태스크 예측

관심 부위 검출과 신선도 예측을 위해 본 연구에서는 VFNet을 기반으로 한 bbox-level 멀티태스크 구조를 설계하였다. 제안 모델은 입력 영상으로부터 어체 전체와 주요 부위를 먼저 검출한 뒤, 검출된 bbox 위치의 특징을 이용하여 신선도 등급과 화학지표인 pH, VBN을 동시에 예측한다. 이때 pH와 VBN 예측 head는 데이터 구축 단계에서 측정된 실측값을 target으로 하여 학습되며, 추론 단계에서는 입력 이미지의 시각 특징으로부터 pH, VBN 예측값을 산출한다. 본 연구에서는 어종별 관능검사 기준과 시각적 특성을 반영하여 주요 관심 부위를 다르게 정의하였다. 구체적으로 고등어, 명태, 연어는 어체 전체, 눈, 아가미를 사용하였고, 가자미는 어체 전체를 중심으로, 아귀는 어체 전체와 지느러미를 주요 관심 부위로 설정하였다. 동일 시료에 대해서도 평가자 간 점수 편차가 존재할 수 있으므로, 본 연구에서는 세 명의 평가자에 대응하는 신선도 예측 head를 독립적으로 구성하여 평가자별 관능평가 성향을 모델링하였다.

그림 5는 제안한 VFNet 기반 관심 부위 검출 구조를

나타낸다. VFNet은 backbone과 feature pyramid network(FPN)^[17]에서 생성된 다중 해상도 특징맵을 이용하여 객체의 위치와 분류 점수를 함께 예측하는 one-stage 검출기이다. 본 연구에서는 어류 신선도 판단에 필요한 국소 부위가 상대적으로 작고 경계가 불분명한 경우가 많다는 점을 고려하여 VFNet을 채택하였다. VFNet은 다중 해상도 특징을 통합하여 작은 부위에 대해서도 비교적 안정적인 후보를 생성할 수 있으며, refinement 구조를 통해 눈, 아가미와 같이 시각적 변화가 미세한 영역에서도 검출 성능을 확보할 수 있다. 입력 이미지는 backbone과 FPN을 거쳐 다중 해상도 특징으로 변환되며, 각 해상도에서 detection head가 관심 부위 후보를 예측한다. 이후 NMS^[18]를 통해 중복 후보를 제거하고, 클래스별 상위 후보를 후속 신선도 및 화학지표 예측에 사용하였다. 학습 시에는 ATSS^[19] 기반 positive assignment를 적용하여 정답 bbox와 잘 정렬되는 위치를 양성 샘플로 선택하였다.

그림 6은 검출된 관심 부위 특징으로부터 신선도와 화학지표를 동시에 예측하는 멀티태스크 구조를 보여준다. 본 연구에서는 기본 VFNet 구조를 그대로 사용하지 않고, VFNet classification tower의 출력 특징을 재사용하여 bbox 단위의 신선도 및 화학지표 예측 head를 추가하였다. 기본 VFNet head는 FPN feature를 입력으로 받아 classification branch와 regression branch를 통해 관심 부위 클래스 점수와 bbox 좌표를 예측한다. 이후 초기 bbox 예측값을 이용하여 star-shaped deformable convolution offset을 계산하고, 이를 통해 refined bbox와 IoU-aware classification score를 산출한다. 검출 손실은 Varifocal Loss, 초기 bbox에 대한

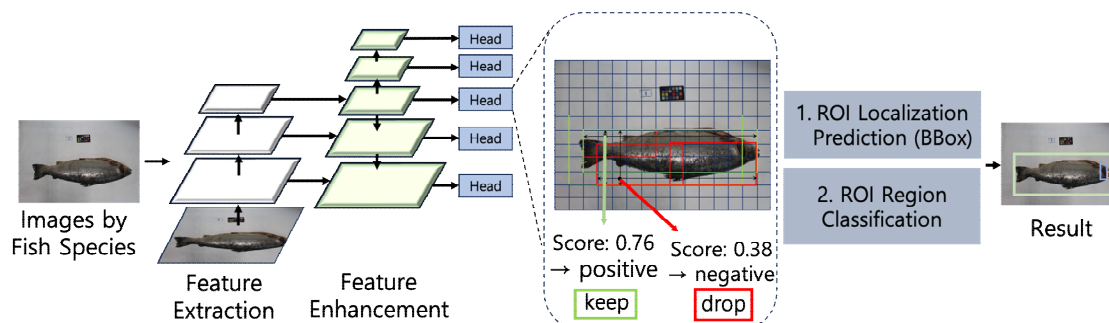


그림 5. VFNet 기반 관심 부위 검출 구조

Fig. 5. VFNet-based Region of Interest Detection Architecture

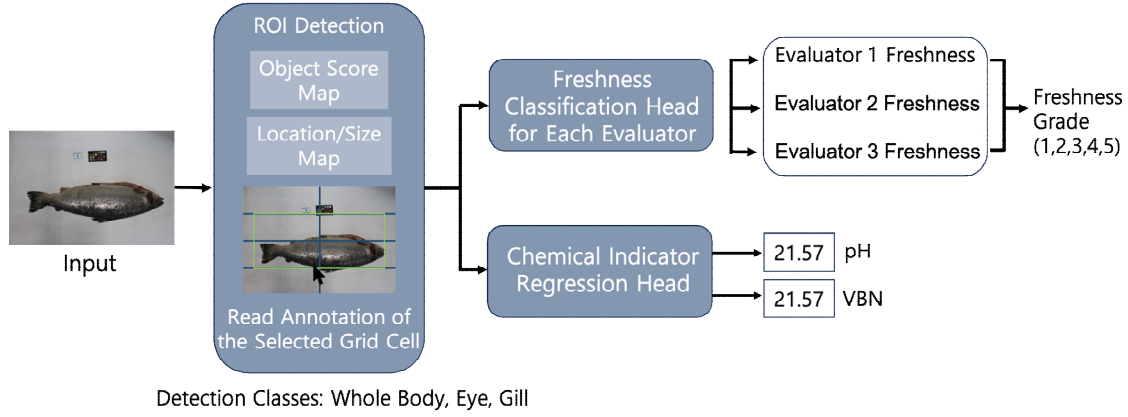


그림 6. 멀티태스크 예측 구조
Fig. 6. Multi-task Prediction Architecture

GIoU Loss, refinement bbox에 대한 GIoU Loss로 구성된다.

추가된 멀티태스크 head는 VFNet의 classification tower에서 얻어진 cls_feat를 입력으로 사용한다. 본 연구에서는 세 명의 평가자에 대응하는 신선도 분류 head와 pH, VBN 회귀 head를 병렬로 추가하였다. 평가자별 신선도 head는 attr_head_r1, attr_head_r2, attr_head_r3로 구성되며, 각 head는 256채널 feature를 입력으로 받아 5개 등급 logits를 출력하는 1×1 convolution layer이다. 각 신선도 head는 서로 다른 평가자의 관능평가 점수를 target으로 학습되며, 이를 통해 평가자 간 점수 차이를 하나의 평균 라벨로 단순화하지 않고 개별 평가자의 판단 경향을 모델링하도록 하였다. pH와 VBN 회귀 head는 각각 ph_head_map, vbn_head_map으로 구성되며, 동일하게 256채널 feature를 입력으로 받아 1채널 회귀 map을 출력하는 1×1 convolution layer이다.

$$a_i^{(r)} = h_{fresh}^{(r)}(z_i), \quad r \in 1, 2, 3 \quad (1)$$

$$\begin{aligned} \hat{p}_i &= \sigma(h_{pH}(z_i)) \cdot S_{pH}, \\ \hat{v}_i &= \sigma(h_{VBN}(z_i)) \cdot S_{VBN} \end{aligned} \quad (2)$$

여기서 z_i 는 positive bbox 위치 i 의 classification tower feature이고, $a_i^{(r)} \in R^5$ 는 r 번째 평가자 head의 5등급 신선도 logits이다. $\hat{p}_i$ 와 $\hat{v}_i$ 는 각각 pH와 VBN 예측값이며, 본

연구에서는 $S_{pH}=14.0$, $S_{VBN}=50.0$ 으로 설정하였다. 실제 구현에서도 pH와 VBN head의 raw output에 sigmoid를 적용한 뒤 scale factor를 곱해 최종 회귀값을 예측한다.

학습 시에는 ATSS에 의해 선택된 positive 위치만을 사용하여 추가 head의 loss를 계산하였다. 먼저 positive bbox와 ground truth bbox 사이의 IoU를 계산하고, 가장 높은 IoU를 갖는 ground truth annotation을 해당 positive sample의 target으로 사용하였다. 평가자별 신선도 target은 각 평가자의 관능평가 점수에서 가져오며, 원래 점수가 1점부터 5점까지의 값을 가지므로 cross-entropy 계산을 위해 0부터 4까지의 class index로 변환하였다. 따라서 학습 단계에서는 평가자 3인의 점수를 평균 또는 중앙값으로 먼저 통합하지 않고, 각 평가자 점수를 독립적인 supervision으로 사용하였다. 기본 구조에서는 최종 신선도 항목을 사용하였고, 부위별 확장 구조에서는 whole body, eye, gill에 대해 각각 다른 관능 항목을 사용할 수 있도록 구성하였다.

평가자별 신선도 loss는 세 개의 head에 대해 각각 cross-entropy loss를 계산한 뒤 평균적으로 반영하였다.

$$L_{fresh} = \frac{\lambda_{fresh}}{3} \sum_{r=1}^3 CE(a_i^{(r)}, y_i^{(r)}) \quad (3)$$

여기서 $y_i^{(r)}$ 는 i 번째 positive sample에 대한 r 번째 평가자의 신선도 target이다. 또한 등급 불균형을 완화하기 위해 평가자별 class weight를 별도로 적용하였다.

pH와 VBN은 연속값이므로 MSE loss를 사용하였다. 최종 학습 loss는 detection loss, 평가자별 신선도 classification loss, pH regression loss, VBN regression loss의 가중합으로 구성하였다.

$$L_{total} = L_{det} + \lambda_{fresh} L_{fresh} + \lambda_{pH} L_{pH} + \lambda_{VBN} L_{VBN} \quad (4)$$

본 연구에서는 $\lambda_{fresh}=1.0$, $\lambda_{pH}=0.5$, $\lambda_{VBN}=0.5$ 로 설정하였다. 이를 통해 관심 부위 검출, 평가자별 신선도 분류, pH 회귀, VBN 회귀가 하나의 네트워크 안에서 동시에 최적화되도록 하였다.

학습되는 parameter는 크게 네 부분으로 구성된다. 첫째, ResNet50 backbone은 입력 영상의 기본 시각 특징을 추출하도록 학습되며, 초기 stage 일부는 고정하였다. 둘째, FPN neck은 다양한 크기의 관심 부위를 안정적으로 탐지하기 위한 multi-scale feature를 생성하도록 학습된다. 셋째, VFNet detection head는 관심 부위 클래스와 bbox 좌표를 예측하도록 학습된다. 넷째, 본 연구에서 추가한 각 head가 신선도 등급과 화학지표를 예측하도록 학습된다. Optimizer는 SGD를 사용하였고, learning rate는 0.005, momentum은 0.9, weight decay는 5×10^{-4} 로 설정하였다. 추가된 평가자별 신선도 head는 빠른 적응을 유도하기 위

해 기본 learning rate보다 10배 큰 learning rate multiplier를 적용하였다.

추론 단계에서는 각 FPN level에서 bbox, class score, 평가자별 신선도 logits, pH, VBN 예측값을 같은 위치 index 기준으로 함께 추출한다. 이후 score thresholding, top-k filtering, NMS를 수행하여 중복 bbox를 제거하고, 남은 bbox에 대응되는 attribute logits와 pH, VBN 예측값을 함께 유지한다. 최종 출력은 bbox 좌표, class label, detection score, 평가자별 신선도 logits, pH 예측값, VBN 예측값을 포함한다. 최종 신선도 등급 산출 시에는 검출된 후보 중 대표 bbox를 선택하고, 세 평가자 head의 예측값을 중앙값 보팅하여 최종 등급으로 사용하였다. 추론 단계에서는 평가자 점수가 입력으로 필요하지 않으며, 입력 이미지만으로 모델이 세 평가자 head의 예측값과 pH, VBN 예측값을 산출한다. 이후 세 평가자 head의 신선도 예측값을 중앙값 보팅하여 최종 신선도 예측 등급을 결정한다.

3. 생성형 AI 기반 최종 판정 보조 및 설명 생성 모듈

설명 생성 모듈은 비전 기반 예측 결과를 관능검사 기준과 연결하여 자연어 설명문을 생성하는 구조로 설계하였다. 본 모듈의 입력은 멀티태스크 예측 모듈에서 산출한 1차

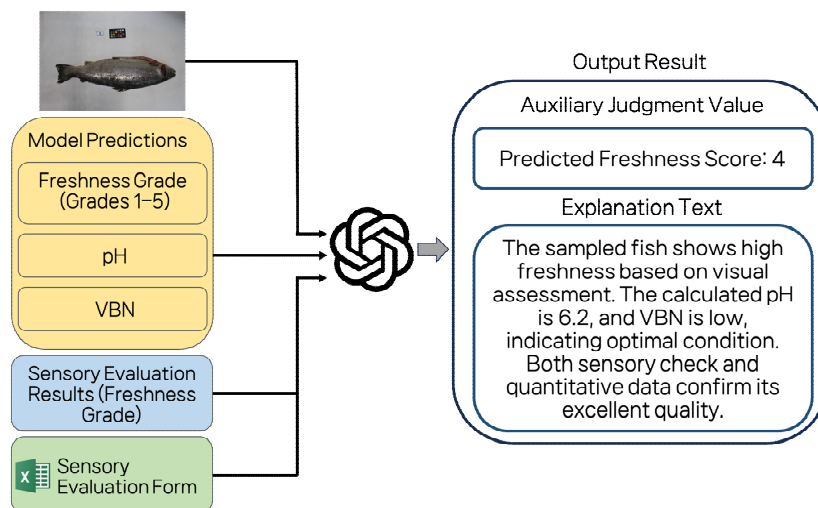


그림 7. 생성형 AI 기반 선도 설명문 생성 구조

Fig. 7. Generative AI-based Freshness Explanation Generation Architecture

신선도 등급, 부위별 신선도 예측값, pH 및 VBN 예측값, 평가자 3인의 관능평가 결과, 그리고 어종별 관능검사 채점표로 구성된다.

그림 7은 생성형 AI 기반 판정 보조 및 선도 설명문 생성 흐름을 나타낸다. 먼저 입력 생선 영상은 관심 부위 검출 및 멀티태스크 예측 모듈을 거쳐 1차 신선도 등급, 부위별 예측값, pH 및 VBN 값으로 변환된다. 이후 생성형 AI 모듈은 비전 모델 예측값과 관능검사 채점표의 기준 문항을 결합하여 보조 판정 정보와 설명문을 생성한다. 본 논문의 설명 생성 실험에서는 비전 모델 예측 결과와 평가자 결과의 일치 여부를 분석하기 위해 평가자 3인의 관능평가 점수를 추가적으로 사용하였다.

표 4는 생성형 AI 설명 모듈의 구현 설정을 나타낸다. 본 연구에서는 GPT-4o를 사용하였으며, 프롬프트는 system prompt와 user prompt로 구분하여 구성하였다. System prompt에서는 모델의 역할을 수산물 관능검사 보조 AI로 제한하고, 입력된 관능검사 채점표와 예측 결과에 근거하여 설명을 생성하도록 지시하였다. 생성 설정에서 temper-

ature는 출력의 무작위성을 조절하는 값으로, 본 연구에서는 0.3으로 설정하여 설명문의 일관성을 높이하고자 하였다. 또한 response_format 옵션을 사용하여 출력 형식을 JSON 객체로 제한함으로써, 생성된 설명문을 구조적으로 parsing 하고 후속 검증을 수행할 수 있도록 하였다. 입력으로 제공되지 않은 보관 이력, 냄새, 식용 가능 여부, 안전성 여부는 임의로 생성하지 않도록 프롬프트에서 제한하였으며, 생성 결과에 대해 JSON parsing과 필수 필드 검증을 수행하여 출력 형식의 일관성을 확인하였다.

구체적으로 생성형 AI 모듈의 입력은 다음과 같이 구성된다. 첫째, 비전 모델이 예측한 1차 신선도 점수와 부위별 예측값을 입력한다. 둘째, 예측한 pH와 VBN 값을 정량 보조 지표로 입력한다. 셋째, 비전 모델 예측 결과와 평가자 결과의 일치 여부를 분석하기 위해, 평가자 3인의 개별 관능평가 점수와 중앙값을 분석용 입력으로 추가하였다. 넷째, 눈, 아가미, 표피, 절단면 등 관능검사 항목별 점수 기준 문장을 함께 제공한다. 생성형 AI는 이러한 정보를 종합하여 비전 모델의 예측값이 평가자 결과 및 관능 기준과 일치

표 4. 생성형 AI 설명 모듈의 구현 설정

Table 4. Implementation settings of the generative AI explanation module

Item	Setting
Model	GPT-4o
Prompt structure	System prompt and user prompt
System prompt	Assigns the model the role of a seafood sensory-inspection assistant, constrains explanations to the sensory inspection rubric, and enforces JSON output format
User prompt	Provides the fish species, vision-model predictions, part-level scores, predicted pH/VBN values, evaluator scores used for analysis, and sensory inspection rubric
Output format	Structured JSON schema based output
Decoding setting	temperature=0.3, max tokens=1024
Hallucination mitigation	Prohibits generation of information not provided in the input, requires reflection of sensory rubric items, and applies JSON parsing and required-field validation

표 5. 생성형 AI 설명 모듈의 입출력 구성

Table 5. Input and output structure of the generative AI explanation module

Category	Field	Description
Input	Vision freshness score	Freshness score predicted by the vision model
Input	Part-level scores	Freshness scores predicted from each region of interest
Input	Chemical indicators	Predicted pH and VBN values from the multitask vision model
Input	Sensory rubric	Criterion texts for each sensory inspection item
Optional input for analysis	Evaluator scores	Individual and median sensory scores from three evaluators used for agreement analysis
Output	Auxiliary freshness score	Decision support information generated by the language model
Output	Vision agreement	Agreement between the vision prediction and evaluator median
Output	Explanation	Natural-language explanation of the freshness decision
Output	Key sensory factors	Main sensory cues used in the explanation
Output	Key chemical factors	Predicted pH and VBN used as supporting indicators
Output	Rubric citations	Sensory rubric items cited in the explanation

하는지 검토하고, 보조 판정 정보와 그 근거를 설명문 형태로 출력한다.

표 5는 생성형 AI 기반 설명 모듈의 입력 정보와 출력 항목을 정리한 것이다. 본 모듈의 기본 입력은 비전 모델이 예측한 신선도 등급, 부위별 예측값, pH, VBN 예측값, 입력 이미지, 그리고 관능검사 항목별 점수 기준 문장으로 구성된다. 평가자 3인의 개별 관능평가 점수와 중앙값은 실제 추론 단계에서 반드시 필요한 입력이 아니라, 본 논문의 설명 생성 실험에서 비전 모델 예측 결과와 평가자 결과의 일치 여부를 확인하기 위한 분석용 정보로 사용하였다. 출력 항목은 보조 신선도 점수, 비전 모델과 평가자 결과의 일치 여부, 한국어 설명문, 주요 관능 근거, 주요 화학 근거, 그리고 설명에 사용된 관능검사 기준 문항으로 구성하였다. 따라서 본 모듈은 비전 모델의 예측 결과를 검사자가 이해할 수 있는 형태로 변환하고, 관능검사 기준에 기반한 설명 가능한 판정 보조 정보를 제공한다.

한편, 절단면 영상에 대해서는 Grad-CAM^[20] 기반 시각화를 적용하여 판정에 영향을 준 영역을 함께 제공하였다. 그림 8은 절단면 영상에 대한 Grad-CAM 시각화 구조를 보여준다. 입력 절단면 이미지는 분류 네트워크를 거쳐 예측 결과를 산출하며, 동시에 activation gradient를 이용해 판정에 기여한 영역을 heatmap 형태로 시각화한다. 이를 통

해 사용자는 생성형 AI가 제공하는 설명문뿐 아니라, 절단면의 어느 영역이 판정에 주요하게 반영되었는지를 함께 확인할 수 있다.

최종 출력은 비전 모델의 신선도 예측값, pH 및 VBN 예측값, 생성형 AI가 작성한 보조 설명문, 절단면 Grad-CAM heatmap, 그리고 필요 시 재검 권고 신호를 포함한다. 이러한 출력 방식은 기존의 단순 등급 출력과 달리, 시각 정보, 정량 지표, 다중 평가자 결과, 관능검사 기준을 함께 반영한 설명 가능한 결과를 제공함으로써 관능검사자의 최종 의사 결정을 지원한다.

IV. 실험 결과 및 분석

1. 어종별 혼동행렬

어류 5종에 대한 최종 신선도 분류 성능을 평가하기 위해 어종별 혼동행렬을 분석하였다. 최종 정확도는 연어 82.98%, 아귀 84.87%, 고등어 87.77%, 가자미 87.33%, 명태 85.42%였으며, 전체 평균 정확도는 85.67%로 나타났다. 이는 제한한 시스템이 서로 다른 외형, 표면 질감, 색상 변화 특성을 가지는 어종에 대해서도 전반적으로 안정적인

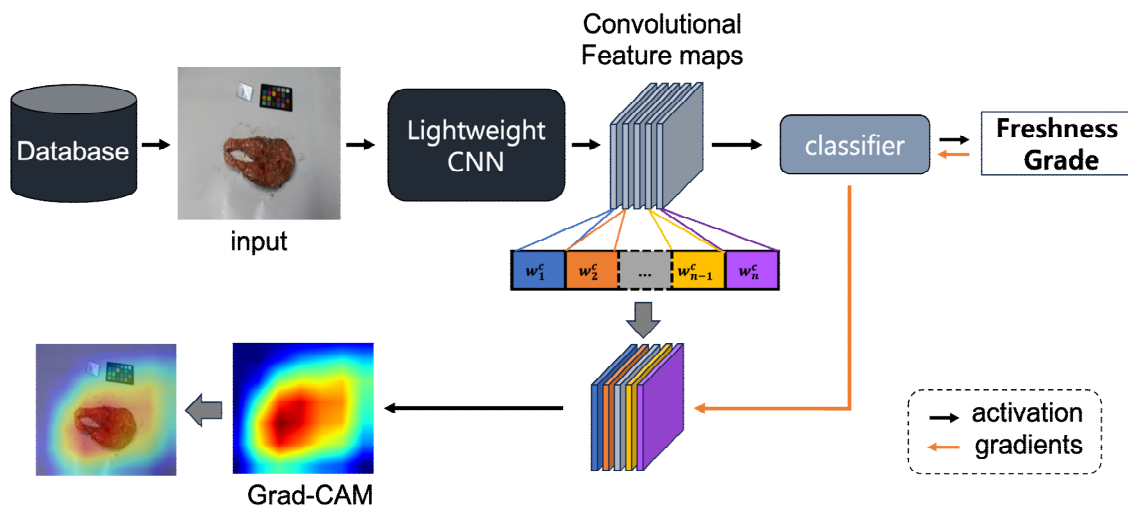


그림 8. 절단면 Grad-CAM 시각화 구조

Fig. 8. Grad-CAM Visualization Architecture for Cut Surface Images

신선도 판별 성능을 보였음을 의미한다. 특히 모든 어종의 혼동행렬에서 대각선 성분이 우세하게 나타나, 모델이 각 신선도 등급의 주요 시각적 특징을 일정 수준 이상 학습하였음을 확인할 수 있었다. 다만 혼동은 대부분 멀리 떨어진 등급 사이보다는 인접 등급 사이에서 발생하였다. 이는 신

선도 저하가 명확히 분리된 이산적 단계로 변화하기보다는 시간에 따라 점진적으로 진행되는 연속적 현상에 가깝기 때문에, 경계 등급에서는 시각적 차이가 작고 관능평가 기준 역시 평가자에 따라 다소 다르게 해석될 수 있기 때문으로 판단된다.

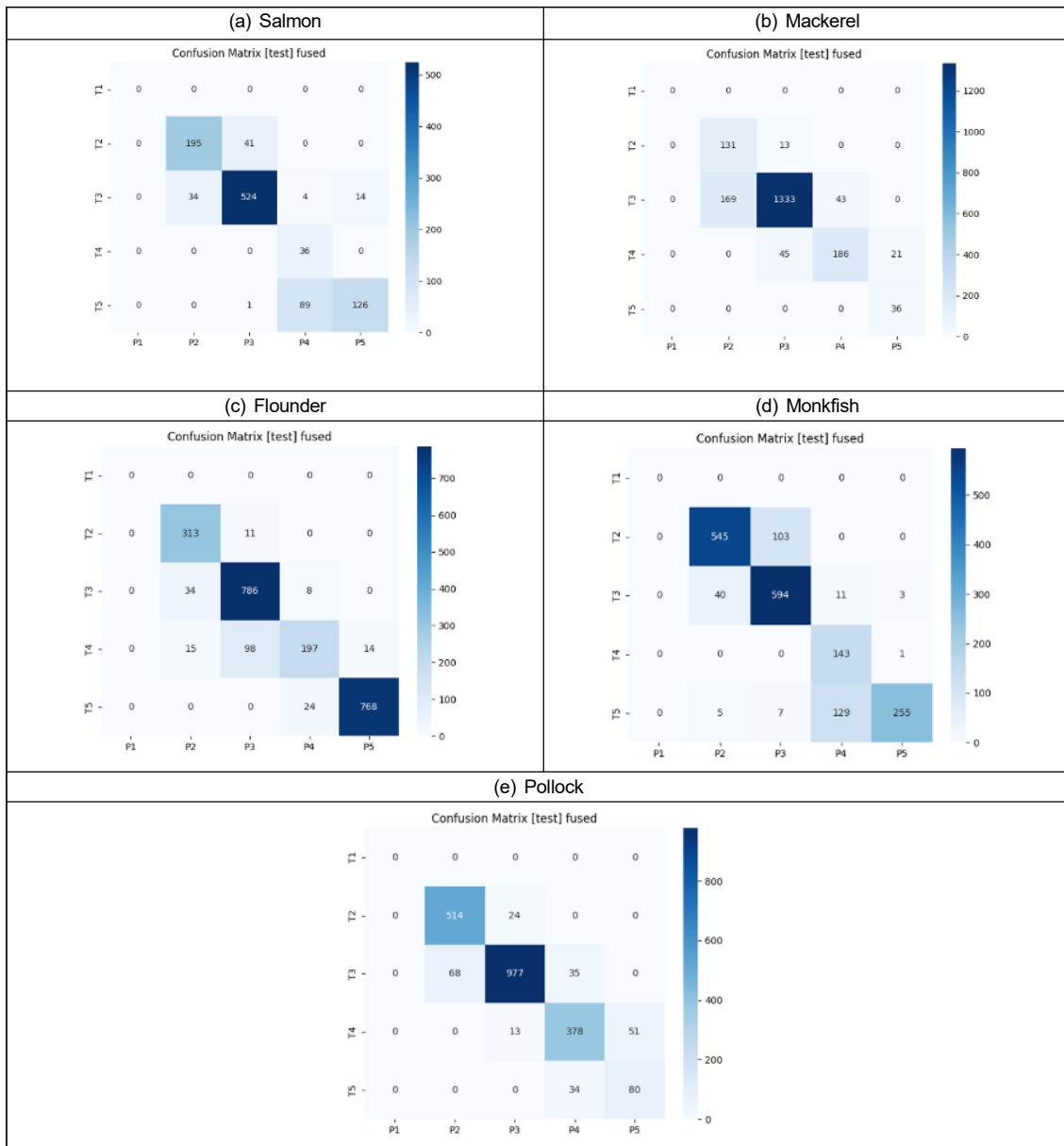


그림 9. 어종별 최종 혼동행렬

Fig. 9. Final Confusion Matrices by Fish Species

그림 9 (a)는 연어의 최종 신선도 분류 혼동행렬을 나타낸다. 연어의 경우 전체 정확도는 82.98%로 나타났으며, 2점과 3점, 4점과 5점 사이에서 혼동이 관찰되었다. 특히 2점 샘플 중 일부가 3점으로 예측되고, 5점 샘플 중 일부가 4점으로 예측되는 경향이 나타났다. 이는 연어의 신선도 변화가 체표 색이나 눈의 탁도, 아가미 상태에서 급격하게 나타나기보다는 완만하게 진행되기 때문으로 해석된다. 특히 비교적 신선한 구간인 4점과 5점은 외관상 차이가 크지 않아, 모델이 가장 신선한 상태와 약간 품질이 저하된 상태를 구분하는 데 어려움을 보인 것으로 판단된다. 반면 낮은 신선도 구간과 높은 신선도 구간 사이의 큰 폭의 오분류는 상대적으로 적어, 모델이 전반적인 신선도 저하 방향은 비교적 잘 포착한 것으로 볼 수 있다.

그림 9 (b)는 고등어의 최종 신선도 분류 혼동행렬을 나타낸다. 고등어는 87.77%로 5종 중 가장 높은 정확도를 보였다. 혼동행렬에서도 대각선 성분이 뚜렷하게 나타나며, 전반적으로 모델이 고등어의 신선도 등급을 안정적으로 판별하였음을 확인할 수 있다. 고등어에서는 주로 2점과 3점, 3점과 4점 사이에서 오분류가 발생하였다. 이는 고등어의 체표 광택, 등쪽 색상, 복부 색 변화와 같은 특징이 부패 진행에 따라 비교적 명확히 변하지만, 중간 등급에서는 변화가 연속적으로 나타나기 때문에 인접 등급 간 경계가 모호해지기 때문으로 보인다. 특히 3점은 신선도 저하가 본격적으로 관찰되기 시작하는 중간 단계이므로, 2점 또는 4점과 시각적 특성이 일부 겹칠 수 있다. 그럼에도 불구하고 고등어에서 가장 높은 정확도가 나타난 것은, 다른 어종에 비해 신선도 변화에 따른 색상 및 표면 상태 변화가 상대적으로 뚜렷하게 관찰되었기 때문으로 해석할 수 있다.

그림 9 (c)는 가자미의 최종 신선도 분류 혼동행렬을 나타낸다. 가자미는 87.33%의 정확도를 보였으며, 고등어와 함께 비교적 높은 분류 성능을 나타냈다. 혼동행렬을 보면 2점, 3점, 5점 구간에서는 대각선 성분이 강하게 나타나 해당 등급의 판별이 안정적으로 이루어진 것을 확인할 수 있다. 그러나 4점 구간에서는 3점 또는 5점으로 일부 오분류되는 경향이 나타났다. 가자미는 체표가 비교적 밝고 평평한 형태를 가지므로, 조명 반사나 표면 건조 상태에 따라 색상 차이가 과장되거나 약화될 가능성이 있다. 따라서 모

델이 실제 신선도 변화와 촬영 조건에 따른 시각적 차이를 완전히 분리하기 어려웠을 수 있다. 이러한 결과는 가자미의 경우 전체적인 등급 분류는 안정적이지만, 경계 등급에서는 화학지표나 부위별 세부 특징을 함께 고려할 필요가 있음을 보여준다.

그림 9 (d)는 아귀의 최종 신선도 분류 혼동행렬을 나타낸다. 아귀의 최종 정확도는 84.87%로 나타났으며, 전체적으로는 대각선 중심의 분류 경향을 보였지만 일부 등급에서 상대적으로 큰 혼동이 관찰되었다. 특히 2점과 3점 사이, 4점과 5점 사이의 혼동이 두드러졌다. 아귀의 경우 전체 바디와 지느러미를 주요 관심 부위로 사용하였기 때문에, 눈이나 아가미처럼 신선도 변화가 비교적 직관적으로 드러나는 부위가 포함된 어종에 비해 세밀한 등급 구분이 더 어려웠을 가능성이 있다. 따라서 아귀에서 나타난 혼동은 단순한 모델 오류라기보다, 어종의 형태적 특성과 관능 기준의 복잡성이 함께 반영된 결과로 해석할 수 있다.

그림 9 (e)는 명태의 최종 신선도 분류 혼동행렬을 나타낸다. 명태의 최종 정확도는 85.42%로 나타났으며, 전체적으로 2점, 3점, 4점 구간에서 대각선 성분이 우세하게 나타났다. 그러나 2점과 3점, 4점과 5점 사이에서 일정 수준의 혼동이 확인되었다. 특히 4점 샘플이 5점으로 예측되거나, 5점 샘플이 4점으로 예측되는 경향은 비교적 신선한 구간에서 시각적 차이가 작다는 점을 보여준다. 명태는 냉동 및 해동 과정에서 표면의 색상 변화, 수분감, 성에, 건조 정도 등이 함께 나타날 수 있으며, 이러한 요소들은 실제 신선도 변화와 독립적으로 이미지 특징에 영향을 줄 수 있다. 따라서 명태의 경우 모델이 신선도에 따른 본질적인 시각 단서와 냉동 유통 환경에서 발생하는 외부 시각 요인을 동시에 학습했을 가능성이 있다. 그럼에도 대각선 성분이 전반적으로 높게 나타났다는 점에서, 제안한 모델은 명태의 주요 신선도 변화 패턴을 유효하게 학습한 것으로 판단된다.

종합적으로, 5종 어류의 혼동행렬 분석 결과는 제안한 시스템이 어종별 특성이 다른 조건에서도 평균 85.67%의 분류 정확도를 보이며 안정적으로 동작함을 보여준다. 특히 모든 어종에서 정답 등급과 예측 등급이 크게 벗어나는 경우는 제한적이었고, 대부분의 오분류는 인접 등급 사이에서 발생하였다. 이는 모델이 신선도 저하의 전체적인 방향

성은 잘 학습했지만, 관능평가 기준상 경계에 위치한 샘플의 세밀한 차이를 구분하는 데에는 여전히 어려움이 있음을 의미한다. 이러한 경향은 실제 관능검사에서도 나타나는 특성과 유사하다. 예를 들어 매우 신선한 상태와 명확히 부패가 진행된 상태는 비교적 쉽게 구분할 수 있지만, 중간 단계나 인접 등급 경계에서는 평가자 간 점수 차이가 발생할 수 있다. 따라서 혼동행렬에서 나타난 인접 등급 혼동은 모델 성능의 단순한 한계라기보다, 신선도 평가 문제 자체가 가지는 연속적이고 주관적인 특성이 반영된 결과로 볼 수 있다.

또한 어종별 혼동 양상은 서로 다르게 나타났다. 고등어와 가자미는 상대적으로 높은 정확도를 보이며 대각선 성분이 뚜렷하게 나타났고, 이는 해당 어종에서 신선도 변화에 따른 시각적 단서가 비교적 안정적으로 학습되었음을 의미한다. 반면 연어, 아귀, 명태에서는 일부 경계 등급에서 혼동이 더 크게 나타났다. 연어는 색상 변화가 완만하고, 아귀는 형태와 표면 질감이 복잡하며, 명태는 냉동 및 해동 과정에서 발생하는 성예와 표면 변화의 영향을 받을 수 있다. 이러한 어종별 차이는 단일한 기준으로 모든 어종의 신선도를 판단하기 어렵다는 점을 보여주며, 어종별 관심 부위 설정과 멀티태스크 예측 구조가 필요한 이유를 뒷받침한다.

결과적으로, 그림 9의 혼동행렬은 제안한 시스템이 단순히 전체 정확도 수치만 높은 것이 아니라, 실제 관능검사에서 중요한 등급별 판별 구조를 일정 수준 반영하고 있음을 보여준다. 동시에 인접 등급에서 반복적으로 발생하는 혼

동은 향후 개선 방향을 제시한다. 경계 등급의 분류 성능을 높이기 위해서는 단일 이미지 기반 시각 특징만 사용하는 것보다, 어체 전체와 주요 부위의 특징을 함께 활용하고, pH 및 VBN과 같은 화학지표를 보조 정보로 결합하는 방식이 필요하다. 또한 생성형 AI 기반 설명 모듈을 통해 모델의 예측값, 부위별 시각 단서, 화학지표를 함께 제시하면, 검사자는 단순한 등급 결과뿐 아니라 해당 결과가 도출된 근거를 확인할 수 있다.

2. 신선도 등급별 화학 지표 분포

신선도 등급별 화학 지표 분포를 분석하기 위해, 데이터 구축 단계에서 화학 분석을 통해 측정된 pH와 VBN 값을 어종별 신선도 등급에 따라 box plot으로 시각화하였다. 본 연구에서 신선도 등급은 관능평가 결과를 기반으로 구성되었으며, 등급 값이 높을수록 상대적으로 신선한 상태를 의미한다. 따라서 화학 지표가 관능평가 기반 신선도 등급과 일관된 경향을 보이는지 확인하는 것이 중요하다. 특히 VBN은 어류 단백질 분해 및 부패 진행과 관련된 대표적인 화학 지표이므로, 신선도가 낮은 등급에서 높은 값을 보이는지 확인하였다. pH는 어종별 생리적 특성과 저장 과정에 따라 변화 양상이 달라질 수 있으므로, VBN과 함께 보조적인 지표로 해석하였다.

그림 10은 연어의 신선도 등급별 pH 및 VBN 분포를 나타낸다. 연어의 pH는 등급 1에서 중앙값 6.19, 등급 2에서 6.21, 등급 3에서 6.25, 등급 4와 5에서 약 6.42로 나타나

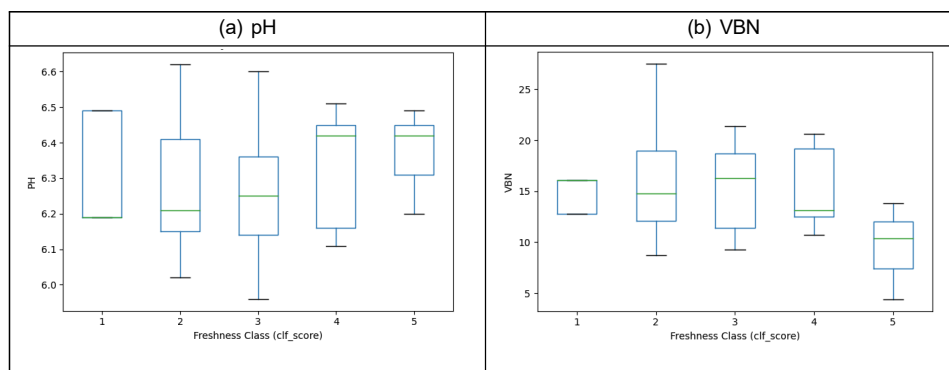


그림 10. 연어의 신선도 등급별 화학 지표 분포. (a) pH, (b) VBN

Fig. 10. Distribution of Chemical Indicators by Freshness Grade in Salmon. (a) pH, (b) VBN

등급이 높아질수록 완만하게 증가하는 경향을 보였다. 그러나 전체 pH 범위가 대체로 6.0 - 6.6 부근에 집중되어 있으며, 인접 등급 간 사분위 범위가 상당 부분 겹치는 것을 확인할 수 있다. 이는 연어의 경우 pH만으로 신선도 등급을 세밀하게 구분하기 어렵다는 것을 의미한다. 반면 VBN은 등급 1 - 4에서 중앙값이 약 13.1 - 16.3 mg% 범위에 분포하였으나, 가장 신선한 등급인 등급 5에서는 중앙값이 약 10.35 mg%로 낮게 나타났다. 즉, 연어에서는 VBN이 신선도가 높은 등급에서 감소하는 경향을 보였으며, 특히 등급 5가 다른 등급에 비해 낮은 VBN 분포를 보였다. 다만 등급 1 - 4 사이에서는 VBN 분포가 서로 중첩되어, VBN만으로 모든 신선도 단계를 명확히 분리하기에는 한계가 있었다.

그림 11은 고등어의 신선도 등급별 pH 및 VBN 분포를 나타낸다. 고등어의 경우 등급 2, 3, 4, 5에 대한 분포가 확

인되었으며, 등급 1 샘플은 포함되지 않았다. pH는 등급 2에서 중앙값 6.87, 등급 3에서 7.00, 등급 4에서 6.66, 등급 5에서 6.54로 나타났다. 등급 3에서 pH가 상대적으로 높게 나타났으나, 등급이 증가함에 따라 단조롭게 감소하거나 증가하는 형태는 아니었다. 특히 등급 3의 pH 분포 범위가 넓고, 등급 2 - 4 사이의 분포가 중첩되어 pH 단독 지표의 판별력은 제한적이었다. 반면 VBN은 신선도 등급과 비교적 뚜렷한 관계를 보였다. 등급 2의 VBN 중앙값은 20.8 mg%, 등급 3은 16.6 mg%, 등급 4와 5는 각각 9.7 mg%로 나타났다. 즉, 신선도 등급이 높아질수록 VBN이 전반적으로 낮아지는 경향이 관찰되었다. 이는 고등어에서 VBN이 부패 진행 정도를 반영하는 유의미한 보조 지표로 작동할 수 있음을 보여준다. 다만 등급 4와 5의 VBN 중앙값이 유사하게 나타났기 때문에, 신선한 구간 내에서의 세부 등급

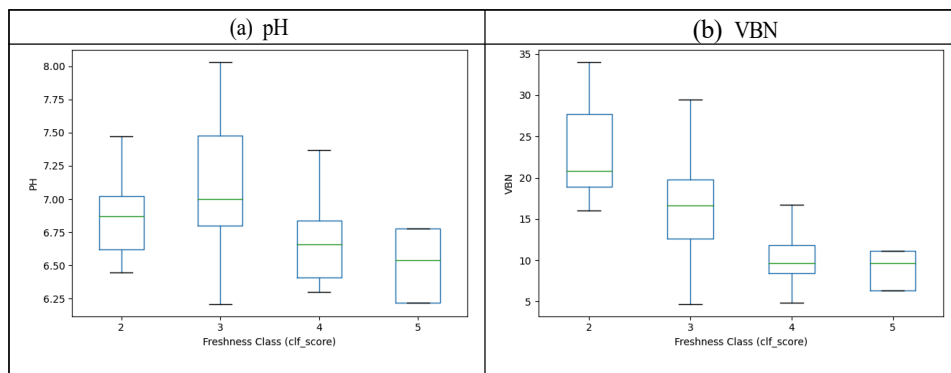


그림 11. 고등어의 신선도 등급별 화학지표 분포. (a) pH, (b) VBN

Fig. 11. Distribution of Chemical Indicators by Freshness Grade in Mackerel. (a) pH, (b) VBN

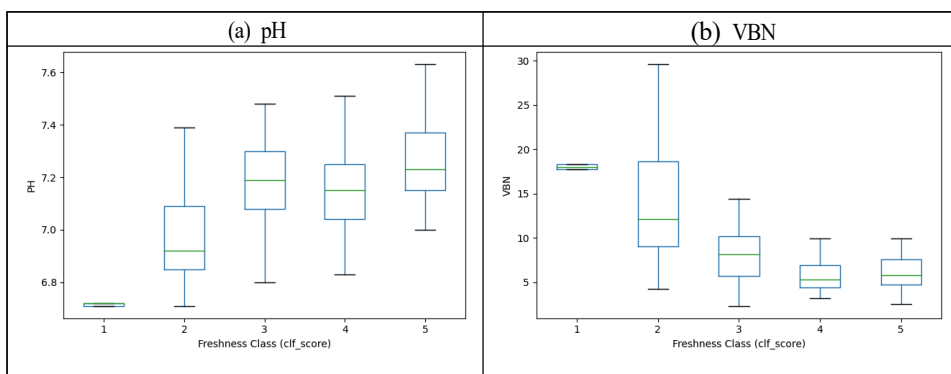


그림 12. 가자미의 신선도 등급별 화학지표 분포. (a) pH, (b) VBN

Fig. 12. Distribution of Chemical Indicators by Freshness Grade in Flounder. (a) pH, (b) VBN

구분에는 시각 정보나 부위별 특징이 함께 필요하다.

그림 12는 가자미의 신선도 등급별 pH 및 VBN 분포를 나타낸다. 가자미의 pH는 등급 1에서 중앙값 6.72, 등급 2에서 6.92, 등급 3에서 7.19, 등급 4에서 7.15, 등급 5에서 7.23으로 나타났다. 전반적으로 신선도 등급이 높아질수록 pH가 증가하는 경향을 보였으며, 특히 등급 1과 등급 5 사이에는 pH 중앙값 차이가 확인되었다. 그러나 등급 3-5 사이에서는 중앙값 차이가 크지 않고 사분위 범위가 서로 겹치기 때문에, pH만으로 상위 신선도 등급을 구분하기는 어렵다. VBN은 등급 1에서 중앙값 18.0 mg%, 등급 2에서 12.1 mg%, 등급 3에서 8.1 mg%, 등급 4에서 5.3 mg%, 등급 5에서 5.8 mg%로 나타났다. 따라서 가자미에서는 신선도 등급이 높아질수록 VBN이 감소하는 경향이 비교적 명확하게 관찰되었다. 특히 등급 1과 등급 4-5 사이의 VBN

차이가 뚜렷하여, VBN이 저신선도와 고신선도 상태를 구분하는 데 유용한 지표가 될 수 있음을 확인하였다. 그러나 등급 4와 5는 VBN 분포가 매우 유사하므로, 최상위 신선도 구간에서는 화학지표만으로 세밀한 등급 구분이 어렵다.

그림 13은 아귀의 신선도 등급별 pH 및 VBN 분포를 나타낸다. 아귀의 pH는 등급 1에서 중앙값 7.575로 가장 높게 나타났고, 등급 2-5는 대체로 6.9-7.03 부근에 분포하였다. 등급 1의 경우 샘플 수가 상대적으로 적고 pH 범위가 매우 좁게 나타나므로, 해당 등급의 pH 경향을 일반화하기에는 주의가 필요하다. 등급 2-5에서는 pH 분포가 넓게 중첩되어 pH가 신선도 등급을 안정적으로 구분하는 지표로 작동하기는 어려웠다. VBN은 등급 1에서 중앙값 24.4 mg%, 등급 2에서 22.85 mg%로 높게 나타났고, 등급 3-5에서는 각각 9.4 mg%, 9.8 mg%, 9.8 mg% 수준으로 낮게

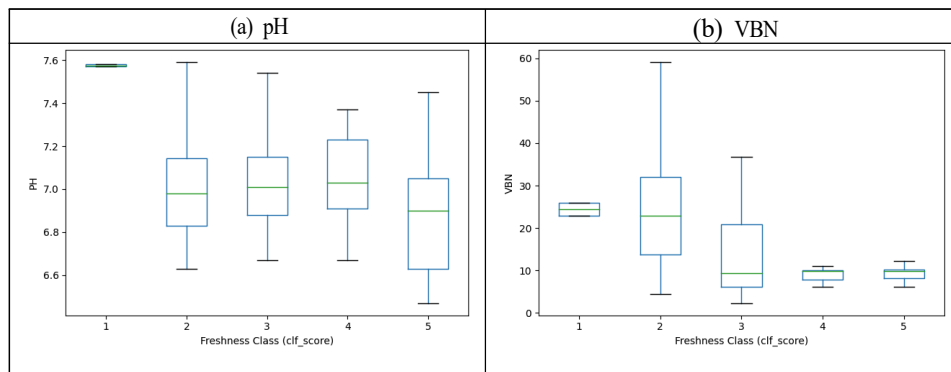


그림 13. 아귀의 신선도 등급별 화학지표 분포. (a) pH, (b) VBN

Fig. 13. Distribution of Chemical Indicators by Freshness Grade in Monkfish. (a) pH, (b) VBN

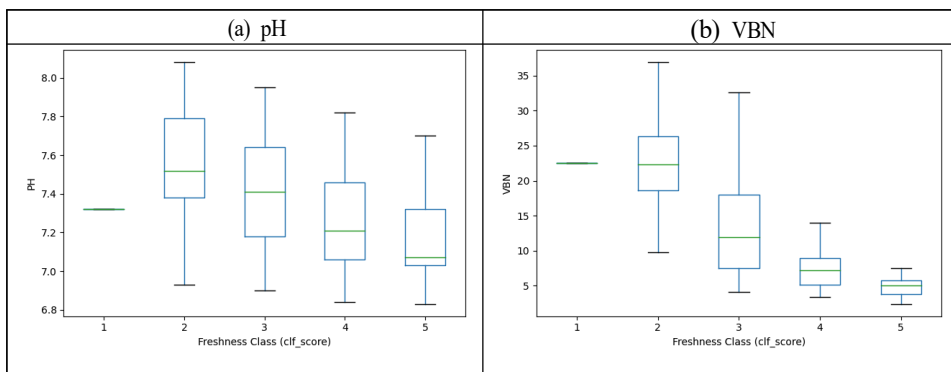


그림 14. 명태의 신선도 등급별 화학지표 분포. (a) pH, (b) VBN

Fig. 14. Distribution of Chemical Indicators by Freshness Grade in Pollock. (a) pH, (b) VBN

나타났다. 즉, 아귀에서는 낮은 신선도 등급인 등급 1-2와 상대적으로 신선한 등급인 등급 3-5 사이에서 VBN 차이가 뚜렷하게 나타났다. 그러나 등급 3의 경우 VBN 최대값이 매우 크게 나타나 표준편차가 크게 증가하였으며, 이는 일부 샘플에서 화학지표 측정값이 다른 샘플과 크게 다른 이상치 또는 경계 상태가 존재했을 가능성을 시사한다. 따라서 아귀에서는 VBN이 저신선도 구간을 구분하는 데에는 유용하지만, 일부 이상치와 등급 간 중첩을 고려할 때 단독 판별 기준으로 사용하기보다는 시각적 특징과 함께 해석하는 것이 적절하다.

그림 14는 명태의 신선도 등급별 pH 및 VBN 분포를 나타낸다. 명태의 pH는 등급 1에서 중앙값 7.32, 등급 2에서 7.52, 등급 3에서 7.41, 등급 4에서 7.21, 등급 5에서 7.07로 나타났다. 등급 2와 3에서 pH가 상대적으로 높고, 등급 4와 5에서 낮아지는 경향이 관찰되었다. 이는 명태의 경우 부패 진행 또는 저장 상태에 따라 pH가 변화할 수 있음을 보여주지만, 등급 2-4 사이의 pH 범위가 일부 겹치므로 pH만으로 신선도 등급을 명확히 구분하기는 어렵다. 반면 VBN은 등급 1과 2에서 각각 22.5 mg%, 22.3 mg%의 중앙값을 보였고, 등급 3에서는 11.9 mg%, 등급 4에서는 7.2 mg%, 등급 5에서는 5.0 mg%로 감소하였다. 즉, 명태에서는 신선도 등급이 높아질수록 VBN이 단계적으로 낮아지는 경향이 가장 뚜렷하게 나타났다. 특히 등급 2와 등급 5 사이의 VBN 차이가 크며, 이는 명태에서 VBN이 신선도 저하를 반영하는 강한 보조 지표가 될 수 있음을 보여준다. 다만 등급 1은 샘플 수가 제한적이고 단일 값에 가깝게 분포하므로, 해당 등급에 대한 해석은 제한적으로 보는 것이 적절하다.

종합적으로, 어류 5종 모두에서 VBN은 신선도 등급과 비교적 일관된 관계를 보였다. 연어, 고등어, 가자미, 아귀, 명태 모두에서 신선도 등급이 낮은 구간은 상대적으로 높은 VBN 값을 보였고, 신선도 등급이 높은 구간은 낮은 VBN 값을 보이는 경향이 확인되었다. 특히 고등어, 가자미, 명태에서는 등급이 높아질수록 VBN 중앙값이 감소하는 경향이 뚜렷하게 나타났다. 아귀 역시 등급 1-2와 등급 3-5 사이에서 VBN 차이가 명확하게 나타나, VBN이 저신선도 상태를 구분하는 데 효과적일 수 있음을 보였다. 연어

의 경우 다른 어종에 비해 등급 간 VBN 차이가 상대적으로 완만했으나, 가장 신선한 등급인 등급 5에서 낮은 VBN 분포가 확인되었다.

반면 pH는 어종에 따라 변화 양상이 다르게 나타났다. 가자미는 신선도 등급이 높아질수록 pH가 증가하는 경향을 보였고, 명태는 등급 2-3에서 상대적으로 높은 pH를 보인 뒤 등급 4-5에서 감소하는 양상을 보였다. 고등어와 아귀는 pH 분포가 등급 간에 넓게 중첩되어 신선도 등급과의 단조로운 관계가 뚜렷하지 않았다. 연어 역시 pH 변화 폭이 크지 않아 등급 구분 지표로서의 독립적인 판별력은 제한적이었다. 이러한 결과는 pH가 어종별 특성, 저장 조건, 시료 상태에 따라 다르게 변화할 수 있으며, 모든 어종에서 동일한 기준으로 신선도 등급을 구분하기에는 한계가 있음을 의미한다.

또한 화학지표 분포에서는 인접 등급 간 중첩이 반복적으로 관찰되었다. 예를 들어 연어의 등급 1-4는 VBN 분포가 상당 부분 겹쳤고, 고등어의 등급 4와 5는 VBN 중앙값이 거의 동일하였다. 가자미에서도 등급 4와 5의 VBN 분포가 유사하게 나타났으며, 아귀는 등급 3-5에서 VBN 중앙값이 거의 비슷하였다. 이는 화학지표가 신선도 상태를 반영하는 중요한 보조 정보임에도 불구하고, 화학지표만으로 모든 등급을 세밀하게 분리하는 데에는 한계가 있음을 보여준다. 특히 관능평가 기반 신선도 등급은 눈의 혼탁도, 아가미 색상, 체표 상태, 절단면 변화, 이취 등 여러 요소를 종합하여 결정되므로, 단일 화학지표와 일대일로 완전히 대응되기 어렵다.

이러한 분석 결과는 본 연구에서 제안한 멀티태스크 구조의 필요성을 뒷받침한다. 제안 시스템은 단순히 신선도 등급만을 예측하는 것이 아니라, 관심 부위 기반 시각 특징을 활용하여 신선도 등급과 pH, VBN을 함께 예측하도록 설계되었다. 실험 결과 VBN은 전반적으로 신선도 변화와 관련된 유의미한 경향을 보였으나, 인접 등급에서는 분포 중첩이 존재하였다. pH 역시 일부 어종에서 등급 변화와 관련된 양상을 보였지만, 어종별 차이가 크고 단독 판별력은 제한적이었다. 따라서 실제 현장에서 활용 가능한 신선도 평가 시스템을 위해서는 화학지표를 단독 기준으로 사용하는 것보다, 어체 전체, 눈, 아가미, 절단면 등 부위별

시각 정보와 함께 통합적으로 해석하는 구조가 필요하다. 결과적으로, 혼동행렬 분석에서 확인된 인접 등급 간 오분류는 화학지표 분포 분석에서도 유사하게 확인되었다. 즉, 인접 등급은 시각적으로도 차이가 미세할 뿐 아니라 pH와 VBN 분포 역시 중첩되는 경우가 많아, 모델이 경계 등급을 구분하는 데 어려움을 겪을 수 있다. 반면 낮은 신선도 등급과 높은 신선도 등급 사이에서는 VBN 차이가 비교적 뚜렷하게 나타나, 화학지표가 전체적인 신선도 저하 방향을 설명하는 데 유효함을 확인하였다. 따라서 제안 시스템은 시각 기반 예측 결과와 화학지표 예측 결과를 함께 제공함으로써, 단순한 등급 출력보다 더 해석 가능한 관능검사 보조 정보를 제공할 수 있다. 특히 생성형 AI 기반 설명 모델들은 이러한 시각 정보와 모델이 예측한 pH, VBN 결과를 함께 활용하여, 왜 해당 신선도 등급으로 판단되었는지에 대한 자연어 설명을 생성할 수 있다는 점에서 실제 검사자의 의사결정 지원에 활용될 수 있다.

3. 구성 요소별 Ablation Study

제안 모델의 주요 구성 요소가 신선도 분류 성능에 미치는 영향을 확인하기 위해 고등어 데이터를 대상으로 ablation study를 수행하였다. 모든 실험은 동일한 고등어 데이터 분할을 사용하였으며, train set은 7,920장, validation set은 1,977장으로 구성하였다. 데이터 분할은 개체 풀더 단위로 수행하여 동일 개체의 영상이 train set과 validation set에 동시에 포함되지 않도록 하였다. 신선도 라벨은 평가자 3인의 점수 중 중앙값을 사용하여 1 - 5등급으로 구성하였다. Validation set의 등급 분포는 2등급 144장, 3등급 1,545장, 4등급 252장, 5등급 36장으로, 3등급 시료가 다수를 차지하는 불균형 구조를 보였다. 따라서 본 실험에서는 전체 정답률인 accuracy보다 각 등급별 recall을 평균한 balanced

accuracy를 주요 지표로 사용하였다. Balanced accuracy는 다수 등급에 치우친 예측의 영향을 줄이고, 소수 등급에 대한 분류 성능까지 함께 확인할 수 있는 지표이다. 이와 함께 macro-F1과 quadratic weighted kappa(QWK)를 보조 지표로 제시하였다.

표 6은 ablation study에 사용한 비교 모델 간 주요 차이를 나타낸다. Baseline과 RoI-based 모델은 모두 ImageNet 사전학습 ResNet50^[21] 기반 5-class 분류 모델이며, 두 모델의 차이는 입력 영상의 형태이다. Baseline은 전체 이미지를 입력으로 사용하였고, RoI-based 모델은 whole body bounding box를 기준으로 crop한 이미지를 입력으로 사용하였다. Proposed model은 VFNet 기반 구조로, whole, eye, gill 영역을 검출하고 검출된 관심 부위 feature를 기반으로 신선도 등급을 예측한다. 또한 평가자 3인의 관능평가 점수를 각각 독립적인 head로 학습한 뒤, 중앙값 보팅을 통해 최종 신선도 등급을 산출한다.

표 7은 고등어 데이터셋에서 제안 모델의 구성 요소별 기여도를 확인하기 위한 ablation study 결과를 나타낸다. 실험 결과, RoI-based 모델은 Baseline과 비교하여 모든 지표에서 향상된 성능을 보였다. Whole body crop을 적용한 RoI-based 모델의 macro-F1은 0.492에서 0.593으로, QWK는 0.622에서 0.723으로, balanced accuracy는 0.588에서 0.636으로 증가하였다. 이는 전체 이미지보다 생선 영역을 중심으로 입력을 제한하는 것이 배경 영향을 줄이고 신선도 분류에 필요한 시각 정보를 더 안정적으로 활용하는 데

표 7. 고등어 데이터셋에 대한 구성 요소별 ablation study
Table 7. Component-wise ablation study on the mackerel dataset

Method	Macro-F1	QWK	Balanced Acc
Baseline	0.492	0.622	0.588
RoI-based	0.593	0.723	0.636
Proposed model	0.834	0.872	0.877

표 6. Ablation study 비교 모델의 주요 차이

Table 6. Main differences among comparison models in the ablation study

Method	Input	Model structure	RoI usage	Evaluator-specific heads	Final Prediction
Baseline	Whole image	ResNet50	X	X	5-class classification
RoI-based	Wholebody crop	ResNet50	GT whole body bbox	X	5-class classification
Proposed model	RoI features	VFNet detector-classifier	Wholebody, eye, gill	O	Median fusion of evaluator heads

도움이 되었음을 의미한다.

제안 모델은 macro-F1 0.834, QWK 0.872, balanced accuracy 0.877로 가장 높은 성능을 보였다. 특히 RoI-based 모델과 비교했을 때 macro-F1은 0.593에서 0.834로, balanced accuracy는 0.636에서 0.877로 크게 향상되었다. 이는 제안 모델이 단순히 whole body crop을 사용하는 것에 그치지 않고, VFNet 기반 관심 부위 feature와 평가자별 head 구조를 함께 활용함으로써 등급별 불균형 상황에서도 더 균형 잡힌 분류 성능을 보였음을 보여준다. 또한 QWK가 0.872로 향상된 것은 제안 모델이 신선도 등급의 순서성을 고려한 예측에서도 더 안정적인 결과를 보였음을 의미한다.

따라서 ablation study 결과는 단순한 RoI crop도 baseline 대비 일정 수준의 성능 향상에 기여하지만, 최종 성능 향상에는 VFNet 기반 관심 부위 feature 활용과 평가자별 head 구조가 더 크게 기여함을 보여준다.

4. 절단면 Grad-CAM 분석

절단면 기반 신선도 분류 모델의 판단 근거를 확인하기 위해, 어종별 정분류 사례를 대상으로 Grad-CAM 시각화 결과를 분석하였다. 모든 예시는 실제 신선도 등급과 모델 예측 등급이 일치한 사례이며, 모델이 각 신선도 등급을 예측할 때 절단면의 어느 영역을 주요 근거로 활용했는지 확인하기 위해 사용하였다.

그림 15는 아귀 절단면에 대한 Grad-CAM 시각화 결과를 나타낸다. 아귀의 경우 중심 뼈를 기준으로 양쪽 근육 조직에 activation이 비교적 넓게 분포하였다. 2점 사례에서는 중심부 주변의 어육 색이 탁하고 균일성이 떨어지는 영역에 heatmap이 집중되었으며, 5점 사례에서도 중심 뼈 주변의 근육 조직과 혈색 변화가 나타나는 영역에 강한 반응이 확인되었다. 아귀는 절단면 내에서 눈이나 아가미처럼 명확한 세부 기관이 드러나지 않고 근육 조직의 비중이 크

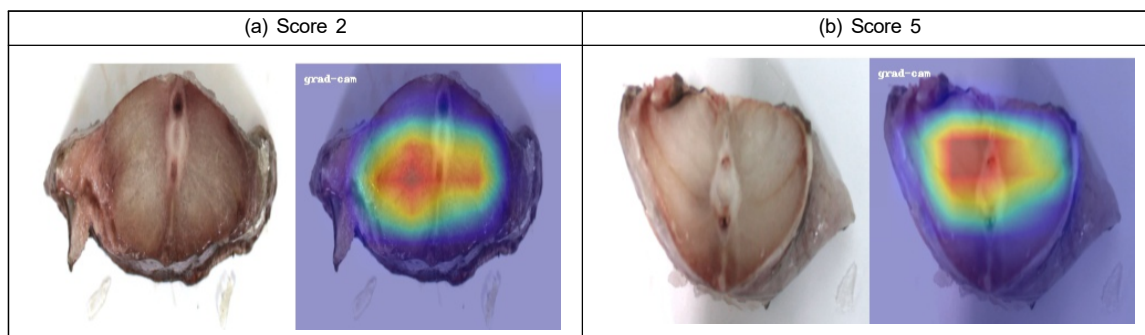


그림 15. 아귀 절단면 사례에 대한 Grad-CAM 시각화 결과. (a) 2점, (b) 5점

Fig. 15. Grad-CAM visualization results for Monkfish cut-surface images. (a) Score 2, (b) Score 5

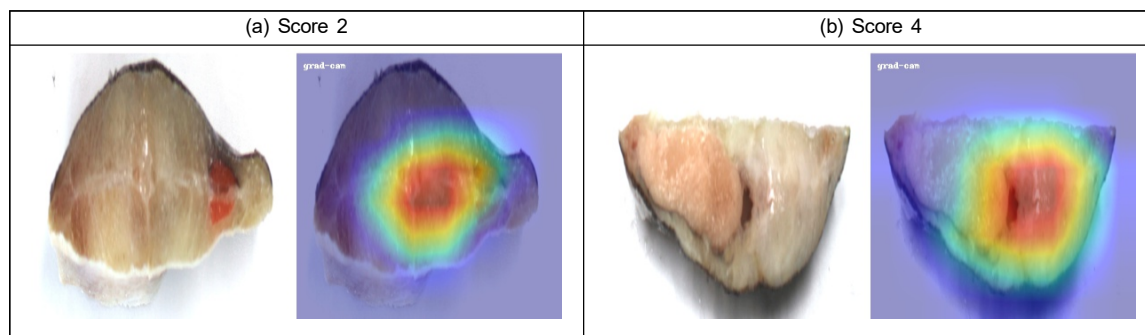


그림 16. 가자미 절단면 사례에 대한 Grad-CAM 시각화 결과. (a) 2점, (b) 4점

Fig. 16. Grad-CAM visualization results for Flounder cut-surface images. (a) Score 2, (b) Score 4

기 때문에, 모델은 절단면 내부의 육색, 중심부 주변의 혈색, 조직 표면의 탁도와 같은 정보를 종합적으로 활용한 것으로 해석된다.

그림 16은 가자미 절단면에 대한 Grad-CAM 시각화 결과를 나타낸다. 가자미의 경우 activation이 절단면 중앙부의 근육 조직을 중심으로 형성되었다. 가자미 절단면은 다른 어종에 비해 밝은 색의 근육 조직이 넓게 분포하고, 표면 광택과 수분감이 비교적 두드러지게 나타난다. 2점 사례에서는 중앙의 진한 색 변화 영역을 중심으로 heatmap이 집중되었으며, 4점 사례에서는 중심부의 절단면 조직과 그 주변 근육 영역에 activation이 형성되었다. 이 결과는 모델이 가자미 절단면에서 장기 조직보다 어육 표면의 색상 균일성, 광택, 수분감, 조직 결 변화를 중심으로 반응했음을 보여준다.

그림 17은 고등어 절단면에 대한 Grad-CAM 시각화 결과를 나타낸다. 고등어의 경우 내장 조직 또는 혈합육 주변

의 어두운 변색 영역에서 activation이 뚜렷하게 나타났다. 고등어 절단면은 흰 근육 조직과 어두운 조직이 함께 존재하며, 신선도 저하에 따라 혈합육 주변의 색 변화나 내장 부위의 변성이 상대적으로 두드러질 수 있다. 2점 사례에서는 어육 중앙부에서 좌측으로 이어지는 어두운 변색 영역을 따라 heatmap이 형성되었고, 4점 사례에서는 절단면 중앙의 붉은 조직과 주변 근육 영역에 activation이 집중되었다. 이를 통해 고등어에서는 혈색 변화, 내장 주변 조직 상태, 중심부 변색 영역이 주요 판별 단서로 작용했음을 확인할 수 있다.

그림 18은 명태 절단면에 대한 Grad-CAM 시각화 결과를 나타낸다. 명태의 경우 장기 조직 또는 내장 잔여물이 포함된 중앙부에 activation이 강하게 집중되었다. 명태는 냉동 및 해동 과정에서 표면 건조, 색상 탁도, 수분감 변화가 함께 나타날 수 있으며, 특히 장기 주변부의 색 변화가 신선도 판단에 중요한 시각 단서가 될 수 있다. 2점 사례에

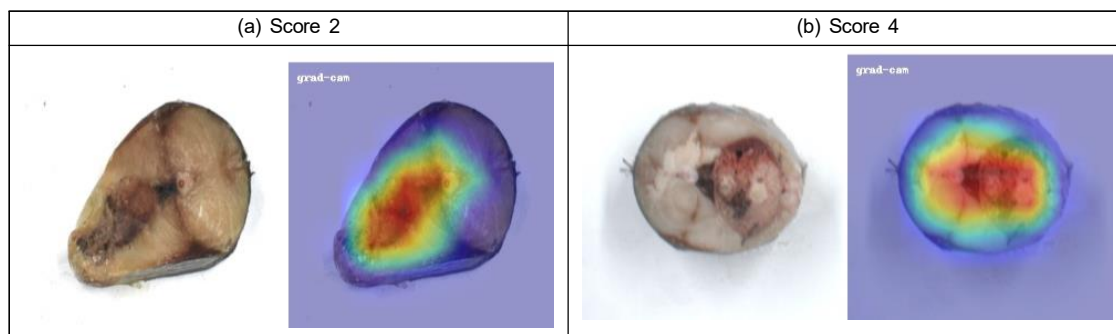


그림 17. 고등어 절단면 사례에 대한 Grad-CAM 시각화 결과. (a) 2점, (b) 4점

Fig. 17. Grad-CAM visualization results for Mackerel cut-surface images. (a) Score 2, (b) Score 4

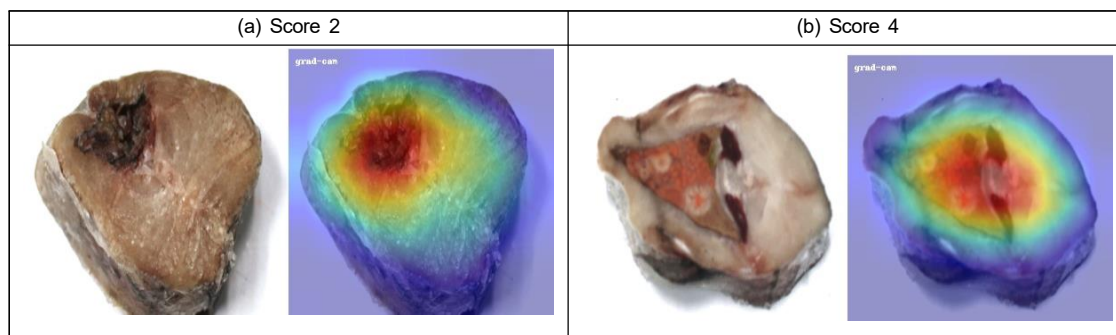


그림 18. 명태 절단면 사례에 대한 Grad-CAM 시각화 결과. (a) 2점, (b) 4점

Fig. 18. Grad-CAM visualization results for Pollock cut-surface images. (a) Score 2, (b) Score 4

서는 중앙의 진한 갈색 장기 조직과 그 주변 어육에 heatmap이 집중되었고, 4점 사례에서는 절단면 내부의 붉은 조직과 주변 근육 영역에 activation이 넓게 형성되었다. 즉, 명태에서는 외곽 형태보다 내부 조직의 색 변화와 장기 주변 상태가 모델 판단에 더 크게 반영된 것으로 보인다.

그림 19는 연어 절단면에 대한 Grad-CAM 시각화 결과를 나타낸다. 연어의 경우 중심 뼈 주변과 붉은 근육 조직의 색상 변화 영역에 activation이 집중되었다. 연어는 어육 자체의 붉은색이 강하고 근섬유 결이 뚜렷하기 때문에, 색상 균일성, 혈색 변화, 중심부 주변 조직의 선명도가 중요한 단서로 작용할 수 있다. 1점 사례에서는 절단면 중앙부의 불균일한 조직과 색 변화 영역에 강한 activation이 나타났으며, 4점 사례에서도 중심 뼈 주변과 인접 근육 조직에 heatmap이 집중되었다. 특히 4점 사례에서는 어육 전반의 붉은색이 비교적 선명하게 유지되는 가운데, 모델의 activation이 중심부 조직에 집중되어 단순한 전체 색상보다 중심

부 조직 상태를 함께 참조하는 것으로 해석된다. 신선도 등급 측면에서 보면, 낮은 등급 사례에서는 변색, 탁도, 조직 손상, 내장 주변의 어두운 영역에 activation이 강하게 나타났고, 높은 등급 사례에서는 어육 색이 비교적 선명하고 조직 경계가 유지된 중심 뼈 주변 또는 주요 근육 조직에 activation이 형성되었다. 이는 모델이 배경, 외곽 형태, 조명 차이에만 의존하지 않고, 어종별 절단면 구조와 등급별 시각 변화를 함께 반영하여 어육 색상, 조직 변성, 수분감, 장기 주변 변색, 중심 뼈 주변의 혈색 변화와 같은 관능검사 관련 시각 단서를 활용하고 있음을 보여준다.

5. 생성형 AI 기반 선도 설명문 생성

생성형 AI 기반 설명 모듈의 동작을 확인하기 위해, 비전 모델의 예측 결과와 부위별 신선도 예측값, pH 및 VBN 화학지표, 평가자 관능평가 결과, 관능검사 기준표를 함께 입

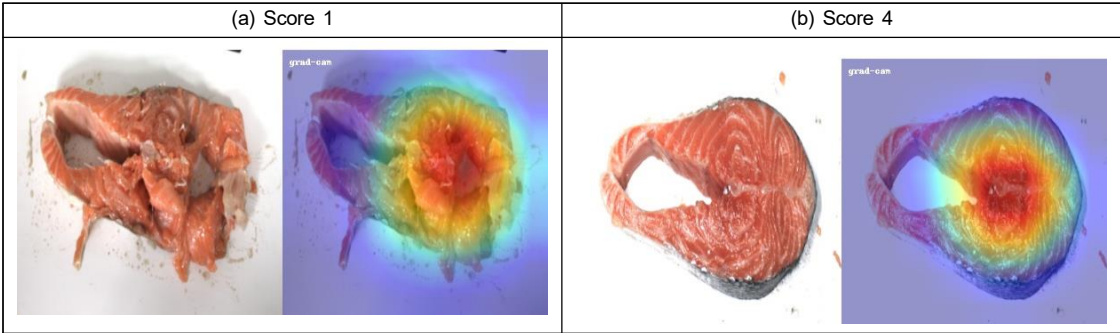


그림 19. 연어 절단면 사례에 대한 Grad-CAM 시각화 결과. (a) 1점, (b) 4점
Fig. 19. Grad-CAM visualization results for Salmon cut-surface images. (a) Score 1, (b) Score 4

표 8. 생성형 AI 기반 신선도 설명문 생성 예시
Table 8. Example of generative AI-based freshness explanation

Category	Content
Initial vision-based freshness prediction	3/5
Part-level median predictions	Eye: 3, Gill: 3, Whole body: 3
Chemical indicators	pH: 6.44, VBN: 16.7 mg/100g
Integrated evaluator GT	3/5
Auxiliary freshness score	3/5
Generated explanation	The vision model's freshness prediction of 3 points agrees with the median score of the evaluators. The eye corresponds to score 3 because it is flat and the cornea appears milky white. The gill is evaluated as score 3 because its color has gradually faded and a large amount of mucus is present. The skin surface is also evaluated as score 3 because the surface color is faded, turbid, and milky. The measured pH of 6.44 and VBN of 16.7 mg/100g are consistent with freshness score 3.

력하여 보조 판정값과 설명문을 생성하였다. 본 실험에서 생성형 AI 모듈은 단순히 신선도 점수를 문장으로 변환하는 방식이 아니라, 입력된 예측값과 관능검사 기준을 함께 참조하여 보조 판정점수, 비전 모델과 평가자 결과의 일치 여부, 주요 관능 근거, 주요 화학 근거, 자연어 설명문을 구조화된 형태로 출력하도록 구성하였다.

표 8은 대표 시료에 대한 생성형 AI 설명 결과를 나타낸다. 해당 시료에서 비전 모델의 1차 신선도 예측값은 3점이었으며, 부위별 예측값 역시 눈, 아가미, 어체 전체에서 모두 3점으로 나타났다. 또한 pH는 6.44, VBN은 16.7 mg/100g으로 입력되었고, 평가자 3인의 관능평가 결과도 모두 3점으로 일치하였다. 생성형 AI 모듈은 이러한 정보를 바탕으로 보조 판정 점수를 3점으로 제시하였으며, 비전 모델의 예측과 평가자 통합값이 일치한다고 판단하였다.

이 결과에서 중요한 점은 생성형 AI가 단순히 3점이라는 결과만 출력하지 않고, 해당 판정의 근거를 관능검사 항목과 연결하여 설명했다는 것이다. 예를 들어 눈 항목에 대해서는 평면하고 각막이 유백색을 띠는 기준을 근거로 3점 판단을 설명하였고, 아가미 항목에 대해서는 색이 점차 흐려지고 점액이 많이 묻어남이라는 기준을 제시하였다. 또한 표피 항목에 대해서는 표면 색상이 퇴색되어 혼탁하고 우유빛을 띠는 기준을 활용하여 비전 모델의 3점 예측이 관능 기준과 부합함을 설명하였다. 즉, 생성 결과는 모델 예측값을 그대로 반복하는 수준이 아니라, 부위별 시각 단서와 관능검사 기준표를 연결하여 해석 가능한 판정 근거를 제공하였다.

화학지표 측면에서도 pH 6.44와 VBN 16.7 mg/100g을 함께 제시함으로써, 시각적 판단 근거와 정량 지표가 동일한 신선도 수준을 지지하는지 확인할 수 있도록 하였다. 특히 VBN은 앞선 화학지표 분포 분석에서 신선도 저하와 관련된 경향을 보였기 때문에, 생성형 AI 설명문에 VBN 값을 함께 포함하는 것은 검사자가 모델 판단을 검토하는 데 도움이 된다.

따라서 표 8의 결과는 제안한 생성형 AI 설명 모듈이 비전 모델의 예측값, 부위별 판단 결과, 화학지표, 평가자 결과, 관능검사 기준을 하나의 설명문으로 통합할 수 있음을 보여준다. 이러한 출력 방식은 기존의 단순 분류 모델이 제공하는 등급 결과보다 해석 가능성이 높으며, 실제 검사자

가 모델의 판단 근거를 확인하고 최종 판정을 검토하는 데 활용될 수 있다. 특히 비전 모델 예측값과 평가자 결과가 일치하는 경우에는 모델 판단의 신뢰성을 설명할 수 있고, 향후 두 결과가 불일치하는 경우에는 어떤 부위 또는 화학 지표에서 차이가 발생했는지를 설명하는 방식으로 확장할 수 있다.

대표 생성 사례를 대상으로 평가한 결과, 생성형 AI 설명문은 입력된 신선도 점수 및 관능검사 기준표와 대체로 일관된 설명을 생성하였다. 특히 눈, 아가미, 표피, 질단면 등 부위별 점수와 기준표 문항을 연결하여 설명하는 경향을 보였다. 또한 입력으로 제공되지 않은 보관 이력이나 식용 가능 여부를 임의로 단정하지 않도록 제한한 결과, 검토한 사례에서는 이러한 표현이 두드러지게 나타나지 않았다. 다만 일부 사례에서는 pH/VBN 예측값만으로 등급을 단정하는 듯한 표현이 나타날 수 있어, 최종 설명문에서는 화학 지표를 단독 판정 기준이 아닌 보조 지표로 서술하도록 프롬프트를 제한하였다.

V. 결 론

본 논문에서는 비전 기반 신선도 분석과 생성형 AI 기반 판정 보조 및 설명 생성을 결합한 수산물 관능검사 지원 및 신선도 평가 시스템을 제안하였다. 제안 시스템은 색상 보정판 기반 전처리, VFNet 기반 관심 부위 검출, 신선도 등급 및 화학지표(pH, VBN) 동시 예측을 통합하고, 입력 생선 이미지, 모델 예측값, 다중 평가자 관능평가 결과, 관능검사 채점표를 함께 활용하여 비전 모델의 신선도 예측 결과에 대해 관능검사 기준 기반의 보조 설명문을 생성하도록 구성하였다. 이를 통해 기존의 단순 등급 분류 결과 제시를 넘어, 검사 기준에 부합하는 자연어 설명과 시각적 근거를 함께 제공할 수 있는 관능검사 지원 구조를 구현하였다.

실험 결과, 제안 시스템은 연어 82.98%, 아귀 84.87%, 고등어 87.77%, 가자미 87.33%, 명태 85.42%의 최종 정확도를 보였으며, 5종 평균 85.67%의 신선도 분류 성능을 달성하였다. 어종별 혼동행렬 분석에서는 모든 어종에서 대각선 성분이 우세하게 나타나 전반적으로 안정적인 등급

판별이 가능함을 확인하였다. 다만 오분류는 주로 2점과 3점, 4점과 5점과 같은 인접 등급 사이에서 발생하였다. 이는 신선도 변화가 명확히 분리된 이산적 단계라기보다 연속적으로 진행되는 품질 변화이며, 실제 관능검사에서도 경계 등급에서 평가자 간 해석 차이가 발생할 수 있음을 보여준다.

또한 신선도 등급별 pH 및 VBN 분포를 분석한 결과, VBN은 전반적으로 신선도 저하와 관련된 경향을 보였으며, 특히 고등어, 가자미, 명태에서 신선도 등급이 높아질수록 VBN이 감소하는 양상이 비교적 뚜렷하게 나타났다. 반면 pH는 어종별로 변화 양상이 다르게 나타났고, 일부 등급에서는 인접 등급 간 분포가 크게 중첩되었다. 이러한 결과는 pH와 VBN이 신선도 판단에 유의미한 보조 지표로 활용될 수 있으나, 화학지표만으로는 세밀한 등급 구분에 한계가 있음을 의미한다. 따라서 어체 전체, 눈, 아가미, 절단면 등 부위별 시각 정보와 정량 지표를 함께 활용하는 멀티태스크 구조가 실제 관능검사 지원에 적합함을 확인하였다.

본 연구의 의의는 관능평가자 기반 다중 평가자 라벨, 부위 annotation, pH 및 VBN 화학지표를 함께 포함하는 데이터 구축 절차를 정립하고, 이를 기반으로 시각 정보와 정량 지표를 동시에 활용하는 통합형 신선도 평가 구조를 제시하였다는 점에 있다. 특히 생성형 AI를 활용하여 모델 예측값, 관능검사 기준, 다중 평가자 결과를 연결하는 설명 계층을 도입함으로써, 판정 결과를 검사자가 이해하고 검토할 수 있는 형태로 제공할 수 있음을 보였다. 이는 향후 수입 수산물 관능검사 현장에서 검사자의 최종 의사결정을 보조하는 설명 가능한 인공지능 시스템으로 활용될 가능성을 보여준다.

향후 연구에서는 화학지표의 예측값과 분포 양상을 함께 활용하여 경계 등급에서의 최종 신선도 값을 보정하는 방향으로 시스템을 확장하고자 한다. 구체적으로는 시각 기반 모델의 신선도 예측 결과와 pH, VBN 예측값 간의 일관성을 함께 고려하여, 인접 등급 사이에서 불확실성이 높은 샘플에 대해 보정 규칙 또는 후처리 모듈을 적용할 수 있다. 또한 생성형 AI가 제공하는 설명문의 일관성, 검사 기준 부합성, 실제 검사자 활용성을 정량적으로 평가하는 절차를 추가하여, 단순 예측 성능뿐 아니라 현장 적용 가능성까지

검증하는 방향으로 연구를 고도화할 예정이다.

참 고 문 헌 (References)

- [1] H. Zhang, Y. Wang, F. Dayoub, and N. Sunderhauf, "VarifocalNet: An IoU-Aware Dense Object Detector," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.8514-8523, June 2021.
doi: <http://doi.org/10.1109/CVPR46437.2021.00841>
- [2] H. H. Huss, *Quality and Quality Changes in Fresh Fish*, FAO Fisheries Technical Paper No. 348, Food and Agriculture Organization of the United Nations, Rome, 1995. URL: <https://www.fao.org/4/v7180e/v7180e00.htm>
- [3] Codex Alimentarius Commission, *Guidelines for the Sensory Evaluation of Fish and Shellfish in Laboratories*, CXG 31-1999, FAO/WHO, Rome, 1999. URL: https://www.fao.org/fao-who-codexalimentarius/sh-proxy/ar/?lnk=1&url=https%253A%252F%252Fworkspace.fao.org%252Fsites%252Fcodex%252FStandards%252FCXG%2B31-1999%252FCXG_031e.pdf
- [4] Ministry of Food and Drug Safety, *Food Standards Codex Online Service*, Ministry of Food and Drug Safety, Cheongju, Republic of Korea, accessed May 18, 2026. URL: <https://www.foodsafetykorea.go.kr/foodcode/index.jsp>
- [5] J. Fang, X. Qiu, Z. Wan, Q. Zou, K. Su, N. Hu, and P. Wang, "A Sensing Smartphone and Its Portable Accessory for On-Site Rapid Biochemical Detection of Marine Toxins," *Analytical Methods*, Vol.8, No.38, pp.6895-6902, 2016.
doi: <http://doi.org/10.1039/C6AY01384H>
- [6] Y. Zhang, Q. Luo, K. Ding, S. G. Liu, and X. Shi, "A Smartphone-Integrated Colorimetric Sensor of Total Volatile Basic Nitrogen (TVB-N) Based on Au@MnO₂ Core-Shell Nanocomposites Incorporated into Hydrogel and Its Application in Fish Spoilage Monitoring," *Sensors and Actuators B: Chemical*, Vol.335, Art. no. 129708, 2021.
doi: <http://doi.org/10.1016/j.snb.2021.129708>
- [7] L. Jiang, H. Ye, D. Ma, J. Rodrigues, R. Sheng, and D. Min, "A Smartphone-Adaptable Fluorescent Sensing Tag for Non-Contact and Visual Monitoring of the Freshness of Fish," *Analyst*, Vol.147, No.5, pp.923-931, 2022.
doi: <http://doi.org/10.1039/D1AN02191E>
- [8] T. Mastnak, G. J. Mohr, and M. Finšgar, "The Use of a Novel Smartphone Testing Platform for the Development of Colorimetric Sensor Receptors for Food Spoilage," *Analytical Methods*, Vol.15, No.13, pp.1700-1712, 2023.
doi: <http://doi.org/10.1039/D2AY02082C>
- [9] H. Mohammadi Lalabadi, M. Sadeghi, and S. A. Mireei, "Fish Freshness Categorization from Eyes and Gills Color Features Using Multi-Class Artificial Neural Network and Support Vector Machines," *Aquacultural Engineering*, Vol.90, Art. no. 102076, 2020.
doi: <http://doi.org/10.1016/j.aquaeng.2020.102076>

- [10] A. Taheri-Garavand, A. Nasiri, A. Banan, and Y.-D. Zhang, "Smart Deep Learning-Based Approach for Non-Destructive Freshness Diagnosis of Common Carp Fish," *Journal of Food Engineering*, Vol.278, Art. no. 109930, 2020.
doi: <http://doi.org/10.1016/j.jfoodeng.2020.109930>
- [11] M. B. Yildiz, E. T. Yasin, and M. Koklu, "Fisheye Freshness Detection Using Common Deep Learning Algorithms and Machine Learning Methods with a Developed Mobile Application," *European Food Research and Technology*, Vol.250, No.7, pp.1919-1932, 2024.
doi: <http://doi.org/10.1007/s00217-024-04493-0>
- [12] M. Jeong, S. Kim, J. Lee, H. Choo, H. Lee, and W.-S. Cheong, "Comparison of Artificial Intelligence Multitask Performance using Object Detection and Foreground Image," *Journal of Broadcast Engineering*, Vol.27, No.3, pp.308-317, May 2022.
doi: <http://doi.org/10.5909/JBE.2022.27.3.308>
- [13] S. Lim and G. Park, "Development of Python-based Annotation Tool Program for Constructing Object Recognition Deep-Learning Model," *Journal of Broadcast Engineering*, Vol.25, No.3, pp.386-398, May 2020.
doi: <http://doi.org/10.5909/JBE.2020.25.3.386>
- [14] J. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.8, No.6, pp.679-698, 1986.
doi: <http://doi.org/10.1109/TPAMI.1986.4767851>
- [15] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed., Pearson, New York, 2018. URL: <https://books.google.co.jp/books?id=0F05vgAACAAJ>
- [16] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge University Press, Cambridge, UK, 2004. URL: <https://www.robots.ox.ac.uk/~vgg/hzbook/>
- [17] T.Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature Pyramid Networks for Object Detection," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.2117-2125, July 2017.
doi: <http://doi.org/10.1109/CVPR.2017.106>
- [18] A. Neubeck and L. Van Gool, "Efficient Non-Maximum Suppression," *Proceedings of the 18th International Conference on Pattern Recognition (ICPR)*, pp.850-855, Aug. 2006.
doi: <http://doi.org/10.1109/ICPR.2006.479>
- [19] S. Zhang, C. Chi, Y. Yao, Z. Lei, and S. Z. Li, "Bridging the Gap Between Anchor-Based and Anchor-Free Detection via Adaptive Training Sample Selection," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.9759-9768, June 2020.
doi: <http://doi.org/10.1109/CVPR42600.2020.00978>
- [20] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp.618-626, October 2017.
doi: <http://doi.org/10.1109/ICCV.2017.74>
- [21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770 - 778.
doi: <https://doi.org/10.1109/CVPR.2016.90>

저 자 소 개



백 채 운

- 2025년 2월 : 가천대학교 AI학과 학사
- 2025년 ~ 현재 : 중앙대학교 가상융합대학원 석사과정
- ORCID : <https://orcid.org/0009-0002-5219-935X>
- 주관심분야 : 자율주행, 로봇 비전, 멀티모달 AI



윤 종 수

- 2023년 8월 : 중앙대학교 AI학과 석사
- 2023년 ~ 현재 : 중앙대학교 첨단영상대학원 영상공학 박사과정
- ORCID : <https://orcid.org/0009-0006-7346-4182>
- 주관심분야 : 멀티모달, 컴퓨터 비전, AI+X

저 자 소 개



배 재 형

- 2025년 8월 : 중앙대학교 영상학과 석사
- 2025년 9월 ~ 현재 : 중앙대학교 영상학과 박사과정
- ORCID : <https://orcid.org/0009-0002-4809-327X>
- 주관심분야 : 의료영상, 멀티모달, VLM



황 준 엽

- 2025년 8월 : 한양대학교 연극영화학 학사
- 2025년 ~ 현재 : 중앙대학교 가상융합대학원 석사과정
- ORCID : <https://orcid.org/0009-0009-1467-5029>
- 주관심분야 : 컴퓨터 비전, 퍼지컬 AI, 생성형 AI



최 종 원

- 2020년 3월 ~ 2024년 2월 : 중앙대학교 첨단영상대학원 조교수
- 2024년 3월 ~ 현재 : 중앙대학교 첨단영상대학원 부교수
- ORCID : <https://orcid.org/0000-0001-9753-8760>
- 주관심분야 : 딥페이크, 컴퓨터 비전