

# 성능 평가 결과서

## 용왕의 한 수

이미지 기반 낚시 환경 분석과 음성 기반 조행 기록을 결합한 스마트 낚시 플랫폼

소속 : LG U+ why not sw camp 8기  
팀명 : NAVIS  
팀장 : 권혜린  
팀원 : 박시은, 오원석, 윤예빈, 조승아

---

&lt; 프로젝트 기획서 &gt;

작성일자 : 2026-01-08

---

<b>I. 성능 평가 기준 및 결과.....</b>	<b>3</b>
1. 성능 지표 설명.....	3
2. 실제 성능.....	3
<b>II. 성능 평가 결과 분석.....</b>	<b>8</b>
1. 결과 분석.....	8
2. 기술적 특징 요약.....	8
<b>III. 결론.....</b>	<b>9</b>
1. 최종 모델.....	9
2. 향후 과제.....	9

< 프로젝트 기획서 >

작성일자 : 2026-01-08

## I. 성능 평가 기준 및 결과

### 1. 성능 지표 설명

- **BERTScore (Precision, Recall, F1):** sLLM 모델의 성능을 평가하기 위해 사용됨. 문맥적 유사성을 바탕으로 정밀도(Precision), 재현율(Recall), F1 Score를 측정함.
- **Accuracy (정확도):** 분류 모델(예기색, 물색)이 정답 레이블을 얼마나 정확하게 맞혔는지 측정하는 지표
- **Recall(재현율):** 실제 정답인 것들 중 모델이 정답이라 예측 한 것을 맞춘 비율
- **F1-score:** Precision(정밀도)와 Recall(재현율)의 조화 평균

&lt; 프로젝트 기획서 &gt;

작성일자 : 2026-01-08

---

## 2. 실제 성능

- yolo모델 버전 별 성능 비교:

모델 버전	precision	recall	mAP50
yolov8n(채택)	57.4	45.3	48.3
yolov8s	54.9	39.6	43.3
yolov11s	56.9	44.3	45.3
yolov11m	53.2	40.1	43.2

&lt; 프로젝트 기획서 &gt;

작성일자 : 2026-01-08

- **CNN 분류 모델 성능**

- **예기색 분류:** 이미지만 사용했을 때보다 이미지와 메타데이터를 **Concat**하여 병렬 처리했을 때 더 높은 성능을 기록함.

- **이미지만 사용한 결과**

	precision	recall	f1-score	support
<b>blue</b>	0.00	0.00	0.00	5
<b>brown</b>	0.75	0.50	0.60	12
<b>green</b>	0.50	0.94	0.66	90
<b>orange</b>	0.00	0.00	0.00	4
<b>pink</b>	0.00	0.00	0.00	8
<b>purple</b>	0.35	0.30	0.32	27
<b>rainbow</b>	0.00	0.00	0.00	6
<b>red</b>	0.00	0.00	0.00	47
<b>yellow</b>	0.00	0.00	0.00	1
<b>accuracy</b>			0.49	200
<b>macro avg</b>	0.18	0.19	0.18	200
<b>weighted avg</b>	0.32	0.49	0.37	200

&lt; 프로젝트 기획서 &gt;

작성일자 : 2026-01-08

---

### ■ 이미지와 메타데이터를 사용한 결과

	<b>precision</b>	<b>recall</b>	<b>f1-score</b>	<b>support</b>
<b>blue</b>	0.20	1.00	0.33	5
<b>brown</b>	0.46	0.92	0.61	12
<b>green</b>	0.71	0.49	0.58	90
<b>orange</b>	0.36	1.00	0.53	4
<b>pink</b>	0.33	0.62	0.43	8
<b>purple</b>	0.69	0.74	0.71	27
<b>rainbow</b>	0.27	0.50	0.35	6
<b>red</b>	0.69	0.19	0.30	47
<b>yellow</b>	0.00	0.00	0.00	1
<b>accuracy</b>			0.51	200
<b>macro avg</b>	0.41	0.61	0.43	200
<b>weighted avg</b>	0.64	0.51	0.51	200

&lt; 프로젝트 기획서 &gt;

작성일자 : 2026-01-08

- 
- **물색 분류** : ResNet 기반 모델이 EfficientNetB0보다 질감 표현에서 우수한 성능을 보임.

	precision	recall	f1-score	support
<b>clear</b>	0.69	0.72	0.71	97
<b>medium</b>	0.49	0.60	0.54	97
<b>muddy</b>	0.71	0.51	0.59	89
<b>accuracy</b>			0.61	283
<b>macro avg</b>	0.63	0.61	0.61	283
<b>weighted avg</b>	0.63	0.61	0.61	283

&lt; 프로젝트 기획서 &gt;

작성일자 : 2026-01-08

- sLLM 모델 성능 비교 (BERTScore):

성능 비교 항목	Polyglot-1.3B (Epoch 3)	Polyglot-1.3B (Epoch 5)	Llama-3.2-3B
Precision	0.6480	0.6513	<b>0.6570</b>
Recall	0.7427	0.7472	<b>0.7618</b>
F1	0.6919	0.6957	<b>0.7054</b>

## II. 성능 평가 결과 분석

### 1. 결과 분석

- **sLLM: Llama-3.2-3B**가 모든 수치에서 미세하게 가장 높게 나왔지만, 한국어 전용이 아니기 때문에 다양한 언어로 답변하는 특성이 나타남. 따라서 한국어 성능이 안정적인 **Polyglot-1.3B (Epoch 5)**를 최종 모델로 선정함.
- **CNN(물색)**: 물색 모델은 **ResNet**이 데이터와 적합하지 않았기 때문에, 레이블이 적었음에도 성능이 낮게 측정됨. 이를 개선하기 위해 **Linear Layer**를 1~3개로 확장하여 실험을 진행함.
- **CNN(에기색)**: 이미지(비정형데이터)와 메타데이터(정형데이터)를 Concat하여 병렬 처리했을 때, **이미지만 사용했을 때(F1-score 0.37)**보다 더 높은 **성능(F1-score 0.51)**을 기록함. 이미지 데이터만으로는 구분이 어려운 특징들을 **메타데이터**가 효과적으로 보완하여 분류 정확도를 유의미하게 향상시켰기에, 이를 최종 모델 아키텍처로 채택함.

### 2. 기술적 특징 요약

- **Data Flow**: CNN 결과물인 **Flatten** (16, 1024) 데이터를 **Linear Layer**를 통해 최종 출력 형태인 (16, 3)으로 변경하여 분류를 수행함.
- **Overfitting 방지**: **Dropout**은 모델이 특정 값에 집착하지 않고 일반적인 규칙을 배우도록 강제하여 테스트 데이터에서의 성능을 높여줌.
- **Feature 압축**: Dense 64를 통해 중요 이미지(비정형데이터)와 기상요소(정형데이터)를 결합한 144 레이어를 64개로 중요 정보만 압축시켜 성능 향상

### III. 결론

#### 1. 최종 모델

- 모바일기기와 같은 저사양 임베디드 장치에서 실시간 추론이 가능하기에 **yolov8n** 채택
- 정형 데이터(CSV)와 비정형 데이터(이미지)를 동시에 학습할 수 있는 **하이브리드 CNN 구조**를 자체 설계함
- 바닷물의 색감과 질감을 고려하여 Pretrained 된 **Resnet50**을 채택
- 한국어 특화 성능과 효율성을 고려하여 **Polyglot-1.3B (Epoch 5)**를 채택

#### 2. 향후 과제

- **ResNet**과 같은 전이 학습 모델을 사용할 때는 도메인 데이터에 더 최적화된 **Fine-tuning**과 레이어 구조 확장이 필요함.
- 추가적인 데이터로 **LoRA 튜닝**을 통해 두족류 낚시 뿐만 아니라 **다양한 바다 낚시로 확장 가능**