



다금바리를 찾아서

이미지 딥러닝 기반의 어종 분석 프로젝트
- 빅데이터 시각화 프로젝트 3팀

김경준 / 양상욱 / 임수진 / 정민욱 / 채지연

목차

Step 01

프로젝트 개요

- 주제 선정 배경
- 팀 구성원 및 역할

Step 02

프로젝트 수행 절차 및 개발 환경

- 프로젝트 수행 절차
- 개발 환경

Step 03

프로젝트 수행 및 결과

- 시스템 구성도
- 이미지 처리 프로세스
- ML 시스템 구성도
- 학습 결과
- 시스템 구현
- 어플 시연

Step 04

프로젝트 평가

- 개선방향
- 프로젝트를 마치며



Step 01

프로젝트 개요

- 주제 선정 배경
- 팀 구성원 및 역할

1.1 주제 선정 배경



[단독]tvN '서울춘놈' 다금바리 가짜 논란.."유의하겠다"

또 어류 칼럼니스트 김지민은 자신의 유튜브 채널 '입질의 추억 TV'를 통해 '서울춘놈'에서 등장한 다금바리가 가짜라는 제보가 있었다면서 진위 여부 가리기에 나...

- | | | |
|------------------------------------|--------|---------------------|
| '서울춘놈' 다금바리 가짜 의혹→tvN 측 "사실 확인..." | 셀럽미디어 | 2020.08.18. |
| '서울춘놈' 다금바리 진위여부 논란 왜 불거... | 머니S | 2020.08.18. 네이버뉴스 |
| 어류 칼럼니스트 김지민 '서울춘놈' 다금바... | 뉴스엔 | 2020.08.18. 네이버뉴스 |
| '서울춘놈' 가짜 다금바리 의혹..., tv... | 스포티비뉴스 | 2020.08.18. 네이버뉴스 |

제주 다금바리, 속지않고 먹는 확실한 방법(능성어 구별해내기)

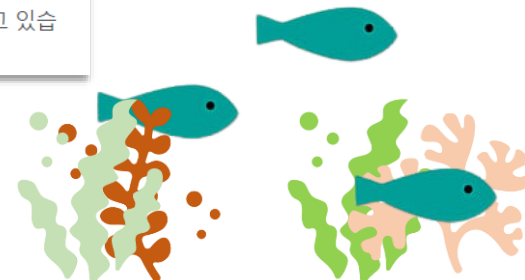
2013. 2. 1. — 보시다시피 체형을 보면 능성어쪽이 좀 더 뚱뚱하며 제주 다금바리는 약간 날씬한 편입니다. 가장 큰 차이점은 무늬에 있는데요. 능성어는 일곱줄의 가로 ...

제주 다금바리와 능성어의 확실한 구별법 - 블로그

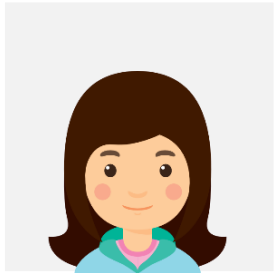
2016. 1. 11. — 다금바리(표준명 자바리)를 사칭하는 대표적인 유사 어종은 능성어(제주 방언 구문쟁이)와 민다금바리가 있는데 아래 내용을 참고하자. . 1) 능성어 - ...

전문가도 헷갈려 하는 고급 생선회 다금바리의 실체

우리가 알고 있는 다금바리, 언론사에서 언급되는 다금바리는 모두 표준명 자바리를 뜻하고 있습니다. 다금바리란 생선을 들어본 사람은 많아요



1.2 팀 구성원 및 역할



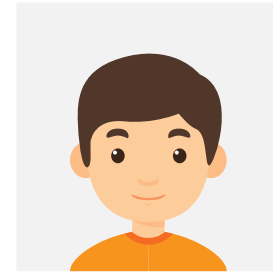
팀장
임수진

- 데이터 수집
- 데이터 정제/가공
- 분석 결과 시각화



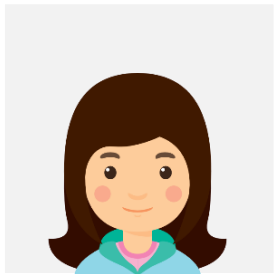
팀원
정민욱

- 데이터 수집
- 딥러닝 Model 구축
- 분석 결과 시각화



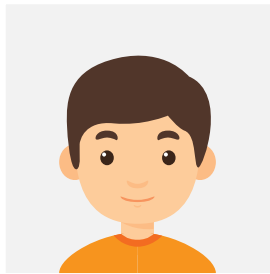
팀원
양상욱

- 데이터 수집
- 분석 결과 시각화
- 서비스 구현



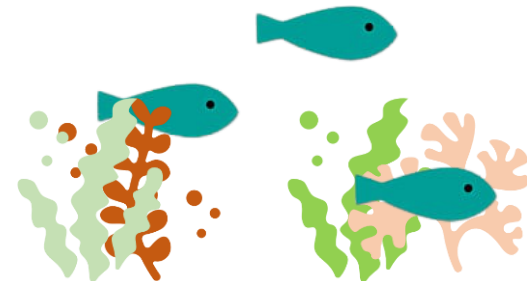
팀원
채지연

- 데이터 수집
- 데이터 정제/가공
- 분석 결과 시각화



팀원
김경준

- 데이터 수집
- 딥러닝 Model 구축
- 분석 결과 시각화





Step 02

프로젝트 수행 절차 및 개발 환경

- 프로젝트 수행 절차
- 개발 환경

2.1 프로젝트 수행 절차

2021
12

SUN	MON	TUE	WED	THU	FRI	SAT
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	



데이터 수집
데이터 가공
데이터 적재

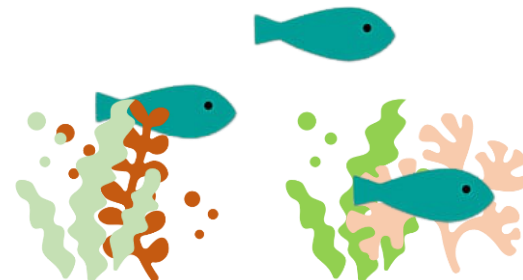
분석 및 결과 도출



분석 및 결과 도출
시각화



시각화



2.2 개발 환경



Language
Python / Kotlin



Deep Learning
TensorFlow /
TensorFlow Lite



Platform
Jupyter / Colab



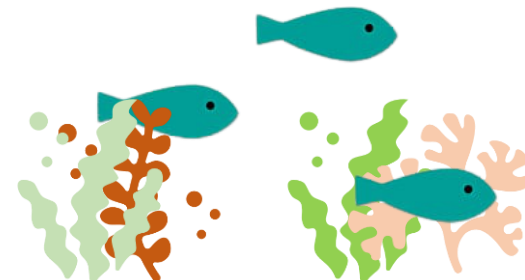
Mobile App
Android Studio



Database
Android Studio



Repository
GitHub



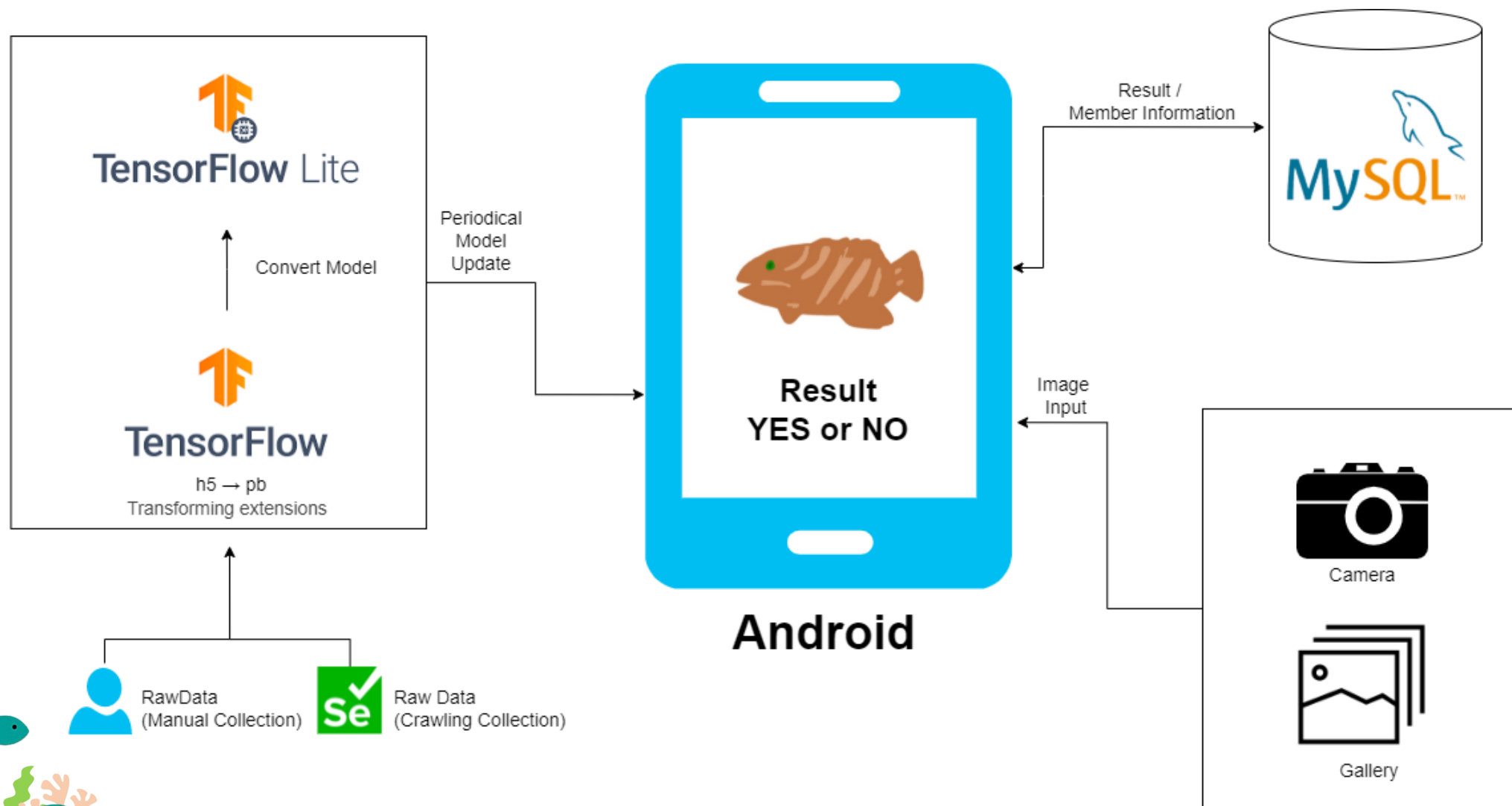


Step 03

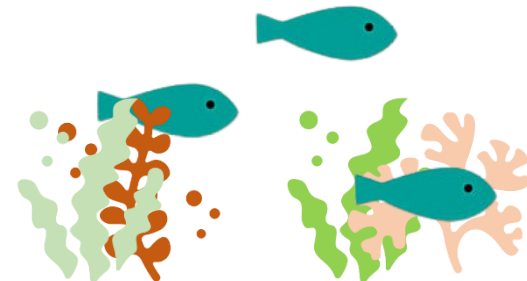
프로젝트 수행 및 결과

- 시스템 구성도
- 이미지 처리 프로세스
- ML 시스템 구성도
- 학습 결과
- 시스템 구현
- 어플 시연

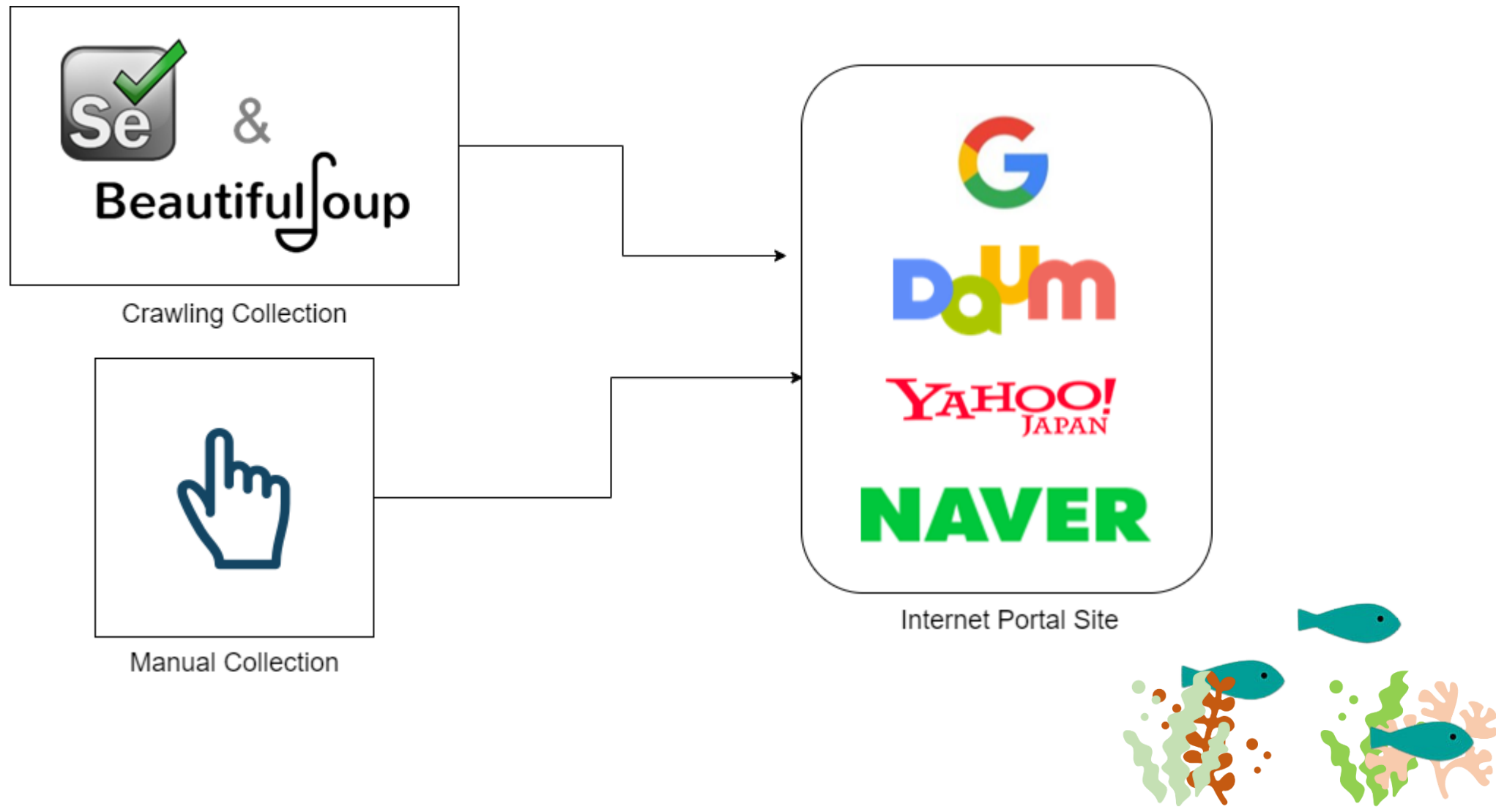
3.1 시스템 구성도



3.2 이미지 처리 프로세스



3.2.1 이미지 수집

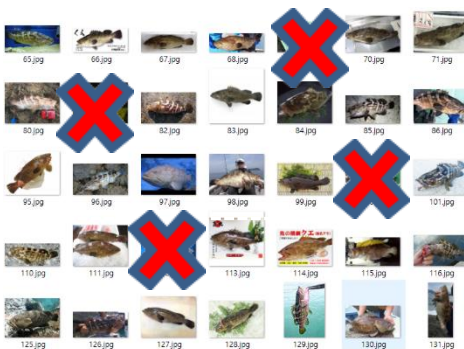


3.2.2 전처리

Raw Data



다금바리, 능성어,
회, 낚시 홍보 이미지,
그 외 물고기

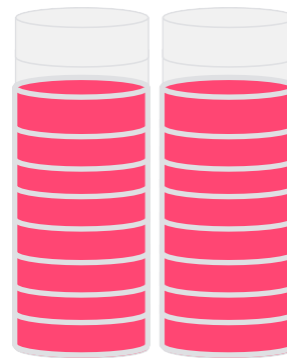


Check



Point

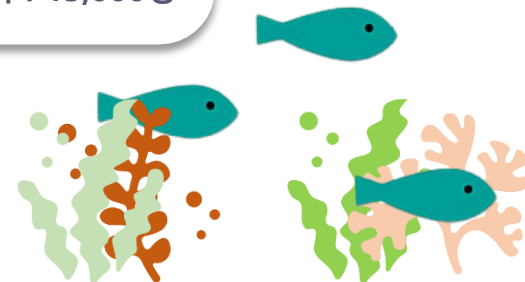
다금바리 / 능성어 분류 데이터



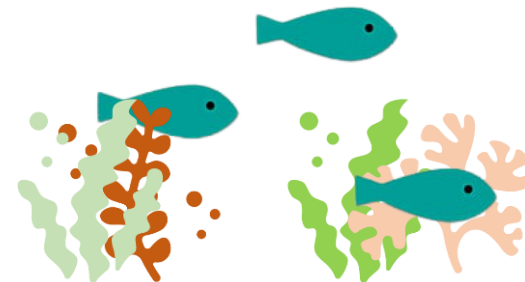
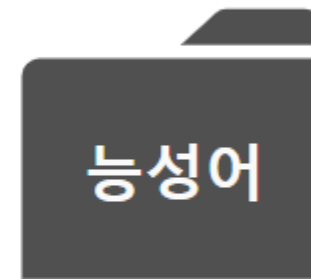
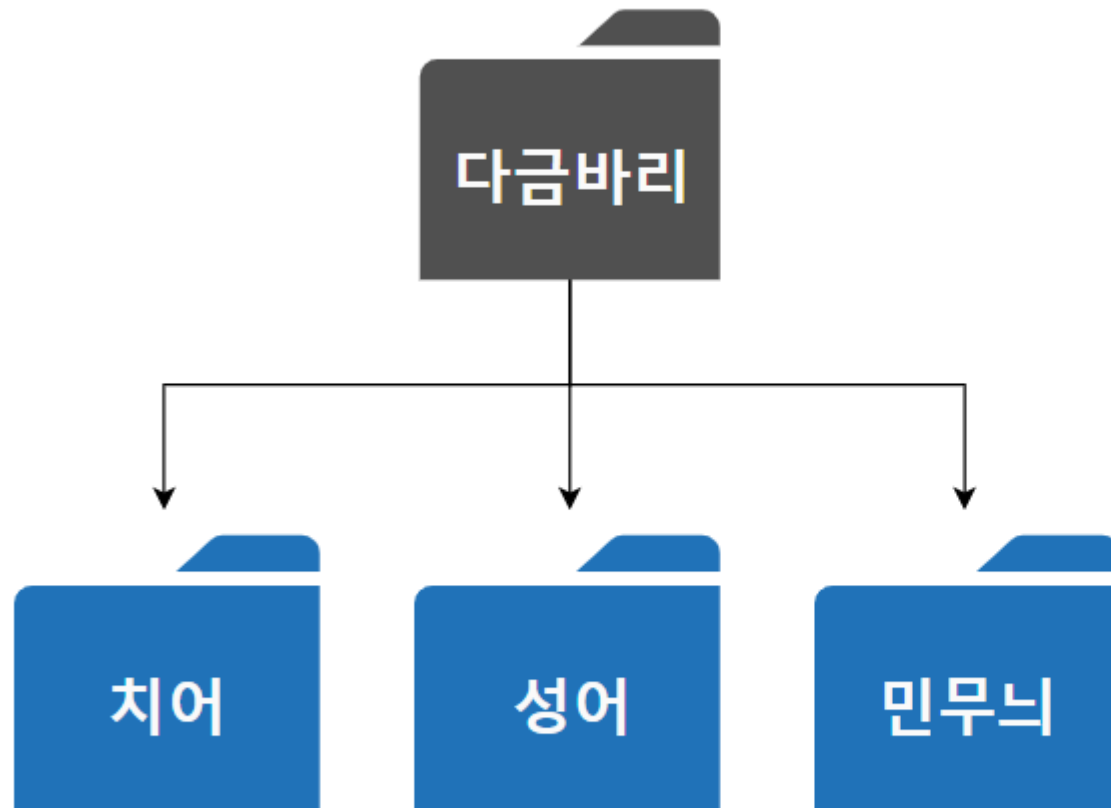
다금바리 : 820장 / 능성어 : 500장
좌우 반전 / 기울이기 / 노이즈 추가



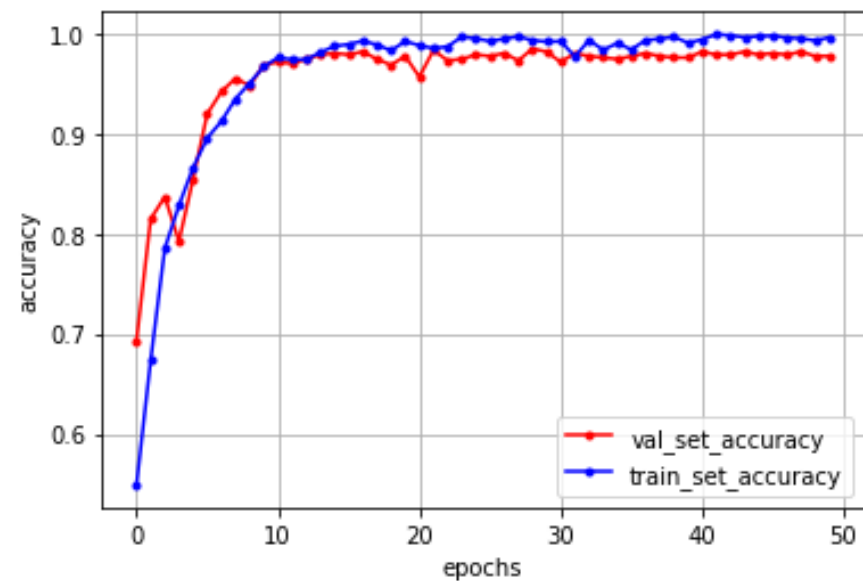
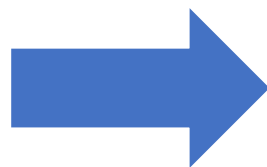
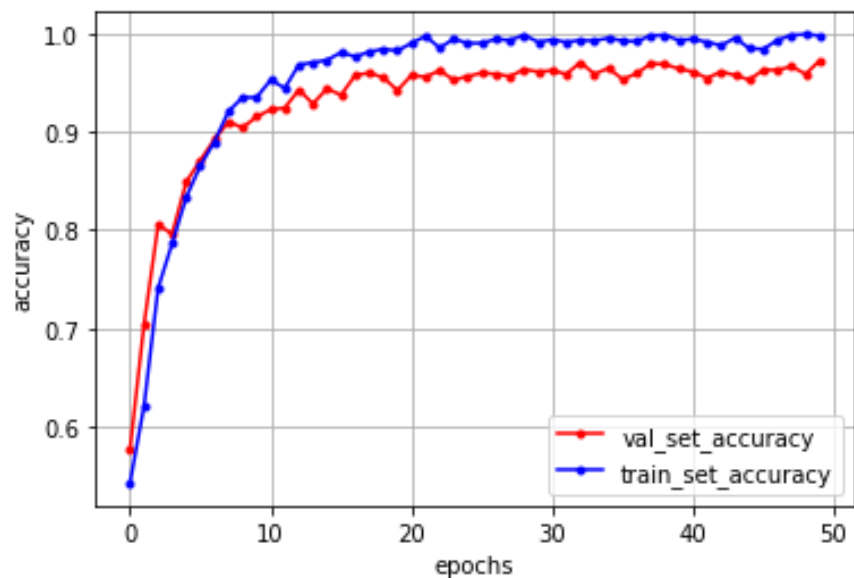
다금바리 : 30,000장 / 능성어 : 15,000장



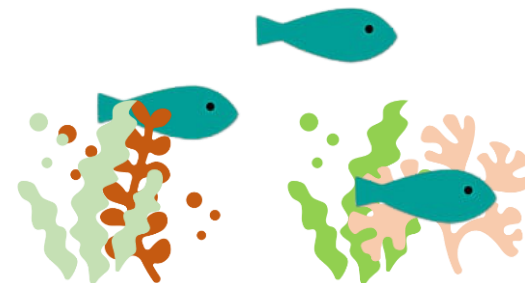
3.2.3 데이터셋



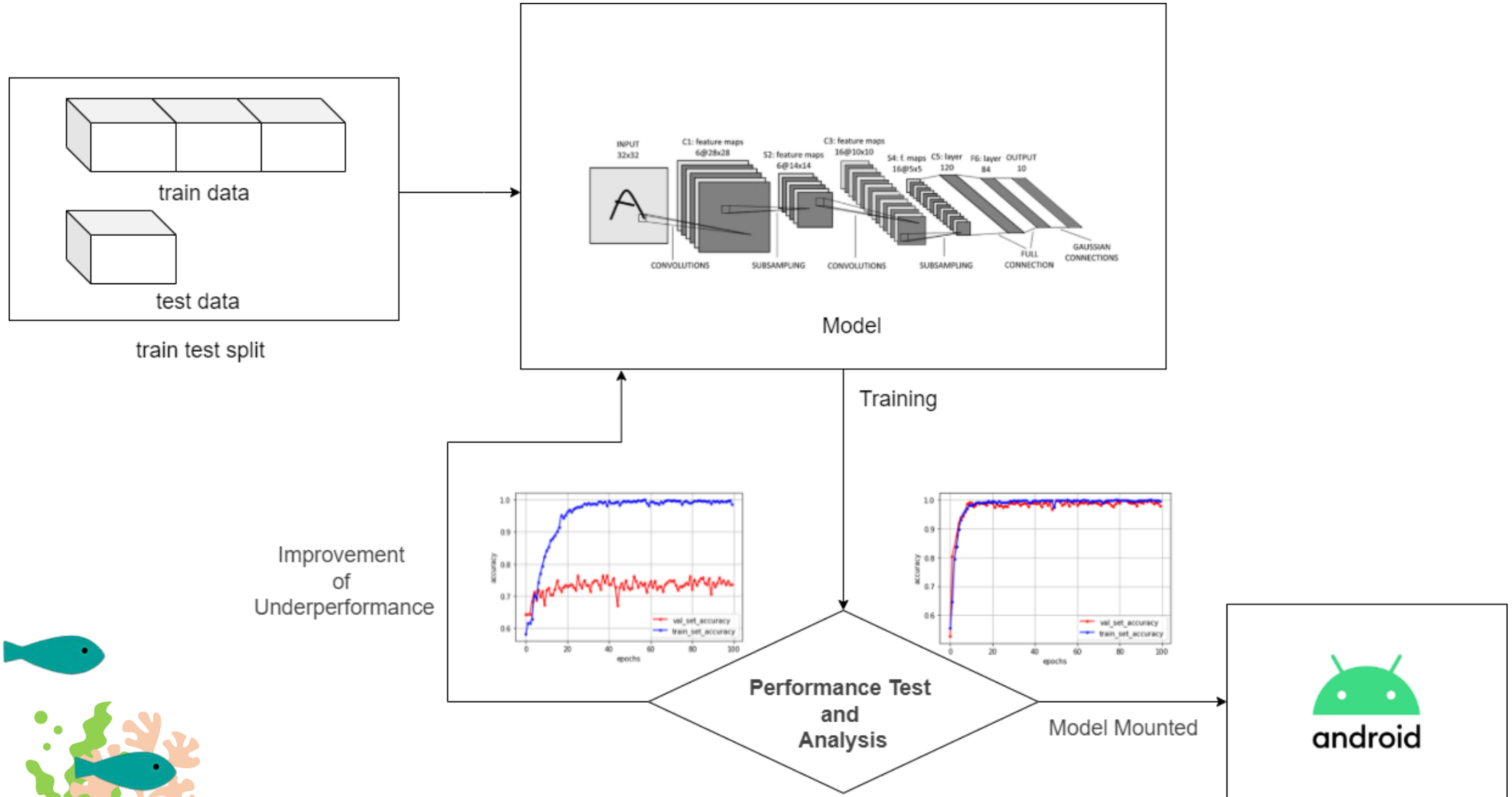
3.2.3 데이터 분류 결과



데이터 분류 전후 정확도 비교

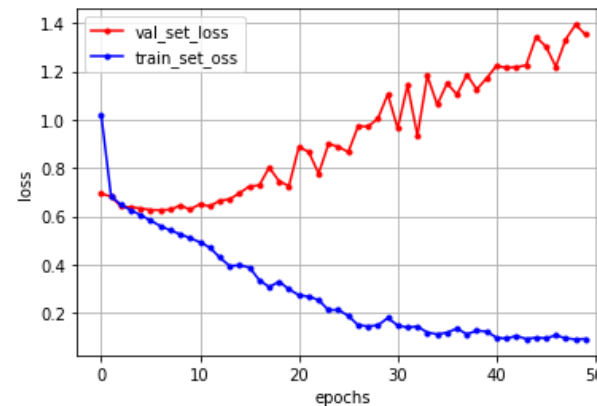
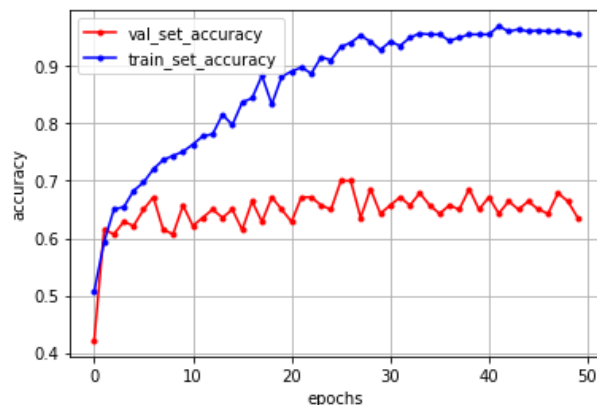


3.3 ML 시스템 구성도

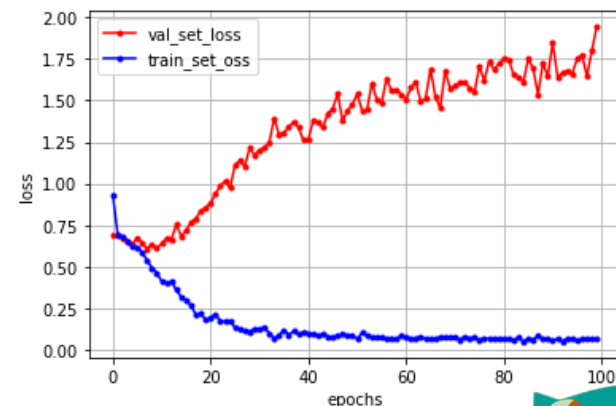
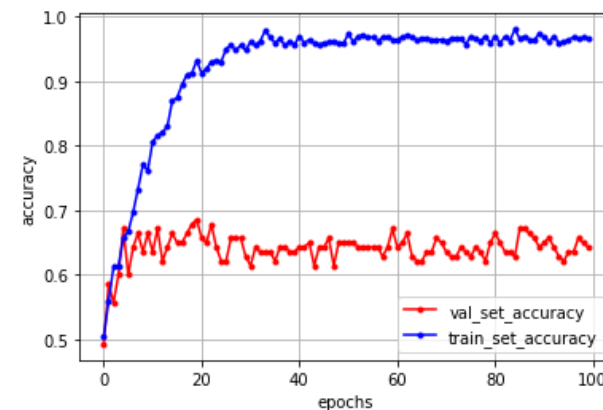


3.4.1 학습 결과

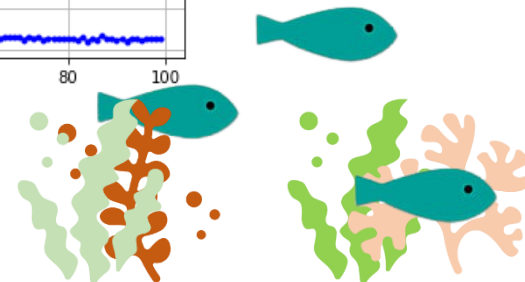
50번 학습



100번 학습



첫 데이터로 학습한 결과

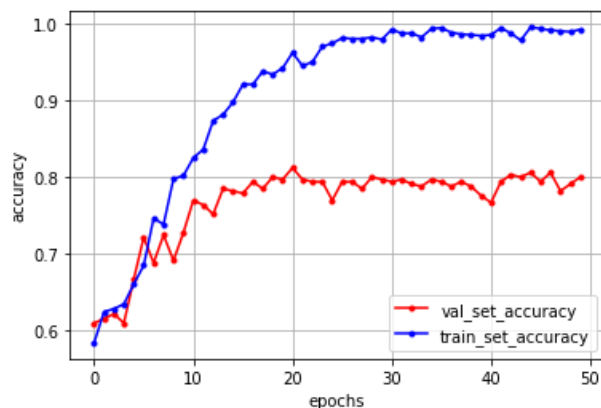


3.4.2 학습 결과

50번 학습

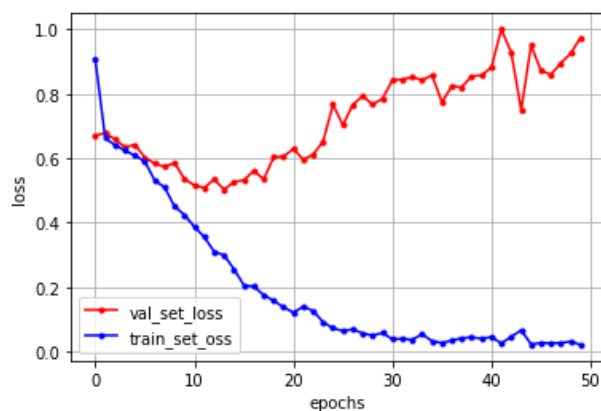
—●— train data 정확도

—●— test data 정확도

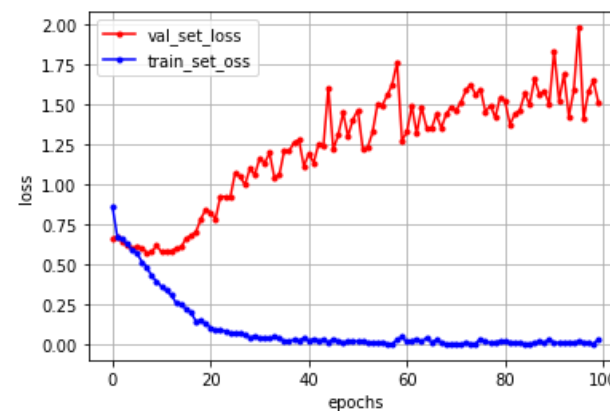
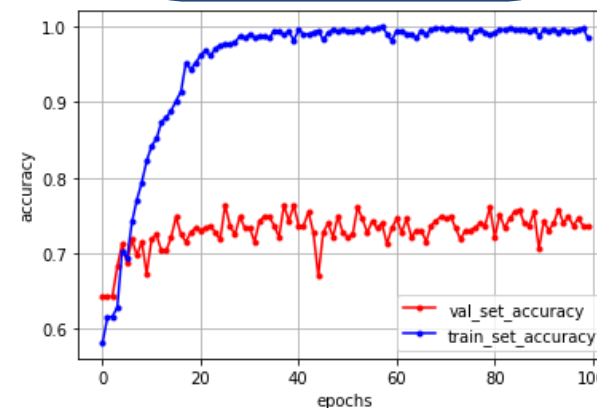


—●— train data 손실률

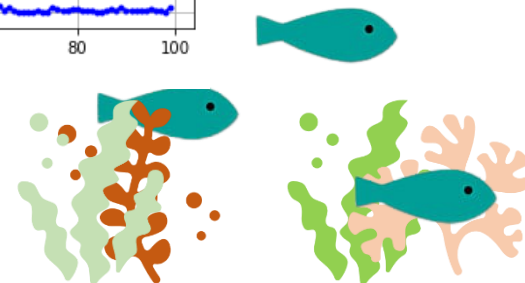
—●— test data 손실률



100번 학습

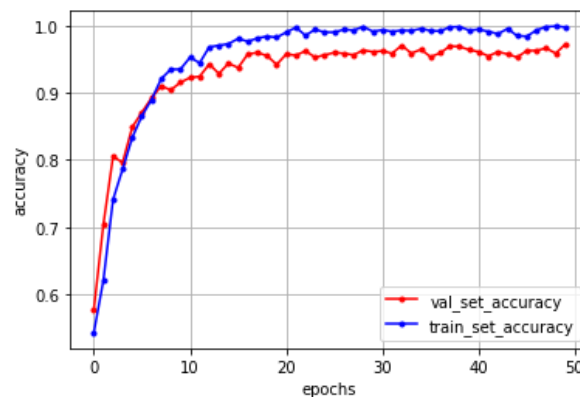


더 많은 사진을 학습 시킨 결과

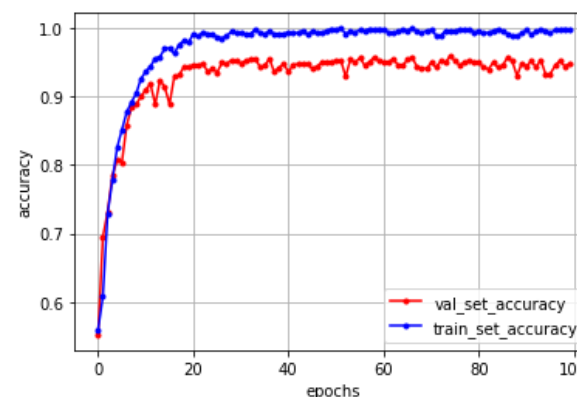


3.4.3 학습 결과

50번 학습



100번 학습

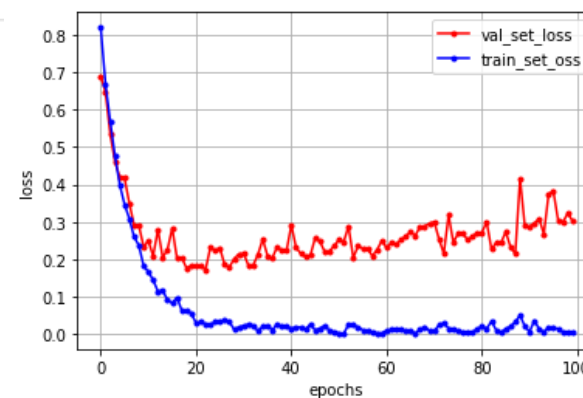
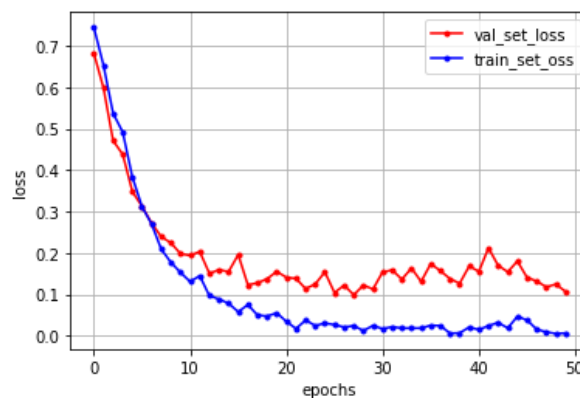


train data 정확도

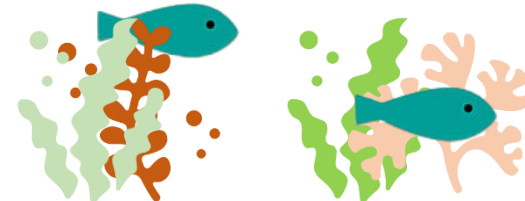
test data 정확도

train data 손실률

test data 손실률



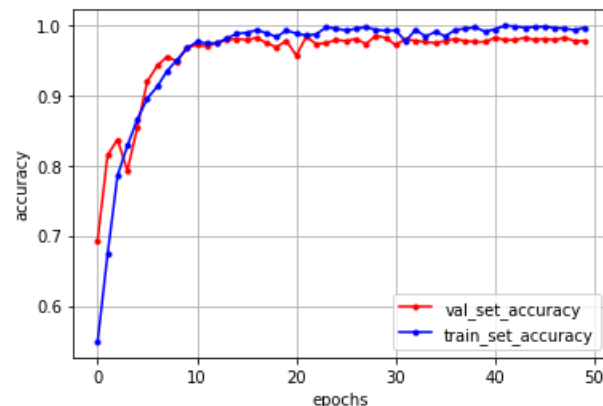
양질의 사진을 선별하여 학습



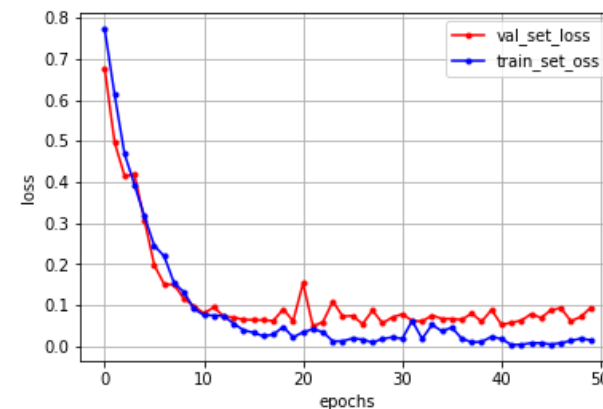
3.4.4 학습 결과

50번 학습

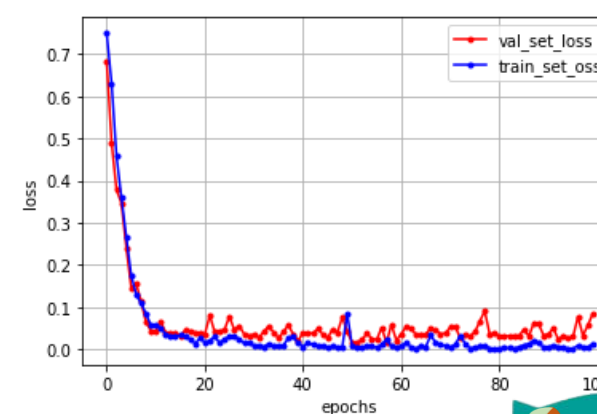
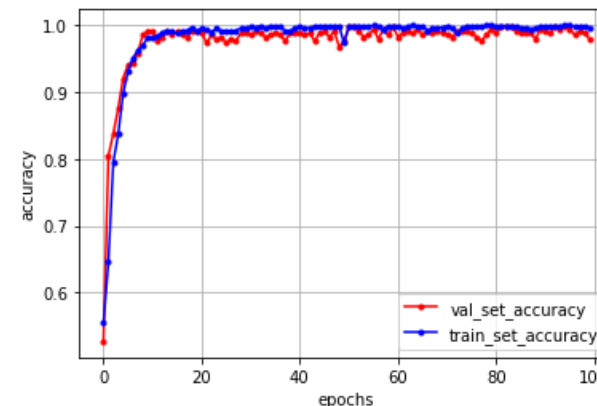
—●— train data 정확도
—●— test data 정확도



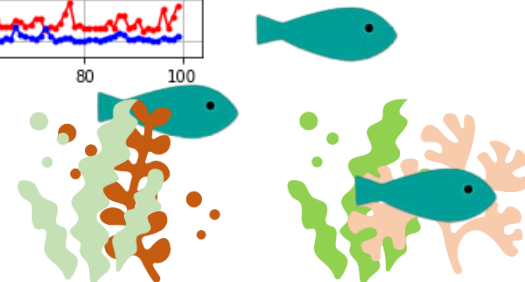
—●— train data 손실률
—●— test data 손실률



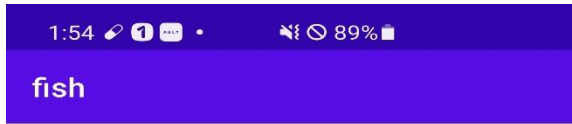
100번 학습



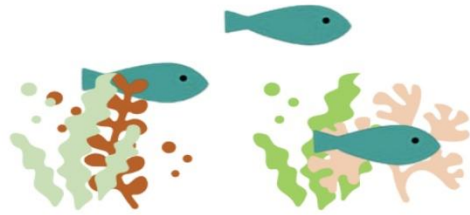
데이터를 더 세세하게 분류하여 학습



3.5.1 시스템 구현

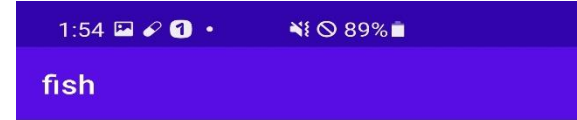


다금바리를 찾아서



START

Main

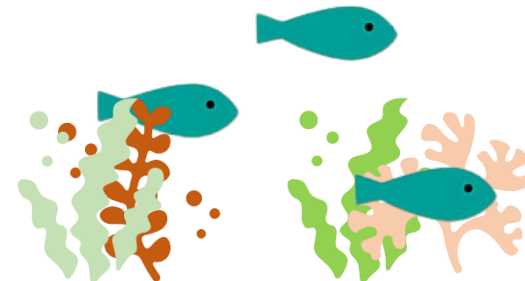


이미지 넣기

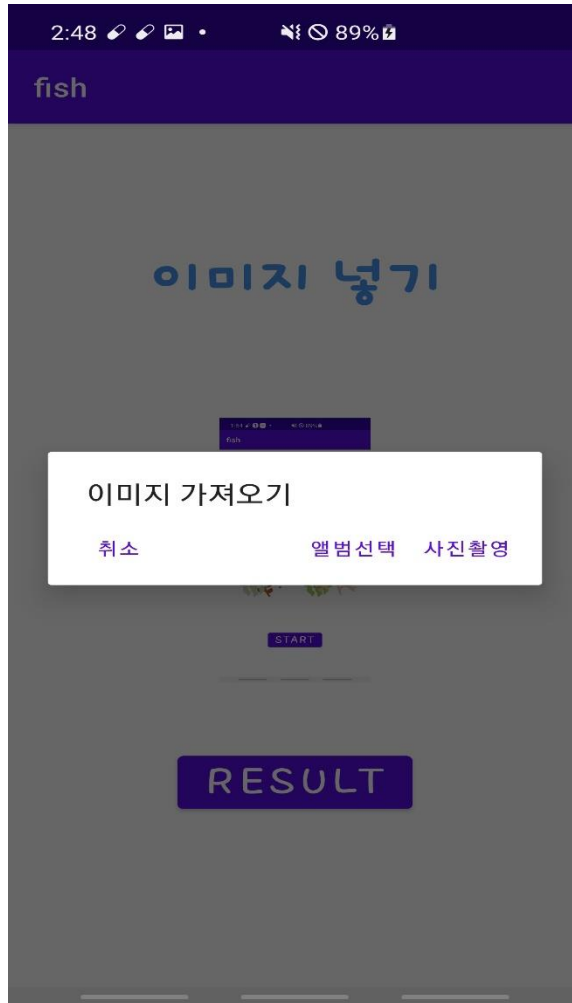


RESULT

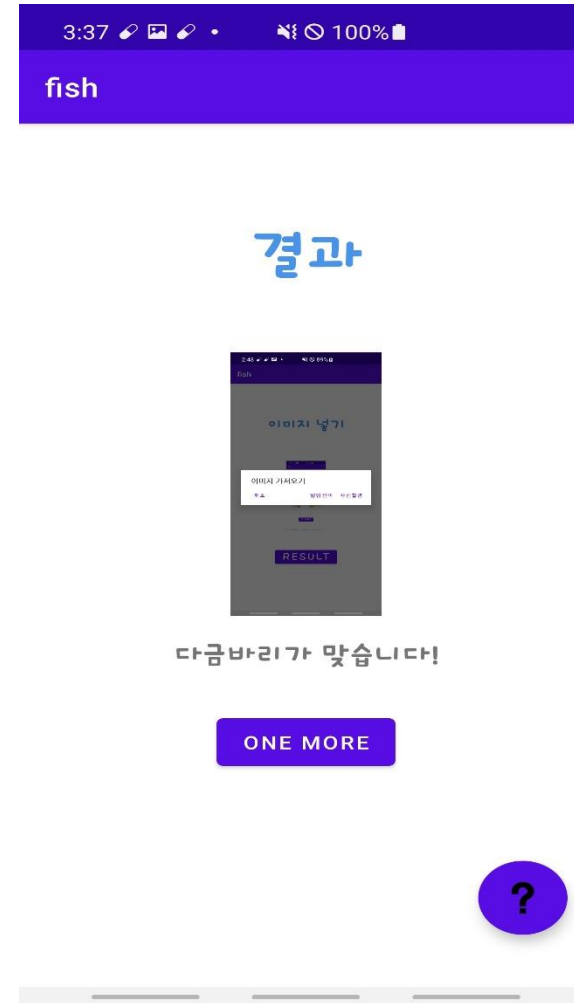
Get Image



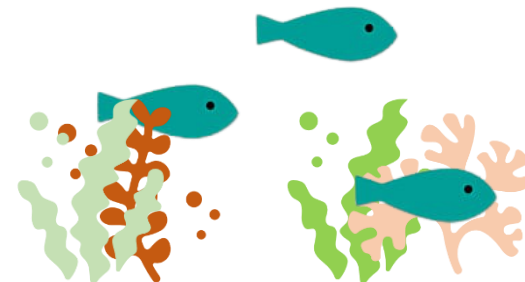
3.5.2 시스템 구현



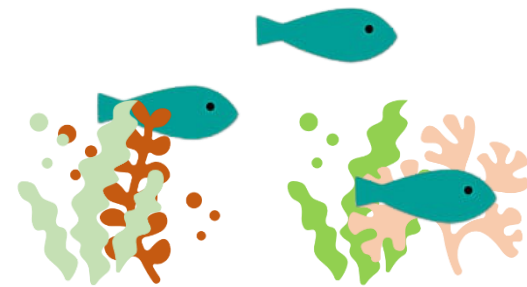
How to bring up
The Image



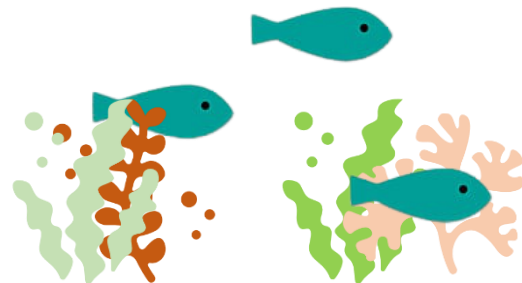
Result




3.5.3 시스템 구현



3.6 어플 시연





Step 04

프로젝트 평가

- 개선 방향
- 프로젝트를 마치며

4.1 개선 방향

- 제3의 데이터 예외 처리

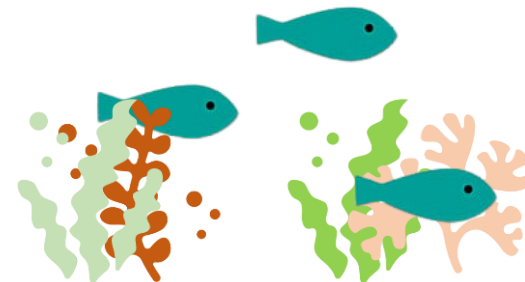
```
img_array = keras.preprocessing.image.img_to_array(img)
img_array = tf.expand_dims(img_array, 0) # Create a batch

predictions = model.predict(img_array)
score = tf.nn.softmax(predictions[0])

print(
    "This image most likely belongs to {} with a {:.2f} percent confidence."
    .format(class_names[np.argmax(score)], 100 * np.max(score))
)
```

```
Downloading data from https://mblogthumb-phinf.pstatic.net/20110101_259/spp080
8%AE%C0%DA%B9%D9%B8%AE%B1%B8%BA%B0%B9%FD1.jpg?type=w2
163840/157188 [=====] - 0s 0us/step
172032/157188 [=====] - 0s 0us/step
This image most likely belongs to nungsung with a 73.11 percent confidence.
```

- 타 어종 학습



4.2 프로젝트를 마치며

- 김경준 모델 제작이 생각대로 되지 않을 때가 많았지만 서로 의논하여 문제를 해결하며 많은 것을 배웠습니다. 빅데이터 분석에 있어서 데이터 자료의 중요함을 느꼈고 분석 과정 또한 재미있어 좋은 경험이었습니다.
- 양상욱 이미지 딥러닝 과정을 전체적으로 이해할 수 있어서 좋았습니다.
- 임수진 모델 학습에 필요한 사진을 대량으로 구하는 것에 어려움이 있었습니다. 그렇게 처음 수집한 자료들로 도출된 결과가 만족스럽지 못해 실망하기도 했습니다. 그러나 팀원들과 정확도를 높이기 위한 방법을 끊임없이 고안하여 높은 정확도의 모델을 만들어 냈고, 그때의 성취감은 잊을 수 없습니다. 이번 프로젝트를 통해 팀원과의 소통 및 협업 방법과 그 가치에 대해 깨닫는 좋은 시간이 되었습니다.
- 정민욱 평소 AI 삼성카메라, 자율주행 기술 등을 접하면서 이러한 기능들은 어떻게 구현한 것인가에 흥미를 가지고 있었습니다. 때마침 이번 프로젝트를 통해 이러한 기술들은 CNN 기술을 활용하여 만들어진다는 것을 알게 되었고, CNN 기술을 활용하여 모델 제작과 학습을 시키는 과정에서 많은 어려움이 있었지만, 팀원들과 머리를 맞대며 하나하나 해결해 나가면서 모델을 학습시킬 때에는 양질의 자료가 중요하다는 것을 몸소 느끼게 되었고, 좋은 경험이 되었습니다.
- 채지연 다금바리를 구별하는 모델을 만들어 안드로이드에 탑재하는 과정에서 처음에는 pytorch로 모델을 구현했더니 이를 tensorflowlite로 변환하는 과정이 복잡해 애를 먹었지만 결과적으로는 tensorflow로 모델을 만들어 보다 쉽게 탑재할 수 있었습니다. 하나 하나 팀원들과 얘기를 나누며 수정했고 이미지 딥러닝 전반적인 과정을 배울 수 있는 기회가 되었습니다.

감사합니다

맛있겠다...

