\* 붉은색 이 포함된 내용은 발표 시 시간 부족으로 설명하지 못한 부분을 작성한 것임.

### Welcome to Coral Guard



**Full Codes on GitHub:** 

https://github.com/c-cys/CoralGuard

## Problem Problem Definition

#### 산호(Coral)



▲ <mark>해양시민과학센터 파란,</mark>

https://greenparan.org/24/?bmode=view&idx=19392912

산호는 해양 생태계에서 중요한 역할을 하는 해양 생물

다양한 해양 생물의 서식지로서 풍부한 생물 다양성을 제공

- 그러나 산호는 특정 온도와 pH 조건에 민감하게 반응하고, 온도 상승이나 해양 산성화가 발생하면 산호는 백화 현상을 겪을 수 있다.

#### **Bleached Coral**



산호가 공생 조류인 조절렌타를 방출하여 색을 잃고 하얗게 변하는 현상

온도 상승, 자외선 증가, 해양 산성화 등 환경 스트 레스로 인해 백화됨.

이는 산호의 영양 감소로 이루어지고, 이는 생물 다양성의 붕괴로 이어지는 심각한 문제

▲ 서울신문, "산호초 백화현상도 속도·정도 다 달라요", 유용하 기자, https://m.go.seoul.co.kr/news/society/enviroment/2022/04/11/20220411021014?cp=en

#### 어떤 모델이 필요할까?

산호의 백화 여부를 탐지할 수 있어야 함.

모든 지역에 광범위하게 적용 가능해야 함.

최종적으로 해당 모델을 통해 산호의 백화 현상을 방지하고, 이미 백화된 산호의 회복에 기여하는 바가 있어야 함.

### Detecting

그렇다면 산호의 백화를 막을 수 있는 방법은 무엇이고, 미연에 이를 방지할 수 있는 방안은 어떤 것이 있을까?

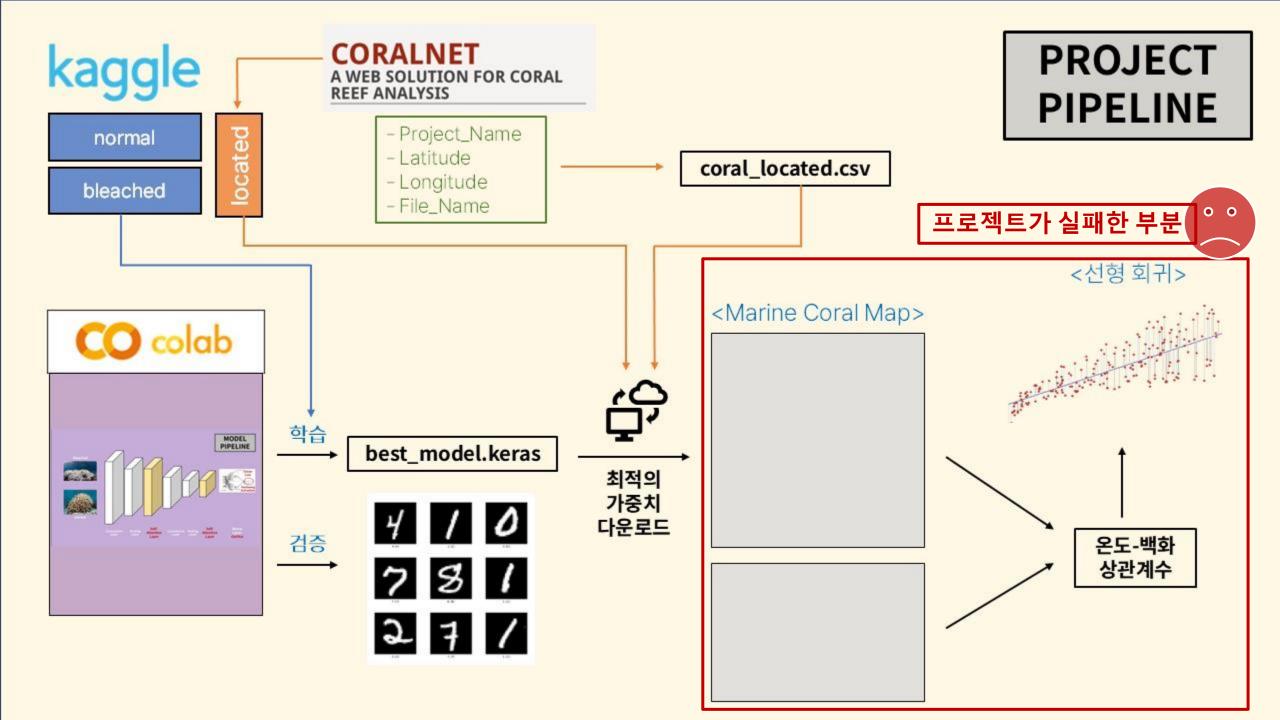
우선 산호의 백화가 심하게 진행된 해안의 경우, 자연적으로 확산되는 산호의 유생 만으로 개체수 회복이 어려울 수 있다. 이때 인공적으로 산호충을 뿌려 산호 생태계를 회복하는 기법을 '유충회복' 이라 함.

(Mumby, Peter J, Robert AB Mason, and Karlo Hock. "Reconnecting reef recovery in a world of coral bleaching." Limnology and Oceanography. Methods 19.10 (2021): 702-713.)

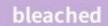
백화된 산호가 생태계를 이룰 가능성이 보이는 지역을 미리 예측 가능하다면, 프로 바이오틱스균을 사용하여 산호의 스트레스 저항성을 증가시킬 수 있다.

Banin, E, Israely, T, Kushmaro, A, Loya, Y, Orr, E, and Rosenberg, E. (2000) Penetration of the coral-bleaching bacterium Vibrio shiloi into Oculina patagonica. Appl Environ

## Model Building



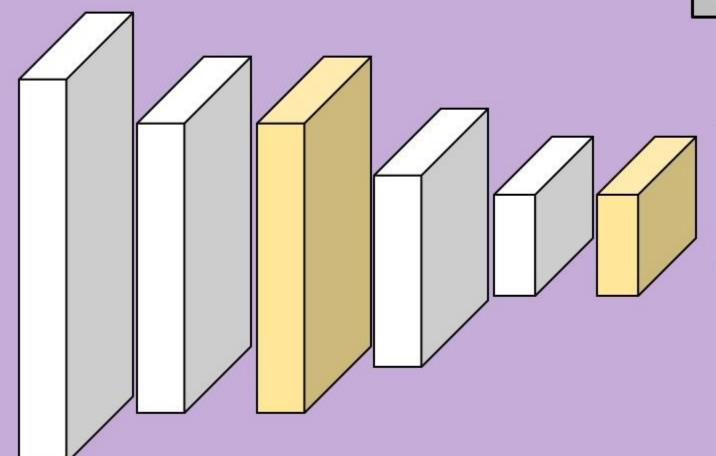
#### MODEL **PIPELINE**

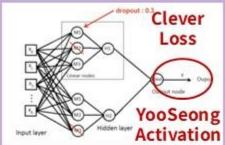






normal





Convolution Layer

Layer

Self-Attention Layer

Convolution Pooling Layer

Layer

Self-Attention Layer

Dense Layer: **OutOut** 

Model: "sequential\_4"

Layer (type)	Output Shape	Param #
conv2d_8 (Conv2D)	(None, 62, 62, 32)	896
max_pooling2d_8 (MaxPooling2D)	(None, 31, 31, 32)	0
<pre>self_attention_6 (SelfAttention)</pre>	(None, 31, 31, 32)	3,168
conv2d_9 (Conv2D)	(None, 29, 29, 64)	18,496
max_pooling2d_9 (MaxPooling2D)	(None, 14, 14, 64)	0
<pre>self_attention_7 (SelfAttention)</pre>	(None, 14, 14, 64)	12,480
flatten_4 (Flatten)	(None, 12544)	0
dense_32 (Dense)	(None, 6912)	86,711,040
lambda_4 (Lambda)	(None, 3456)	0
dense_33 (Dense)	(None, 2)	6,914
yoo_seong_activation_4 (YooSeongActivation)	(None, 2)	0

Total params: 86,752,994 (330.94 MB)

Trainable params: 86,752,994 (330.94 MB)
Non-trainable params: 0 (0.00 B)

### O. Dataset





#### CoralGuard\_bleached

Add a subtitle Edit

Data Card Code (0) Discussion (0) Suggestions (0) Settings

C-CHOIYS AND 2 COLLABORATORS - UPDATED 7 DAYS AGO

#### CoralGuard\_normal

Add a subtitle Edit

Data Card Code (0) Discussion (0) Suggestions (0) Settings

CoralGuard\_located ☆ ⑤ ⊘
파일 수정 보기 삽입 서식 데이터 도구 확장프로그램 도움말 Q 5 ♂ 등 등 100% ▼ | ₩ % .0 .0 123 | 기본값... ▼ | - 10

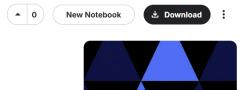
	A	В	С	D	Е	
1	Project	Latitude	Longitude	Country	FileName	
2	Norfolk Island	29.048	167.954	일본, 한국	ydlvkp8dcl.png.800x0_q8	
3	Norfolk Island	29.048	167.954	일본, 한국	25vh717atl.png.800x0_q	
4	Norfolk Island	29.048	167.954	일본, 한국	zcyjkxe3vh.png.800x0_qt	
5	Colby Bermuda	32.4	-64.7	미국	Colby Bermuda 2017 - C	
6	Leah Harper	16.80526	88.08314	인도, 미얀마	LeahHarper1.jpg	
7	Leah Harper	16.80526	88.08314	인도, 미얀마	LeahHarper2.jpg	
8	Leah Harper	16.80526	88.08314	인도, 미얀마	LeahHarper3.jpg	
9	Punta Cana Ree	18.51619	68.3792	파키스탄, 인도	CRDR1.jpg	
10	Punta Cana Ree	18.51619	68.3792	파키스탄, 인도	CRDR2.jpg	
11	Punta Cana Ree	18.51619	68.3792	파키스탄, 인도	CRDR3.jpg	
12	BTS_DEF	-12.7	-38.5	브라질	BTS1.jpg	
13	BTS_DEF	-12.7	-38.5	브라질	BTS2.jpg	
4.4	DTC DEE	40.7	20.5	버기지	DTC0:	



C-CHOIYS AND 2 COLLABORATORS · UPDATED A DAY AGO · PRIVATE

#### CoralGuard\_located

Add a subtitle Edit



### Model Tuning

#### 활성화함수 재정의

-완전연결층에서 쓰이는 함수와 출력층에 쓰이는 함수를 모두 재정의하여 사용.

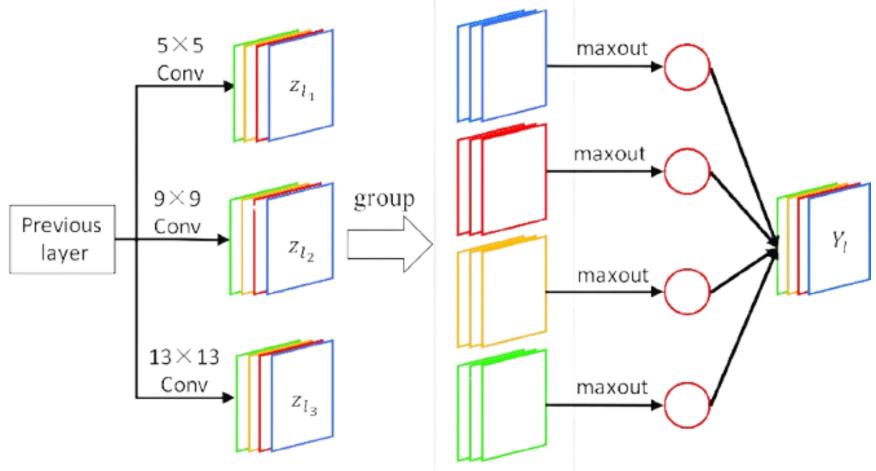
- '모델의 '정확도'나 '속도'보다는 '기능의 독창성'에 집중함.
- 4가지의 새롭고 독창적인 은닉층 활성화함수 OutOut / 출력층 활성화함수 YooSeong / 손실함수 CleverLoss / Self-Attention Layer를 개발함.

-FC layer - OutOut, output - yooseong 로 정의하여 구현

### 1. Activation Function

# Activation Function (1) - OutOut

- outout = max'out' + drop'out'



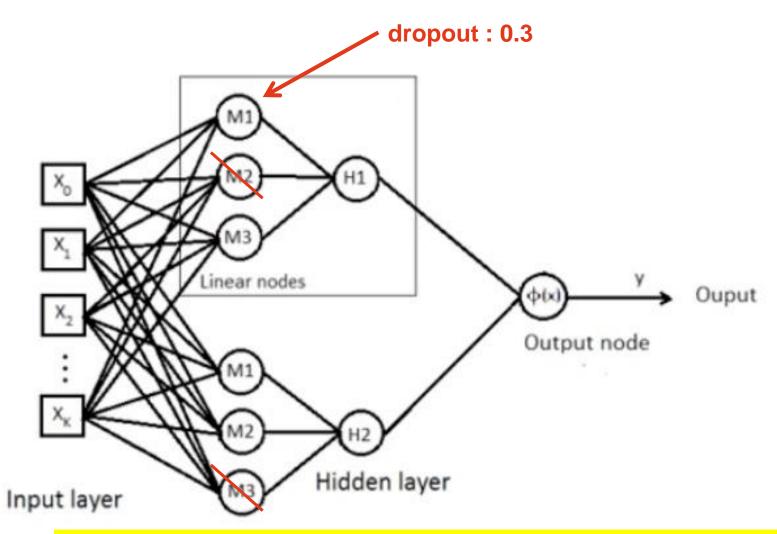
▲ "The proposed module which produces the maxout results of multi-scale", ResearchGate, https://www.researchgate.net/figure/The-proposed-module-which-produces-the-maxout-results-of-multi-scale-convolutional\_fig2\_323595776

- 기존의 maxout은 '활성화함수도 학습을 할 수 있을까' 라는 궁금증에서 출발한 활성화함수이다.
- 함수의 개형이 정형화되어있는 것이 아닌, Maxout만을 위한 새로운 가중치를 정의하여 사용하는 형식임.
- 정확도 측면에서 현재 유망하다고 평가받는 ReLU와 그 아종의 활성화 함수를 압도하는 성능을 가지고 있으나, 과도한 parameter 증가 문제로 인해 계산복잡도가 증가한다는 단점을 가짐.

- 계산 복잡도 문제를 해결하기 위해, 과적합 방지에 쓰이는 dropout의 구조를 본떠 함수 정의에 병합함.

- 학습이 진행되는 각 epoch마다, 마치 dropout처럼 무작위적인 뉴 런을 잠시 비활성화시켜 계산량을 줄이는 방식.

- maxout의 이점인 높은 정확도는 살리고, 계산복잡도는 감소시킴.



▲ "[Deep Learning] Deep Neural Network (2)", Tous Les Jours STUDY, 티스토리,

https://janghan-kor.tistory.com/669

- '활성화함수' 그 자체로서의 기능도 뛰어나지만, dropout의 영향으로 과적합 방지 효과도 지님.
- 기존에는 뉴런의 활성화를 제한하는 dropout과 가중치의 범위를 제한하는 weight regularization을 혼합하여 사용할 수 없었음. (dropout으로 인해 규제할 가중치가 소실되기 때문)
- OutOut을 거치면 Dropout을 활성화 함수 단계에 적용한 효과를 얻으므로, 이후 손실함수에 weight regularization을 사용하여 두 방법 모두를 조화롭게 사용할 수 있음.

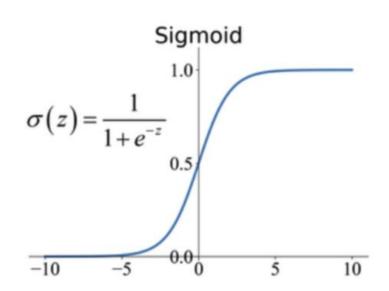
test 횟수	1	2	10	평균 accuracy	max_accuracy	min_accuracy
기존 모델	0.9709	0.973	0.9749	0.97428	0.9791	0.9709
outout 적용	0.9831	0.9836	0.9787	0.97952	0.9843	0.9756

- MNIST dataset을 통해 Test accuracy를 도출한 결과.
- 10 epoch로 시행한 10번의 모두
  Restoring model weights from the end of the best epoch: 10.
  즉 epoch가 높을수록 뛰어난 정확도를 기대할 수 있음.
- 높은 epoch로 인한 과적합 문제는 dropout과 가중치 규제의 중첩 사용으로 방지.

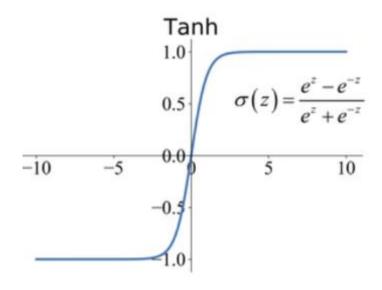
# Activation Function (2) - YooSeong

- 기존의 softmax를 대체할 출력층의 activation function

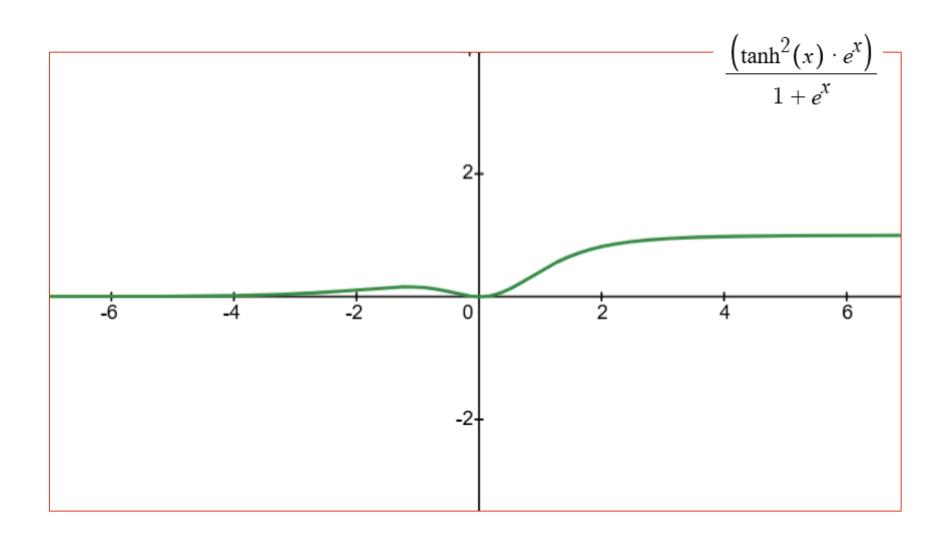
$$\frac{tanh^{2}(x)}{1+e^{-x}}$$
 의 개형을 가짐.



▲ "시그모이드(Sigmoid) 활성화 함수에 대해 알아보고 파이썬으로 구현해보기", rgbitcode.com, <a href="https://rgbitcode.com/blog/senspond/55">https://rgbitcode.com/blog/senspond/55</a>



▲ "Tanh Activation", paperswithcode.com, <a href="https://paperswithcode.com/method/tanh-activation">https://paperswithcode.com/method/tanh-activation</a>



- 양의 무한대에서 1에 수렴/음의 무한대에서 0에 수렴
- -> 기존 시그모이드 개형 유지 & softmax 대체용 총 적분값=1 위함
- 미분 가능함
- 시그모이드의 비선형성 강화
- -> 복잡한 데이터 학습에 유리
- 양수에서 안정한 시그모이드의 성질을 최대한 유지

test 횟수	1	2	10	평균 accuracy	max_accuracy	min_accuracy
기존 모델	0.9709	0.973	0.9749	0.97428	0.9791	0.9709
yoo_seong 적용	0.9086	0.9816	0.9942	0.89573	0.9942	0.8233

- 특정 한번의 test에 대해서는 매우 높은 고점을 보이나, 편차가 매우 큰 경향이 존재함.

### 2. LOSS Function

## Loss Function - CLEVER

#### **CLEVER**

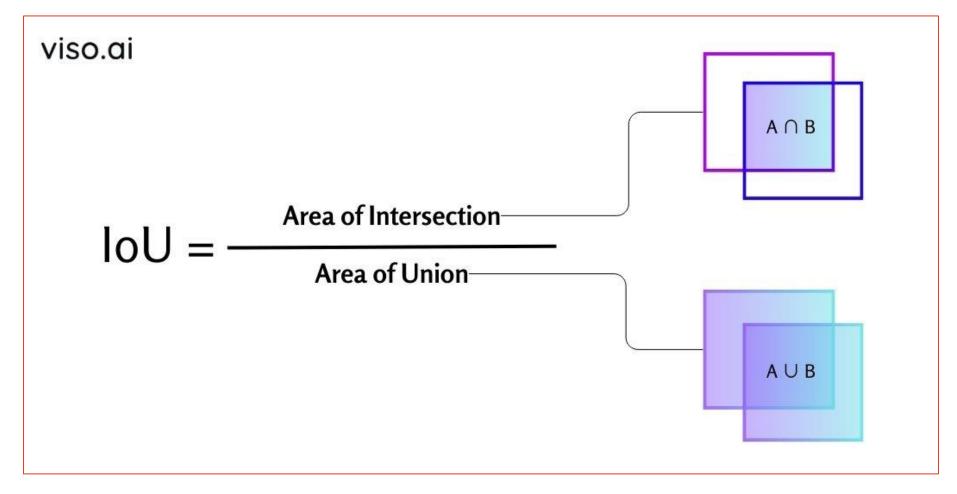
- 기존에는 object detection 모델의 성능 평가에 주로 쓰이던 IOU(Intersection over Union)



- 초기에는 IOU를 모델 개발 후 산호 백화와 온도 간의 상관관계를 지도(map)에서 산출할 때 사용하고자 함 -> but 타당성의 문제 존재(42-43페이지에서 설명) -> IOU를 손실함수에 재활용!

- 이를 착안하여 손실함수에 도입함.
- 계산에 요구되는 파라미터가 레이블의 최대, 최소로 한정되기 때문에 계산량을 대폭 줄일 수 있음.

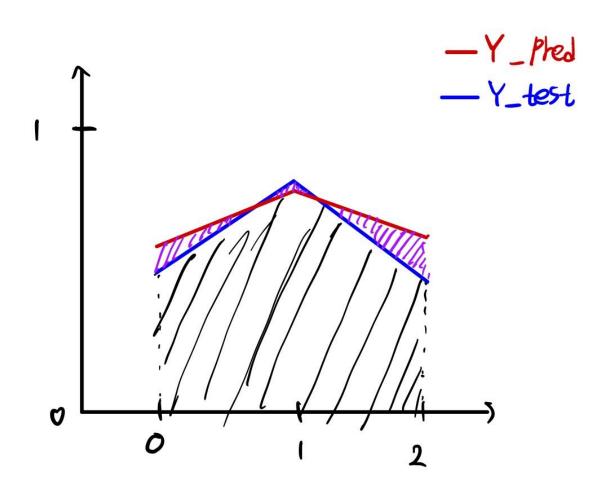
#### IOU



▲ "What is Intersection over Union (IoU)?", viso.ai,

https://viso.ai/computer-vision/intersection-over-union-iou/

#### **CLEVER**



#### **CLEVER**

test 횟수	1	2	10	평균 accuracy	max_accuracy	min_accuracy
기존 모델	0.9709	0.973	0.9749	0.97428	0.9791	0.9709
clever 적용	0.9134	0.9033	0.9033	0.90159	0.9134	0.8903

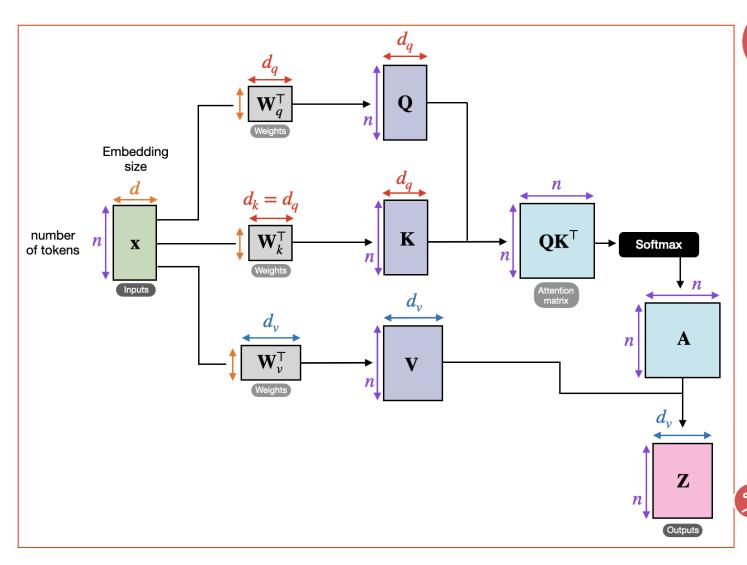
test 횟수	1	2	10	평균 loss	min_loss	max_loss
기존 모델	0.0961	0.0869	0.0842	0.08461	0.0696	0.0961
clever 적용	0.0878	0.0967	0.0967	0.09892	0.0878	0.1112

- 기존 모델보다 부분적으로 낮은 정확도 & 높은 손실을 보임.
- aross entropy에 비해 낮은 모델 성능을 나타냈음.

### 3. Special Hidden Layer

# Special Hidden Layer -Self-Attention

### **Self Attention**





- Self-Attention을 이용하면 input의 임베 딩 벡터들에 대해서 유사도를 기준으로 중요한 정보만을 강조할 수 있음.
- 또한 모든 요소들에 대해서 유사도를 파악할 수 있으므로 지역적 특성을 파악하는 Convolution 이후에 배치하여 전역적인 특성을 파악하고자 함.

▲ "Understanding and Coding the Self-Attention Mechanism of Large Language Models From Scratch", SebastianRaschka, <a href="https://sebastianraschka.com/blog/2023/self-attention-from-scratch.html">https://sebastianraschka.com/blog/2023/self-attention-from-scratch.html</a>

Self-Attention Layer 구현 시에는 생성형 AI 인 ChatGPT의 도움을 받음.



### 독창적으로 추가한 함수들의 성능 평가

	Α	В	С	D	E	F	G	Н	1	J	K	L	M	N
1	Accuracy													
2	test 횟수	1	2	3	4	5	6	7	8	9	10	평균 accuracy	max_accuracy	min_accuracy
3	기존 모델	0.9709	0.973	0.9717	0.9728	0.9767	0.9783	0.9735	0.9719	0.9791	0.9749	0.97428	0.9791	0.9709
4	iou_loss 적용	0.9134	0.9033	0.9009	0.8962	0.9033	0.8917	0.9102	0.8903	0.9033	0.9033	0.90159	0.9134	0.8903
5	yoo_seong 적용	0.9086	0.9816	0.9501	0.8647	0.8765	0.8233	0.8435	0.8737	0.8411	0.9942	0.89573	0.9942	0.8233
6	outout 적용	0.9831	0.9836	0.9768	0.9843	0.9785	0.9778	0.9756	0.977	0.9798	0.9787	0.97952	0.9843	0.9756
7														
8	Loss													
9	test 횟수	1	2	3	4	5	6	7	8	9	10	평균 loss	min_loss	max_loss
10	기존 모델	0.0961	0.0869	0.093	0.0898	0.0778	0.0737	0.0891	0.0859	0.0696	0.0842	0.08461	0.0696	0.0961
11	iou_loss 적용	0.0878	0.0967	0.0991	0.1038	0.0967	0.1082	0.0923	0.1112	0.0967	0.0967	0.09892	0.0878	0.1112
12	yoo_seong 적용	0.0984	0.0872	0.0895	0.1459	0.1121	0.172	0.1643	0.1118	0.1099	0.0845	0.11756	0.0845	0.172
13	outout 적용	0.0684	0.0626	0.0852	0.0619	0.0852	0.0946	0.0919	0.0911	0.079	0.0898	0.08097	0.0619	0.0946



#### Test 데이터셋에 대한 예측 결과

True: 1, Pred: 0



True: 0, Pred: 0



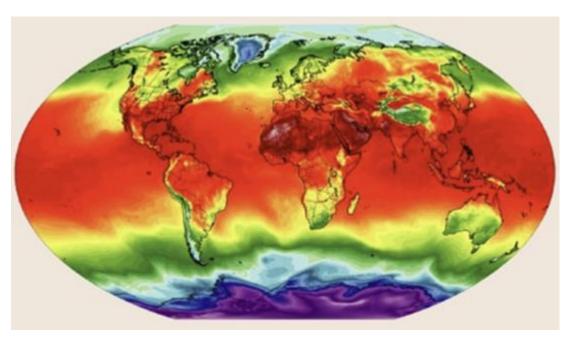
True: 0, Pred: 0



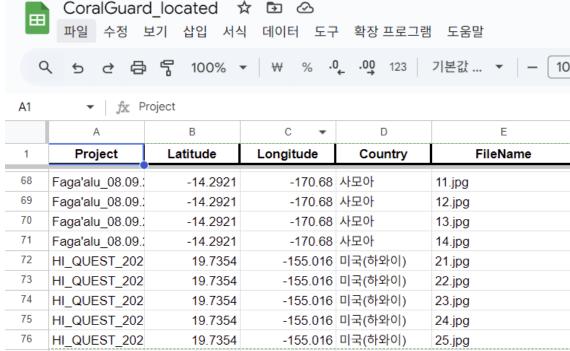
모델의 모든 하이퍼파라미터는 학습을 10번 반복함으로써 얻은 Accuracy와 Loss를 정리하여 최적의 성능을 갖는 하이퍼파라미터를 선택함.

# Correlation Between Bleached Coral & Temperature

## 1. 수치적인 데이터

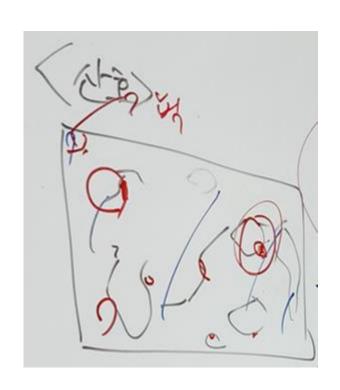


▲ "[기후위기] 관측 이래 지구가 가장 뜨거웠던 날", 내외신문, 김시월 기자, https://www.naewaynews.com/303200



수온 데이터는 무한함에 비해 직접 수집 한 위치 정보 포함 산호 데이터는 75개뿐 -> 데이터셋 불균형 문제 발생

### 2. 지도 -> IOU







- 1) 산호 백화 확률에 따라 원의 반지름을 비례시키는 것은 타당한가?
- 2) 타당하다고 하더라도, 그것이 온도에 대한 지도와의 상관계수를 구할 때 사 용할 수 있는 옳은 방법인가?

에 대한 의문점이 아직 해결되지 않음.



# Correlation and Regression

## 상관계수와 회귀계수를 조금 더 수학적으로



- 상관계수를 통해 회귀계수를 산출하는 식에 대한 증명

회귀계수=상관계수\*분산의 비율



#### 상관계수와 화귀계수를 조금더수학적으로

- 회귀 계수 : '모형의 다른 예측 변수를 유지하면서 예측 변수의 한 단위에 대한 반응 변수의 평균 변화'

즉 두 변수의 관계를 평균적으로 제일 잘 설명할 수 있는 기울기의 값을 말함.

-산호의 백화 비율과 수온의 관계를 회귀로 예측하기 위해 회귀 계수를 도출했는데, 이를 산호 MAP과 수온을 통해 구현하고자 함.

- 이를 위해 상관계수와 회귀계수의 수학적 관계를 명확히 할 필요가 있었는데, 이를 설명하는 논문이 마땅치 않아 자체적으로 직접 증명함.



#### \* अरामिक असमि एन सक्स एक्स एक्स असमिन





#### 1.1A 128조보

$$(N+1) \& x_{i,y} = \underbrace{(N_{i,y} - \overline{X}_{i,k})}_{=\underbrace{(N_{i,y} - \overline{X}_{i,k})}_{=\underbrace{(N_{i,y}$$

#### 监别 #.2

$$(N-1)6\chi^{2} = \Xi(X_{i}-X)^{2}$$

$$= \Xi(X_{i}^{2}-2X_{i}X+X^{2})$$

$$= \Xi(X_{i}^{2}-2X_{i}X+X^{2})$$

$$= \Xi(X_{i}^{2}-2X_{i}X+X^{2})$$

$$= \Xi(X_{i}^{2}-1X_{i}X+X^{2})$$

# CONCLUSION



#### 모델의구현에대한고찰

- <u>산호의 이미지들에 대해 산호의 백화 유무를 구별하는 것까지는 완벽</u> 하게 성공함.
- <u>하지만 이후 지도를 만드는 작업과 회귀 단계까지는 완전히 구현하지</u> <u>못함</u>.
  - <u>지도를 만들고자 하는 과정에서, 상관계수를 구하는 방법 두 가지를</u> 고민하였으나 모두 문제점(의문점)이 존재하였음.
- 그러나 지도를 만드는 간단한 코드 구현과,
- 상관계수를 얻었을 때 이를 회귀계수로 변환하는 식에 대한 증명은 구현함.



### 모델의구현에대한고찰

• 본 프로젝트는 "실용적인 모델 구현의 완성"보다 "아이디어"에 집중함.

- 초기 목표(2024.11.06)는 단순히 "산호의 백화 여부 분류 및 예측"에 그쳤음.
- 팀원 중 한 명이 산호 백화 지도(<a href="https://allencoralatlas.org/">https://allencoralatlas.org/</a>)를 발견함. 이에 보다 새롭고 독창적인 메커니즘을 포함하고자 '수온으로 인한 산호의 백화'라는 문제를 '수온을 통한 산호의 백화 확률 예측'으로 증대함.
- 두번째 목표(2024.11.07)는 "산호의 백화-수온 지도 간의 상관계수 구하기"였음.
- 세번째 목표(2024.11.14.)는 <mark>4가지의 독창적인 메커니즘 추가</mark>(OutOut, YooSeong, Clever, Self-Attention)였음.
- 최종적인 목표(2024.11.19.)는 상관계수를 구하는 <mark>방안을 '수치적인 데이터'로서 비</mark>교가 아닌 '지도의 영역 비교'으로 변경하는 것이었음.

#### 모델의 필요성





I have a question regarding coral bleaching.

2503김영리

11월 20일 (수) 오후 3:27 (1

Hello, we are high school students at Gyeonggi Buk Science High School who are interested in coral bleaching in Korea and want to develop machine learning for t

CORAL Info <info+noreply@coral.org>

11월 20일 (수) 오후 3:27 (10시간 전)

나에게 🕶

Hello,

Thank you for reaching out! Please see below for answers to some of our more frequently asked questions.



대표적인 세계 산호 구호 단체인 'CORAL REFF ALLIANCE' 에 문의하여 현재 백화된 산호 탐지 모델에 대한 수준을 파악하고, 전체적인 모델에 대한 필요성을 검증받음.

#### 결론

- □ 설계한 모델은 현재 특정 지역의 산호가 백화된 비율을 빠르게 판단하고, 미래 수온 예측 데이터를 이용하여 백화된 산호 생태계를 이룰 지역을 미리 알 수 있음.
- □ 이를 통하여 산호 구호 협회 등과 협력하여 <mark>산호가 백화되는 것을 미연에</mark> 방지할 수도 있고, 특정하기 어려운 현재 백화된 산호의 위치를 쉽게 찾아내어 정보를 전달할 수 있음.
- □ 모델을 정의하는데 필요한 여러가지 부분들이 모두 유의미하게 작동하였고, 이를 통해 모델 성능의 향상을 볼 수 있었음.

#### Reference

□ 이미지 Reference: 각 이미지 하단에 캡션으로 표기.

□ 데이터셋

https://www.kaggle.com/datasets/cchoiys/coralguard-normal

https://www.kaggle.com/datasets/cchoiys/coralguard-bleached

https://www.kaggle.com/datasets/cchoiys/coralguard-located

□ Self-Attention Layer 구현: ChatGPT 사용



**Full Codes on GitHub:** 

https://github.com/c-cys/CoralGuard