

2021년도

# 반려동물 학습데이터 유효성 검증모델 개발과업 완료보고서

2021년 12월

(주)비투엔

# 목차

---

1. 개요

2. 모델 및 데이터 업그레이드 방안

3. 학습 결과

4. 진행 간 이슈

5. 소감 및 마무리

# 목차

---

## 1. 개요

## 2. 모델 및 데이터 업그레이드 방안

## 3. 학습 결과

## 4. 진행 간 이슈

## 5. 소감 및 마무리

# 모델 및 데이터 업그레이드 방안

모델 및 데이터 업그레이드 방안으로써 학습 파라미터 최적화, 클래스 불균형 해소, 라벨 데이터 개선 방안을 도입

## 모델 및 데이터 업그레이드 방안

## 내용

## 상세 방안

모델 파라미터 최적화

학습 모델이 해당 데이터에 대해 최적화  
할 수 있는 모델 파라미터 탐색

- Epoch 늘리기
- Optimizer 변경

클래스 불균형 해소

정상:질환 데이터 비율 불균형으로 인해  
올바른 학습이 방해되는 점을 해소

- Weighted Loss
- Data Augmentation

# 목차

---

1. 개요

**2. 모델 및 데이터 업그레이드 방안**

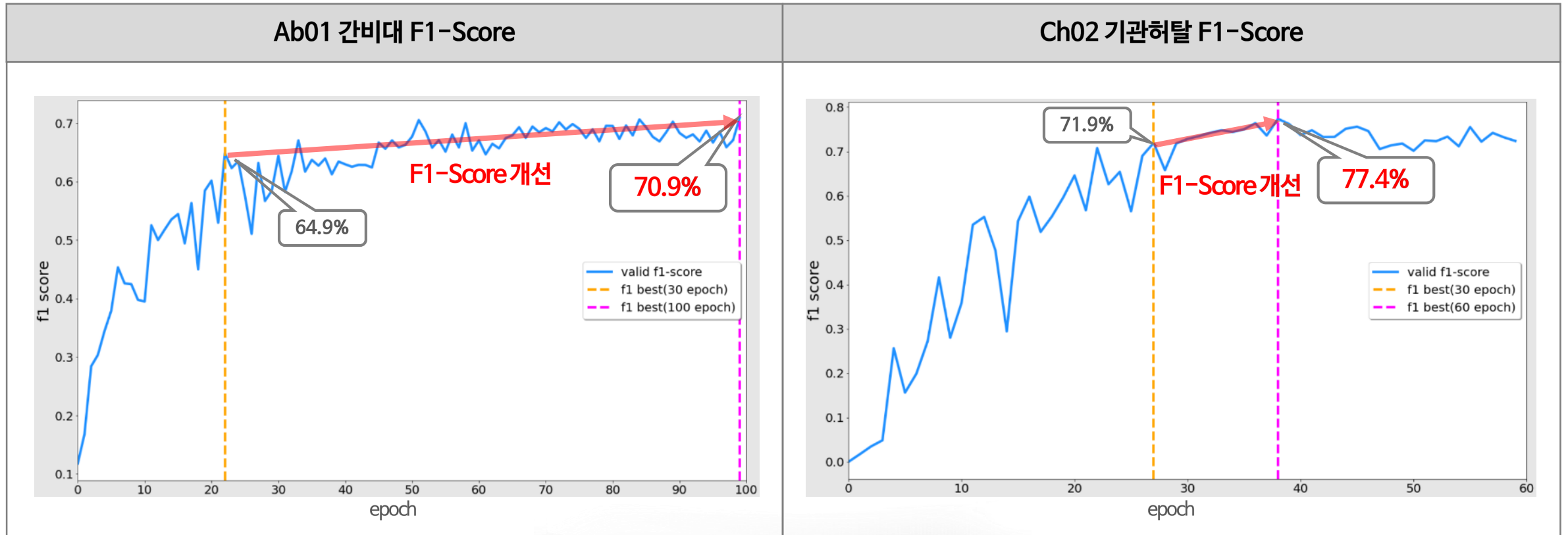
3. 학습 결과

4. 진행 간 이슈

5. 소감 및 마무리

## 모델파라미터 최적화> Epoch 조절하기

모델 정확도를 높이기 위해 Ab01(간비대) 질환은 100 epoch, Ch02(기관허탈) 질환은 60 epoch으로 조정하는 등 기존 모델 학습에 참고한 Vanilla 모델의 Epoch 보다 늘려서 학습 효과를 개선



기존 모델의 Epoch을 늘려 더 많이 학습 수행함으로써 학습 효과 개선

# 모델파라미터 최적화>Optimizer변경

모델 학습 시 최적의 Loss값(Global Optima)로 모델을 업데이트하는 방법인 Optimizer를 다양하게 적용하여 최적의 성능을 보인 Optimizer를 사용

### Optimizer 계보와 원리

모든 자료를 다 검토해서 내 위치의 산기울기를 계산해서 갈 방향을 찾겠다.

**GD**

**SGD** 전부 다 봐야 한걸음은 너무 오래 걸리니까 조금만 보고 빨리 판단한다 같은 시간에 더 많이 간다

**Momentum** 스텝 방향  
스텝 계산해서 움직인 후, 아까 내려 오던 관성 방향 또 가자

**NAG** Nesterov Accelerated Gradient  
일단 관성 방향 먼저 움직이고, 움직인 자리에 스텝을 계산하니 더 빠르더라

**Nadam** Adam에 Momentum 대신 NAG를 붙이자.

**Adam** RMSProp + Momentum  
방향도 스텝사이즈도 적절하게!

**RMSProp** 보폭을 줄이는 건 좋은데 이전 맥락 상황봐가며 하자.

**Adagrad** 안가본곳은 성큼 빠르게 걸어 올고 많이 가본 곳은 잘아니까 갈수록 보폭을 줄여 세밀히 탐색

**AdaDelta** 종종걸음 너무 작아져서 정지하는걸 막아보자.

### Ch02 기관허탈 Optimizer 비교

#### 1) 질적 성능 비교

Image	SGD	SGDP	Adam	AdamW	AdamP	RAdam	참값
정상							
질환							
	IoU : 73.3%	IoU : 0.0%	IoU : 83.1%	IoU : 72.4%	IoU : 75.1%	IoU : 76.7%	

\* 빨강 : 배경, 초록 : 정상, 파랑 : 질환

#### 2) 양적(수치) 성능 비교

Optimizer	배경 IoU	정상 IoU	질환 IoU	F1@0.25
SGD	99.3	75.0	57.5	68.3
SGDP	94.9	0.0	0.0	0.0
Adam	99.5	80.0	62.8	71.4
AdamW	99.3	75.0	52.1	59.3
AdamP	99.4	75.2	53.6	60.3
RAdam	99.5	79.5	60.1	66.4

초기값

루트1 Adam

루트2 Momentum

루트3 SGD

Global Optima

Local Optima

좋은 Optimizer : 단기간에 Global Optima에 수렴

∴ Adam > Momentum > SGD

# 클래스 불균형 해소 > Weighted Loss

정상과 질환 이미지 개수가 불균형한 문제를 해소하기 위해 Segmentation 모델 학습에 Weighted Loss를 적용하여, 모델이 상대적으로 적은 개수의 질환 이미지 및 질환 클래스를 집중적으로 학습할 수 있도록 학습방식 개선

01

데이터 구성

02

문제 분석

03

해결 방안

유효성 검사용 데이터의 불균형 문제

질환	배경 : 정상 : 질환 *픽셀 수	배경 : 정상 : 질환 픽셀 비율
Ab01	1,064,508,477 : 24,493,430 : 15,935,053	96 : 2 : 1
Ch02	3,416,516,862 : 36,765,730 : 27,727,584	98 : 1 : 1

\* Segmentation 분류 단위 : Pixel

간비대 Vanila 모델 예측 결과

정상 이미지

질환 이미지

Image

예측

참값

질환, 정상 픽셀에 비해  
배경 픽셀이 너무 많아  
별다른 조치 없이 학습한 모델은  
배경으로만 분류한다.

\* 빨강 : 배경, 초록 : 정상, 파랑 : 질환

Weighted Loss 적용

① 클래스별 Loss Weight (CW) 적용

② Focal Loss (FL) 적용

Ab01 간비대 Weighted Loss 적용 결과

1) 질적 성능 비교

Image

Vanila

\*\*CW

\*\*\*FL

CW + FL

\*\*\*\*CW' + FL

참값

정상

IoU : 0.0%

IoU : 57.4%

IoU : 54.2%

IoU : 42.2%

IoU : 65.9%

질환

IoU : 0.0%

IoU : 64.1%

IoU : 82.0%

IoU : 63.9%

IoU : 78.7%

\* 빨강 : 배경, 초록 : 정상, 파랑 : 질환

\*\* CW : 클래스별 Weight 적용. 배경 : 정상 : 질환 = 0.01 : 0.2 : 0.79

\*\*\* FL : Focal Loss (Focal Cross Entropy)

\*\*\*\* 클래스별 Weight가 배경 : 정상 : 질환 = 0.51 : 0.72 : 1.74

2) 양적(수치) 성능 비교

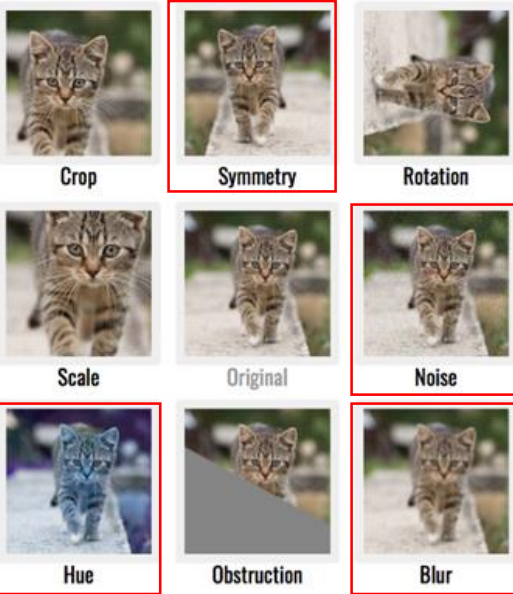
모델	배경 IoU	정상 IoU	질환 IoU	F1@0.5
Vanila	96.4	0	0	0
CW	97.6	58.4	57.5	67.0
FL	98.5	62.8	58.1	63.6
CW+FL	95.5	44.1	42.7	35.1
CW' + FL	98.7	67.8	73.0	74.3



# 클래스 불균형 해소) Data Augmentation

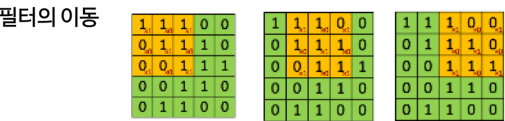
원본데이터에 약간의 변형을 줌으로써 다른 데이터로 가공하는 기법인 Data Augmentation을 질환 이미지에만 적용하여, 모델이 학습할 수 있는 질환 이미지 개수를 늘릴 수 있도록 데이터 전처리 수행

## Data Augmentation 기법 소개



1. Spatial Transform (공간정보 변형):  
- 자르기, 세로축/가로축 전환, 회전, 확대
2. Pixel Transform (픽셀단위정보 변형):  
- 색상 보정, 선명도 조정, 노이즈 추가

딥러닝 모델은 [작게 나누어진 이미지 조각들로부터 시각적 패턴을 학습하는 방식]이다.



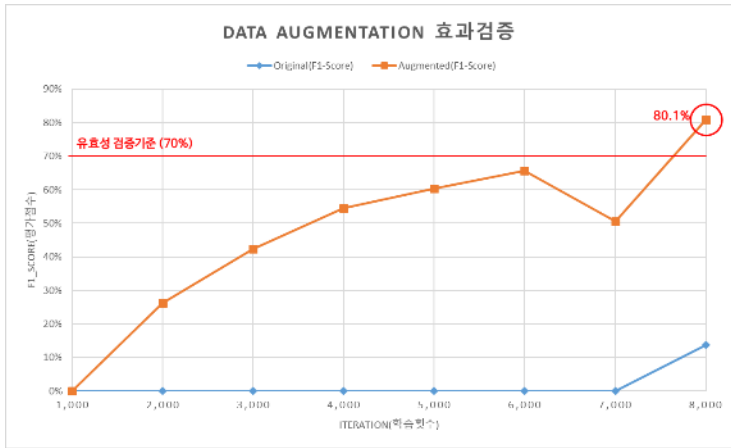
따라서,  
기존 이미지와 Data Augmentation을 거친 이미지를 아예 다른 이미지로 인식할 수 있다.

## Data augmentation 효과 검증

- ✓ 데이터를 변경하고 조작하여 데이터양을 늘리는 데이터 증강 기법은 모델 학습에 도움이 될 수도 있고, 학습을 방해할 수도 있다.
- ✓ 보편적으로는 절대적 데이터양이 어느정도 확보되어야 하므로 Data augmentation을 통한 데이터 보충은 학습에 도움이 된다는 레퍼런스가 많다.
- ✓ Ab04 복부종양, Ch02 기관허탈, Ch05 기흉 세 질병의 경우, 질환에 대한 데이터양이 매우 적었기 때문에 Data augmentation을 진행하여 클래스별 비율을 맞추어 주었다.

ex) Ch05 기흉

	Train 데이터 수	Valid 데이터 수	Test 데이터 수
NOR (정상)	13,799	1,726	1,726
ABN (질환)	(original) 320 → (Data Aug) 14,080	41	41
NOR/ABN	(original) 43.1 → (Data Aug) 1.0	42.1	42.1



학습 진행되는 모습을 보임

	원본이미지	Data Augmentation
1000회	0.0000	0.0000
2000회	0.0000	0.2632
3000회	0.0000	0.4242
4000회	0.0000	0.5455
5000회	0.0000	0.6038
6000회	0.0000	0.6557
7000회	0.0000	0.5057
8000회	0.1379	0.8085

## 학습용 데이터 질환 클래스에 Data Augmentation 기법 적용



# 질적분석(테스트 이미지) - Ab01 간비대

간비대 질환은 간이 비정상적으로 커지는 증상을 가르키며, 간이 횡격막에서 늑골궁 라인 사이의 영역을 벗어난 경우 질환으로 분류

간비대 (Ab01) 라벨링 가이드

정상 이미지

질환 이미지

테스트셋 모델 분류 결과

Ab01_간비대		정답값 (Ground Truth)	
		질환	정상
예측값 (Prediction)	질환	127	57
	정상	31	339

〈분류 평가점수〉

- Accuracy : 0.84
- Precision : 0.69
- Recall : 0.80
- F1-Score : 0.74

#1. 정확히 분류해낸 사례

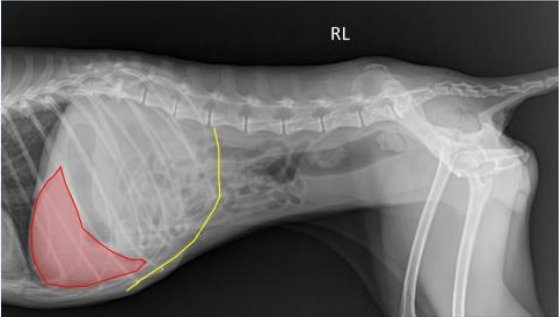
이미지

모델 예측

참값

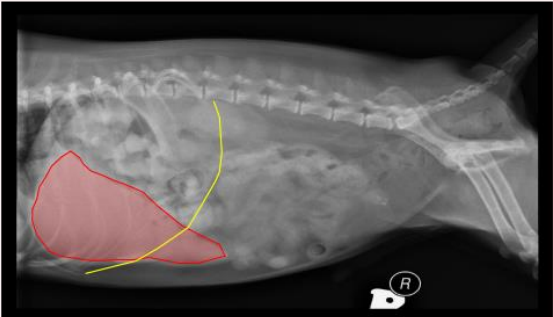
RL

횡격막 경계와 늑골궁 라인 사이에 삼각형의 간 음영이 있는 경우 **정상**입니다.



R

횡격막 경계에서 늑골궁 라인 밖으로 삼각형의 간 음영이 있는 경우 **질환**입니다.



정상 이미지

질환 이미지

테스트셋 모델 분류 결과

Ab01_간비대		정답값 (Ground Truth)	
		질환	정상
예측값 (Prediction)	질환	127	57
	정상	31	339

〈분류 평가점수〉


- Accuracy : 0.84
- Precision : 0.69
- Recall : 0.80
- F1-Score : 0.74

이미지


모델 예측

참값


R

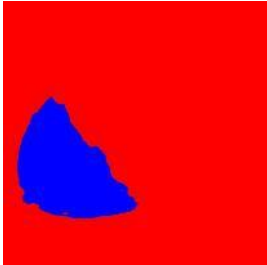


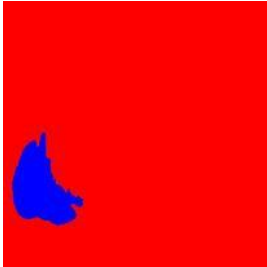
R

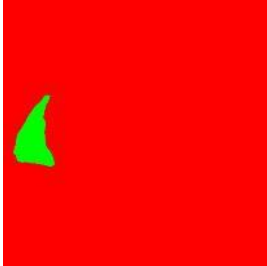


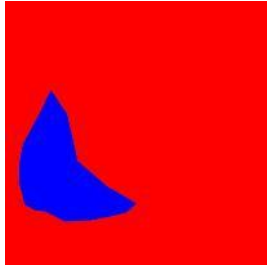
R

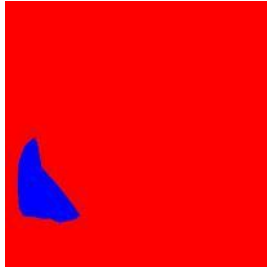


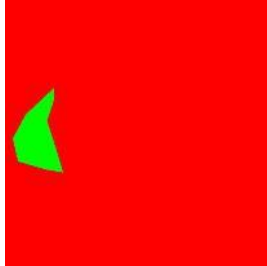










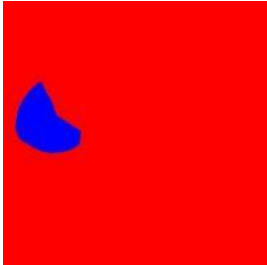
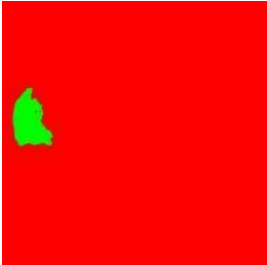

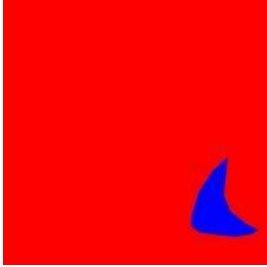
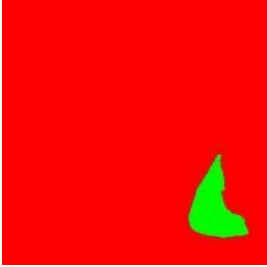

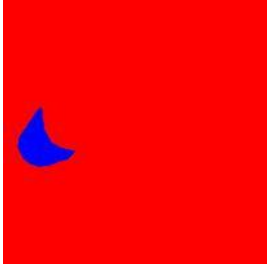


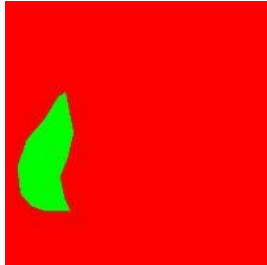


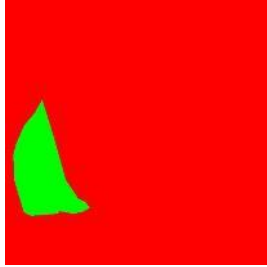
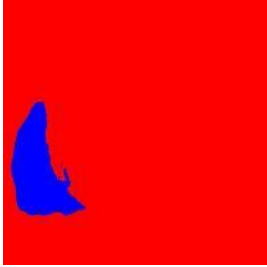

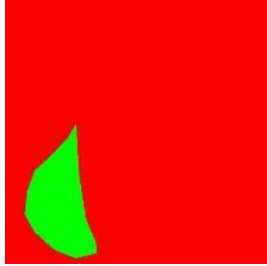




\* 빨강 : 배경, 초록 : 정상, 파랑 : 질환

10

# 질적분석(테스트 이미지) - Ab01 간비대

간비대 학습 모델이 제대로 분류하지 못한 사례를 분석한 결과, GT라벨의 정상과 질환이 반대로 라벨링 되어 있는 것으로 의심되는 데이터를 발견

#2. 분류 실패 사례 - (1) 질환 간이 작은 경우	#2. 분류 실패 사례 - (2) 정상 간이 큰 경우
<div><div>이미지</div><div>모델 예측</div><div>참값</div><div></div><div></div><div><div>R</div></div></div> <div>* 빨강 : 배경, 초록 : 정상, 파랑 : 질환</div>	<div><div>이미지</div><div>모델 예측</div><div>참값</div><div></div><div></div><div></div></div> <div>* 빨강 : 배경, 초록 : 정상, 파랑 : 질환</div>

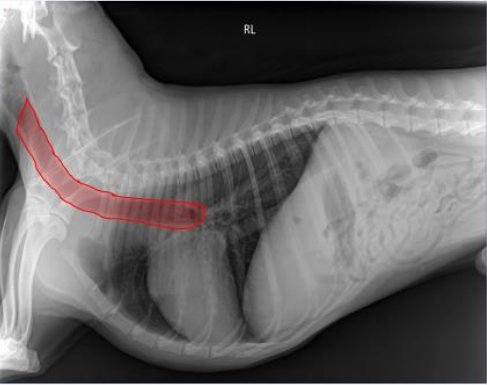
# 질적분석(테스트이미지) - Ch02기관허탈

기관허탈 질환은 기관이 비정상적으로 좁아지는 증상을 말하며, 목부터 가슴부위로 이어지는 기관의 두께가 일정하지 않고 좁아지는 경우 질환으로 분류

기관허탈 (Ch02) 라벨링 가이드

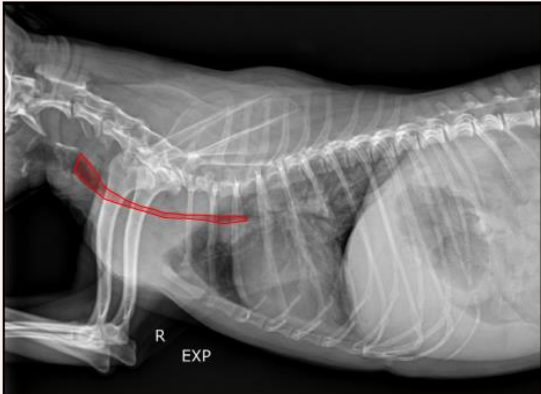
정상 이미지

목부터 가슴부위에 있는 기관이 좁아지지 않고 일정할 경우 정상입니다.



질환 이미지

목부터 가슴부위에 있는 기관이 좁아지는 경우 질환입니다.



테스트셋 모델 분류 결과

Ch02_기관허탈		정답값 (Ground Truth)	
		질환	정상
예측값 (Prediction)	질환	80	33
	정상	31	801

〈분류 평가점수〉

- Accuracy : 0.93
- Precision : 0.71
- Recall : 0.72
- F1-Score : 0.71

#1. 정확히 분류해낸 사례

이미지

모델 예측

참값





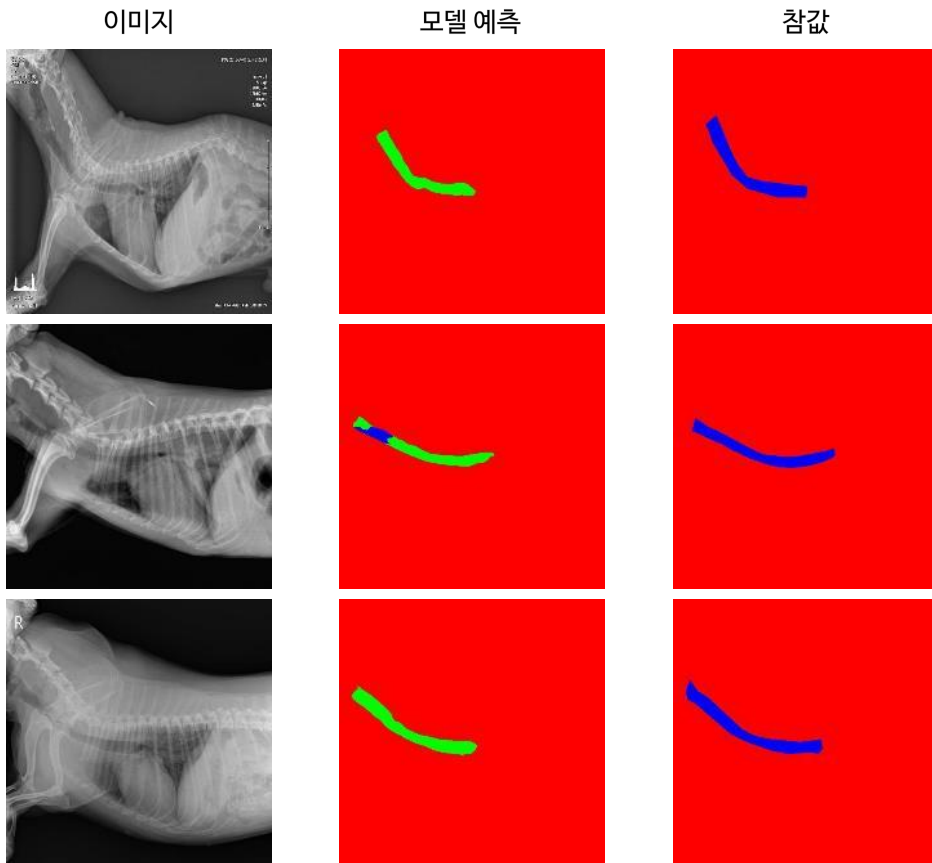


\* 빨강 : 배경, 초록 : 정상, 파랑 : 질환

# 질적분석(테스트 이미지) - Ch02기관허탈

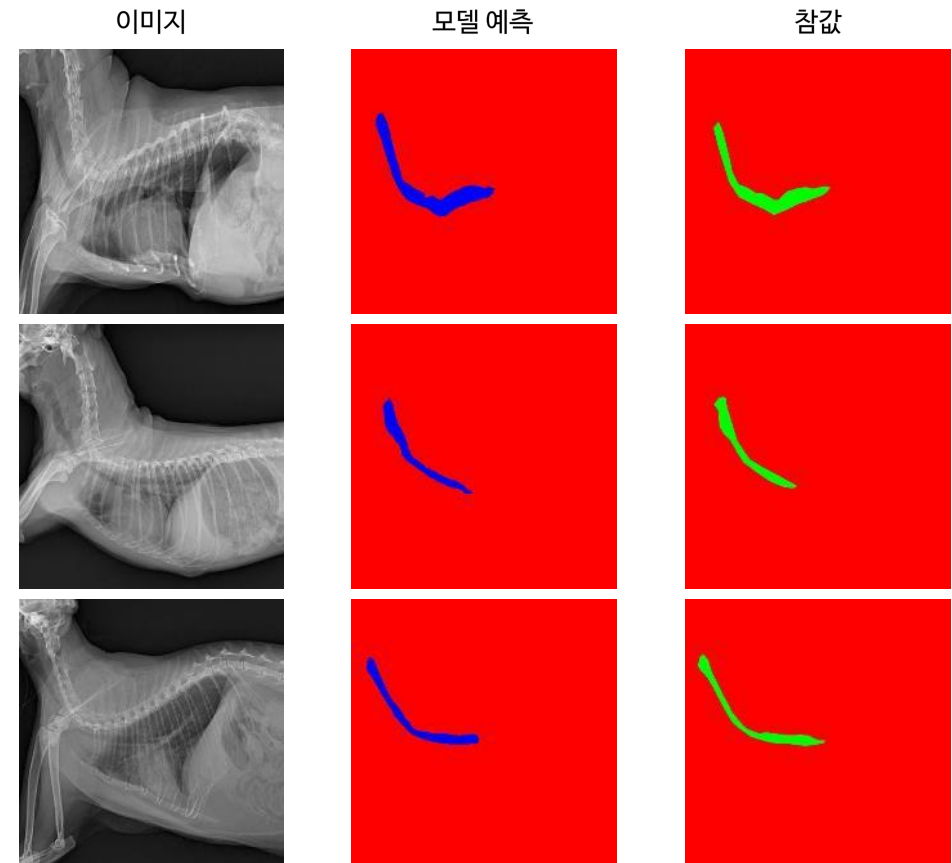
기관허탈 학습 모델이 제대로 분류하지 못한 사례를 분석한 결과, GT라벨의 정상과 질환이 반대로 라벨링 되어 있는 것으로 의심되는 데이터를 발견

## #2. 분류 실패 사례 - (1) 질환 기관의 두께가 일정한 경우



\* 빨강 : 배경, 초록 : 정상, 파랑 : 질환

## #2. 분류 실패 사례 - (2) 정상 기관의 두께가 일정하지 않은 경우



\* 빨강 : 배경, 초록 : 정상, 파랑 : 질환

# 목차

---

1. 개요

2. 모델 및 데이터 업그레이드 방안

3. 학습 결과

**4. 진행 간 이슈**

5. 소감 및 마무리



# 평가지표 (F1-score+IoU)

주어진 평가지표는 Image Classification 모델을 평가하는 F1-Score 평가방식과 Segmentation, Object Detection 모델을 평가하는 IoU 평가방식이 혼합된 지표로써  
보편적으로 사용되는 지표가 아니기 때문에 실제 사용에 있어 맹점이 있으나, 실제 서비스의 경우 보편적 지표보다 사례에 적합한 지표를 개발해야 한다는 시사점 발견

평가 지표 비교				*TTA 평가 방식	
구분	지표	분류 단위	대표 적용 과제		
보편 지표	F1-score	이미지	Image Classification		
	IoU	객체 (픽셀, 바운딩 박스)	Segmentation, Object Detection		
	mAP	객체 (바운딩 박스)	Object Detection		
TTA 지표	F1-score + IoU	이미지 + 객체	-		

평가지표의 맹점	평가지표의 시사점
<div>1. 분류단위가 서로 다른 F1-score와 IoU를 결합함으로 인해 <b>분류 결과 판단 방식이 모호하다는 점.</b> 한 이미지에 정상과 질환 클래스가 동시에 예측된 경우 해당 이미지는 정상인지 질환으로 분류할 지, 또는 이미지 안의 객체가 질환으로 예측되어 질환 이미지로 분류하더라도 IoU 기준에 맞지 않는 경우 해당 이미지는 정상으로 분류해야 하는지 등 다양한 사례들에 대한 분류 판단 방식이 모호했습니다.</div> <div>2. <b>IoU 기준 설정에 합리적인 근거 및 가이드라인이 없는 점.</b> IoU 0.5를 일괄 적용하도록 했지만, 질환별 질환 탐지의 난이도를 고려하지 않아, Ch02 기관허탈과 같이 특정 질환의 성능이 과소평가되는 경우가 있습니다.</div>	<div>1. 서비스 관점에서 보편적 지표에 얽매이지 않고 <b>서비스에 더 적합한 지표를 개발해야 합니다.</b> 질병 진단 목적을 고려하면, 이미지 상의 질병 여부 뿐 아니라 질환 위치에 대한 정확한 예측 여부도 중요할 것입니다.</div> <div>2. 앞선 맹점을 해결하면 <b>F1-score와 IoU를 결합한 지표도 유의미한 성능 평가 지표로 활용할 수 있습니다.</b></div>

\*보다 상세한 설명은 부록(TTA 평가기준 변경(12/17) - 질환여부 평가방식 변경 (2)) 참조

# 목차

---

1. 개요

2. 모델 및 데이터 업그레이드 방안

3. 학습 결과

4. 진행 간 이슈

**5. 소감 및 마무리**



# 소감및마무리

이번 반려동물 유효성 검증모델 개발사업을 통해 배운 점과 아쉬운 점, 그리고 한 줄 소감

## 〈배운 점〉

- 이론이 아닌 실제 모델 학습 수행
- 다양한 모델 학습 개선방안(Optimizer, LR Scheduler) 탐구
- GPU 분산 학습 경험
- AI 모델의 성능은 데이터의 품질이 가장 중요하다는 점
- Data augmentation 등을 통한 클래스 비율의 균형의 중요성
- 파이썬 모듈/클래스 구조 및 관계성 파악

## 〈아쉬운 점(개선방안)〉

- 동시에 코드 분석 및 수정작업 시, 공유작업의 어려움 (다음 번엔 Git 같은 협업을 위한 버전 관리 도구를 사용해볼 것)
- 수시로 바뀌는 이미지/라벨데이터에 대한 관리방안 부재(이미지/라벨데이터의 DB화, 특정 공간에 이미지/라벨데이터를 저장해두고 DB처럼 관리해보는 방안)
- 계획적인 실험이 필요 (이를 위해 모든 실험의 결과를 정리하고, 새로운 가설을 세워 검증하는 방식 도입필요)
- Detectron2 에서 validation loss값에 대한 신뢰성이 부족해 학습 정도 파악이 어려웠음. (1 iteration = 1 batch학습 (not 1 epoch))

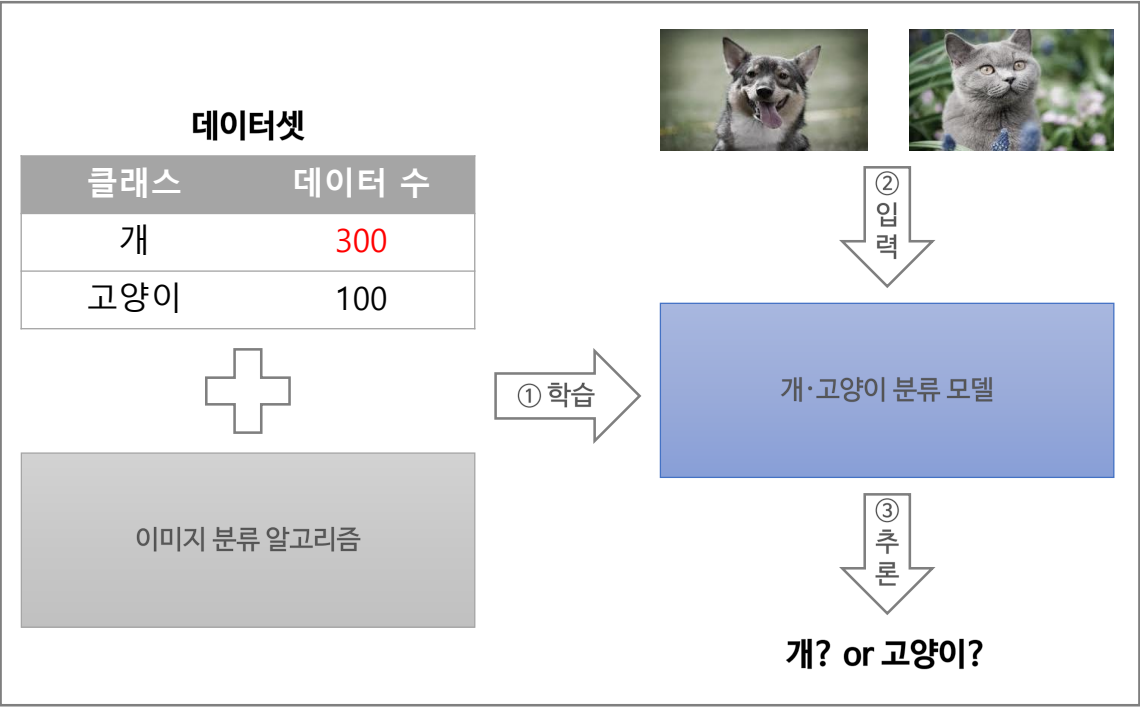
## 〈한 줄 소감〉

- 실전을 통해 머리로만 이해하던 이론을 체득하기도 했지만, 머신러닝 연구는 계획적으로 실험하고 결과를 기록하는 것이 중요하다는 것을 경험했습니다.
- 첫 비전 딥러닝 모델링 프로젝트로써 데이터 이슈, 학습성능 이슈, 시간적 압박 등 많은 어려움을 이겨내고 목표치를 달성해냈다는 점에서 기쁘고 앞으로 유사한 프로젝트에서 더 좋은 퍼포먼스를 보여줄 수 있다는 자신감이 듭니다.
- 도커 환경 이해하는 것부터 모델 시뮬레이션해보고 이론 개념을 모델에 적용해보는 과정을 경험해볼 수 있는 좋은 기회였다.

부록>데이터 불균형

분류 문제에서 클래스 비율이 불균형한 데이터셋으로 학습한 모델은 편향된 예측을 하는 문제가 발생합니다.

불균형 데이터 학습 문제점



개 이미지가 많은 데이터셋을 학습한 분류 모델은  
어떤 이미지를 입력해도 개로 예측하는 경우가 더 많음

유효성 검사용 데이터의 클래스 비율

질환명	정상데이터 수	질환데이터 수	정상 : 질환
Ab01 간비대	3954	1579	2.5 : 1
Ab04 복부종양	9140	370	24.7 : 1
Ch02 기관허탈	8336	1106	7.5 : 1
Ch05 기흉	17251	402	42.9 : 1
Mu02 사지골절	4214	897	4.7 : 1
Mu04 엉덩관절탈구	3915	886	4.4 : 1

정상데이터에 비해 질환데이터가 부족해 이 데이터셋으로 학습한 모델은  
거의 모든 이미지에 대해 정상이라 예측할 가능성이 높다.

## 부록 > Weighted Loss

데이터 불균형 문제를 해결하기 위해 ① 데이터를 추가 확보하거나, ② Weighted Loss, ③ Data Augmentation를 활용할 수 있습니다. 이 중 클래스별 Weight를 적용하는 방법이 가장 간단하고 쉽게 적용할 수 있습니다.

### 데이터 불균형 문제 해결을 위한 Weighted Loss



### 유효성 검사용 데이터셋의 Loss Weight 구하기

Weight는 임의의 숫자를 정할 수 있다. 데이터 불균형 문제를 해결하는 목적으로 Weight를 적용한다면, 간단한 방법 중 하나는 (1-해당 클래스의 데이터 비율)을 Weight로 하는 것이다.

- Segmentation 모델 Loss Weight (데이터 단위 = Pixel)

질환	클래스	이미지 수	픽셀 수	픽셀 비율	Weight
Ab01	배경	-	1,064,508,477	0.9634	0.0366
	정상	3162	24,493,430	0.0222	0.9778
	질환	1262	15,935,053	0.0144	0.9856
Ch02	배경	-	3,416,516,862	0.9815	0.0185
	정상	5780	36,765,730	0.0106	0.9894
	질환	*6022	27,727,584	0.008	0.992

- Object Detection 모델 Loss Weight (데이터 단위 = Bounding Box 객체)

질환	정상:질환 이미지 수	정상:질환 객체 수	정상:질환 객체 비율	정상:질환 Weight
Ab04	7312 : *7400	7312 : 7450	0.4953 : 0.5047	0.5047 : 0.4953
Ch06	13799 : *14080	13799 : 14080	0.495 : 0.505	0.505 : 0.495
Mu02	3371 : 716	0 : 735	-	-
Mu04	3170 : 717	5594 : 1190	0.8246 : 0.1754	0.1754 : 0.8246