

검정고무신 이미지 생성 모델을 활용한 서비스

부제: 클라우드를 이용한 시스템 아키텍처 효율적 구축



201918298 김지성

201911895 정예성

201911838 황예찬

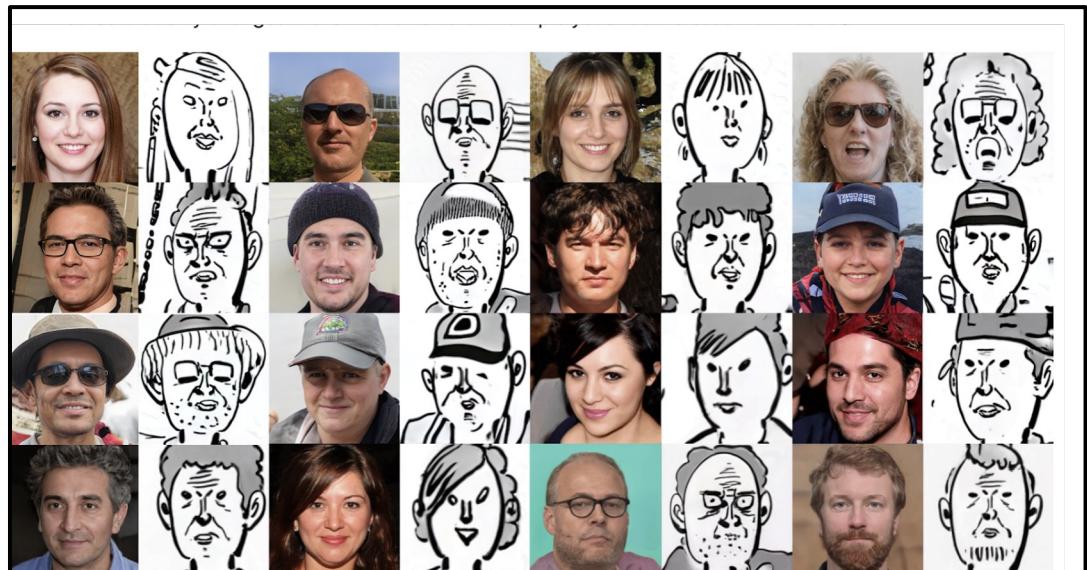
201911752 강봉구

목차

- 개요
- 주제 및 목표
- 시스템 아키텍처
 - LoadBalancer, AutoScalingGroup를 활용한 인스턴스(모델 + 웹) 서버
 - Lambda를 활용한 웹 서버
- 실험 결과
 - 비용 측정
 - 속도 테스트
 - 트래픽 한계 테스트
 - 서비스 사용 예제
- 한계 및 개선방안

개요

생성형 AI



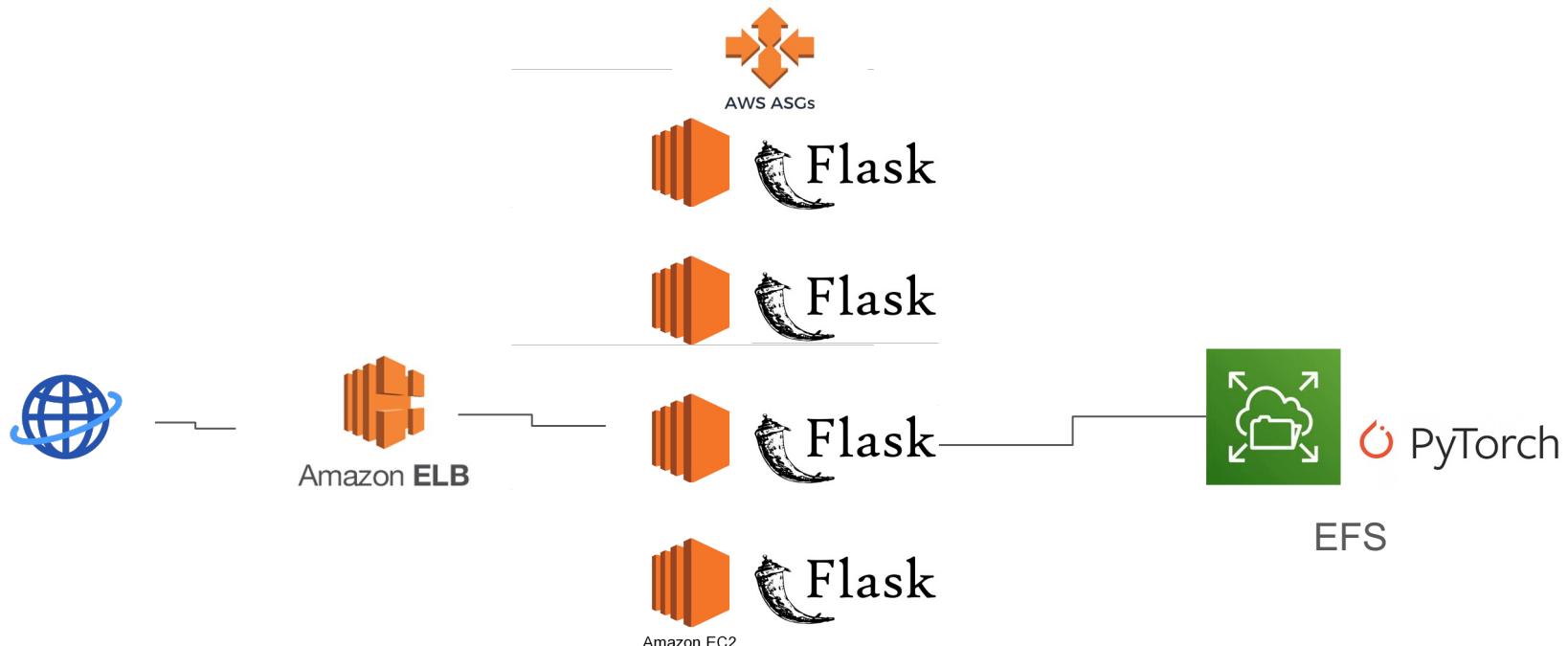
개요



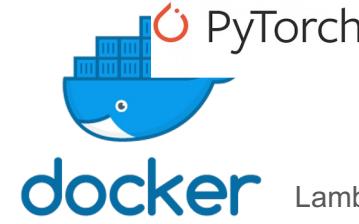
개요

- 주제 : 클라우드를 이용한 시스템 아키텍처 효율적 구축
 - 클라우드를 이용한 검정고무신 이미지 생성 모델
- 목표 : 인스턴스 활용 아키텍처와 Lambda 활용 아키텍처를 비교하여 장단점을 분석
- 동작 방식: 웹에 변환을 원하는 인물 이미지를 업로드 후 생성된 이미지를 확인한다.
- 활용 기술
 - 사용 모델
 - UGATIT-pytorch (<https://github.com/znxlwm/UGATIT-pytorch>)
 - 데이터셋 구축
 - 얼굴 인식: <https://github.com/ultralytics/yolov5/releases/tag/v2.0>
 - 원본 검정고무신 이미지:
<https://www.youtube.com/playlist?list=PLrNFI43wt6gCEdWfQjzPf2Dnza7liRcpL>
 - 원본 얼굴 데이터셋: <https://github.com/JingchunCheng/All-Age-Faces-Dataset>

ALB + ASG 인스턴스 서버



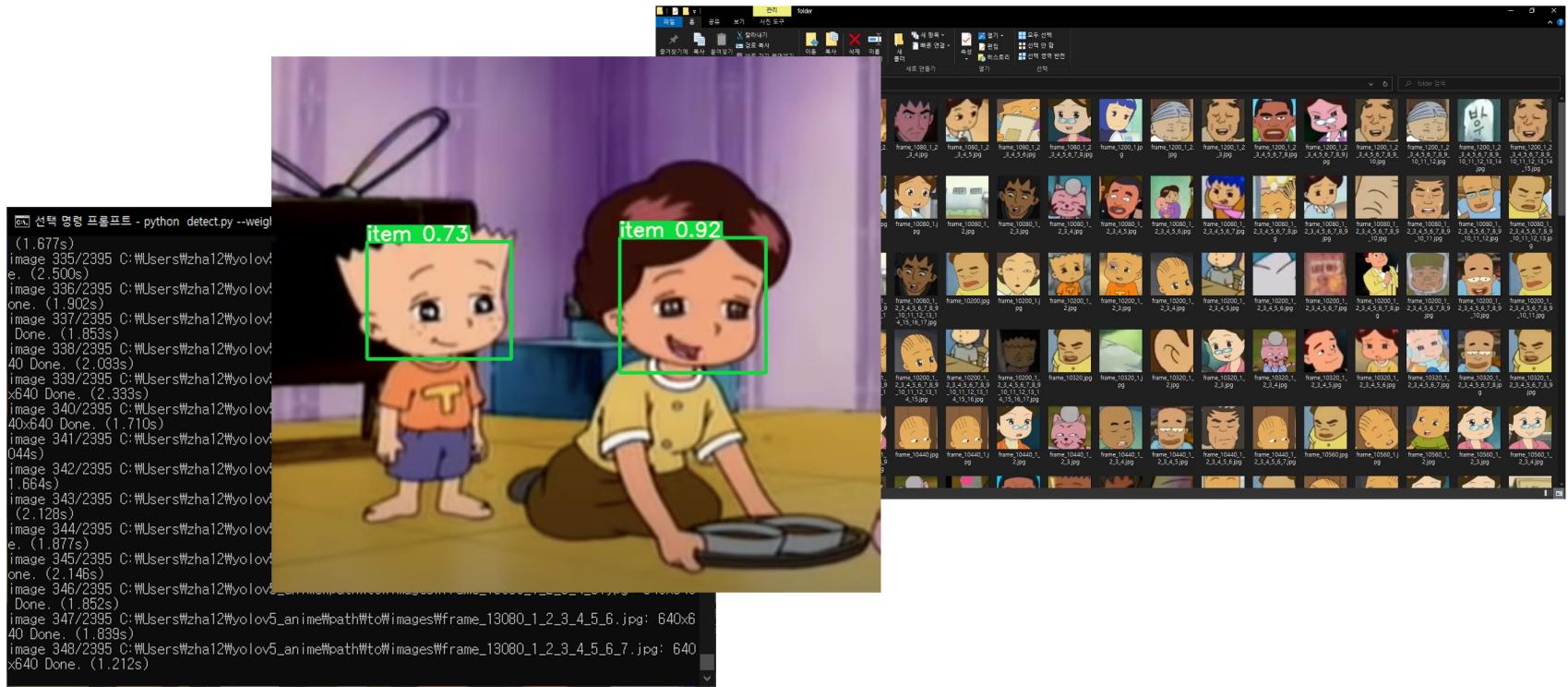
Lambda + 웹 서버



Lambda는 250MB로 제한

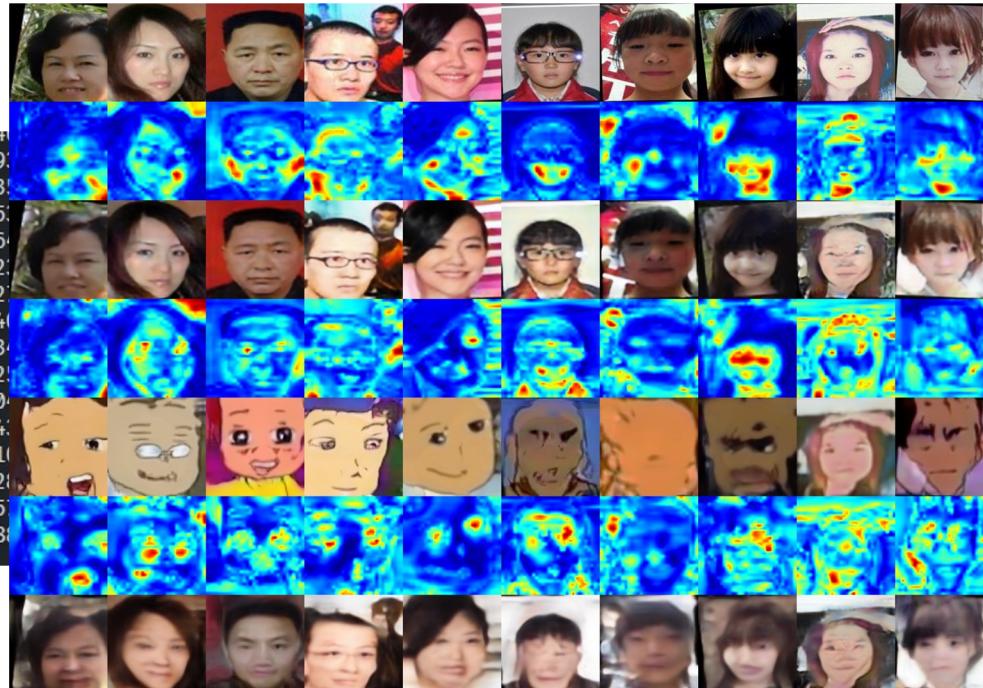


모델 학습 데이터셋 생성 - yolo5 활용



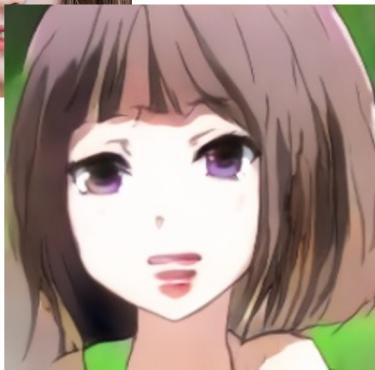
모델 학습

```
[ 5825/250000] time: 5395.8575 d_loss: 2.526584  
[ 5826/250000] time: 5394.7844 d_loss: 2.825349  
[ 5827/250000] time: 5395.7132 d_loss: 2.265783  
[ 5828/250000] time: 5396.6418 d_loss: 2.335025  
[ 5829/250000] time: 5397.5707 d_loss: 2.670626  
[ 5830/250000] time: 5398.4995 d_loss: 2.644992  
[ 5831/250000] time: 5399.4291 d_loss: 2.463432  
[ 5832/250000] time: 5400.3583 d_loss: 2.888154  
[ 5833/250000] time: 5401.2868 d_loss: 2.457223  
[ 5834/250000] time: 5402.2161 d_loss: 3.002572  
[ 5835/250000] time: 5403.1437 d_loss: 2.774020  
[ 5836/250000] time: 5404.0737 d_loss: 3.154744  
[ 5837/250000] time: 5405.0016 d_loss: 2.825731  
[ 5838/250000] time: 5405.9318 d_loss: 2.537362  
[ 5839/250000] time: 5406.8599 d_loss: 2.809345  
[ 5840/250000] time: 5407.7900 d_loss: 2.881838
```



초기 결과

기존 서비스



Selfie2Anime アニメ

HOME BLOG ABOUT PORTFOLIO CONTACT

English / 中文

WHAT DO YOU LOOK LIKE IN ANIME?

Upload a Selfie

(Photos you upload will NOT BE PUBLISHED)

SHARE TWEET SEND

The screenshot shows the homepage of the Selfie2Anime website. At the top, there's a navigation bar with links for HOME, BLOG, ABOUT, PORTFOLIO, and CONTACT. Below that is a language selection bar with English and Chinese options. The main feature is a large, rounded rectangular box with a dashed red border containing the text "WHAT DO YOU LOOK LIKE IN ANIME?". Below this text is a pink cloud icon with an upward arrow and the instruction "Upload a Selfie". Underneath the box, it says "(Photos you upload will NOT BE PUBLISHED)". At the bottom of the box are three social media sharing buttons for Facebook, Twitter, and Email. To the left of the main content area is a vertical sidebar with a grid of anime faces, and to the right is another grid of anime faces. A file explorer window is visible on the far right side of the screen.

프로젝트 시연



Google Chrome 12월 15일 (금) 00:19

이미지 업로드 및 Drag & Drop

BRS 테마 변경

이미지를 드래그 앤 드롭하거나, 클릭하여 파일을 선택하세요



이미지 처리

♂ ♀

이미지 처리

이미지 업로드 및 Drag & Drop

BRS 테마 변경

이미지를 드래그 앤 드롭하거나, 클릭하여 파일을 선택하세요



이미지 처리

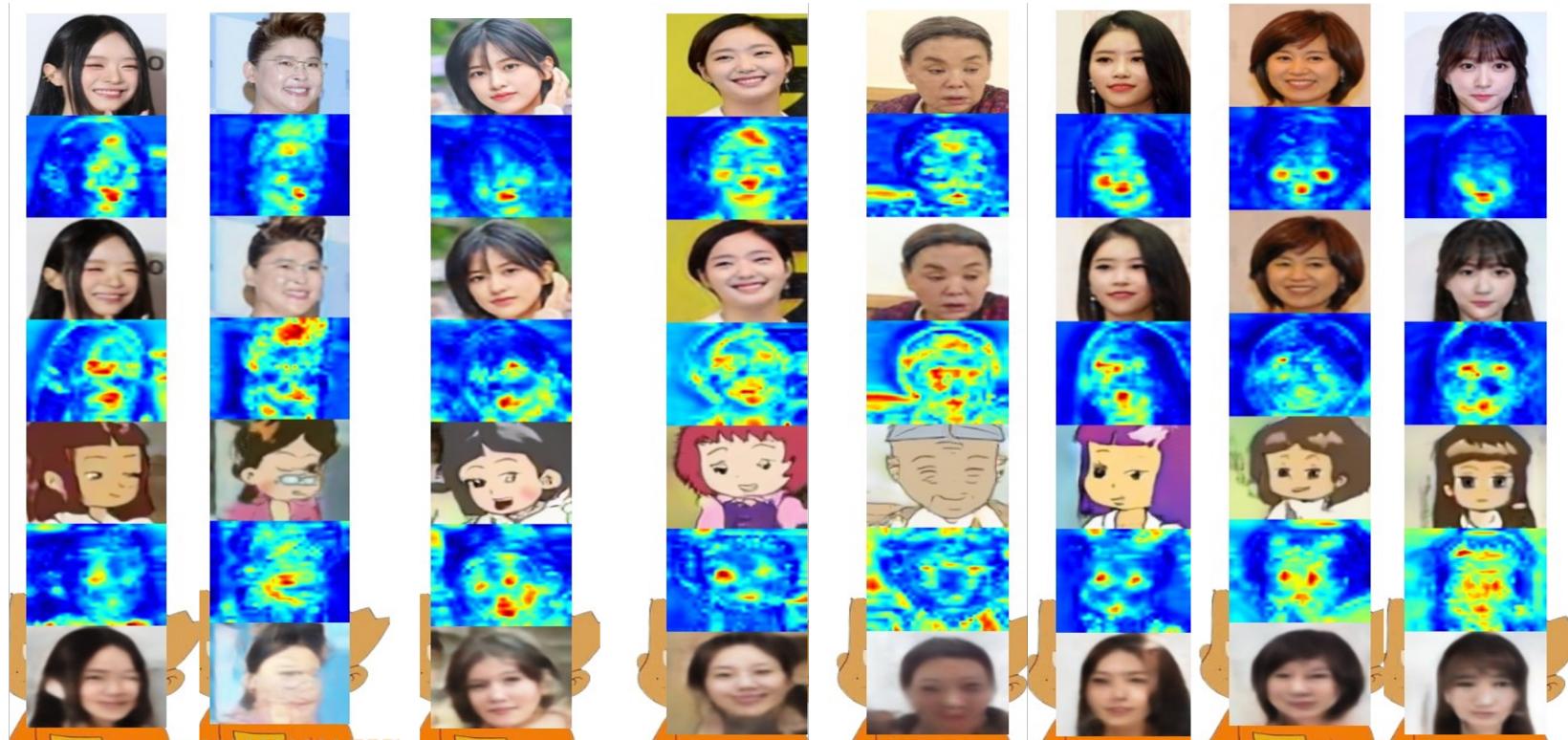
♂ ♀

이미지 처리

이미지 업로드 및 Drag & Drop

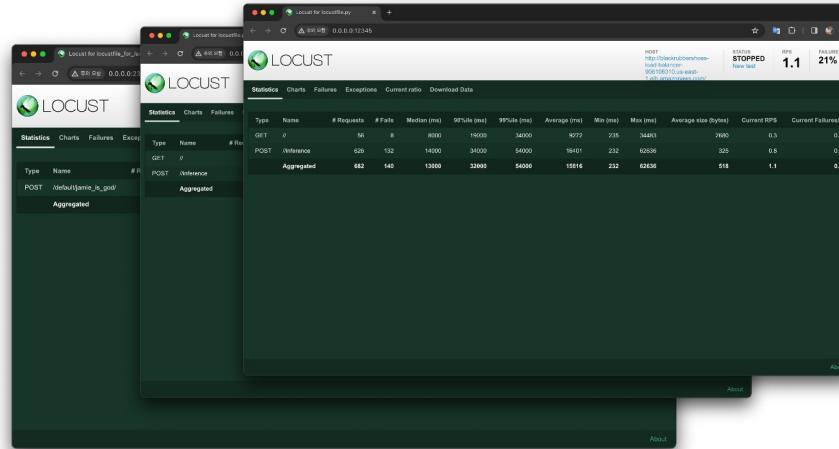
BRS 테마 변경

프로젝트 시연 결과



속도 및 트래픽 비교

- 측정 환경: Locust 사용하여 측정
- 사용자가 이미지를 업로드한 시점부터 출력 이미지를 전달받는 시점까지 계산



속도 및 트래픽 비교

- 각 실험은 15분 간 측정

요청수/ 단위 시간	Instance 서버					Lambda + 웹 서버				
	Req.	Fail	Mean (ms)	Min (ms)	Max (ms)	Req.	Fail	Mean (ms)	Min (ms)	Max (ms)
(1회/10초)	115	6	5899	248	20657	62	0	15107	9051	23191
(10회/10초)	372	67	22171	222	60654	681	1	11639	384	21216
(15회/10초)	682	140	15816	232	62636	1409	617	6858	242	20905

비용 측정

AWS Calculator

구성 AWS Lambda 정보

설명
예상 비용에 대한 설명 입력

위치 유형 선택 정보
리전

Lambda 함수 - 프리 티어 포함
Lambda는 콘솔에서 테스트 호출을 포함하여, 이벤트 알림이나 호출에 대한 응답으로 요청 실행이 시작될 때마다 요청 수를 계산합니다. 그리고 요금은 모든 함수에서 총 요청 수에 대해 청구됩니다. 요금은 함수에 할당한 메모리 크기에 따라 달라집니다. Lambda 프리 티어에는 월별 1백만 개의 무료 요청과 총 400,000GB-초의 컴퓨팅 시간이 포함됩니다.

Lambda 함수
Lambda는 콘솔에 모든 함수에서

리전 선택
미국 동부(오하이오)

서비스 설정 정보

아키텍처
x86

요청 수
금액 입력

유닛
월별

각 요청의 기간(ms)
소요 시간은 코드가 실행되기 시작한 시간부터 반복 또는 종료될 때까지 계산됩니다.

소요 시간(ms) 입력

할당된 메모리 크기
100MB이 100B 사이의 크기 이다

총 결제 비용: 0.00 USD | 세부 정보 표시 ▾
월별 총비용: 0.00 USD

제공 메인화면 | 구성 | 제작 메인화면

비용 비교

요청수/ 단위 시간	Instance 서버			Lambda + 웹 서버		
	EC2 (t2.medium)	ALB	합계	EC2 (t2.micro)	Lambda (3000MB)	합계
Toy Project (680회/1일)	33.87USD (1대)	18.77USD (1 대)	52.64USD	8.47USD (1대)	5.25USD	13.72USD
기존 모델 이 용자 서비스 (3300회/1일)	33.87USD (1대)	18.77USD (1 대)	52.64USD	8.47USD (1대)	20.78USD	29.25USD
(7200회/1일)	33.87USD (1대)	18.77USD (1 대)	52.64USD	8.47USD (1대)	51.13USD	59.6USD
네이버 Snow (5만 회/1일)	270.9USD (8 대)	19.35USD (1 대)	290.33USD	8.47USD (1대)	381.38USD	389.85USD

한계 및 개선방안

- 모델
 - 성능 개선
- 램다
 - 콜드 스타트
- S3 활용(IAM 이슈)
 - EFS 대신 아카이빙
 - 한 번 처리가 된 이미지는 캐싱하여 처리