



2020 DreamAI Healthcare Hackathon

TASK 1

Medical Diagnosis challenge for Covid19

Team Nekaravuru

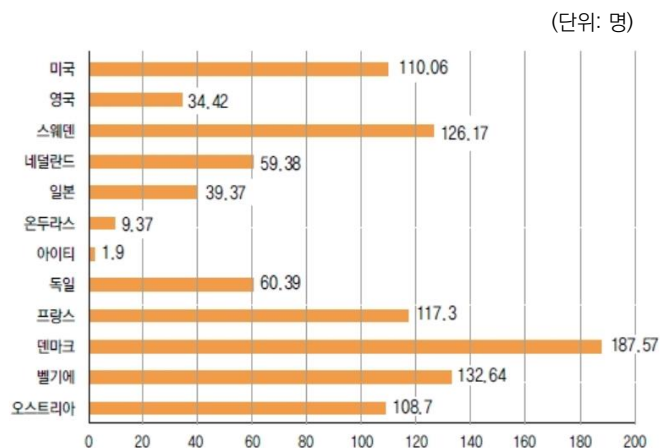
목차

1. 문제 인식 및 개발동기
2. 아이템 소개
3. 아이템의 기술적 설명
4. 사업화 계획
5. 시장 분석 및 진입 전략
6. 자금 계획
7. 사업 확장 계획

01 아이템의 필요성

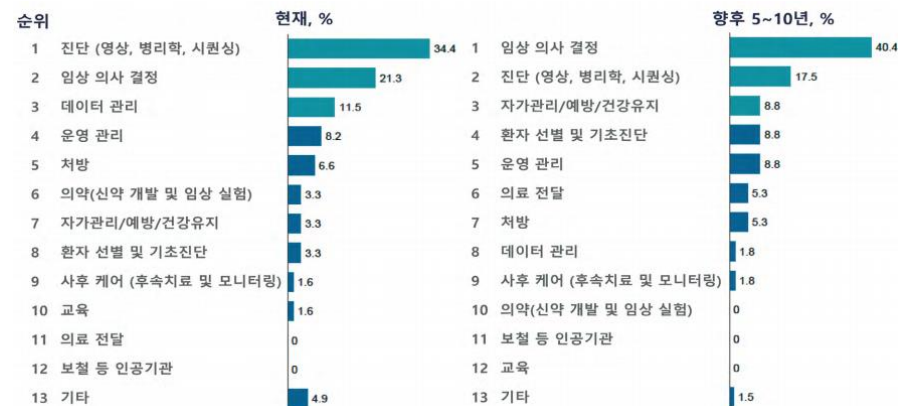
■ 효율적인 COVID-19 보조진단 도구의 필요성

인구 100만명 당 영상의학과 의사 수



출처: Journal of American College of Radiology

의료전문인들이 예상하는 중요한 AI 적용 분야



출처: EIT health and McKinsey & Company(2020.3)

01 아이터의 필요성

■ 팬데믹 장기화로 인한 개인의 불안감 급증

HOME > 사회 > 경기

코로나로 우울·불안·울분 22만명 정신적 고통 호소

이경훈 | 승인 2020.10.05 18:37 | 수정 2020.10.05 18:36 | 2020.10.06 1면 | 댓글 0

정신건강복지센터 상담 유형을 보면 '기침이 나는데 코로나19 증상이 아닌지 불안하다'와 같은 단순한 증상 관련 문의부터 '나도 코로나19에 걸리지 않았는지 무섭다', '집 밖에 나가지 못하겠다', '사람들을 만나기가 무섭다' 등 불안감을 호소하는 사례가 많았다.

마음마져 갇아먹는 코로나19...불안·대인기피 등 상담 줄이어

게시일 2020-03-05 조회수 426

'코로나19'에 심리적 불안도 심해졌다...상담 사례 잇따라

입력 2020.03.05 07:44 | 수정 2020.03.05 07:44

5일 국가트라우마센터에 따르면 코로나19 사태가 본격화한 1월 29일부터 이달 3일까지 1개월여 동안 트라우마센터와 전국 각 지역 정신건강복지센터에는 확진자와 자가격리자는 물론 일반인들의 코로나19와 관련한 정신적 스트레스 등 심리적 문제 상담이 줄을 잇고 있다.

같은 기간 전국 정신건강복지센터가 자가격리자와 일반인 요청을 받아 진행한 코로나19 관련 심리상담은 1만8천60건에 달했다. 확진 판정을 받지 않았더라도 일상에서 코로나19에 대한 불안감이 심해 관계기관 상담까지 받으려는 수요가 그만큼 많은 것으로 풀이된다.

01 개발 동기

- 특별 장비를 필요로 하지 않는 자가진단 방법을 일반 사람들에게도 제공
→ 어느 정도 불안감 해소 가능



기침을 이용한 개인용 자가진단 모바일 앱



- 비용 및 시간 효율적인 진단보조 방법 도입
→ 더 빠르고 효율적인 진단 가능
→ 의료진이 치료 등 다른 업무에 집중할 수 있음



병원용 진단보조 웹사이트



02 아이템 소개



■ 개인용 자가진단 모바일 앱

- 기침을 통한 COVID-19 간편한 자가 진단
- 양성인 경우 선별진료 권유
- 사용자 동의 하에 익명화된 데이터 수집

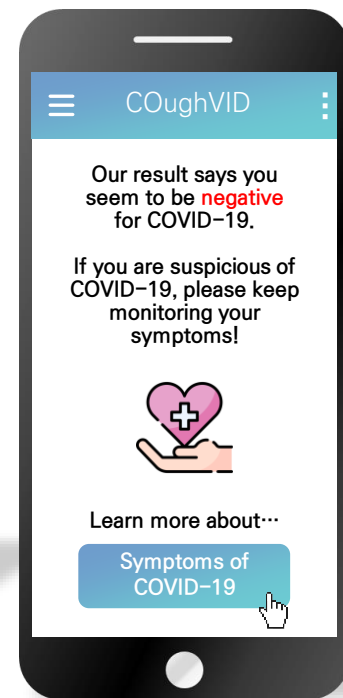
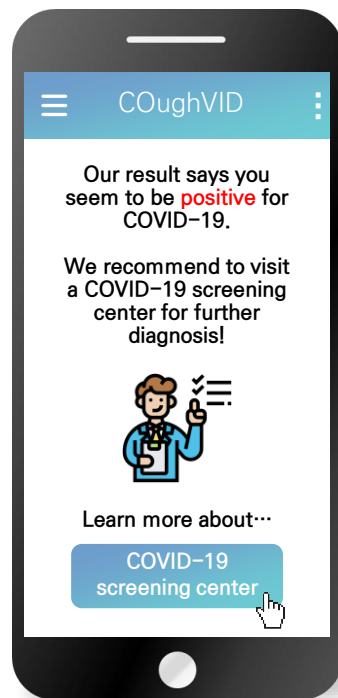
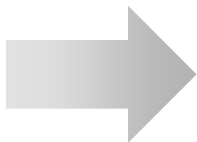
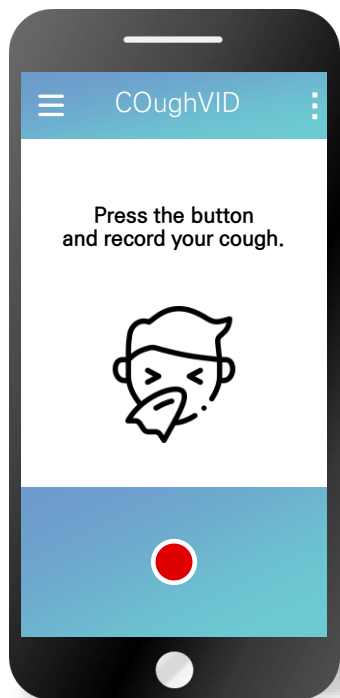


■ 병원용 진단보조 웹사이트

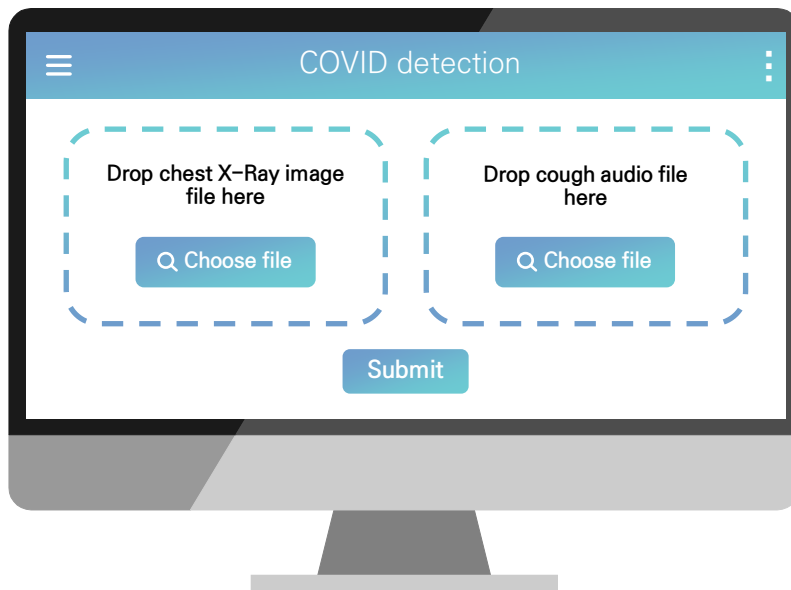
- 의료영상 및 기침 데이터 활용해 진단보조
- 더 많은 정보로 진단 성능을 높임
- 의심 병변에 대한 정보 제공
- 같은 환자의 의료영상-기침 데이터 쌍 수집

02 MVP

- 데모 동영상



- 데모 동영상



03 기술 설명

- 목차

1

사용한 데이터셋

2

데이터 전처리 과정

3

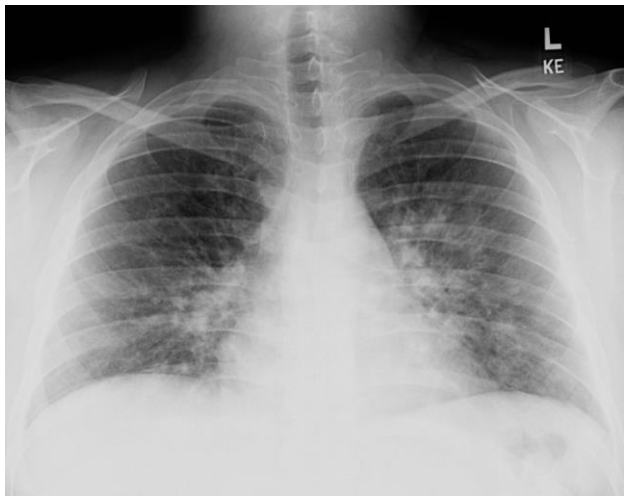
제안 모델의 구조

4

학습 결과 (검증 성능)

03 데이터셋

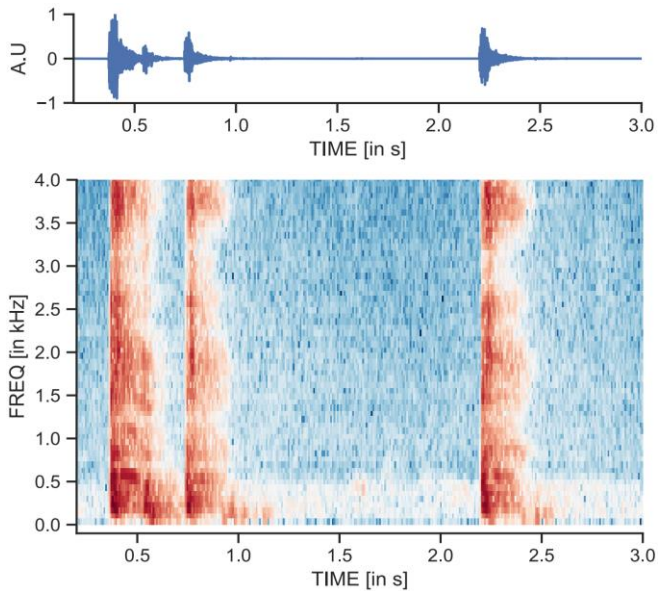
- COVIDx Dataset



데이터 종류	흉부 X-Ray 영상
데이터 수	13,958 (Non-COVID: 13,441, COVID-19: 617)
포맷	jpg, png (multisite)
크기	Site마다 다름
수집 기관	<ul style="list-style-type: none">• COVID chest X-ray dataset• COVID19 radiology dataset• RSNA pneumonia detection challenge

03 데이터셋

- Coswara Dataset



데이터 종류	기침 음성 데이터
데이터 수	1,442 (Healthy: 1,158, COVID-19: 98)
포맷	.wav
크기	2 ~ 10초
수집 기관	Indian Institute of Science

03 데이터 전처리

■ Image dataset



- Multisite dataset
- 같은 크기로 resize
- 밝기 값 normalization

■ Audio dataset



- 음성 데이터의 길이가 수 초에서 10초 정도로 다름
- Peak point를 기준으로 전후 2초씩 총 4초로 데이터를 분할
- Mel-spectrogram을 추출하여 사용

03 데이터 전처리

■ Class balance

- Weight balancing 방법과 healthy class undersampling 방법 시도
- 더 좋은 결과를 보인 weight balancing을 적용

■ Data augmentation

- 영상 데이터
 - Horizontal flip
- 음성 데이터
 - Adding Gaussian noise, Time stretch

03 데이터 전처리

▪ Paired scenario for

- 쌍으로 구성된 cough audio & X-ray image dataset은 현재 존재하지 않음
- Paired dataset을 가정하기 위하여 unpaired dataset에서 같은 라벨의 cough audio & X-ray image 로 구성된 paired dataset을 구축
- 음성 데이터의 수가 더 적기때문에 음성 데이터의 수에 맞춤

	Negative	Positive
Train set	1753	133
Validation set	439	33

03 제안 모델

■ 사용된 pre-trained 모델

■ Image encoder



- ResNet50
- Pretrained using ImageNet dataset

■ Audio encoder



- ResNet22
- Pretrained using large-scale AudioSet dataset

■ 개인용 자가진단 모델

- [Reference paper] Motiian, Saeid, et al. "Unified deep supervised domain adaptation and generalization." Proceedings of the IEEE International Conference on Computer Vision. 2017.
- 같은 라벨을 가지는 서로 다른 도메인의 데이터를 이용해 분류 수행
- Domain adaptation을 수행하기 위한 contrastive semantic alignment loss 개념을 소개

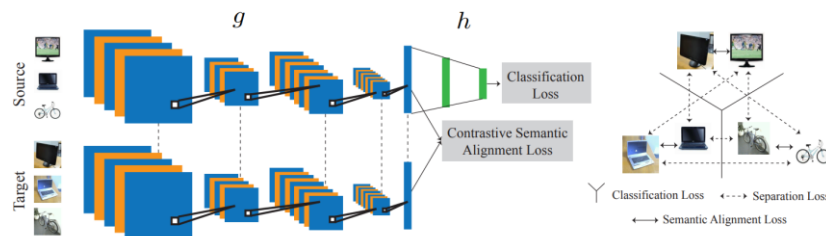


Figure 1. **Deep supervised domain adaptation.** In training, the semantic alignment loss minimizes the distance between samples from different domains but the same class label and the separation loss maximizes the distance between samples from different domains and class labels. At the same time, the classification loss guarantees high classification accuracy.

■ 개인용 자가진단 모델

- The semantic alignment loss: 같은 라벨, 다른 도메인의 샘플들 간의 거리를 최소화
- The separation loss: 다른 라벨, 다른 도메인의 샘플들 간의 거리를 최대화

- 영상 데이터를 이용해 음성 데이터 기반의 모델의 낮은 성능을 보완하려는 목적

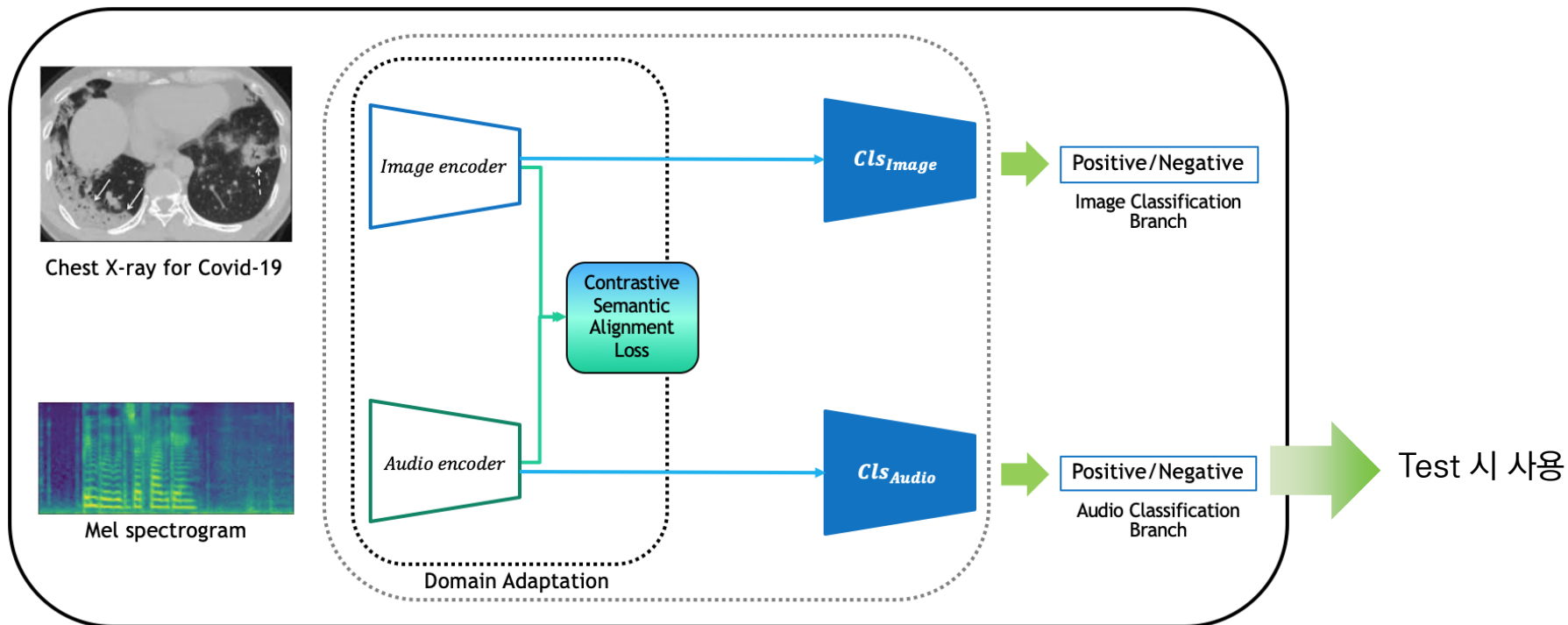


Image domain = source domain

Audio domain = target domain

- Domain-adapted feature가 분류기에 입력으로 주어져 결과 도출

03 제안 모델



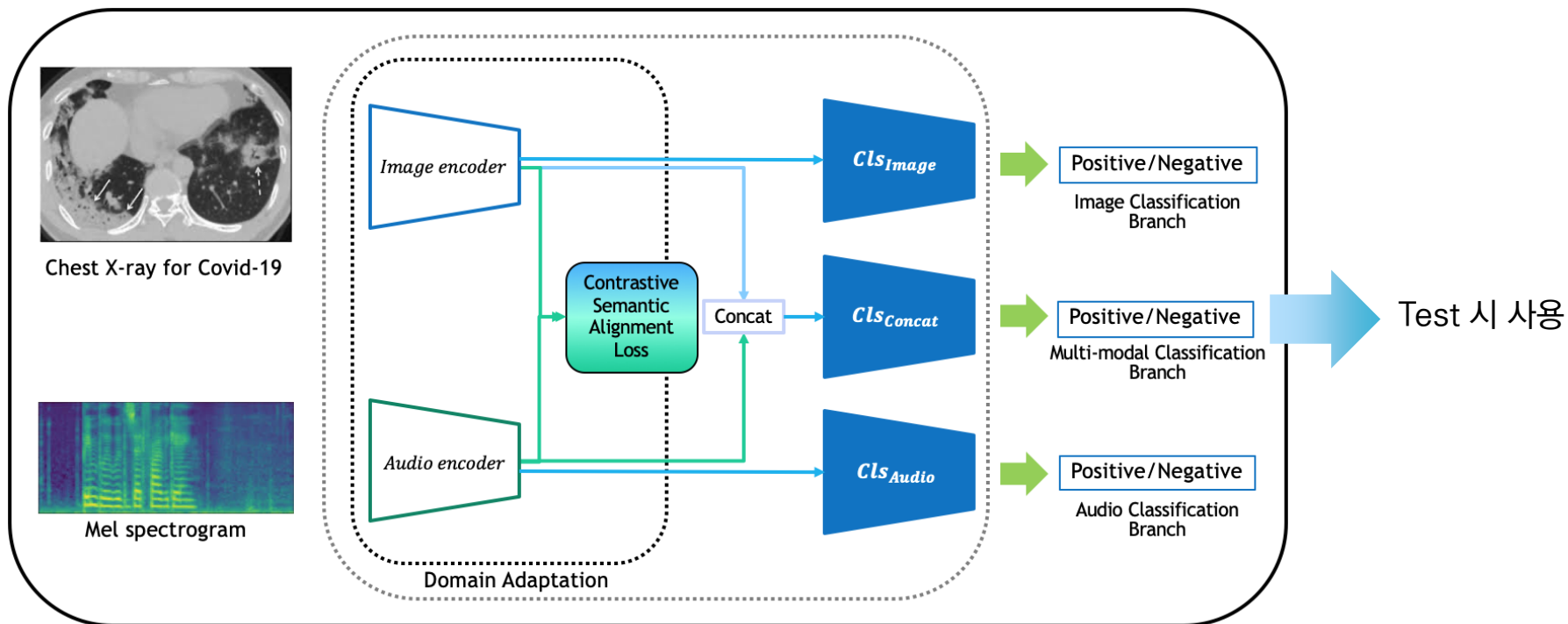
개인용 자가진단 모델의 구조

■ 병원용 진단보조 모델

➤ 기침 데이터를 이용한 성능 향상과 병변 정보의 시각화가 주요 강점

- 협력 병원과 paired dataset을 수집한 이후 모델 학습 및 서비스 배포 계획
- Paired scenario로 성능 확인한 결과, multi-modal feature를 이용하여 성능 향상되는 것 확인
- 현재는 팬데믹 특성 상 labelled data 수집 인력의 부족으로 feature activation map을 이용해 병변 정보 제공
→ 추후 segmentation task 추가하여 더 정확한 병변 정보 제공 예정

03 제안 모델



병원용 진단보조 모델의 구조

03 학습 결과

■ 개인용 자가진단 모델의 성능

Baseline (MEL spectrogram using audio data only)

Validation		Predicted	
		Positive	Negative
Actual	Positive	12	19
	Negative	21	420

- Accuracy: 91.5
- Sensitivity: 36.4
- Specificity: 95.7



Validation		Predicted	
		Positive	Negative
Actual	Positive	23	17
	Negative	10	422

- Accuracy: 94.5
- Sensitivity: 69.7
- Specificity: 96.1

03 학습 결과



■ 병원용 진단보조 모델의 성능

1. 적은 수의 영상 데이터

Validation with Image input		Predicted	
		Positive	Negative
Actual	Positive	27	13
	Negative	6	426

- Accuracy: 96.0
- Sensitivity: 81.8
- Specificity: 97.0

2. Multi-modal 데이터

Validation with multi-modal input		Predicted	
		Positive	Negative
Actual	Positive	29	14
	Negative	4	425

- Accuracy: 96.2
- Sensitivity: 87.9
- Specificity: 96.8

3. DA (Domain adaptation)

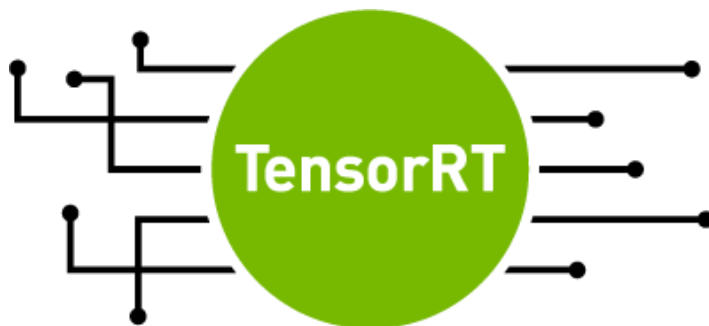
Validation with multi-modal input		Predicted	
		Positive	Negative
Actual	Positive	30	13
	Negative	3	426

- Accuracy: 96.6
- Sensitivity: 90.1
- Specificity: 97.0

03 추론 결과

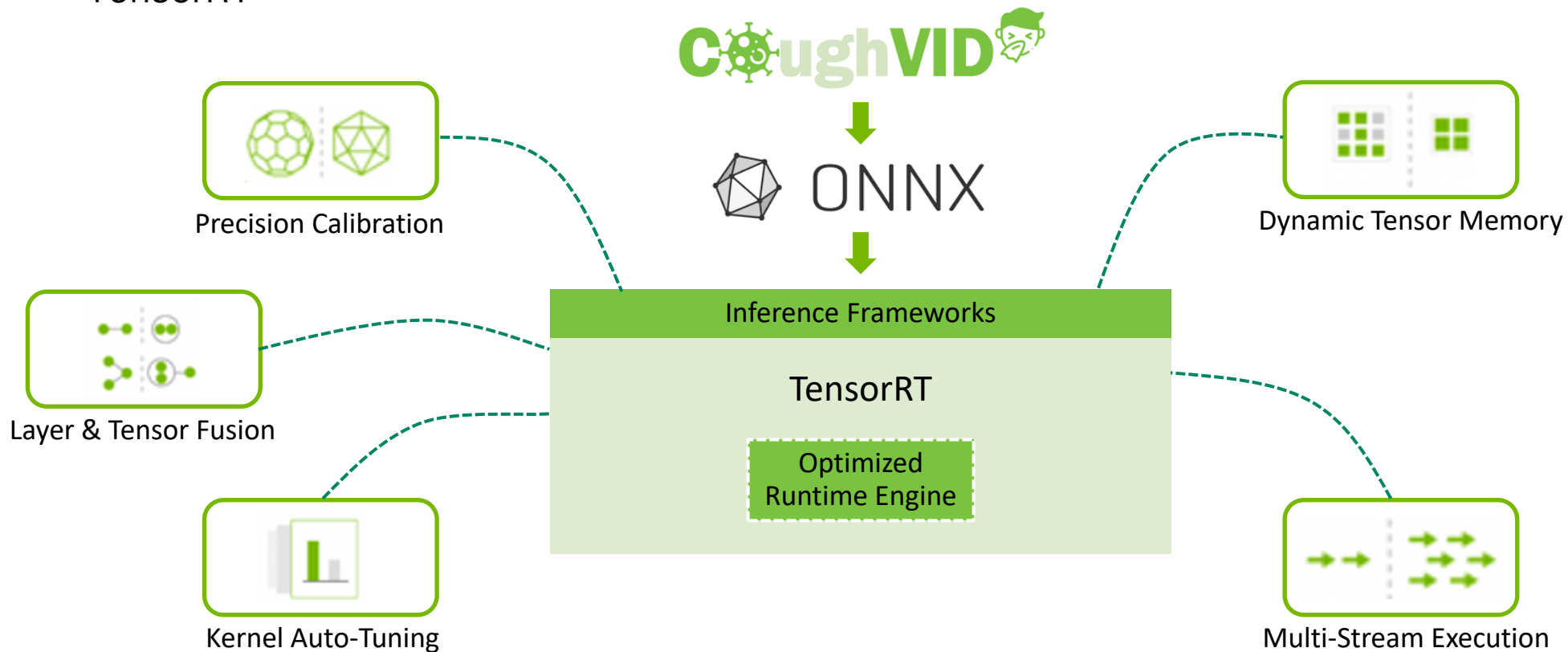
- TensorRT

- 학습된 DL 모델을 최적화하여 NVIDIA GPU 상에서의 inference 속도를 수배 ~ 수십 배 까지 향상시켜 DL 서비스의 TCO (Total Cost of Ownership)를 개선하는데 도움을 줄 수 있는 모델 최적화 엔진



03 추론 결과

- TensorRT



03 추론 결과

- TensorRT 사용으로 인한 효과
 - Inference 속도 및 성능 비교

	Throughput	Latency (ms)	Accuracy
Pytorch	382.68	105.92	94.5
TensorRT	12.66	21.05	94.45

03 추론 결과

- TensorRT 사용으로 인한 효과

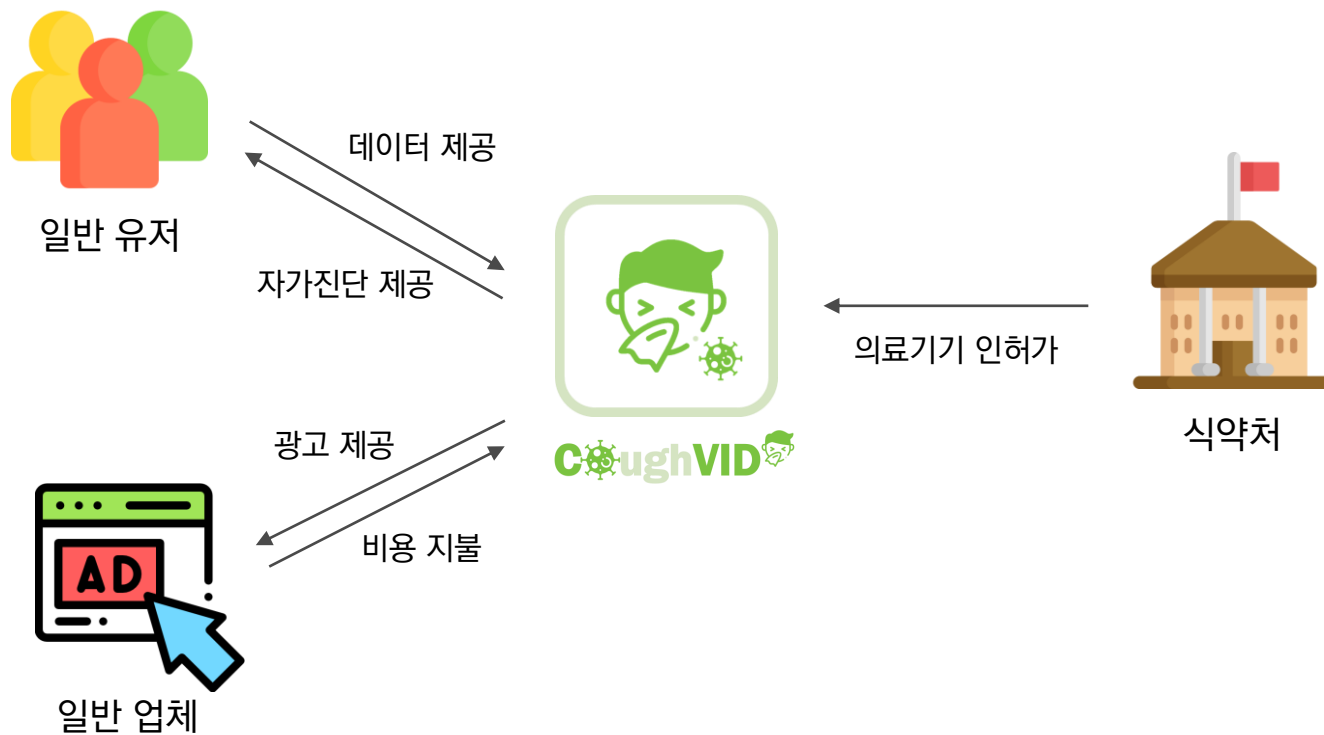
- Inference 속도 및 성능 비교

	Throughput	Latency (ms)	Accuracy
Pytorch	382.68	105.92	94.5
TensorRT	12.66	21.05	94.45

- 방사선 전문의의 성능

- 300건의 영상을 판독하는 데 약 2시간 30분에서 3시간 30분 소요
- 76~81%의 정확도

04 사업화 계획



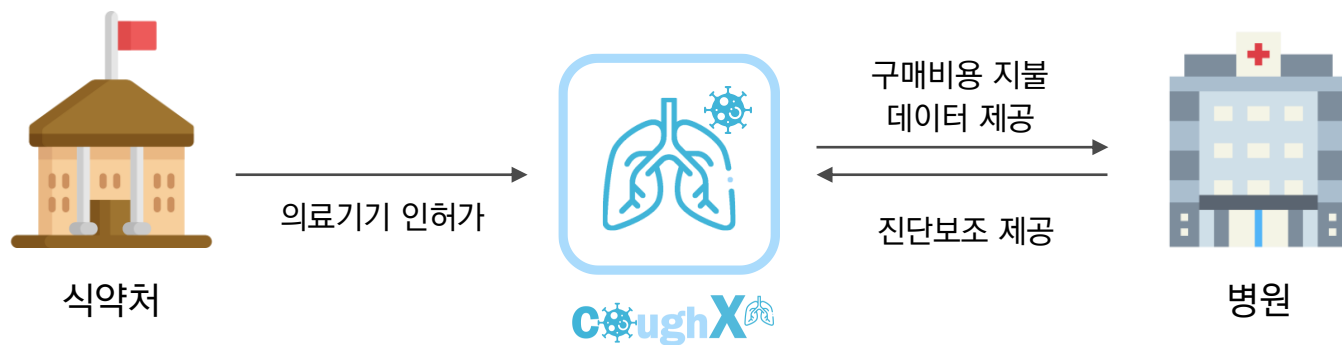
04 사업화 계획



Lean Canvas

문제 <ul style="list-style-type: none"> COVID-19 증상의 모호성 개인의 불안감 	솔루션 <ul style="list-style-type: none"> 기침을 이용한 자가진단 핵심 지표 <ul style="list-style-type: none"> 모바일 앱 다운로드 수 검출 수행 횟수 데이터 수집량 	가치 제안 <ul style="list-style-type: none"> 개인의 불안감 해소 높은 접근성 	경쟁 우위 <ul style="list-style-type: none"> 비용/시간 효율성 채널 <ul style="list-style-type: none"> 모바일 어플리케이션 	고객군 <ul style="list-style-type: none"> COVID-19 감염을 의심하고 스마트폰을 사용하는 누구나
비용 구조 <ul style="list-style-type: none"> 서버 및 데이터 저장장치 비용 앱 개발 및 유지 비용 온/오프라인 마케팅 비용 			수입원 <ul style="list-style-type: none"> 어플리케이션 광고 (COVID-19 팬데믹 관련 제품/서비스/앱) 정부지원금 	

04 사업화 계획



04 사업화 계획



Lean Canvas

문제 <ul style="list-style-type: none">COVID-19의 빠른 전파진단 장비 부족의료 인력의 부족시간 효율적인 진단보조의 필요성	솔루션 <ul style="list-style-type: none">기침/의료영상 데이터를 이용한 딥러닝 기반 진단보조 시스템	가치 제안 <ul style="list-style-type: none">진단 소요시간 감소더 정확한 진단이 가능정보 제공으로 의료진의 진단을 수월하게 함	경쟁 우위 <ul style="list-style-type: none">더 많은 정보로 성능 향상병변 정보 제공	고객군 <ul style="list-style-type: none">진단의 효율성을 높이길 원하는 의사인력 부족을 체감하는 병원
	핵심 지표 <ul style="list-style-type: none">검출 수행 횟수구매 병원 수데이터 수집량		채널 <ul style="list-style-type: none">웹사이트	
비용 구조 <ul style="list-style-type: none">서버 및 데이터 저장장치 비용웹 개발 및 유지 비용온/오프라인 마케팅 비용			수입원 <ul style="list-style-type: none">병원 대상 어플 판매 비용정부지원금	

04 팀 구성 및 역량

■ 팀 내부 역량



- 전자전기/인공지능 학과 석사 (예정)
- 의료영상처리 및 딥러닝 전공
- 관련 기술/동향에 대한 지식
- 진단 모델 자체 개발 가능

■ 외부 역량



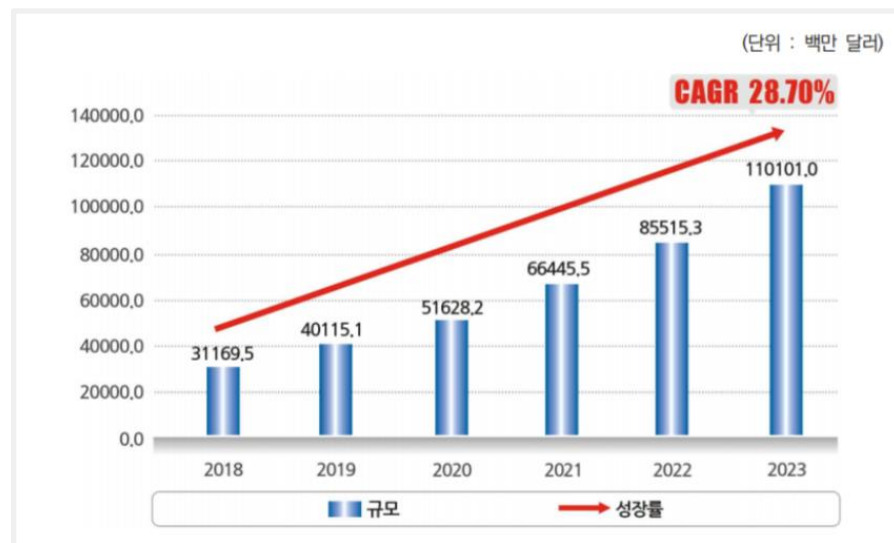
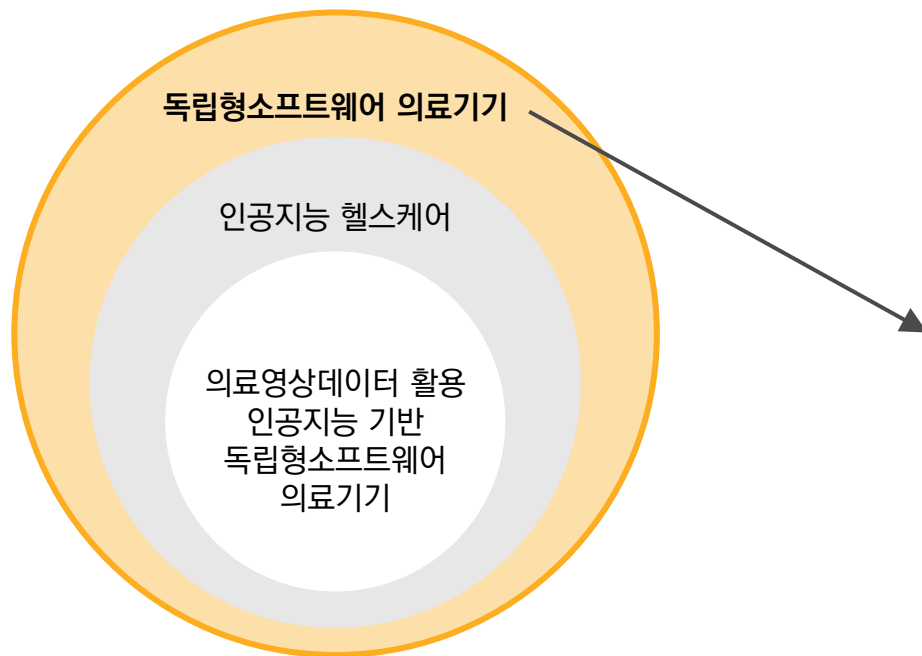
- 삼성병원, 성빈센트병원, 과천길병원 영상의학과 의료진
- 활발한 협력 연구 진행중
- 데이터 수집 협업
- 진단모델 개발 시 자문 제공
- 병원 유치 시 연결



- 킹고스프링 소속 투자전문가
- 성균DX협동조합 소속 경영전문가 및 사업가
- 사업전략/재무/마케팅 분야 자문
- 개인투자자, 기업 및 VC들과 연계
- 기업의 해외진출 지원 및 마케팅 대행업무

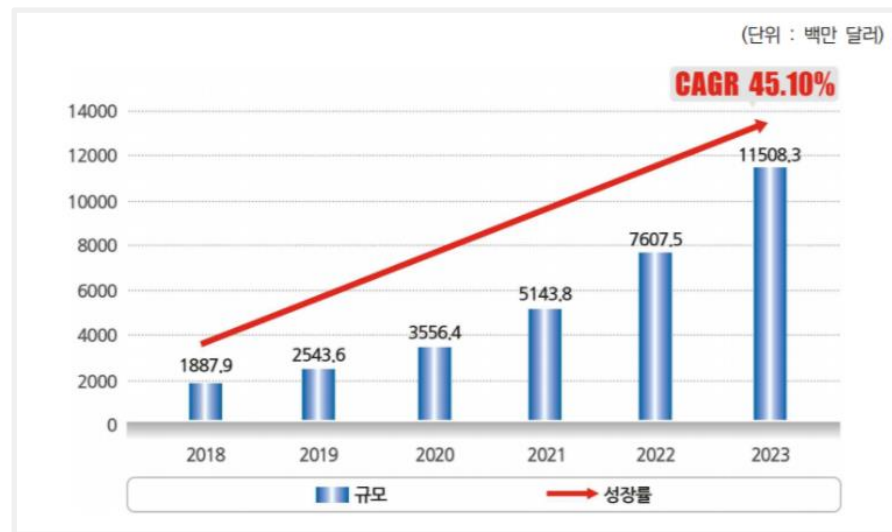
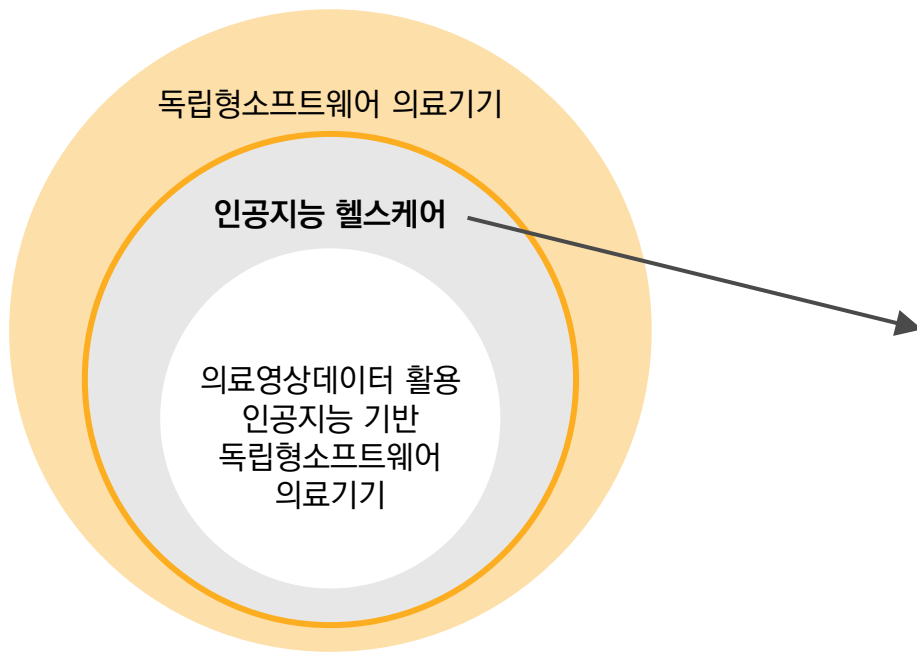
05 시장분석

■ Tam-Sam-Som



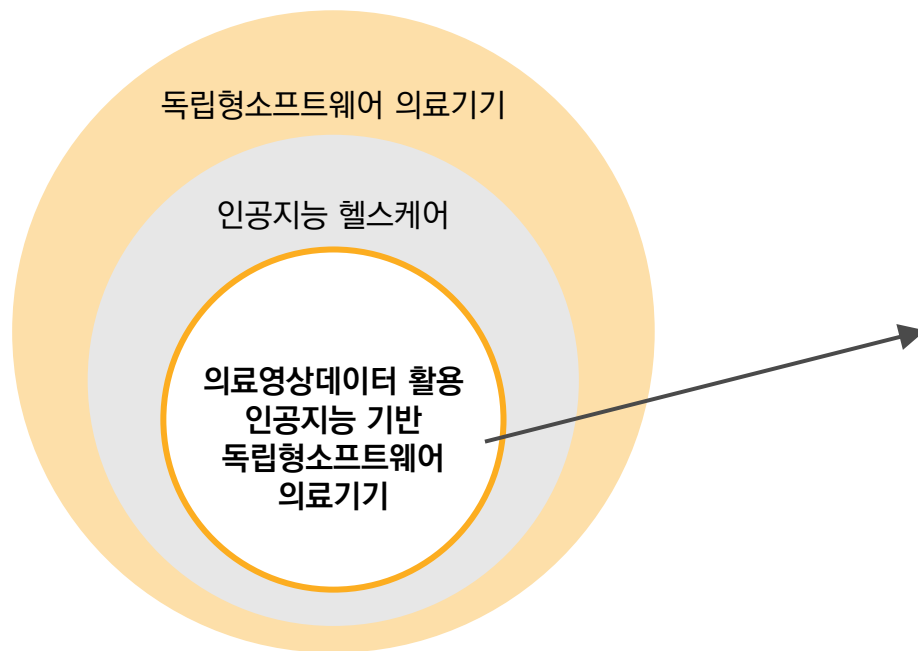
05 시장분석

■ Tam-Sam-Som

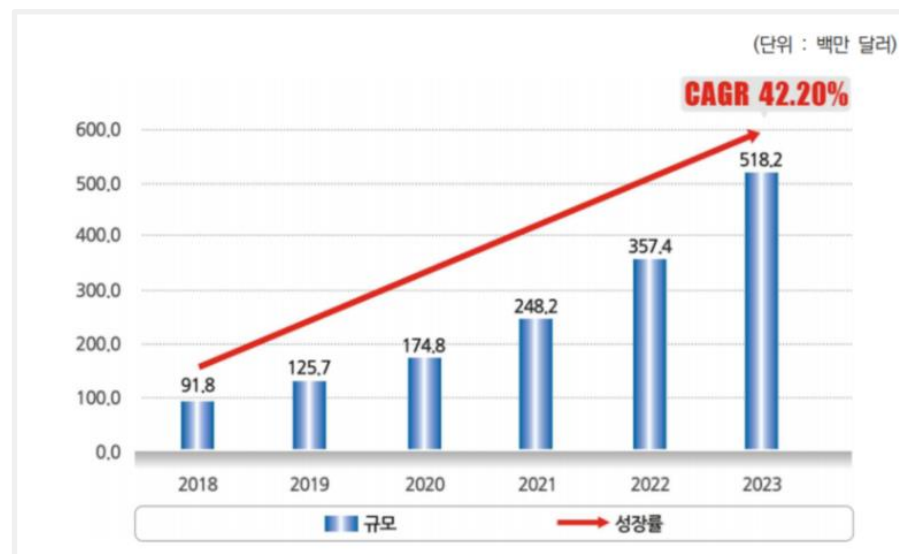


05 시장분석

■ Tam-Sam-Som



“1차 년도에 의료영상 기반 Covid-19 진단시장 10% 점유 목표”



05 시장분석



COVID-19 자가진단 시장

- 현재로서는 개인이 수행할 수 있는 COVID-19 자가진단 앱이 존재하지 않음
- 자가진단 키트와 비교해 별도의 용품이 필요 없다는 장점



딥러닝 기반 COVID-19 진단보조 시장

[AI 기반 진단 사례 및 연구]

해 외	중국 알리바바	- 5,000명의 코로나 바이러스 환자로부터 얻은 데이터를 기반으로, 흉부 CT 촬영에서 코로나 바이러스를 검출할 수 있는 AI 시스템을 개발 - 20초 안에 판독 수행(인간 전문가는 15분 소요), 96%의 정확도 주장
	이스라엘 카훈(Kahun)	- 이스라엘 기반 의료기술 스타트업인 Kahun은 코로나19 진단 도구인 Kahun knowledge graph를 출시하여, 의료진의 신속한 결정을 지원 - 이 무료 AI 엔진은 의학도서관인 PubMed에서 제공한 2,000개 이상의 논문 등을 기반으로 하며, 환자가 위급 단계로 넘어갈 위험이 높은지 여부를 결정할 수 있게 해줌
	미국 매사추세츠 대학	- 매사추세츠 대학(UMass Amherst) 연구진은 기침 소리를 분석하여 바이러스성 호흡기 질환의 잠재적 확산 가능성이 있는지를 평가하는 AI 장치인 플루센스(FluSense)를 개발 중
	네덜란드 델프트 대학	- 네덜란드의 델프트 대학(University of Delft) 연구진은 흉부 X-레이 영상으로부터 코로나19를 진단하는 AI 모델(CAD4COVID) 발표
국 내	씨젠	- 코로나19 진단 키트 개발 과정에 AI를 활용하여 개발 기간을 2주 내외로 단축
	뷰노 JLK인스펙션	- 뷰노와 JLK인스펙션은 AI를 활용하여 수초 안에 X-레이를 판독함으로써 확진자의 폐 질환 여부 확인 기술 개발
	루닛	- 흉부 X-레이 AI 영상 분석 SW(루닛 인사이트 CXR)가 국내외 코로나19 진단에 활용

05 시장진입 전략



■ 국내/외 시장 진입 전략

- 국내/외 관련기관 허가
- COVID-19 관련 콘텐츠를 업로드하는 유튜버 / SNS 등 온라인 마케팅 진행
- 발빠른 초기 진입으로 사용자 확보가 목표

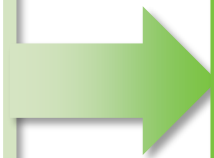


■ 국내 시장 진입 전략

- 식약처로부터 의료영상진단보조소프트웨어 승인
- 보유한 의료진 커넥션 기반으로 초기 진입
- 병원/의료진 대상 마케팅
- 기침데이터의 활용성 및 병변 정보 제공을 무기로 경쟁우위 획득

■ COVID-19 자가진단 시장

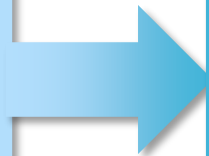
- 현재 기침을 이용한 자가진단 어플이 없음
- 데이터를 수집할수록 성능 개선
- 자가진단 서비스는 일회성에 그침



- 시장에 초기 진입해 마켓 점유
- 빠른 이용자 확보로 성능을 개선해 후발주자에 대응
- COVID-19 증상, 선별진료소, 현황 정보 제공 및 관련 제품 입점
- 앱 체류시간 및 방문 횟수 늘림

■ 딥러닝 기반 COVID-19 진단보조 시장

- 많은 기업들이 이미 서비스를 제공 중
 - ✓ 대부분이 의료영상 (X-Ray / CT)만을 사용



- 기침 데이터를 활용해 성능 개선
- 병변 분석 정보 제공

06 자금계획

■ 자금조달 계획

➤ 교내/외 창업지원금



기관명	주요 지원업무	금액	기타 지원
킹고스프링	엑셀러레이터, SBA펀드(개인투자조합)	5천만원	엔젤투자 연계
LINC+ 사업단	스마트 및 바이오 신산업 지원, 창업기업 성장 지원	2천만원	
창업지원단	예비창업자 지원 및 교육	3천만원	사무실 제공
경기도경제과학진흥원	WINGS (민간투자연계형 기술창업지원사업)	2억원	

➤ 2021년까지 엔젤투자로 1.6억원 유치 예정 (모교 창업지원 네트워크 적극 활용)

06 자금계획

■ 자금조달 계획

➤ 창업지원 트랙

연도	기관	지원사업명	세부과제	기간	금액
2021	창업진흥원	예비창업패키지	4차산업 분야	9개월	1억원
2022	창업진흥원	초기창업패키지	특화(연구실) 기술 트랙	7개월	1.5억원
2022	중소벤처기업부	창업성장기술개발 디딤돌 창업과제	첫걸음/여성참여과제	1년	1.5억원
2024	중소벤처기업부	창업성장기술개발 전략형 창업과제	4IR (인공지능)	2년	4억원
2027	중소벤처기업부	창업성장기술개발 TIPS 과제	4IR (인공지능)	2년	5억원

06 자금계획

■ 자금조달 계획

➤ 4IR (바이오헬스, 인공지능) / R&D 창업 관련 정부지원금

기관명	지원사업명	분야	금액
중소벤처기업부	비대면 청년창업 아이디어 발굴 지원	스마트헬스케어 부문	7천만원
	혁신분야 창업패키지(3대 신산업분야)	바이오헬스 (의료기기)	1.3억원
한국보건산업진흥원	인공지능 활용 체외진단의료기기 개발 지원사업	인공지능 분석 솔루션 도입	5백만원
	바이오헬스 투자인프라 연계형 R&D	의료기기	2천만원
	기술혁신형 창업기업 지원사업	스마트의료기기	5천만원
보건복지부	의료기술상용화지원센터 신규지원 대상과제	바이오헬스 산업 혁신	2천만원

06 자금계획

■ 1차년도 사업비 계상

구분		내용	금액
직접비	내부인건비	운영진 및 개발자	1.2억원
	연구장비·재료비	데이터 저장장치 및 서버 구매 비용	2천만원
		데이터 구매 비용	1천만원
	연구활동비	전문가 활용비 (자문, 웹/앱 개발 및 홍보)	3천만원
		기술정보수집비 (임상시험비)	1천만원
간접비	성과활용지원비	지식재산권 출원·등록비	1천만원
		기술개발 결산 및 사업화 성과 평가, 회계감사비용	2천만원

07 발전 방향

■ 성능 향상 도모

- 지속적인 데이터 수집
- 수집 데이터를 기반으로 성능 향상
- 지속적인 모델 개발 연구 및 피드백

■ 사업 확장성

- 진단 범위를 다양한 호흡기 질환으로 확대
- 호흡기 질환 진단을 넘어 케어 솔루션 제공
- 가족의 호흡기 질환 진단까지 확장

07 발전 방향

RESEARCH

Open Access

A prospective multicentre study testing the diagnostic accuracy of an automated cough sound centred analytic system for the identification of **common respiratory disorders in children**

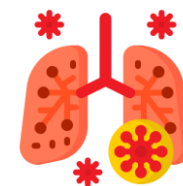


Articles

Stratifying **asthma** severity in children using cough sound analytic technology

Vinayak Swarnkar , BE, PhD, Udantha Abeyratne , BScEE, MEng, GradCertPaedSleepSc, GradCertHigherEducation, PhD, Jamie Tan , MBBS, GradDipNeo, FRACP, Ti Wan Ng , MBBS (Hons), DCH, FRACP, Joanna M. Brisbane , BSc(Hon), GradDipGenetCouns, MPH, Jennifer Choveaux , BSc(Nursing), Grad Cert Midwifery, Grad Cert Intensive Nursing Science & ...show all
Received 09 Jul 2019, Accepted 20 Oct 2019, Accepted author version posted online: 22 Oct 2019, Published online: 25 Nov 2019

■ 사업 확장성



진단 범위를 다양한 호흡기 질환으로 확대

- 호흡기 질환 진단을 넘어 케어 솔루션 제공
- 가족의 호흡기 질환 진단까지 확장

07 발전 방향

Physiological Measurement



PAPER

Predicting **spirometry readings** using cough sound features and regression

- ✓ 폐기능 기록 및 관리
- ✓ 증상 기록 및 관리
- ✓ 복용 정보 기록 및 관리
- ✓ 기록된 건강상태에 따른 관리법 제공

■ 사업 확장성



- 진단 범위를 다양한 호흡기 질환으로 확대
- **호흡기 질환 진단을 넘어 케어 솔루션 제공**
- 가족의 호흡기 질환 진단까지 확장

07 발전 방향

Computers and Electronics in Agriculture 129 (2016) 15–26



ELSEVIER

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag



Original papers

Early recognition of **bovine** respiratory disease in calves using automated continuous monitoring of cough sounds



Hindawi
Journal of Sensors
Volume 2020, Article ID 2985478, 7 pages
<https://doi.org/10.1155/2020/2985478>



Research Article

A Novel Method for **Broiler** Abnormal Sound Detection Using WMFCC and HMM

■ 사업 확장성

- 진단 범위를 다양한 호흡기 질환으로 확대
- 호흡기 질환 진단을 넘어 케어 솔루션 제공

가축의 호흡기 질환 진단까지 확장



07 사업확장 계획

“No1. 호흡기 질환 진단 및 케어 서비스 업체로 도약”

