3차원 인체 스캔 데이터와 이중 에너지 엑스선 흡수기(DXA)데이터를 이용한 XGBoost기반 비만 판단 방법론 제안

# **Content**

# 1. 서론

- 연구의 필요성
- related work

# 2. 본론

- 데이터 수집
- 연구 프레임워크
- 데이터 전처리
- 분류구성 단계, 기본특징 학습단계
- 평가지표

# 3. 결론

- 기대효과, 한계점

# 1. 서론

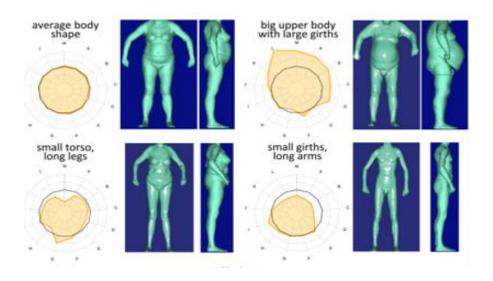
#### 연구의 필요성

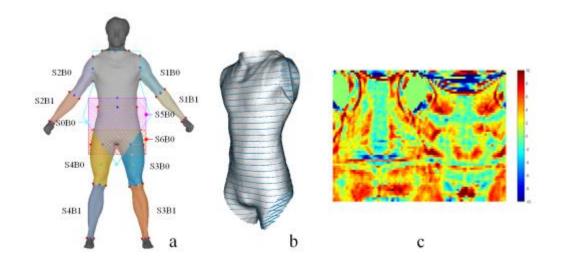
- 기존 비만 판단 지표 한계
- 체중에 기반을 둔 BMI(Body Mass Index)는 근육과 지방을 구분하지 않음
- 신체 부위별 정보를 제공하지 않음
- → 비만이나 건강 상태를 평가하는 지수로 정확한 판단 어려움
- 기존 데이터셋의 한계
- 기존의 3차원 인체 데이터는 각각에 대응되는 객관적인 건강 정보가 부족함
- 체형진단과 그에 따른 헬스케어를 위한 정확한 치수와 3차원 데이터를 얻을 수 없음
- 대부분의 데이터셋이 서양인이 피실험자임
- → 한국인에 맞는 데이터셋 구축이 필요함

한국인 3차원 인체 데이터셋과 이에 대응되는 건강 정보를 수집하고, 이에 대한 연구가 필요함

# 1. 서론

#### 관련 기존 연구들





3차원 스캐너로 인체 분류 혹은 비만도 판단 선행 연구

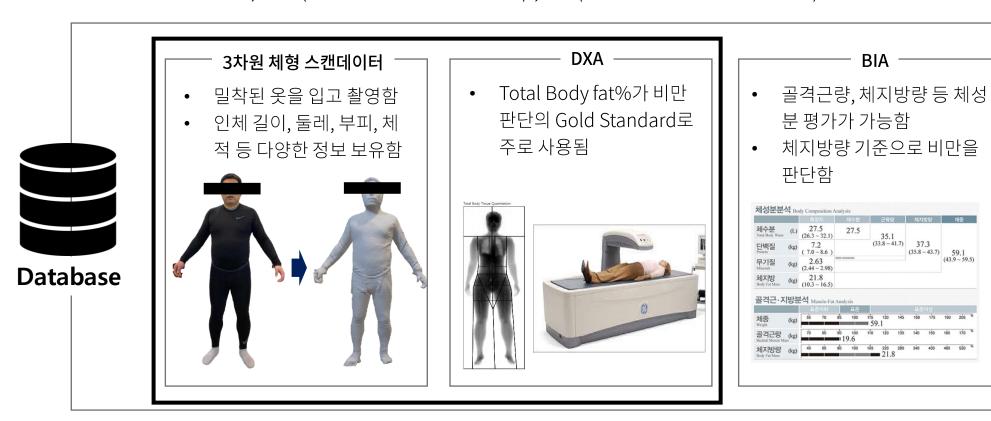
- Novel Anthropometry Based on 3DBodyscans Applied to a Large Population Based Cohort
- ⇒ 객관적인 의료데이터의 비교 판단 불가
- 3D Shape-based Body Composition Prediction Model Using Machine Learning
- ⇒ 서양인 데이터 이고, 데이터 수가 적음, 강건성에 한계가 있음

#### 데이터 수집

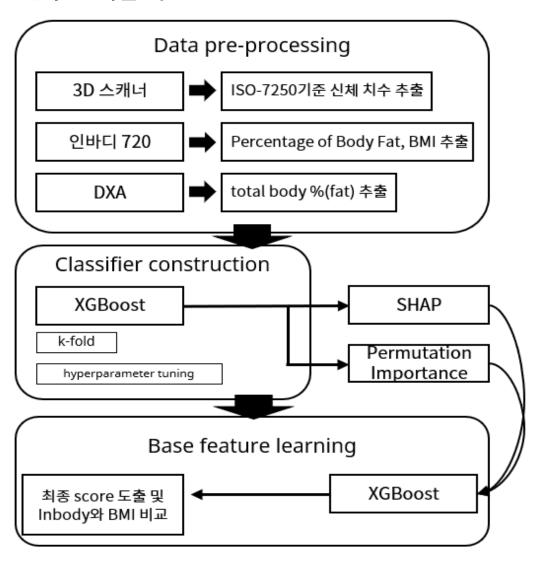
수집 기간: 2021-04-05 ~ 2021-04-17 (약 12일)

• 피실험자 : 62명(여성 32명, 남성 30명/20~30대)

• 수집 데이터: 3차원 체형 스캔데이터, DXA(이중에너지엑스선흡수기), BIA(생체 전기 임피던스 분석기)



#### 연구 프레임 워크



- 데이터 전처리

3차원 스캐너 mesh 데이터로 부터 ISO-7250 랜드마크 기준 치수 추출

BIA

인바디 720로 부터 인바디 체지방량 추출

DXA

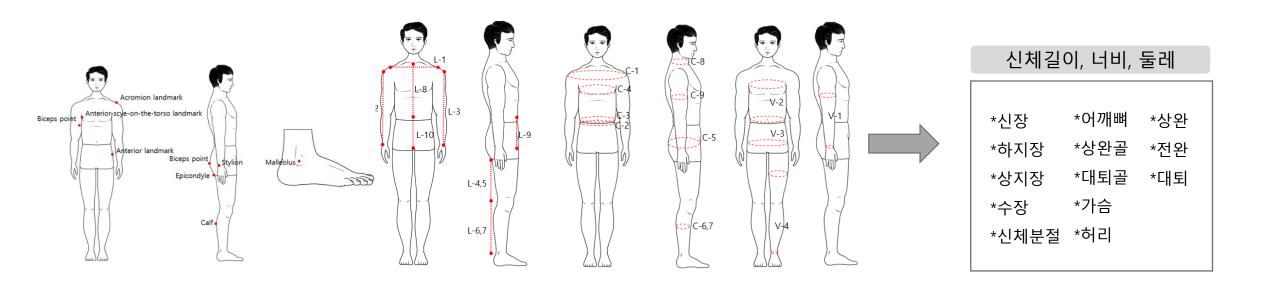
덱사 체지방량 추출

- 분류 구성 단계
- K-fold 적용하여 XGBoost 학습
- 기본 특징 학습 단계 SHAP와 Permutation Importance를 비교하여 변수 추출 후 다시 학습

### 데이터 전처리

- 3차원 체형 스캔데이터
- ISO-7250에서 정의한 신체 랜드마크를 기준으로 60가지 항목의 신체 치수 도출
- ✔ 신체 기본정보 7개(신장, 체중, WHrR, WCR, WHR, WHtR, THR), ✔ 단면적정보 13개(허리, 가슴, 엉덩이 등)
- ✔ 길이정보 13개(어깨, 가슴, 허리, 다리 등)
- ✓ 둘레정보 14개(배, 허리, 엉덩이 등)

- ✔ 체적정보 8개(어깨, 상복부, 하복부 등)
- ✔ 각도정보 5개(어깨, 견갑, 허리, 엉덩이 등)



### 데이터 전처리

#### DXA

- 방사선이 인체를 투과할 때의 투과물질의 방사선 투과율 차이로 측정함
- 체성분 분석의 Gold Standard로써 주로 사용됨
- 전신 부위별 골밀도, 면적, 지방, 제지방량 등의 데이터를 추출 가능함
- 그 중 Total body Fat(%)을 이용하여 체형을 정의하고 본 논문의 Label로 사용함

Total body Fat(%) (남성)	Total body Fat(%) (여성)	비만 판단
~ 15%	~ 18%	저체중
15% ~ 20%	18% ~ 28%	보통
20% ~ 25%	28% ~ 33%	과체중
25% ~	33% ~	비만

저체중	정상	과체중	비만
2	17	10	33

Facility ID:
Referring Physician: cspa
Measured: 2021-04-05 오후 12:50:32 (13.60)
Analyzed: 2021-04-05 오후 12:51:33 (13.60)

Bone Density Trend Reference Chart: No reference data for Pediatric Total Body [Total] region.

No patient or valid information available for display.

	Total Body Bo	ne Density Tre	nd .
•		1	9

Region	BMD (g/cm²)
Head	2.319
Arms	0.857
Legs	1.301
Trunk	0.963
Ribs	0.694
Pelvis	1.252
Spine	1.082
Total	1.217

COMMENTS

Image not for diagnosis

Printed: 2021-04-05 ♀≢ 12:51:11 (13:60)76:0.15:153.04:31.4 0.00:-1.00

4.81x13.01 13.0%Fat=28.2% 0.00:0.00 0.00:0.00

Filename: 권우\_aqn2rq6ia8.dfb Scan Mode: Standard 0.4 µGy 1 - Statistically 68% of repeat scans fall within 1SD (± 0.010 g/cm² for Total Body Total)

#### 데이터 전처리

- BIA(생체 전기 임피던스 분석)
- 신체의 체수분량을 측정하여 인체에 미약한 전류를 흐르게 한 뒤, 생체 전기 저항을 측정하여 체성분 분석에 이용함
- 체수분, 단백질, 무기질, 골격근량, 체지방량 등의 신체 체성분 정보를 얻을 수 있음
- 체지방률과 BMI를 이용한 기존 비만판단 지표와 비교

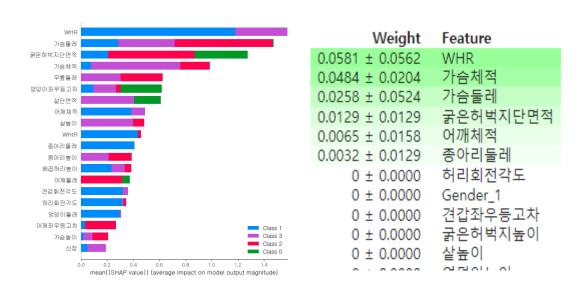
체지방율 (남성)	체지방율 (여성)	ВМІ	비만 판단
~ 15%	~ 18%	~ 18.5	저체중
15% ~ 20%	18% ~ 28%	18.5 ~ 23	보통
20% ~ 25%	28% ~ 33%	23 ~ 25	과체중
25% ~	33% ~	25 ~	비만



#### 분류구성 단계

- 변수와 정답 값 설정
- ISO 7240 landmark기준으로 추출한 60개의 치수를 변수로 설정
- DXA 기준으로 정답 값 설정
- 클래스의 불균형 & 데이터 수 적음 => K-fold 교차검증 적용
- 10번의 교차검증 실시 평균값을 정확도로 사용하였음
- SHAP와 Permutation importance를 통해 중요 변수 도출
- 변수가 목적변수에 미치는 영향을 기준으로 선택
- SHAP와 Permutation importance를 비교 분석하여 기본특징 학습 단계에서 사용할 변수 설정





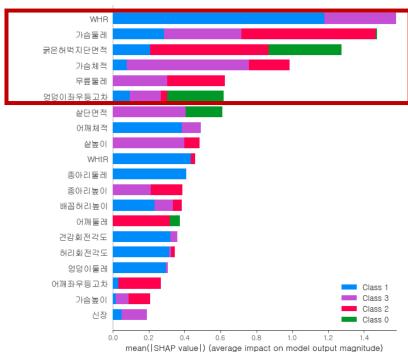
<Class 0, 1, 2, 3은 저체중, 정상, 과체중, 비만을 뜻함>

#### 기본 특징 학습 단계

- SHAP
- 변수가 목적변수에 미치는 영향을 기준으로 선택
- Permutation Importance
- 특정 Feature를 쓰지 않았을 때 성능 손실에 미치는 영향 파악

→ WHR, 가슴체적, 가슴둘레, 굵은 허벅지 단면적, 어깨 체적, 종아리 둘레를 기본특징 변수로 설정

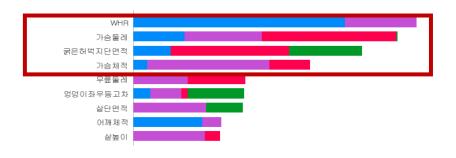




#### 기본 특징 학습 단계

- 변수 추출 후 다시 학습시킨 XGBoost 모델
- 17%포인트 증가
- 학습시킨 XGBoost 모델과 BIA, BMI 정확도 비교
- BMI보다 19%포인트 증가
- BIA보다 11%포인트 증가

	정확도
신체치수 XGBoost 모델	68%
기본특징 XGBoost 모델	81%
BIA (생체 전기 임피던스 분석)	70%
BMI (체질량지수)	62%



- → BMI와 BIA보다 3차원 인체 데이터를 사용 한 머신러닝 방법론이 비만 판단을 내릴 때 적합 할 수 있는 가능성을 보여줌
- → 비만의 정도에 따라 하나 혹은 두개의 신체 치수에 의존하는 BMI나 WHR같은 지표보다 여러 신체 치수를 반영하는 것이 비만을 판단하는 것에 유리함

# 3. 결론

#### 기대 효과

- 3차원 인체 데이터와 DXA 데이터를 쌍으로 수집하여 헬스케어 데이터로서 객관적 지표가 될 수 있음
- 머신러닝 기법을 통해 데이터를 기반으로 한 새로운 비만 판단 Index 제시
- 일반적으로 수집하기 어려운 인체 데이터를 수집 및 분석함으로써 보다 실용적인 연구 방향성 제시
- 비만을 판단하는 Feature를 도출함으로써 신체의 부위별 비만 정도 및 비만의 세분화 분석 가능

#### 한계점

- 신체를 측정하는데 필요한 3차원 스캐너와 DXA는 개인이 사용하기 어려움
- 적은 데이터로 인해 균등하게 클래스가 분포되지 않음
- DXA에서 얻을 수 있는 데이터들을 부분적으로 밖에 사용하지 않음
- Landmark를 통해 추출한 치수로 실험을 한 것이므로 3차원 데이터를 온전히 사용했다고 하기 어려움

# 감사합니다

Q&A