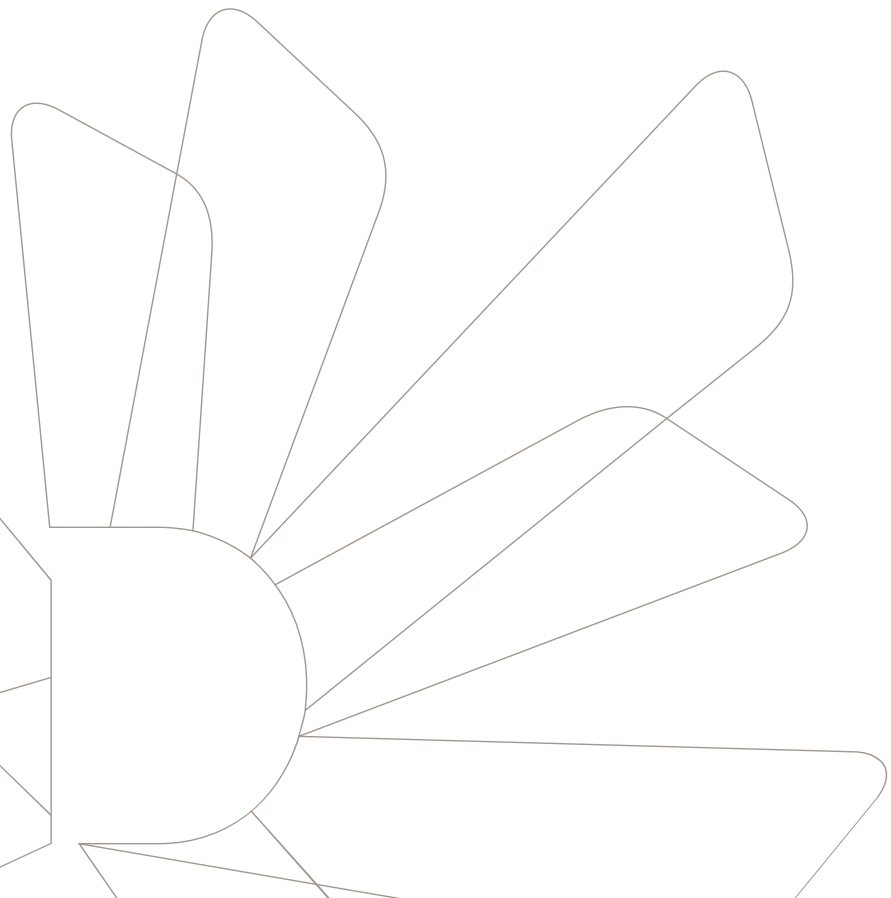


# 한 장의 이미지에서 2.5차원 표현 추출을 통한 BMI 추론 방법론

2022.06.02

Industrial and Systems Engineering  
Seungjin Jeon  
M.S Student

Supervisor *prof.* Sekyoung Youm  
Co-work Minji Kim



# Index

## 1. Introduction

- The Importance of BMI
- The Importance of Visual based BMI Estimation
- Related Works
- Main contribution

## 2. Method

- Main Idea
- Overall Architecture
- 2.5D Representation Estimator
- BMI Predictor

## 3. Experiments

- 2D Image to BMI Dataset
- Data preprocessing
- Implementation details

## 4. Results

- Comparison

## 5. Conclusion

- Limitation
- Expected Effect

# Introduction


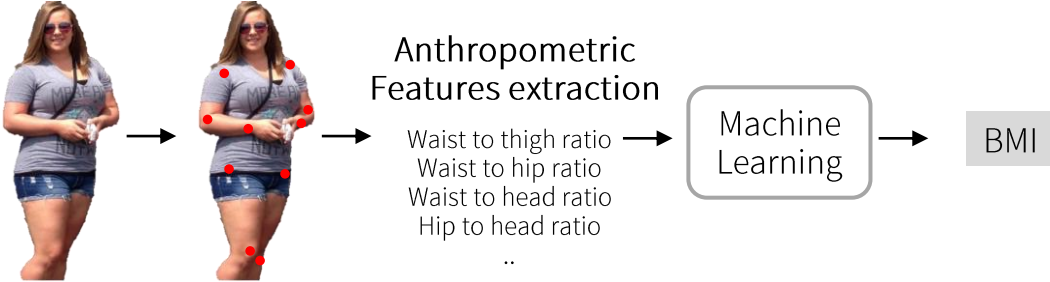
---

- The Importance of BMI
  - 키와 몸무게만 이용하여 간편한 건강의 지표로 WHO(세계보건기구)에서 비만판단의 지표로써 사용함
  - 메디컬 분야에서는 BMI와 질병 사이의 연관성을 현재까지도 계속해서 연구 중임
    - 대표적으로 당뇨병, 고혈압은 BMI와 높은 상관관계가 있음이 밝혀짐
  - 국가차원에서 인구의 비만수준에 대한 평가는 질병 유병율 및 발병율을 모니터링 하는 데에 중요함
    - 현재 대규모 인구조사에서 많은 사람의 BMI를 얻기 위해 직접 측정하는 것은 시간 및 공간적인 한계가 존재함
- The Importance of Visual - based BMI Estimation
  - 체중을 재기 위한 보편적인 방법은 체중계를 이용하는 것임
    - 일반적으로 바르게 선 자세로 측정하며, 다른 자세에서 체중 측정은 불가능함
  - 최근 원격의료에 대한 관심이 증가함에 따라, BMI 정보는 진료의 중요한 요소로 사용 가능함
  - 체중계 등의 추가적인 기기 없이 간편하게 BMI 예측이 가능하며, 시간 및 공간적인 한계 극복 가능함

헬스케어 및 메디컬 분야의 BMI 중요성에 따라, 쉽게 적용가능한 BMI 측정 기술이 필요함

# Introduction

- Related Works – Visual based BMI Estimation

RGB-D Based Method	RGB Based Method
<p><b>Pros</b></p> <ul style="list-style-type: none"><li>- RGB Based Method보다 BMI 예측 성능이 좋음</li><li>- 부피에 대한 정보를 함유 할 수 있음</li></ul> <p><b>Cons</b></p> <ul style="list-style-type: none"><li>- Kinect와 같은 RGB-D를 수집하기 위한 추가적인 센서가 필요함</li><li>- RGB-D에서 Depth에 주로 의존적인 경향을 보이는데, RGB-D 센서 특성상 광원에 민감하여 야외에서 예측하지 못함</li><li>- 카메라와 사람 사이의 거리에 민감함</li></ul> <div><pre>graph LR; Kinect[Kinect] --&gt; Image[Image]; Image --&gt; Silhouette[Depth Map / Silhouette]; Silhouette --&gt; FE[Features extraction]; FE --&gt; ML[Machine Learning]; ML --&gt; BMI[BMI]</pre></div>	<p><b>Pros</b></p> <ul style="list-style-type: none"><li>- 추가적인 장비 필요없이 쉽고 간편하게 BMI 예측이 가능함</li><li>- BMI를 측정하는 것에 있어서 공간에 제약 받지 않음</li></ul> <p><b>Cons</b></p> <ul style="list-style-type: none"><li>- 이미지에 대한 정보가 부족하여 정확한 BMI 추정에 한계 존재</li><li>- 신체정보(Anthropometric)를 Feature로써 사용하여 Handcraft 방식의 연구가 주로 진행됨</li></ul> <div><pre>graph LR; Image[Image] --&gt; AF[Anthropometric Features extraction]; AF --&gt; ML[Machine Learning]; ML --&gt; BMI[BMI]; AF --&gt; Ratios["Waist to thigh ratio&lt;br/&gt;Waist to hip ratio&lt;br/&gt;Waist to head ratio&lt;br/&gt;Hip to head ratio&lt;br/&gt;.."]</pre></div>

# Introduction

## BMI 예측의 중요성

- 여러 질병과 BMI 사이의 상관관계가 높음
- 국가의 질병 유병율 및 발병율 모니터링의 중요한 요소임
- 헬스케어, 메디컬 분야의 중요한 지표로써 사용됨

## 현 BMI 예측 방법의 문제점

- 체중계 : 추가적인 장비, 시간과 공간의 제약
- RGB Based : 정보의 부족으로 좋은 성능을 기대하기 힘들
- RGB-D Based : 추가적인 장비가 필요함

## Contribution Summary

1. 2.5차원(Normal, Depth map)을 이용하여 3차원의 인체 특징을 반영한 BMI Estimator 개발
  - 기존 RGB 방식보다 정확한 BMI 예측이 가능함
  - 추가적인 장비 없이 BMI 예측이 가능함
2. End-to-end 방식으로 한 장의 신체 이미지를 이용하여 BMI 추정이 가능함
  - *Not handcraft* and End-to-end 네트워크 구조를 가짐
3. BMI를 구함에 있어서 의미론적인 접근을 하여 문제를 해결하였음
  - BMI의 내포 되어있는 부피의 정보를 2.5D 정보를 통해 네트워크를 학습시킴으로서 더 나은 성능을 갖게 됨

# Method

## Main Idea

사람의 체중은 인체의 3차원 표현을 가지고 있음

$$BMI = \frac{W}{H^2}$$

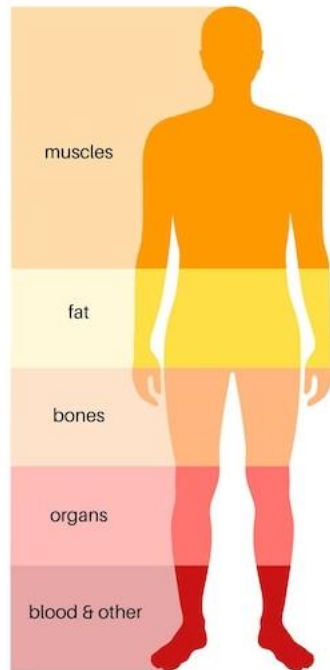
$W$  = 체중(kg)

$H$  = 신장(m)

✓ 인체를 구성하는 물질들은 알려져 있으며,  
밀도를 구할 수 있음

➔ 따라서 체중은 부피를 통해 역추정이 가능한 요소임

COMPOSITION OF THE HUMAN BODY



## Problem

Limitations when using only 2D Image

- ① 인체의 3차원적 표현을 제대로 반영하지 못함
- ② 기존 연구에서 RGB-D를 사용하여 BMI를 예측하였을 때, 깊이 값에 의존적인 경향을 보이므로 이미지만으로 한계가 존재함을 알 수 있음

How can we use a 3D representation?



2D image → 3D representation

- 실제 사용하기에 Memory Efficiency에 문제가 있음

2D image → 2.5D representation

- Monocular depth 이미지는 많은 발전을 이루었음

# Method

## 2.5D Representation

2차원이지만 3차원의 표현을 포함하고 있음

### Depth Map

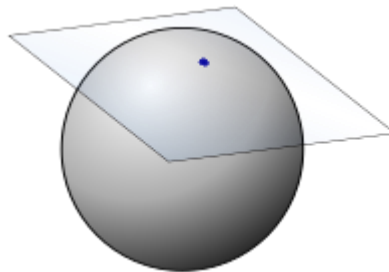
- 관찰 시점(viewpoint)으로부터 물체 표면과의 거리와 관련된 정보가 담긴 하나의 채널임

$$P = \begin{bmatrix} \frac{2}{\text{right-left}} & 0 & 0 & -\frac{\text{right+left}}{\text{right-left}} \\ 0 & \frac{2}{\text{top-bottom}} & 0 & -\frac{\text{top+bottom}}{\text{top-bottom}} \\ 0 & 0 & \frac{-2}{\text{far-near}} & -\frac{\text{far+near}}{\text{far-near}} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$z' = 2 \cdot \frac{z - \text{near}}{\text{far} - \text{near}} - 1$$

### Normal Map

- 일반적으로 RGB 구성 요소가 표면 법선의 X, Y 및 Z 좌표에 각각 대응하는 일반 RGB 이미지로 표현됨



## Monocular estimation

Adapt 2.5D representation

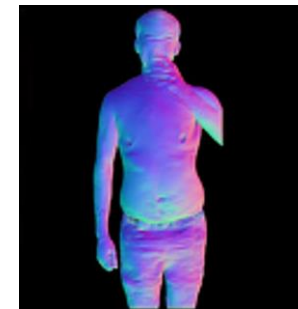
### Depth Map

- 단순 RGB이미지에 비하여 3차원 표현인 카메라와 물체로부터의 깊이 내용을 포함 가능함
- 카메라 시점에 의존적인 경향이 존재함



### Normal Map

- 물체기준의 법선 벡터의 정보를 가지고 있으므로 카메라 시점의 의존성을 낮춤
- Depth map 보다 세밀한 입체 정보 표현가능



# Method

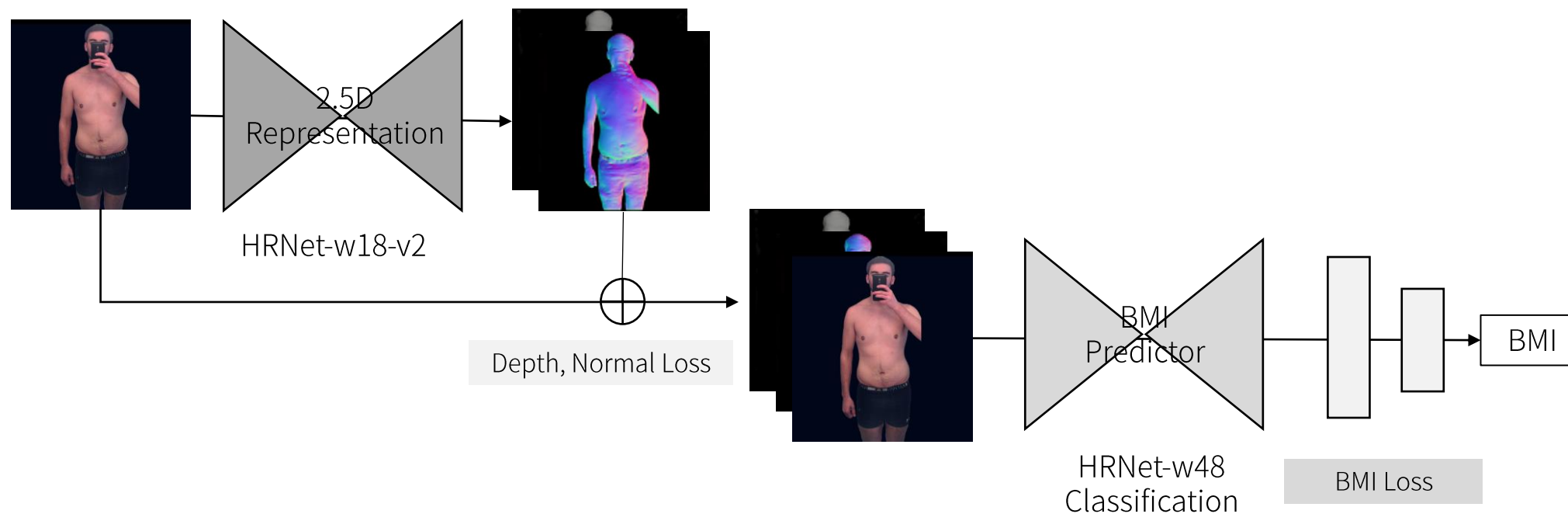
- Overall Architecture

- Backbone : HRNet-v2

→ 고해상도의 Feature map을 유지하여 최근 Backbone 모델 중 성능이 우수한 장점이 존재함

- 크게 2.5D Representation을 추론하는 부분과 BMI를 예측하는 부분으로 이루어짐

- Loss는 Depth, Normal Loss(L2)와 BMI Loss(L2)의 합으로 이루어짐





# Method

- 2.5D Representation estimator
  - THUMAN2.0 데이터셋을 이용하여 2.5D representation estimator를 Pretrained하고, BMI Prediction을 위해 Fine-tuning함

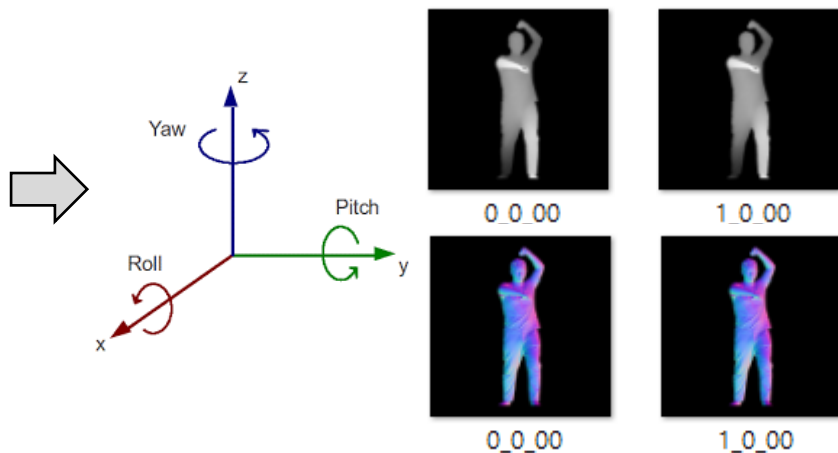
## THUMAN 2.0 Dataset

- 총 300명의 데이터를 Mesh Yaw를 기준으로 1도씩 360도 랜더링하여 한 사람 당 360개의 랜더링된 RGB, Depth Map, Normal Map GT를 도출함

3D Data



Depth Map, Normal Map GroundTruth



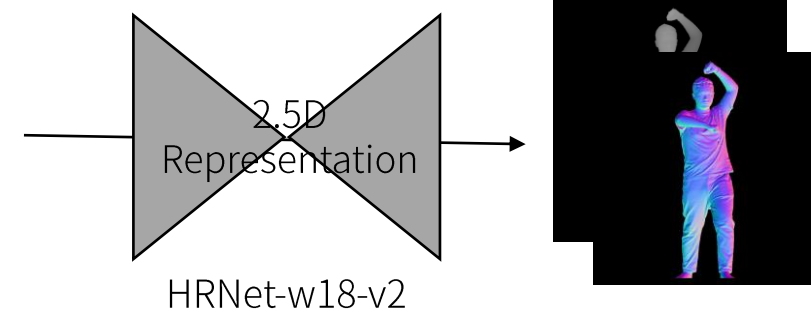
## Pretrained

- 랜더링된 이미지를 이용해 Normal map과 Depth map을 추론함
- $500 \times 360 = 180000$  쌍의 이미지, Normal map, Depth map 의 쌍을 갖게 됨
- HRNet-w18-v2를 이용함

랜더링 이미지

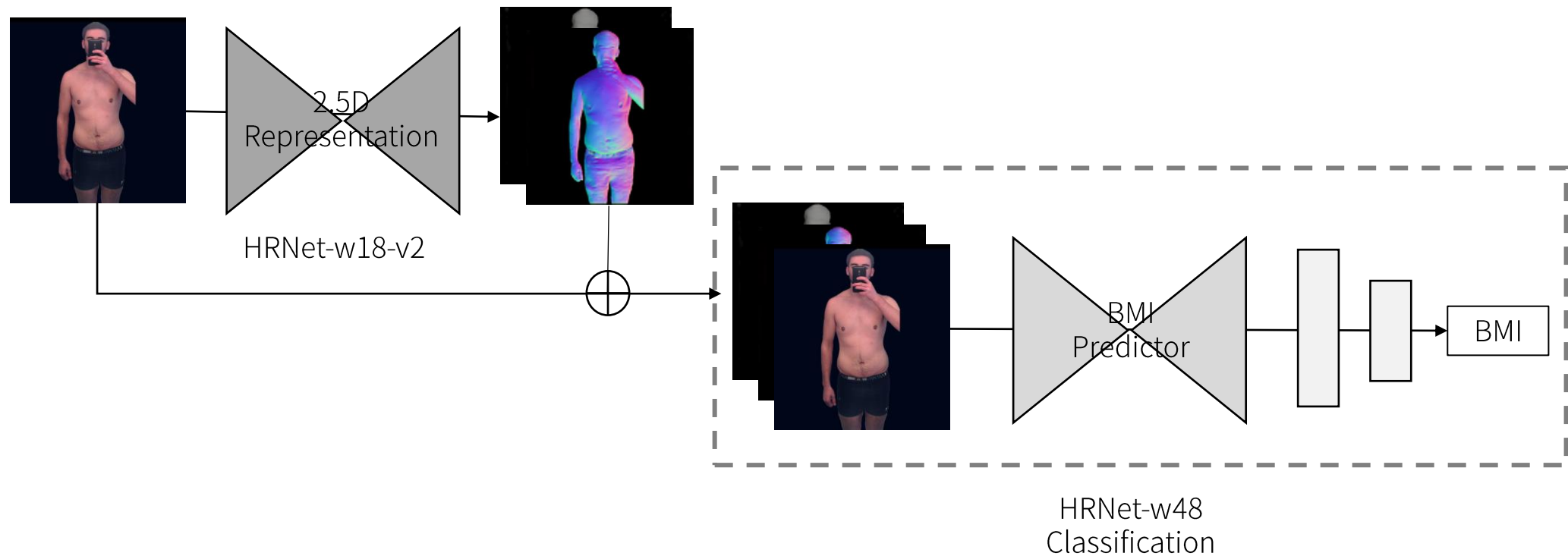


Normal map, Depth map



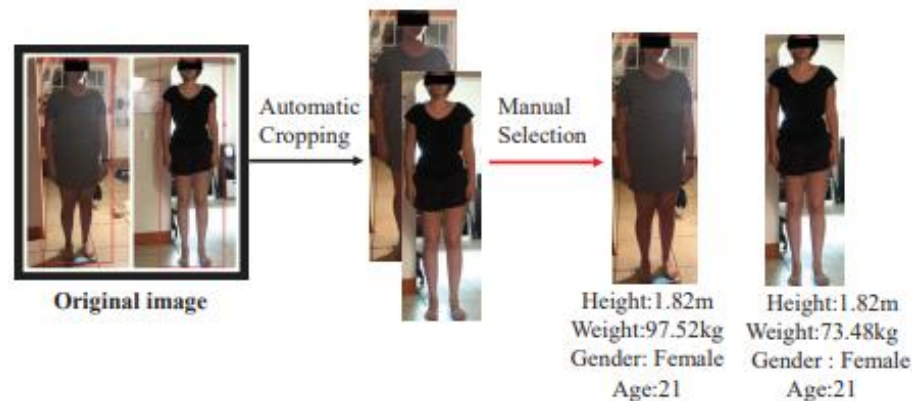
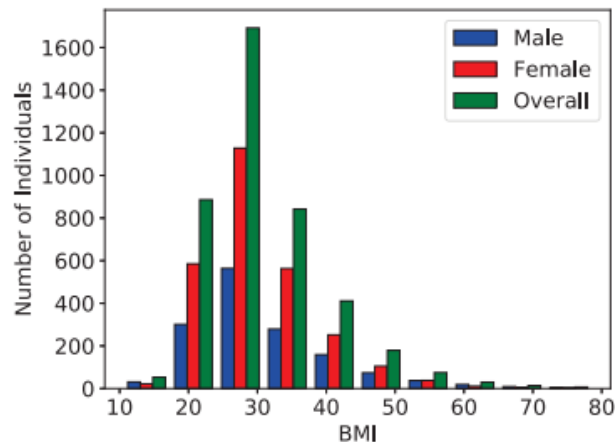
# Method

- BMI Predictor
  - Backbone 뒤에 Fully Connected Layer를 추가함
  - FC 1024  $\rightarrow$  Leaky ReLU  $\rightarrow$  FC 512  $\rightarrow$  ReLU  $\rightarrow$  FC 1



# Experiments

- 2D image to BMI Dataset
  - 한 웹사이트에서 수집한 데이터로 성별, 키, 나이, 몸무게 속성이 포함되어 있음
  - Train Data : 2917개, Test Data : 1254개
  - (BMI 기준) 저체중 : 1.77%, 보통 : 27.63%, 과체중 : 29.63%, 비만 : 41.47%



- Data Augmentation
  - BMI < 25, BMI > 35 에 대해 -30~30 각도의 rotation, flip을 진행하였음
  - Train Data : 7535개, Test Data : 1254개

# Experiments

---

- Implementation Training Details

- 2.5D Representation Estimator

- › Optimizer : Adam

- › Epoch : 70

- › Learning rate :  $10^{-4}$

- › Batch size : 32

- › Single Nvidia RTX 3090 GPU

- BMI Predictor

- › Optimizer : Adam

- › Epoch : 120

- › Learning rate :  $10^{-3}$

- › Batch size : 16

- › FC 1024 → Leaky ReLU → FC 512 → ReLU → FC 1

- › Single Nvidia RTX 3090 GPU

# Results – Ablation study

- Comparison of results according to backbone

	Test data	Underweight	Normal	Overweight	Obesity	overall
Ours	MAE	9.54	3.14	3.32	5.61	5.40
	MAPE	60.29%	13.88%	12.15%	14.19%	14.10%
HRNet-W18	MAE	13.64	1.24	4.53	5.23	6.16
	MAPE	86%	32.02%	16.58%	13.89%	15.30%

- Comparison of results according to input data

	Test data	Underweight	Normal	Overweight	Obesity	Overall
Ours	MAE	9.54	3.14	3.32	5.61	5.40
	MAPE	60.29%	13.88%	12.15%	14.19%	14.10%
Only RGB	MAE	9.19	3.59	3.43	5.60	5.45
	MAPE	58.09%	15.89%	12.56%	14.87%	14.90%
Normal, Depth	MAE	8.89	4.25	3.55	6.47	5.79
	MAPE	60.29%	13.88%	12.15%	14.89%	16.56%

# Results – other model

- Comparison with others model

	Test data	Underweight	Normal	Overweight	Obesity	Overall
Ours	MAE	9.54	3.14	3.32	5.61	5.40
	MAPE	60.29%	13.88%	12.15%	14.19%	14.10%
Zhi et.al	MAE	67.52	3.80	2.70	5.19	4.13
	MAPE	64.38%	17.22%	9.93%	12.58%	13.81%
Jiang et.al	MAE					5.94
	MAPE					19.20%

# Conclusion

---

- Limitation

- Data preprocessing으로 배경제거를 해야함 (Masking 필요)
- 데이터셋이 서양인으로 구성되어 있어서 인종에 따라 차이가 있을 것으로 보임
- 검은 옷을 입었거나 문신을 했을 경우 depth estimation을 하지 못하는 경우가 생김

- Expected Effect

- 물리적 공간과 기구에 구애 받지 않고 BMI를 측정 가능 함
- 미래의 헬스케어는 4P(Predictive, Preventive, Personalized, Participatory)로 변화될 것으로 예측됨에 따라, 건강의 지표로 사용되는 BMI를 예측해주는 시스템은 4P를 충족시켜주는 새로운 헬스케어 서비스로 활용 가능함
- 국가차원 혹은 대규모 인구조사에서 BMI를 통해 질병 유병률 모니터링이 가능할 것으로 기대됨