

深層学習を用いた遮蔽あり 単一画像からの 人体3次元形状推定

工学研究科 情報システム工学分野

指導教員 棕木雅之

T2003531 森慎太郎

2022/1/31

研究背景

- 写真を撮るときに、遮蔽で人体の一部が欠損する時がある。
 - 集合写真、防犯カメラ
- 欠損を推定して、3次元形状推定する技術は重要。

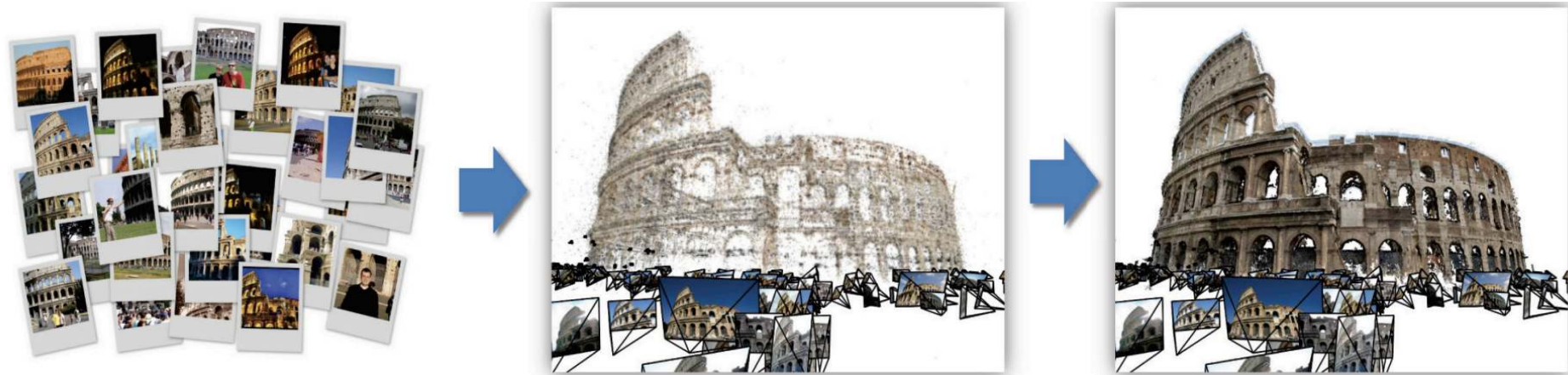


別角度から撮影した画像の生成
防犯カメラ中の人物像の再現



3次元形状推定

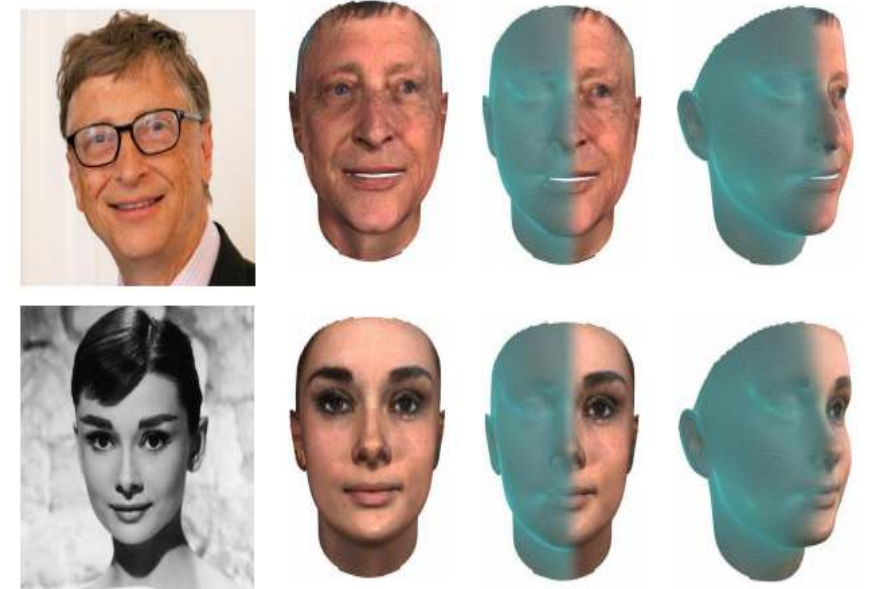
- 画像が与えられたとき、3次元形状を推定すること



[1] Sameer Agarwal, Yasutaka Furukawa, Noah Snavely, Ian Simon, Brian Curless, Steven M. Seitz, and Richard Szeliski. 2011. "Building Rome in a day". Commun. ACM 54, 10 (October 2011), 105-112.

深層学習を用いた 3次元形状推定

- 入力画像が少なく、3次元モデル生成が可能
- 学習データに依存する
- 学習すれば、1枚の画像から推定可能



深層学習による3次元形状推定の現状

- 一般背景下では、3次元形状推定されていない。
- 限定的な画像を用いて推定を行なっている



研究の目的

- 遮蔽で画像が一部分欠損しても、正確に人体の3次元形状推定

- 画像は、単一の画像を使用

- ① 深層学習によるシルエット抽出

- ② 深層学習による遮蔽部分の補完

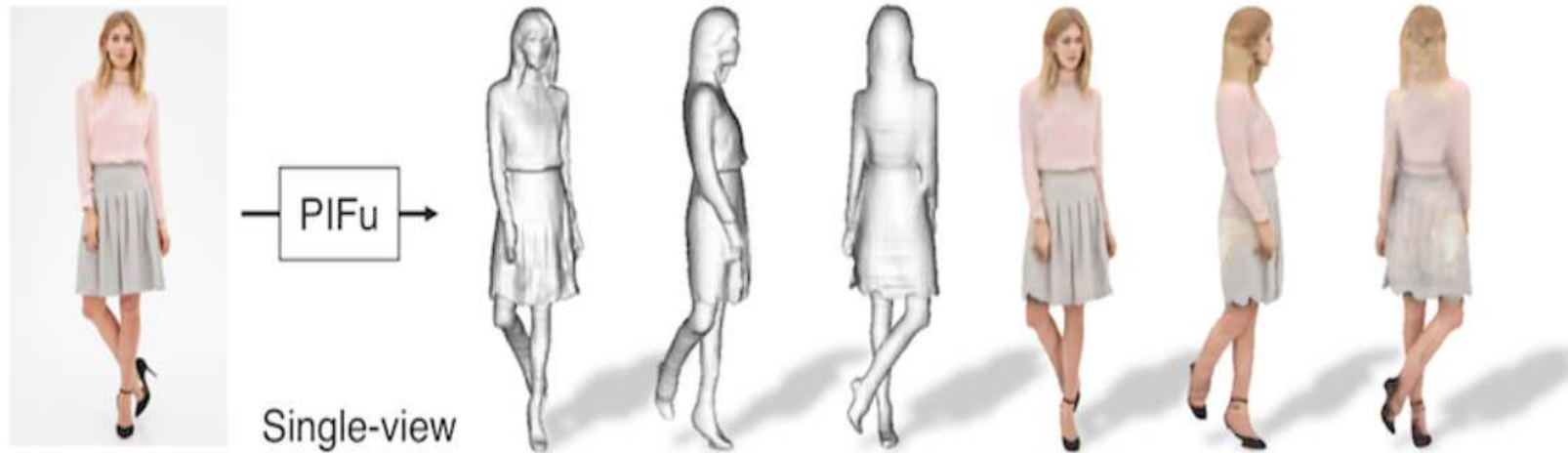
- ③ 深層学習による人体3次元形状推定

の組み合わせで問題解決



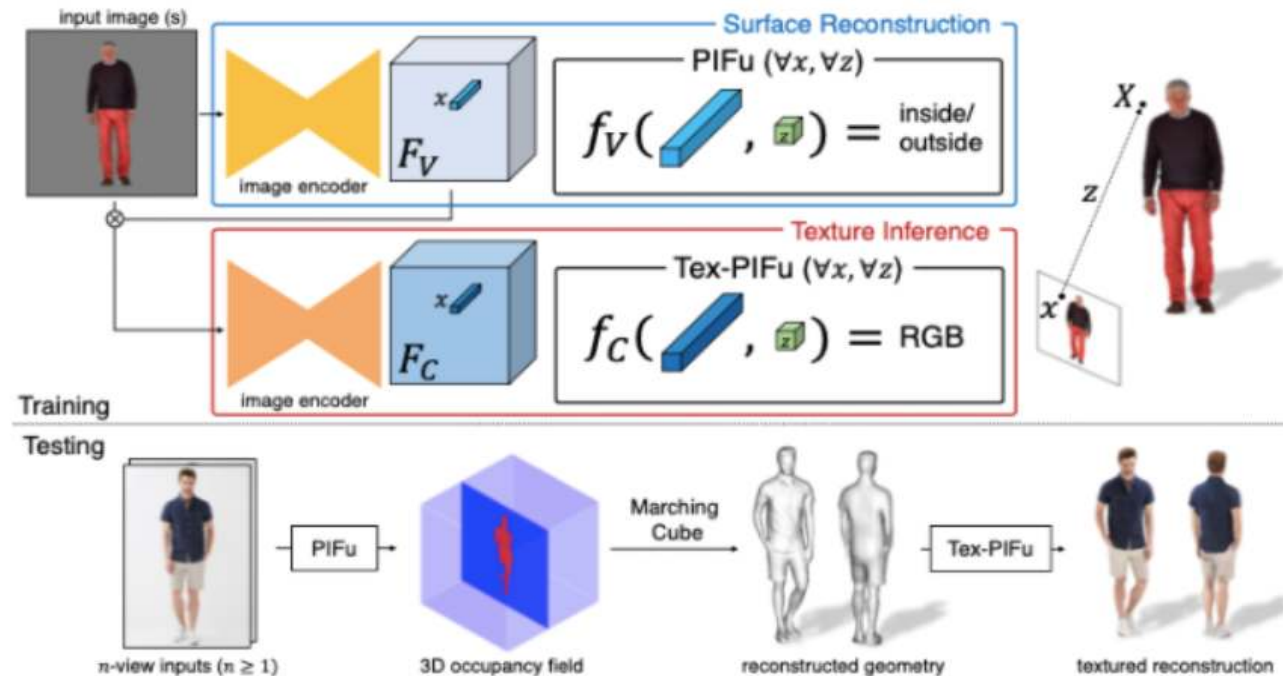
人体 3 次元形状推定

- PIFuを利用する。
- 入力：元画像 + シルエット画像
- 出力：人体の 3 次元モデル



Saito, S., Huang, Z., Natsume, R., Morishima, S., Li, H., & Kanazawa, A. PIFu: Pixelaligned implicit function for high-resolution clothed human digitization.", Proc. International Conference on Computer Vision, pp. 2304-2314, 2019

PIFuの概要

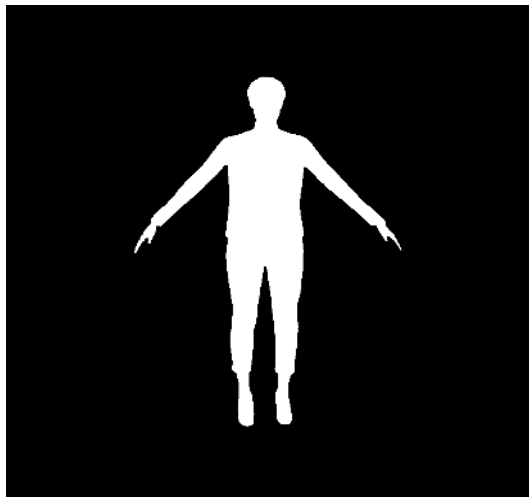


3次元形状推定を行う: Surface Reconstruction
テクスチャ推定を行う: Texture Inference

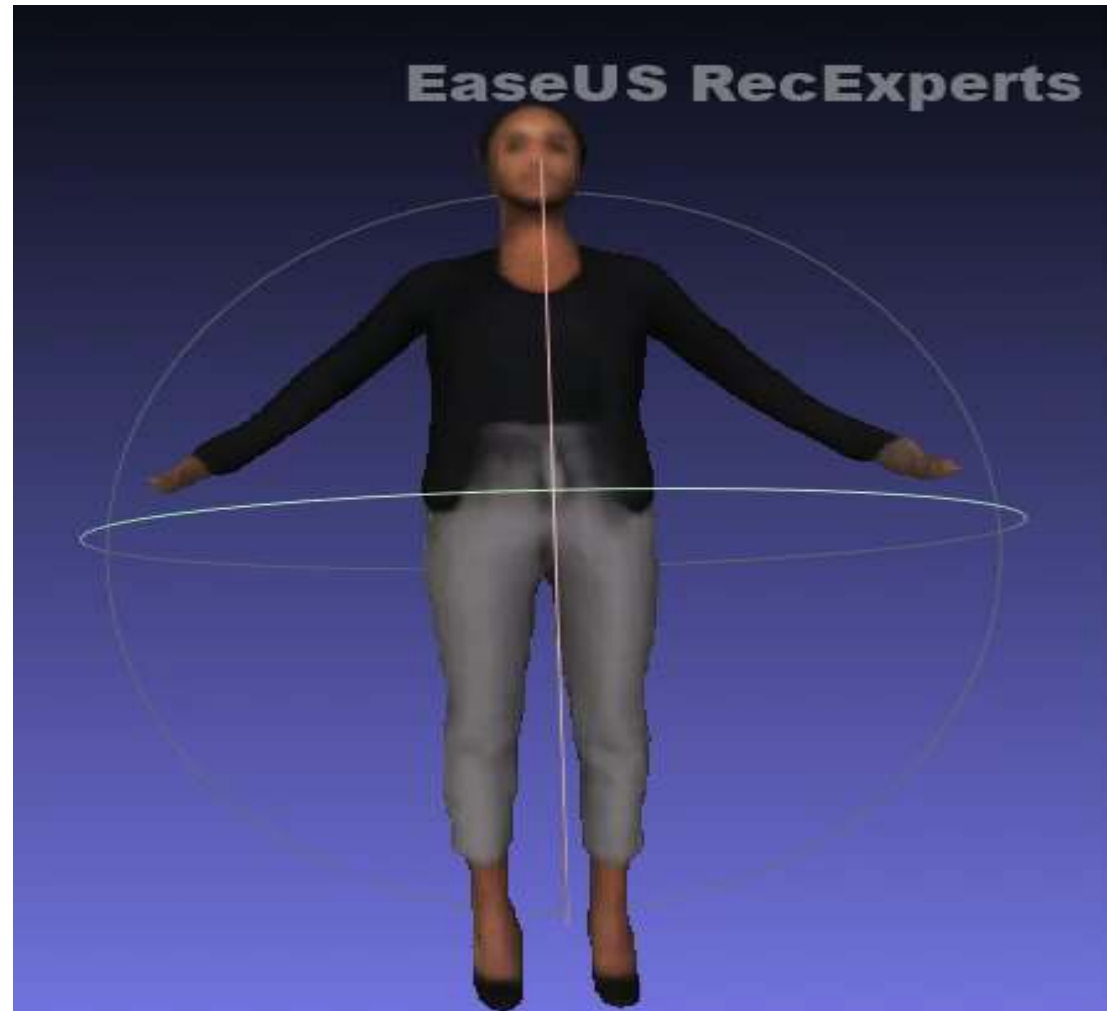
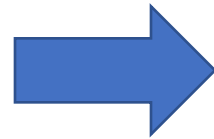
PIFuの実験(欠損なし)



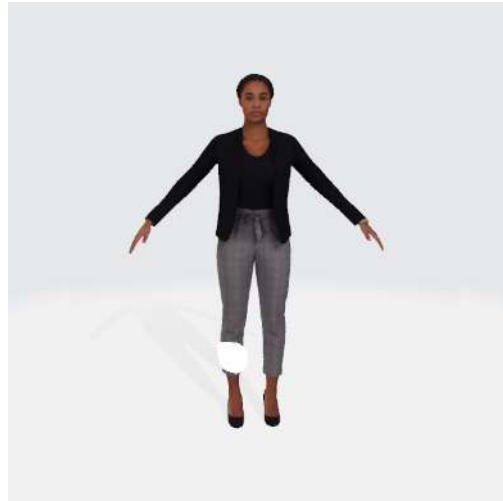
元画像



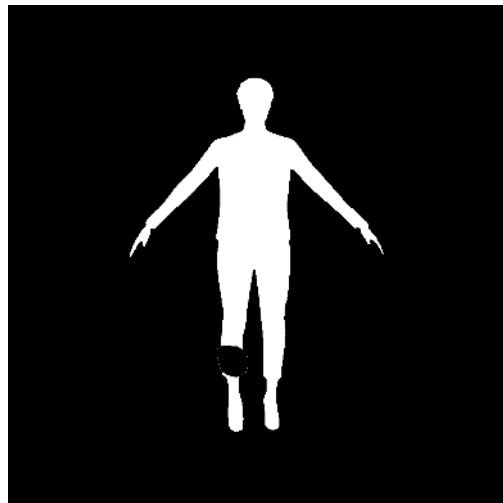
シルエット画像



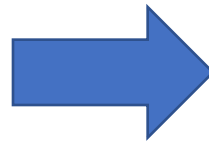
PIFuの実験(欠損あり)



元画像



シルエット画像

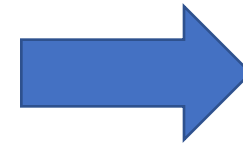
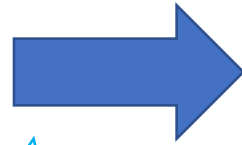
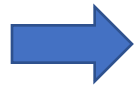


PIFuの問題点

- 元画像とシルエット画像の入力が必要
- 元画像が欠損しているると、形が不自然



処理の流れ



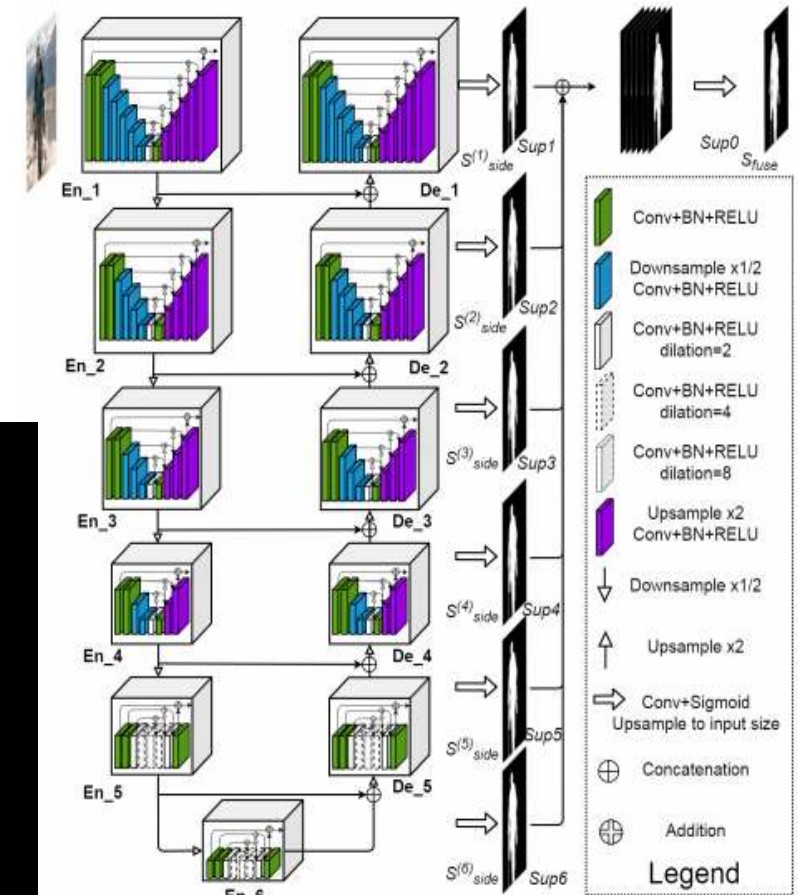
シルエット
抽出

遮蔽部分
の補完

PIFu

シルエット抽出

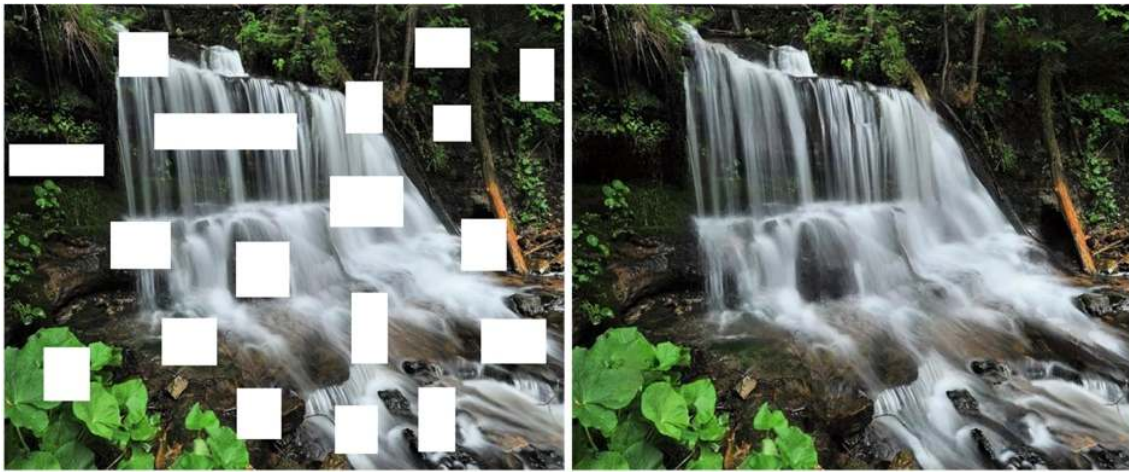
- シルエット抽出では、前景抽出を利用
- 前景抽出は、前景と背景を分ける
- $U^2 - Net$ を利用
- 最後に2値化



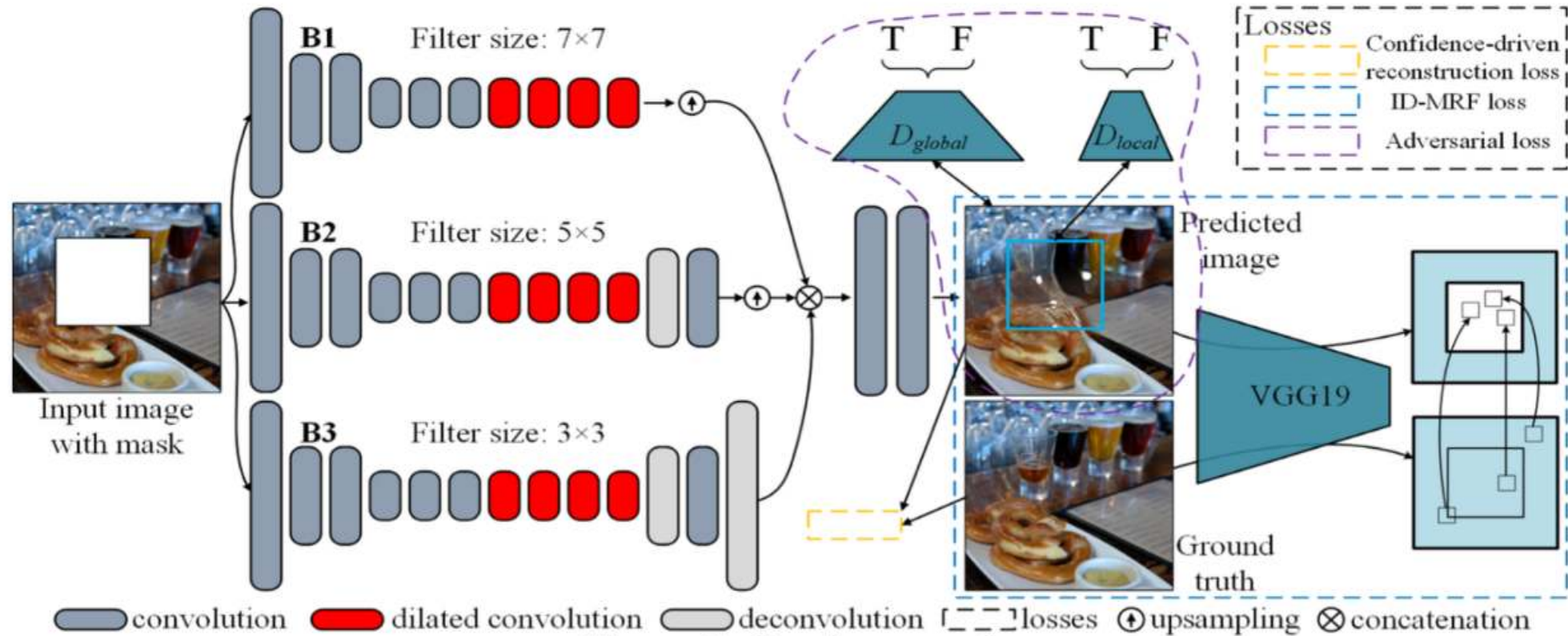
$U^2 - Net$ の概略

遮蔽部分の補完

- 画像補完を利用
- 本来はない画素を画像処理で作り出して補うこと。
- GMCNNを利用



GMCNNの概要



GMCNNには、元画像、シルエット画像を適用

実験

- 1. シルエット抽出の評価実験
- 2. 遮蔽部分の補完の評価実験
- 3. 提案手法の評価実験

を行います。

シルエット抽出の実験

- 実験設定

画像1:背景が単純な室内画像

画像2:背景が複雑な室内画像

画像3:背景が複雑な屋外画像

画像4:背景が単純な屋外画像

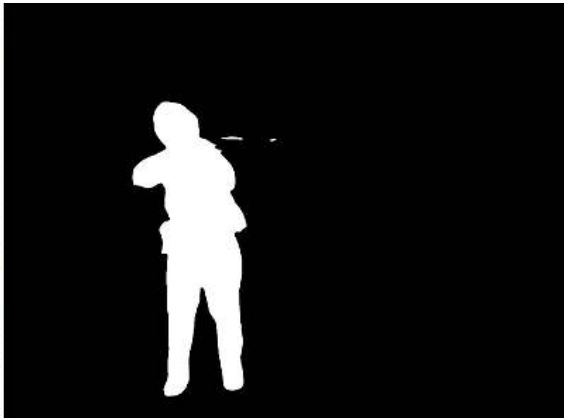
画像5:背景の一部が人体と似ている屋外画像

この **5種類**の画像を使用

シルエット抽出の結果(成功例)



シルエット抽出の結果(失敗例)



結果：シルエット抽出として有効

画像補完の実験

- マスク領域によってどのような結果が出るか検証
- マスク領域
 - 太もも
 - 足先
- 白い部分がマスク領域

GMCNNの結果(1)



(a)元画像とマスク領域



(b)補完元画像



(c)(b)のシルエット画像



(d)補完シルエット画像



GMCNNの結果(2)



(a)元画像とマスク領域



(b)補完元画像



(c)(b)のシルエット画像



(d)補完シルエット画像

考察

- シルエット自体に補完をすると、精度良く補完ができる。
- しかし、マスク領域が手先全体や足先全体になると欠けたまま
- 腕や足のつながっている一部分が遮蔽されていることが条件

提案手法の評価実験

- 深層学習の3つの手法を組み合わせた方法で3次元形状推定
- 定性的に評価
- 様々な方向から評価

実験 1



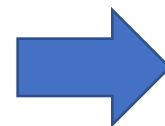
元画像とマスク領域



補完元画像



補完シルエット画像



実験2



元画像とマスク領域



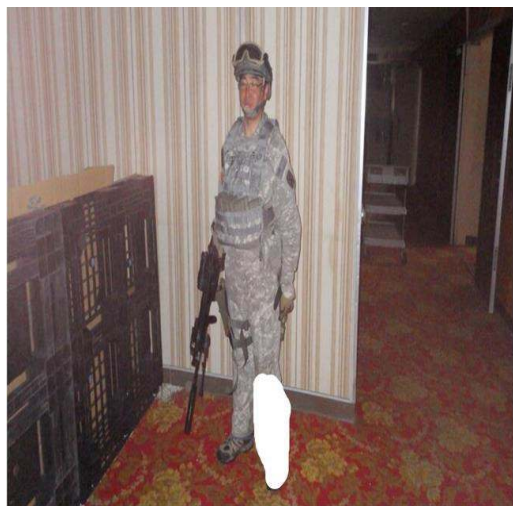
補完元画像



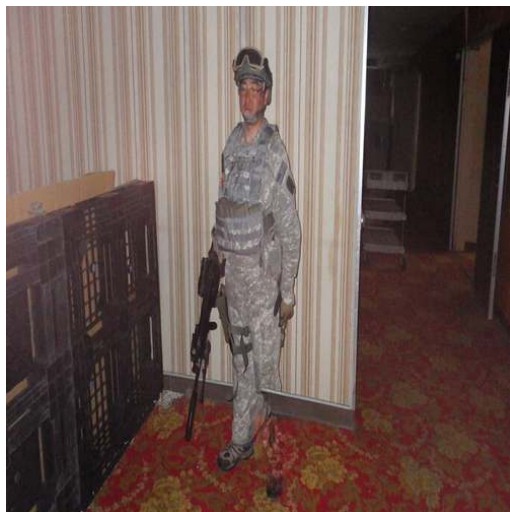
補完シルエット画像



実験3



元画像とマスク領域



補完元画像



補完シルエット画像



補足実験(銃が消える)



元画像とマスク領域



補完元画像



補完シルエット画像



まとめ

- 本研究では、遮蔽あり単一画像からの3次元形状推定の手法の提案を行なった
- 形状推定自体は違和感なく行えた
- 足先や腕先がすべて欠損していると、うまくいかない

今後の課題

- 定量的な評価
- 1つのネットワークに統合
- 精度のよい画像補完の手法を見つける
- 犬や猫・植物などへの違う種類への対応