# BodyMAP - Jointly Predicting Body Mesh and 3D Applied Pressure Map for People in Bed

Abhishek Tandon* Anujraaj Goyal*
Carnegie Mellon University
United States
{abhishektnd, argo}@cmu.edu

Henry M. Clever
NVIDIA
United States
hclever@nvidia.com

Zackory Erickson
Carnegie Mellon University
United States
zackory@cmu.edu

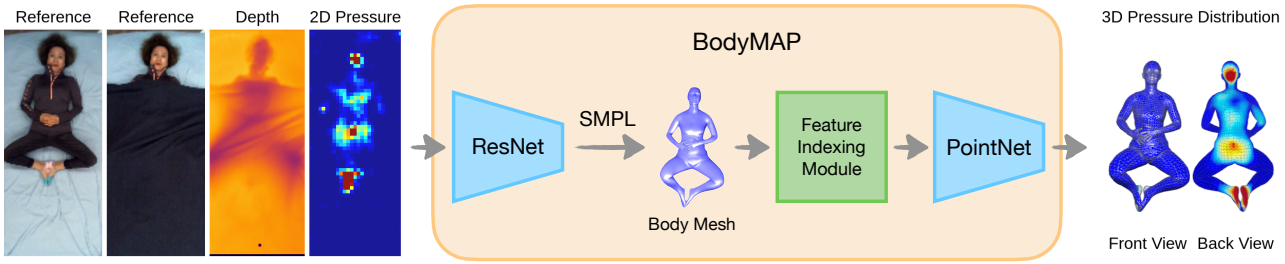


Figure 1. BodyMAP leverages a depth and pressure image of a person in bed covered by a blanket, to jointly predict the body mesh and a 3D pressure map of pressure distributed along the human body.

## Abstract

*Accurately predicting the 3D human posture and the pressure exerted on the body for people resting in bed, visualized as a body mesh (3D pose & shape) with a 3D pressure map, holds significant promise for healthcare applications, particularly, in the prevention of pressure ulcers. Current methods focus on singular facets of the problem—predicting only 2D/3D poses, generating 2D pressure images, predicting pressure only for certain body regions instead of the full body, or forming indirect approximations to the 3D pressure map. In contrast, we introduce BodyMAP, which jointly predicts the human body mesh and 3D applied pressure map across the entire human body. Our network leverages multiple visual modalities, incorporating both a depth image of a person in bed and its corresponding 2D pressure image acquired from a pressure-sensing mattress. The 3D pressure map is represented as a pressure value at each mesh vertex and thus allows for precise localization of high-pressure regions on the body. Additionally, we present BodyMAP-WS, a new formulation of pressure prediction in which we implicitly learn pressure in 3D by aligning sensed 2D pressure images with a differentiable 2D projection of the predicted 3D pressure maps. In evaluations with real-world human data, our method outperforms the current state-of-the-art technique by 25% on both body mesh and 3D applied pressure map prediction tasks for people in bed.*

## 1. Introduction

With 2.5 million cases annually in the U.S. alone, pressure ulcers remain a pressing concern in healthcare systems worldwide [3]. Although body repositioning serves as a common preventive measure [30], concerns persist about whether such pose shifts effectively redistribute applied pressure on the body [37]. Wearable devices [31, 38] while aiding in localizing peak pressure on the body, can disrupt normal physical activities and contribute to pressure injuries themselves [4, 18, 20]. In response, clinics and long-term care facilities are increasingly looking at pressure sensor systems to aid in pressure ulcer prevention [17, 30, 42]. These systems utilize a pressure sensing array, which is placed beneath the person, and measures the pressure applied by the human body onto the sensor array, producing a 2D pressure image. While informative about pressure applied on the body [2, 13, 37], these pressure images have inherent ambiguities due to their 2D depiction of pressure. As illustrated in Fig. 2, distinct body postures can result in remarkably similar pressure images, failing to correctly convey which body parts are under high pressure. These systems thus currently require a human caregiver to deduce the correct 3D distribution of pressure on the human body, based on the 2D pressure image and 3D body posture. This

* Equal contribution.

step not only adds cognitive workload to caregivers but also introduces delays or potential errors in implementing preventive measures. Moreover, individuals in bed are often covered with blankets, making it challenging for nurses or caregivers to accurately identify high-pressure regions on a care recipient's body.

Addressing these challenges, we propose BodyMAP (Fig. 1), a deep-learning model, that jointly estimates a human body mesh (3D pose & shape) and a 3D pressure distribution on the body mesh for people in bed, covered with blankets. Visualizing the pressure map on the 3D human body mesh, as illustrated in Fig. 2, precisely pinpoints body regions under peak pressure. Automatic body mesh and 3D pressure map predictions could reduce the need for caregivers to manually infer them, and offer visual insights into pressure redistribution as caregivers reposition a person's body. This real-time feedback on peak body pressure can transform the current 'blind' body repositioning into an informed process, enhancing care quality and potentially reducing the occurrence of preventable pressure injuries. Additionally, other domains requiring body mesh and pressure knowledge such as assistive robotics [7, 44], sports rehabilitation [16], and eldercare applications [33] could benefit from accurate pressure prediction on the human body.

BodyMAP uses a combination of a depth image and the corresponding 2D pressure image, for an individual in bed, as inputs to predict the body mesh and the applied 3D pressure map. A depth camera, placed vertically above the bed, provides a top-down view of the mattress, while a fabric pressure sensing array placed beneath the person, provides the 'bottom-up view' of the body. We represent the human body using the SMPL parametric model [29] and predict the SMPL parameters related to body shape and joint angles. Our method first predicts the SMPL mesh and then leverages the predicted mesh vertices and latent features to estimate a pressure value at each vertex, producing a 3D pressure map. To train our model, we obtain ground truth applied pressure map data by projecting 2D pressure images onto corresponding ground truth body meshes [9]. Our project page containing code, trained models and the 3D pressure ground truth data for SLP [25, 27] and BodyPressureSD [9] datasets used in our work are available for research purposes at https://bodymap3d.github.io/.

Prior methods either predict pressure maps for a subset of body regions [44] or indirectly approximate 3D pressure [9]. In contrast, BodyMAP uses a unified model to jointly predict the human mesh and 3D applied pressure map across the entire body. This joint training approach offers several advantages, including human-interpretable visualizations of pressure on the human body mesh for caregivers, reduced model inference time, and more accurate predictions of body pose and 3D pressure.

Additionally, we introduce BodyMAP-WS (3D applied pressure prediction without supervision), a variant of our model, designed to implicitly learn the 3D pressure map without direct supervision. BodyMAP-WS constructs a differentiable 2D projection of a predicted 3D pressure map and aligns the 2D projection with the input pressure image. This alignment encourages the model to learn the correct 3D pressure map without relying on ground truth pressure maps on the human body that are difficult to measure in the real world. By removing the need for labels, this technique holds the potential to reduce data collection efforts and allow for training models on unlabeled data.

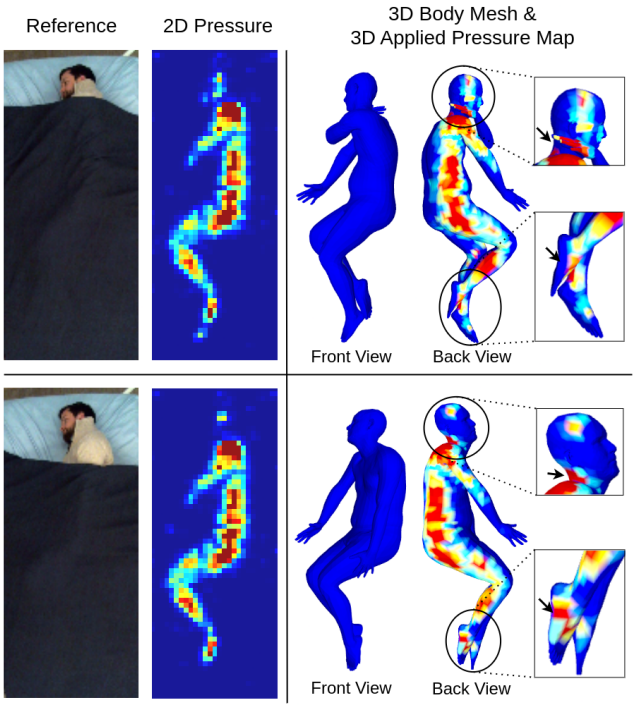


Figure 2. Distinct postures can have similar 2D pressure images. The insets of the 3D pressure map show pressure being applied to different areas demonstrating its use in localizing pressure that is applied on the human body.

Our method sets a new state-of-the-art in predicting body mesh (3D pose & shape) and 3D applied pressure on the human body, surpassing previous works by 25% on both tasks, as evidenced by our model's performance across various 3D pose, 3D shape, and 3D pressure metrics. In summary, our contributions include the following:

- A deep learning method, BodyMAP, that takes as input a depth image and the corresponding 2D pressure image for a person in bed and infers jointly the body mesh and the 3D pressure map of pressure applied on the body.
- An implicit 3D pressure map prediction method, BodyMAP-WS, to support scaling up body pressure prediction models on real-world unlabeled datasets.

## 2. Related Work

### 2.1. Pose estimation using deep learning

Pose estimation is a well-established task in Computer Vision [1, 28, 40, 41]. Parametric human models [29, 34] like SMPL simplify the representation of the human body, forming a dense 3D body mesh from only a small set of parameters for body shape and joint angles.

While supervised methods [23, 24] use 3D ground truth data, recent efforts have explored geometric cues and multi-view consistency to predict body mesh from RGB images [5, 22, 43]. Gong *et al*. [12] use multiple 2D modalities to predict the SMPL body and employ self-correction between the input and 2D projections to train their model. Inspired by this line of research in pose estimation, we developed our method, BodyMAP-WS to learn 3D pressure maps without direct supervision.

### 2.2. In-bed human pose estimation

In-bed human poses present distinct challenges compared to active poses observed in sports activities. Heavy occlusion caused by blankets and self-occlusion present complexities. Additionally, there is limited body visibility in modalities such as pressure images which fail to capture limbs that are not in contact with the mattress. Datasets like the SLP dataset [25, 27] featuring real-world human participants and the synthetic BodyPressureSD dataset [9], which we leverage in this study, have propelled research in this domain.

Yin *et al*. [45] introduce Pyramid Fusion, incorporating modalities like RGB, pressure, depth, and IR images. However, their design necessitates multiple model passes to derive the final predicted body mesh. Considering potential clinical deployments and privacy concerns, we move away from using RGB images in our work. Clever *et al*. propose methods [8, 9] involving multiple models to estimate the body mesh. These approaches only rely on a single modality (pressure image and depth respectively). Our method, BodyMAP, simplifies this process by employing a single pass through a single model to jointly predict both the body mesh and applied 3D pressure map, while utilizing multiple visual modalities as input. Our findings demonstrate that our streamlined approach significantly improves performance over prior methodologies.

### 2.3. 3D applied pressure map prediction

Prior techniques [9, 26] have explored the task of predicting 2D pressure images from alternative modalities like depth and RGB images respectively, offering a cost-effective means of obtaining information about applied pressure. However, as illustrated in Fig. 2, 2D pressure images fail to correctly identify the body regions under high pressure.

In the related domain of pain monitoring, 3D pain drawings have emerged to be more effective visualization tools than their 2D counterparts [11, 39] in analyzing body pain. In this work, we expand on this concept by directly predicting the 3D applied pressure map onto a human body mesh, overcoming the limitations of 2D pressure images.

Only a few prior methods explore predicting pressure on the human body. Wang *et al*. [44] predict pressure applied during cloth interaction, only on specific body regions such as arms. Clever *et al*. [9] proposed the concept of full-body 3D pressure maps. The authors developed a multi-model pipeline trained in multiple stages, to predict both the body mesh and 2D pressure image from a depth image. Finally, they approximate the 3D pressure map by vertically projecting the 2D pressure image over the estimated 3D human body. We note that errors from each model often compound, affecting the accuracy of 3D pressure map predictions.

In contrast, we devise a unified model architecture, BodyMAP, to estimate the body mesh and pressure map jointly. Our model leverages PointNet [35] to predict 3D pressure map at the vertex level across the entire body mesh.

## 3. Method

We train a deep-learning model, BodyMAP, to jointly predict the body mesh $\hat{M}$ (3D pose & shape), along with the 3D applied pressure map $\hat{P}$. Specifically, our model takes the individual's gender $g$, the depth image $d$, and the 2D pressure image $p$ as inputs.

The depth image is captured by a depth camera situated above the bed, while the 2D pressure image is generated by a pressure sensing system positioned beneath the individual. This arrangement captures complementary features regarding the body, illustrated in the appendix Fig. 5, thereby enhancing the context available to the model for accurate predictions of both body mesh and pressure map.

The human body is represented using the SMPL [29] mesh while the 3D pressure map is represented at the vertex level with a pressure value for each vertex of the human mesh, allowing for peak pressure localization on the body.

We train our models on both the BodyPressureSD dataset [9] of simulated humans in bed and the real-world SLP dataset [25, 27]. For both datasets, we have depth and pressure images, aligned with the 3D ground truth mesh for diverse poses (supine, left & right lateral) and multiple blanket thickness configurations.

### 3.1. BodyMAP

In this section, we detail the architecture of BodyMAP, as illustrated in Fig. 3. The depth and pressure images are resized, concatenated and processed together as image channels by the model.

**Body mesh prediction**: The input is first encoded using ResNet18 [15]. The latent features are then passed through a multi-layer perceptron to predict the SMPL parameters

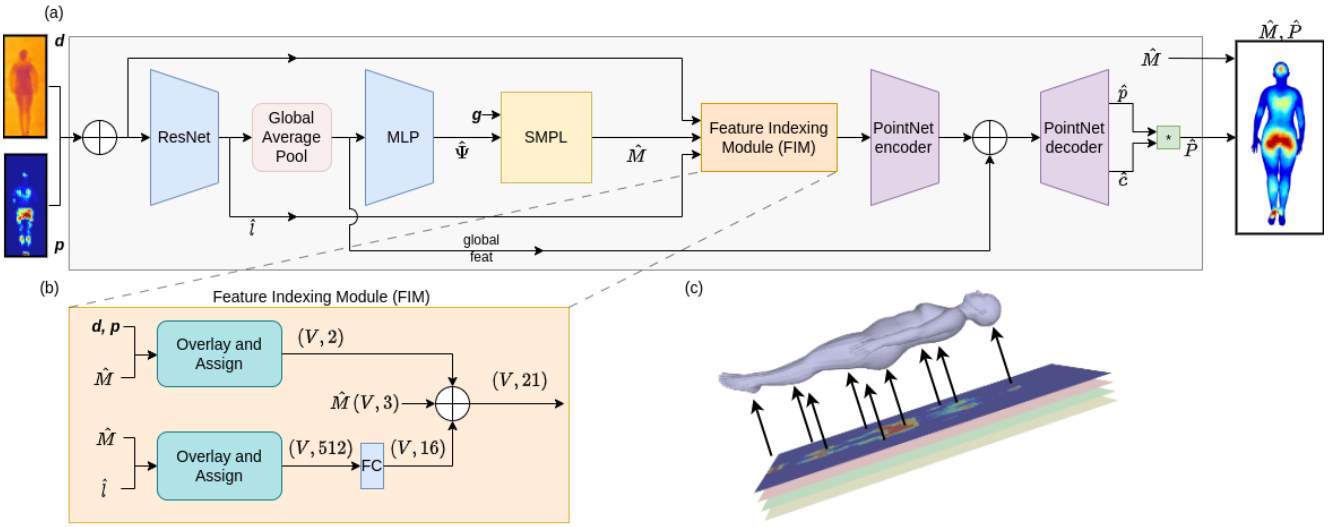


Figure 3. BodyMAP jointly predicts body mesh and 3D applied pressure map for an individual in bed. (a) Model architecture that encodes depth image ***d*** and 2D pressure image ***p*** to predict SMPL [29] parameters $\hat{\Psi}$, used to reconstruct the SMPL mesh $\hat{M}$. The Feature Indexing Module (FIM) accumulates features for the mesh vertices from the input images and ResNet features. PointNet predicts the per-vertex pressure $\hat{p}$ & per-vertex binary contact $\hat{c}$ along the human body using the mesh features as input. Finally, the 3D pressure map $\hat{P}$ is calculated as the product of the per-vertex pressure value and binary contact value. (b) FIM overlays the predicted mesh over the ResNet feature maps ($\hat{l}$) and input images by mapping mesh vertex locations to pixel positions and then assigns features to each vertex (V). These are fused along with the mesh vertex locations and used for 3D pressure map prediction. (c) FIM's 'overlay and assign' step visualization.

$\hat{\Psi}$. These parameters include body shape, joint angles, root-joint translation, and root-joint rotation. The SMPL parameters $\hat{\Psi}$ and gender information $g$, serve as inputs to the SMPL embedding block [19]. This block does not contain any learned parameters and outputs a differentiable human body mesh $\hat{M}$ with vertices $\hat{V}$, and 3D joint positions $\hat{S}$.

**Feature Indexing Module**: We introduce the Feature Indexing Module (FIM), depicted in Fig. 3(b), to accumulate features for each mesh vertex. As mentioned in [9], the mesh predictions are spatially registered with the corresponding input images. This registration provides pixel locations where each mesh vertex would project onto the input images [9]. FIM assigns features to each mesh vertex from both the input images and the latent ResNet features (before global average pooling) using the mentioned pixel locations (Fig. 3(c)). These features are fused with the vertex locations and are utilized for pressure map prediction.

**Pressure map prediction**: To predict a per-vertex pressure map ($\hat{P}$), we employ PointNet [35] utilizing the latent features formed from FIM for each vertex as input. This establishes a strong correlation between the mesh prediction and pressure prediction, ensuring their consistency with each other. In a manner akin to point-based segmentation architectures [35, 36], ResNet features (after global average pooling) are fused with PointNet encoder features. This provides the model with an enhanced contextual understanding. The PointNet decoder then predicts a per-vertex binary contact value and pressure value. The per-vertex contact value serves as an indicator of whether each mesh vertex is in contact with the mattress. We use the predicted contact values to further tune the predicted pressure values. Specifically, the 3D pressure map is finally estimated as a product of the binary contact value with the corresponding pressure value for each vertex.

#### 3.1.1 Training strategy

We train the network to jointly predict body mesh and 3D applied pressure map with the following loss:

$$L = \mathcal{L}_{\text{SMPL}} + \lambda_1 \mathcal{L}_{\text{v2v}} + \lambda_2 \mathcal{L}_{\text{P3D}} + \lambda_3 \mathcal{L}_{\text{contact}} \quad (1)$$

where $\mathcal{L}_{\text{SMPL}}$ minimizes the absolute error on SMPL parameters $\hat{\Psi}$ and squared error on the 3D joint positions $\hat{S}$. $\mathcal{L}_{\text{v2v}}$ (vertex-to-vertex loss) provides extra supervision to better match ground truth body mesh by minimizing squared error on the vertex positions $\hat{V}$. $\mathcal{L}_{\text{P3D}}$ (3D pressure loss) minimizes the squared error on the 3D pressure map values. $\mathcal{L}_{\text{contact}}$ is applied as a cross-entropy loss between predicted and ground truth 3D contact, where the ground truth 3D contact is obtained from all non-zero elements of the ground truth 3D pressure maps. The loss weighting coefficients are set empirically. Further details about the loss function are mentioned in appendix Sec. 8.3.

### 3.2. BodyMAP-WS

BodyMAP-WS is a variant of the BodyMAP model, learning without supervision for the 3D pressure map prediction. This model utilizes a pre-trained mesh regressor to obtain mesh predictions and ResNet image features as its primary inputs. The network architecture follows a similar design as BodyMAP's pressure prediction part (Fig. 3) in its use of FIM and PointNet [35], and is further illustrated in the appendix Fig. 7. To train BodyMAP-WS without supervision, we drop the prediction of binary contact value from BodyMAP and directly predict the pressure value at each vertex. During training, the network constructs a 2D projection of the predicted 3D pressure map and aligns it with the input pressure image, allowing for the implicit learning of the 3D applied pressure map onto the body mesh. To form the differentiable projection, we utilize the co-registration of the mesh with pressure taxels [9], which establishes a direct link between the pressure on vertices $v_j$ and the pressure on the taxel located beneath the vertex. To project the predicted 3D pressure map onto a 2D image, we iterate over all the pressure taxels and compute the average pressure value of the vertices positioned directly above a taxel (vertices with $x$, $y$ coordinates corresponding to the taxel).

#### 3.2.1 Training strategy

To form the pre-trained mesh regressor, we first train the mesh regressor section of BodyMAP using supervision with $\mathcal{L}_{\text{SMPL}}$ and $\mathcal{L}_{\text{v2v}}$. Subsequently, BodyMAP-WS is trained to leverage the frozen pre-trained mesh model's predicted ResNet features and mesh vertex locations as inputs, for predicting the 3D applied pressure map. This prediction is guided by the following loss function:

$$L = \mathcal{L}_{\text{P2D}} + \lambda_1 \mathcal{L}_{\text{Preg}} \tag{2}$$

Here, $\mathcal{L}_{\text{P2D}}$ (2D pressure loss) minimizes the squared error between the 2D projection of the predicted 3D pressure map and the input pressure image. The bed mattress is situated on the $Z = 0$ plane, and vertices predicted to be positioned above this plane ($Z > 0$) should ideally not have any applied pressure on them as they do not make contact with the mattress. To enforce this constraint, we utilize $\mathcal{L}_{\text{Preg}}$, a regularization term that penalizes positive pressure on these vertices by minimizing the norm of predicted pressure values for these vertices. Further details about the loss function are mentioned in appendix Sec. 9.1.

## 4. Evaluation

The training datasets [9, 25, 27] used in this work provide aligned input depth and pressure images. Clever *et al*. [9] provide 3D ground truth meshes for their released synthetic data and SMPL-3D fits for the real SLP dataset [25, 27]. To establish the ground truth 3D pressure map, we utilize the vertical projection method outlined in [9]. This method involves projecting the 2D pressure image onto the 3D-oriented ground truth mesh, forming a ground truth 3D pressure map. Although this method has limitations for self-contact (overlapping body parts), this procedure provides effective approximations for ground truth 3D pressure maps, facilitating comparison with prior methods [9, 45].

We train our network, BodyMAP, on the entire simulated BodyPressureSD dataset [9] and the initial 80 subjects (1-80, with the exclusion of subject 7 due to calibration errors) from the real-world SLP dataset [25, 27]. We pre-train the mesh regressor for BodyMAP-WS on the same training split. However, we train the BodyMAP-WS model on only the real-world SLP training data (1-80 subjects) [25, 27] using only the depth and pressure images and no ground truth 3D pressure map data. We evaluate our methods on the final 22 subjects (81-102) from the real SLP dataset [25, 27], which aligns with the analysis of our baselines [9, 45].

### 4.1. Network Evaluation

We evaluate our approaches using multiple 3D pose, 3D shape and 3D pressure map metrics.

For 3D pose prediction, we evaluate our methods with 3D mean-per-joint position error (MPJPE) and 3D per-vertex error (PVE). These are computed as mean Euclidean error between predicted and ground truth 3D joint positions and 3D vertex positions respectively.

For 3D shape prediction, we assess our methods by computing anatomical measurements for height and circumferences of the chest, waist, and hips of the mesh as proposed by Choutas *et al*. [6]. The measurements are computed on the SMPL [29] rest pose mesh, which is generated using predicted shape parameters and rest pose angle parameters. The 3D shape metrics are calculated as the absolute error between predicted and ground truth anatomical values.

We employ the below metrics to evaluate the 3D pressure map predictions:

- Vertex-to-Vertex Pressure (v2vP) Error: This metric, as proposed by Clever *et al*. [9], assesses the alignment between the predicted 3D pressure map and the ground truth data at the vertex level.
- v2vP 1EA (v2vP one edge away) & v2vP 2EA: The dense structure of the SMPL mesh [29] results in the 3D pressure map predicted at the vertex level to have a very fine granularity. To account for the region size over which pressure is applied, we form and evaluate coarse representations of the 3D pressure map. The coarse representation is obtained by setting the pressure value at each vertex as the average pressure value of its neighboring vertices. This averaging process can be interpreted as a smoothing operator on the vertex-level 3D pressure maps. For v2vP 1EA, we consider the vertices that are one edge away to

calculate the averages. For v2vP 2EA, we include the vertices that are both one and two edges away. The smoothing operator region is illustrated in the appendix Fig. 8. After forming these coarser pressure maps for both the ground truth and predicted 3D pressure maps, we apply the v2vP metric to calculate the mean squared error between predicted and ground truth 3D pressure maps, resulting in v2vP 1EA and v2vP 2EA.

In addition, we assess pressure prediction error on a per-body part basis. To do so, we identify 14 distinct body parts, including left & right heels, toes, elbows, shoulders, hips and spine, head, sacrum and ischium. We compute the v2vP metric on only the vertices belonging to a body part.

We compare our methods, BodyMAP and BodyMAP-WS, against the multimodal Pyramid Fusion model from Yin *et al*. [45] and BPB, BPW models from Clever *et al*. [9]. We compute the above evaluation metrics for the BPW model, for which the model weights are made available by the authors. We report the metrics for the other methods when available in their respective papers.

Furthermore, we conduct a custom training of models on only lateral poses (left & right lateral), and evaluate on supine poses. We compare two BodyMAP-WS models, one trained on only lateral poses and the other trained on both lateral and supine poses, with BodyMAP and BPW [9] trained on only lateral poses. For BodyMAP-WS models, we freeze the mesh model once it is pre-trained on lateral poses. This evaluation allows us to test on an out-of-training-distribution setting and gauge the improvements from training BodyMAP-WS implicitly on the entire data.

Additionally, we compare our methods to a custom baseline, BodyMAP-Conv, that is similarly capable of jointly predicting human body mesh and a 3D applied pressure map. BodyMAP-Conv, illustrated in the appendix Fig. 6, replaces the FIM and PointNet components of BodyMAP with fully connected layers to predict 3D pressure map $\hat{P}$.

**Hyperparameters**: We train our networks, BodyMAP and BodyMAP-Conv, with the loss function defined in Eq. (1) and $\lambda_1 = 0.25$, $\lambda_2 = 0.1$, and $\lambda_3 = 0.1$. We train BodyMAP-WS with loss as defined in Eq. (2) and $\lambda_1 = 500$. We train all networks with a batch size of 64, learning rate of 0.0001, weight decay of 0.0005, and the Adam optimizer [21] for gradient optimization and use random rotation and random erasing image augmentations. We train BodyMAP and BodyMAP-Conv for 100 epochs, while we train BodyMAP-WS for 15 epochs.

## 5. Results and Discussion

Tab. 1 presents the results of BodyMAP compared to all baseline methods across various metrics for 3D pose, 3D shape, and 3D pressure distribution, evaluated on 22 test subjects from the real SLP dataset [25, 27]. The averages are computed over all blanket configurations present in the SLP dataset [25, 27].

BodyMAP-PointNet, our model utilizing PointNet [35] for 3D pressure map prediction, achieves notable improvements over the prior state-of-the-art BPW [9]. Specifically, when trained solely on the depth modality akin to the training approach of BPW [9], BodyMAP-PointNet achieves a 12.5% and 13% enhancement for body mesh and 3D pressure map prediction tasks, respectively. When trained on multiple modalities, the BodyMAP-PointNet model attains a substantial 25% improvement for both tasks over BPW [9]. Visual comparisons between BodyMAP-PointNet and BPW [9] are illustrated in Fig. 4, and in the appendix in Fig. 10, and Fig. 11.

Remarkably, our BodyMAP-WS model demonstrates an 8.8% performance improvement over BPW [9] for 3D pressure map prediction, despite lacking access to ground truth 3D pressure maps during training. Furthermore, our baseline model, BodyMAP-Conv, surpasses BPW [9] in direct comparison, trained solely on the depth modality. We hypothesize that errors introduced at each stage in BPW [9] accumulate, leading to adverse impacts on their final 3D pressure map predictions.

Detailed evaluations over specific blanket configurations are presented in the appendix Tab. 4.

**Elevated precision in 3D pressure map prediction across critical body regions**: Specific anatomical regions, such as the back of the head, heels, sacrum, spine, ischium, and hips, are known to be particularly susceptible to pressure ulcers [30]. Evaluating model performance across multiple body regions, revealed substantial improvements over BPW [9]. The results are summarized in Tab. 2.

BodyMAP-WS utilizes a regularization loss that discourages the model from predicting pressure on body vertices not in contact with the mattress (Eq. (2)). As a possible consequence, the model exhibits slightly lower performance than BPW [9] in areas with smaller contact surfaces, notably the heels and toes. However, in regions like the hips and spine, BodyMAP-WS surpasses BPW [9].

BodyMAP-PointNet achieves the lowest pressure error across all body parts. These advancements bear significant implications for aiding caregivers in localizing peak pressure on critical body regions, potentially enhancing the effectiveness of pressure injury prevention protocols.

**Implicit learning of 3D applied pressure map eliminates the need for ground truth annotations**: BodyMAP-WS, is trained through alignment of the 2D projection of the predicted 3D pressure map with the input pressure image. As outlined in Tab. 3, when trained on lateral poses, and evaluated on supine poses, BodyMAP-WS outperforms the prior state-of-the-art BPW method [9]. Moreover, as presented in Tab. 3, extending the training of BodyMAP-WS to the entire dataset (comprising both supine and lat-


Figure 4. Results of inferring body mesh and 3D applied pressure map for examples from the SLP [25, 27] test set.

| | | 3D Pose Error (mm) ↓ | | 3D Shape Error (cm) ↓ | | | | 3D Pressure Error (kPa$^2$) ↓ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Network | Modalities | MPJPE | PVE | Height | Chest | Waist | Hips | v2vP | v2vP 1EA | v2vP 2EA |
| Pyramid Fusion [45] | RGB-D-PI-IR | 79.64 | - | - | - | - | - | - | - | - |
| BPBnet [9] | D | 74.54 | - | - | - | - | - | 2.87 | - | - |
| BPWnet [9] | D | 69.36 | 84.72 | 3.85 | 3.78 | 4.70 | 3.21 | 2.84 | 1.79 | 1.22 |
| BodyMAP - Conv | PI | 76.38 ± 2.060 | 93.24 ± 2.015 | 4.36 ± 0.125 | **2.86 ± 0.147** | 3.75 ± 0.078 | 3.19 ± 0.254 | 2.58 ± 0.009 | 1.61 ± 0.006 | 1.09 ± 0.005 |
| BodyMAP - PointNet | PI | 76.54 ± 3.170 | 93.6 ± 4.029 | 4.83 ± 0.164 | 3.06 ± 0.323 | 4.1 ± 0.433 | 3.56 ± 0.409 | 2.29 ± 0.014 | 1.39 ± 0.011 | 0.92 ± 0.007 |
| BodyMAP - Conv | D | 61.29 ± 1.978 | 72.97 ± 1.797 | 3.98 ± 0.144 | 3.37 ± 0.159 | 4.39 ± 0.212 | 3.62 ± 0.382 | 2.66 ± 0.018 | 1.68 ± 0.017 | 1.15 ± 0.014 |
| BodyMAP - PointNet | D | 60.66 ± 1.615 | 72.47 ± 1.687 | 3.99 ± 0.569 | 3.2 ± 0.241 | 4.3 ± 0.407 | 3.51 ± 0.229 | 2.47 ± 0.034 | 1.54 ± 0.03 | 1.04 ± 0.026 |
| BodyMAP - Conv | D - PI | 51.79 ± 0.379 | 62.9 ± 0.32 | 3.61 ± 0.251 | 3.09 ± 0.256 | 4.1 ± 0.407 | 3.46 ± 0.484 | 2.56 ± 0.007 | 1.59 ± 0.007 | 1.07 ± 0.008 |
| BodyMAP - PointNet | D - PI | **51.01 ± 1.071** | **61.66 ± 0.919** | **3.49 ± 0.459** | 2.93 ± 0.215 | **3.7 ± 0.071** | **2.8 ± 0.360** | **2.14 ± 0.030** | **1.27 ± 0.025** | **0.83 ± 0.020** |

Table 1. Results for 3D pose, 3D shape, and 3D pressure error metrics evaluated for the 22 test subjects in SLP dataset [25, 27], averaging over all blanket conditions (uncovered, cover1, and cover2). The modalities described are depth (D), pressure image (PI), infrared (IR) and RGB. The metrics for BodyMAP models are shown with the mean and standard deviation computed over 3 random runs.

| | | 3D Pressure Error v2vP (kPa$^2$) ↓ | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Network | Modalities | Left Heel | Right Heel | Left Toes | Right Toes | Left Elbow | Right Elbow | Left Shoulder | Right Shoulder | Spine | Head | Left Hip | Right Hip | Sacrum | Ischium |
| BPW [9] | D | 4.13 | 3.75 | 0.41 | 0.36 | 1.55 | 1.41 | 1.83 | 1.65 | 1.76 | 13.06 | 15.83 | 18.29 | 10.60 | 6.11 |
| BodyMAP-Conv | D | 5.23±0.007 | 4.46±0.014 | **0.32±0.001** | 0.31±0.001 | 1.05±0.005 | 0.77±0.019 | 1.49±0.007 | 1.32±0.021 | 1.63±0.020 | 11.35±0.108 | 14.28±0.136 | 16.65±0.145 | 9.12±0.152 | 5.52±0.064 |
| BodyMAP-WS | D - PI | 4.26±0.164 | 3.74±0.009 | 0.42±0.022 | 0.36±0.02 | 1.46±0.015 | 1.1±0.063 | 1.71±0.075 | 1.6±0.037 | 1.67±0.017 | 11.4±0.361 | 13.13±0.251 | 17.72±0.919 | 7.6±0.252 | 5.15±0.057 |
| BodyMAP-Conv | D - PI | 5.22±0.015 | 4.44±0.009 | **0.32±0.001** | 0.32±0.001 | 1.05±0.021 | 0.77±0.007 | 1.41±0.047 | 1.25±0.043 | 1.54±0.026 | 11.22±0.199 | 13.25±0.07 | 15.76±0.026 | 8.6±0.16 | 5.24±0.116 |
| BodyMAP-PointNet | D - PI | **3.81±0.088** | **3.46±0.016** | 0.33±0.005 | **0.31±0.002** | **0.82±0.009** | **0.61±0.019** | **1.25±0.053** | **1.16±0.019** | **1.52±0.007** | **10.22±0.269** | **10.55±0.233** | **13.44±0.205** | **7.17±0.187** | **4.49±0.180** |

Table 2. Results of 3D pressure map error evaluated over multiple body regions for the 22 test subjects in SLP dataset [25, 27], averaging over all blanket conditions (uncovered, cover1, and cover2). Our method achieves more than 30% improvement for body regions vulnerable to pressure injuries such as left hip & sacrum regions. The metrics are shown with the mean and std. dev. computed over 3 random runs.

eral poses), while continuing to use the mesh model pre-trained on only lateral poses, leads to further performance gains, even surpassing our proposed BodyMAP-PointNet model. This methodology to train models without relying on ground truth labels has the potential to streamline cost-intensive annotation efforts, potentially unlocking training these models on unlabeled data from deployed sensors.

**Multimodal input elevates performance**: The results, as presented in Tab. 1 for BodyMAP-PointNet and BodyMAP-Conv with various modality configurations, underscore the significance of combining pressure and depth modalities for optimal results. Notably, there is a significant reduction in body mesh prediction performance when exclusively using the pressure modality. This decline is attributed to the absence of visual cues from pressure images when limbs are not in contact with the mattress. Conversely, we observe the pressure modality to be more effective for 3D pressure map prediction when compared to the depth modality. This substantiates a preference to leverage both the depth (top-down view) and 2D pressure mattress images (bottom-up view), providing greater context for the model to jointly predict body mesh and 3D applied pressure maps.

Moreover, as illustrated in the appendix Tab. 4, integrating multiple modalities leads to consistent mesh prediction performance across all blanket configurations. This stands in contrast to BPW [9], which experiences a significant drop when individuals in bed are covered with blankets due to its training on only a depth modality.

**Feature Indexing Module improves 3D pressure performance**: The proposed Feature Indexing Module (FIM) creates a comprehensive feature set for each vertex by fusing vertex locations, ResNet features (formed before the global average pool), and input image features. This provides the model with rich contextual information, enhancing the 3D applied pressure map prediction. Notably, fusing global ResNet features (formed after the global average pool) with PointNet encoder features also yields an improvement in predicting 3D pressure maps. We present the ablations of FIM in the appendix Tab. 5.

| Network | Train Data | 3D Pose Error MPJPE (mm) ↓ | 3D Pressure Error v2vP (kPa$^2$) ↓ |
|---|---|---|---|
| BPW [9] | Lateral poses | 122.8 | 3.164 |
| BodyMAP-PointNet | Lateral poses | 100.96 | 2.614 |
| BodyMAP-WS | Lateral (mesh) + Lateral (pressure) | **98.82** | 2.865 |
| BodyMAP-WS | Lateral (mesh) + All (pressure) | **98.82** | **2.513** |

Table 3. Results from training on lateral poses and evaluating on supine poses in the test subjects in the SLP dataset [25, 27], averaging over all blanket conditions (uncovered, cover1 and cover2). Here, BodyMAP-WS models use a mesh model pre-trained solely on lateral poses. All models are trained on only depth modality.

## 6. Conclusion

Our proposed method, BodyMAP, jointly predicts the body mesh and 3D applied pressure map for people in bed. These predictions offer interpretable visualizations of pressure on the body, which could aid caregivers in accurately identifying high-pressure regions and potentially enhance the prevention of pressure injuries. Furthermore, we introduced BodyMAP-WS to implicitly learn a 3D pressure map on the body without relying on labeled ground truth data, opening new avenues for learning on large real-world datasets where obtaining 3D labels is both costly and challenging. We have demonstrated the performance of our methods in a series of ablations and comparisons to prior state-of-the-art models.

## References


[1] Ankur Agarwal and Bill Triggs. Recovering 3d human pose from monocular images. *IEEE transactions on pattern analysis and machine intelligence*, 28(1):44–58, 2005. 3

[2] Robert Behrendt, Amir M Ghaznavi, Meredith Mahan, Susan Craft, and Aamir Siddiqui. Continuous bedside pressure mapping and rates of hospital-associated pressure ulcers in a medical intensive care unit. *American Journal of Critical Care*, 23(2):127–133, 2014. 1

[3] Dan Berlowitz, C VanDeusen Lukas, V Parker, A Niederhauser, J Silver, C Logan, and E Ayello. Preventing pressure ulcers in hospitals: a toolkit for improving quality of care. *Agency for Healthcare Research and Quality*, 2011. 1

[4] Joyce M Black, Janet E Cuddigan, Maralyn A Walko, L Alan Didier, Maria J Lander, and Maureen R Kelpe. Medical device related pressure ulcers in hospitalized patients. *International wound journal*, 7(5):358–365, 2010. 1

[5] Ching-Hang Chen, Ambrish Tyagi, Amit Agrawal, Dylan Drover, Rohith Mv, Stefan Stojanov, and James M Rehg. Unsupervised 3d pose estimation with geometric self-supervision. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5714–5724, 2019. 3

[6] Vasileios Choutas, Lea Müller, Chun-Hao P Huang, Siyu Tang, Dimitrios Tzionas, and Michael J Black. Accurate 3d body shape regression using metric and semantic attributes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2718–2728, 2022. 5

[7] Henry M Clever, Ariel Kapusta, Daehyung Park, Zackory Erickson, Yash Chitalia, and Charles C Kemp. 3d human pose estimation on a configurable bed from a pressure image. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 54–61. IEEE, 2018. 2

[8] Henry M Clever, Zackory Erickson, Ariel Kapusta, Greg Turk, Karen Liu, and Charles C Kemp. Bodies at rest: 3d human pose and shape estimation from a pressure image using synthetic data. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 6215–6224, 2020. 3

[9] Henry M Clever, Patrick L Grady, Greg Turk, and Charles C Kemp. Bodypressure-inferring body pose and contact pressure from a depth image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1):137–153, 2022. 2, 3, 4, 5, 6, 8, 11, 12, 13, 14, 15

[10] Mihai Fieraru, Mihai Zanfir, Elisabeta Oneata, Alin-Ionut Popa, Vlad Olaru, and Cristian Sminchisescu. Learning complex 3d human self-contact. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 1343–1351, 2021. 15

[11] Gheorghita Ghinea, Fotis Spyridonis, Tacha Serif, and Andrew O Frank. 3-d pain drawings—mobile data collection using a pda. *IEEE Transactions on information technology in biomedicine*, 12(1):27–33, 2008. 3

[12] Xuan Gong, Meng Zheng, Benjamin Planche, Srikrishna Karanam, Terrence Chen, David Doermann, and Ziyan Wu. Self-supervised human mesh recovery with cross-representation alignment. In *European Conference on Computer Vision*, pages 212–230. Springer, 2022. 3

[13] Lena Gunningberg, Ulrika Pöder, and Cheryl Carli. Facilitating student nurses' learning by real time feedback of positioning to avoid pressure ulcers: evaluation of clinical simulation. *J Nurs Educ Practice*, 6(1):1–8, 2016. 1

[14] Mohamed Hassan, Vasileios Choutas, Dimitrios Tzionas, and Michael J Black. Resolving 3d human pose ambiguities with 3d scene constraints. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 2282–2292, 2019. 15

[15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 3, 11, 14

[16] Thomas Holleczek, Alex Rüegg, Holger Harms, and Gerhard Tröster. Textile pressure sensors for sports applications. In *SENSORS, 2010 IEEE*, pages 732–737, 2010. 2

[17] Lisa Hultin, Estrid Olsson, Cheryl Carli, and Lena Gunningberg. Pressure mapping in elderly care. *Journal of Wound, Ostomy and Continence Nursing*, 44(2):142–147, 2017. 1

[18] Debra Jackson, Ahmed M Sarki, Ria Betteridge, and Joanne Brooke. Medical device-related pressure ulcers: a systematic review and meta-analysis. *International journal of nursing studies*, 92:109–120, 2019. 1

[19] Angjoo Kanazawa, Michael J Black, David W Jacobs, and Jitendra Malik. End-to-end recovery of human shape and pose. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7122–7131, 2018. 4, 11

[20] Susan A Kayser, Catherine A VanGilder, Elizabeth A Ayello, and Charlie Lachenbruch. Prevalence and analysis of medical device-related pressure injuries: results from the international pressure ulcer prevalence survey. *Advances in skin & wound care*, 31(6):276, 2018. 1

[21] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 6

[22] Muhammed Kocabas, Salih Karagoz, and Emre Akbas. Self-supervised learning of 3d human pose using multi-view geometry. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1077–1086, 2019. 3

[23] Nikos Kolotouros, Georgios Pavlakos, Michael J Black, and Kostas Daniilidis. Learning to reconstruct 3d human pose and shape via model-fitting in the loop. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 2252–2261, 2019. 3

[24] Kevin Lin, Lijuan Wang, and Zicheng Liu. End-to-end human pose and mesh reconstruction with transformers. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1954–1963, 2021. 3

[25] Shuangjun Liu and Sarah Ostadabbas. Seeing under the cover: A physics guided learning approach for in-bed pose estimation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 236–245. Springer, 2019. 2, 3, 5, 6, 7, 8, 13, 14, 16, 17

[26] Shuangjun Liu and Sarah Ostadabbas. Pressure eye: In-bed contact pressure estimation via contact-less imaging. *Medical Image Analysis*, 87:102835, 2023. 3

[27] Shuangjun Liu, Xiaofei Huang, Nihang Fu, Cheng Li, Zhongnan Su, and Sarah Ostadabbas. Simultaneously-collected multimodal lying pose dataset: Enabling in-bed human pose monitoring. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1):1106–1118, 2022. 2, 3, 5, 6, 7, 8, 13, 14, 16, 17

[28] Wu Liu, Qian Bao, Yu Sun, and Tao Mei. Recent advances of monocular 2d and 3d human pose estimation: A deep learning perspective. *ACM Computing Surveys*, 55(4):1–41, 2022. 3

[29] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J Black. Smpl: A skinned multi-person linear model. In *Seminal Graphics Papers: Pushing the Boundaries, Volume 2*, pages 851–866. 2023. 2, 3, 4, 5, 11, 12, 14

[30] Sam Mansfield, Katia Obraczka, and Shuvo Roy. Pressure injury prevention: A survey. *IEEE Reviews in Biomedical Engineering*, 13:352–368, 2020. 1, 6

[31] Danielle M Minteer, Patsy Simon, Donald P Taylor, Wenyan Jia, Yuecheng Li, Mingui Sun, and J Peter Rubin. Pressure ulcer monitoring platform—a prospective, human subject clinical study to validate patient repositioning monitoring device to prevent pressure ulcers. *Advances in wound care*, 9(1):28–33, 2020. 1

[32] Lea Muller, Ahmed AA Osman, Siyu Tang, Chun-Hao P Huang, and Michael J Black. On self-contact and human pose. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9990–9999, 2021. 15

[33] Kaichiro Nishi and Jun Miura. Generation of human depth images with body part labels for complex human pose recognition. *Pattern Recognition*, 71:402–413, 2017. 2

[34] Georgios Pavlakos, Vasileios Choutas, Nima Ghorbani, Timo Bolkart, Ahmed AA Osman, Dimitrios Tzionas, and Michael J Black. Expressive body capture: 3d hands, face, and body from a single image. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10975–10985, 2019. 3

[35] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 652–660, 2017. 3, 4, 5, 6, 11, 14

[36] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *Advances in neural information processing systems*, 30, 2017. 4

[37] Ronald G Scott and Kristen M Thurman. Visual feedback of continuous bedside pressure mapping to optimize effective patient repositioning. *Advances in wound care*, 3(5):376–382, 2014. 1

[38] Devdip Sen, John McNeill, Yitzhak Mendelson, Raymond Dunn, and Kelli Hickle. A new vision for preventing pressure ulcers: wearable wireless devices could help solve a common-and serious-problem. *IEEE pulse*, 9(6):28–31, 2018. 1

[39] Fotios Spyridonis and Gheorghita Ghinea. 3-d pain drawings and seating pressure maps: Relationships and challenges. *IEEE Transactions on Information Technology in Biomedicine*, 15(3):409–415, 2011. 3

[40] Yating Tian, Hongwen Zhang, Yebin Liu, and Limin Wang. Recovering 3d human mesh from monocular images: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023. 3

[41] Alexander Toshev and Christian Szegedy. Deeppose: Human pose estimation via deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1653–1660, 2014. 3

[42] Gurjot S Walia, Alison L Wong, Andrea Y Lo, Gina A Mackert, Hannah M Carl, Rachel A Pedreira, Ricardo Bello, Carla S Aquino, William V Padula, and Justin M Sacks. Efficacy of monitoring devices in support of prevention of pressure injuries: systematic review and meta-analysis. *Advances in skin & wound care*, 29(12):567–574, 2016. 1

[43] Keze Wang, Liang Lin, Chenhan Jiang, Chen Qian, and Pengxu Wei. 3d human pose machines with self-supervised learning. *IEEE transactions on pattern analysis and machine intelligence*, 42(5):1069–1082, 2019. 3

[44] Yufei Wang, David Held, and Zackory Erickson. Visual haptic reasoning: Estimating contact forces by observing deformable object interactions. *IEEE Robotics and Automation Letters*, 7(4):11426–11433, 2022. 2, 3

[45] Yu Yin, Joseph P Robinson, and Yun Fu. Multimodal in-bed pose and shape estimation under the blankets. In *Proceedings of the 30th ACM International Conference on Multimedia*, pages 2411–2419, 2022. 3, 5, 6, 8, 13