

Article

Posture Monitoring for Health Care of Bedridden Elderly Patients Using 3D Human Skeleton Analysis via Machine Learning Approach

Jui-Chiu Chiang^{1,2,3}, Wen-Nung Lie^{1,2,3,*} , Hsiu-Chen Huang^{4,*}, Kuan-Ting Chen¹, Jhih-Yuan Liang¹, Yu-Chia Lo¹ and Wei-Hao Huang¹

¹ Department of Electrical Engineering, National Chung Cheng University (CCU), Chiayi 62102, Taiwan; rachel@ccu.edu.tw (J.-C.C.); eric129979@alum.ccu.edu.tw (K.-T.C.); 607415180@alum.ccu.edu.tw (J.-Y.L.); andy955302@alum.ccu.edu.tw (Y.-C.L.); 608415120@alum.ccu.edu.tw (W.-H.H.)

² Center for Innovative Research on Aging Society (CIRAS), National Chung Cheng University (CCU), Chiayi 62102, Taiwan

³ Advanced Institute of Manufacturing with High-Tech Innovations (AIM-HI), National Chung Cheng University (CCU), Chiayi 62102, Taiwan

⁴ Rehabilitation Department, Ditmanson Medical Foundation Chiayi Christian Hospital, Chiayi 60002, Taiwan

* Correspondence: ieewnl@ccu.edu.tw (W.-N.L.); 03281@cych.org.tw (H.-C.H.)

Abstract: For bedridden elderly people, pressure ulcer is the most common and serious complication and could be prevented by regular repositioning. However, due to a shortage of long-term care workers, repositioning might not be implemented as often as required. Posture monitoring by using modern health/medical caring technology can potentially solve this problem. We propose a RGB-D camera system to recognize the posture of the bedridden elderly patients based on the analysis of 3D human skeleton which consists of articulated joints. Since practically most bedridden patients were covered with a blanket, only four 3D joints were used in our system. After the recognition of the posture, a warning message will be sent to the caregiver for assistance if the patient stays in the same posture for more than a predetermined period (e.g., two hours). Experimental results indicate that our proposed method is capable of achieving a high accuracy in posture recognition (above 95%). To the best of our knowledge, this application of using human skeleton analysis for patient care is novel. The proposed scheme is promising for clinical applications and will undertake an intensive test in health care facilities in the near future after redesigning a proper RGB-D (Red-Green-Blue-Depth) camera system. In addition, a desktop computer can be used for multi-point monitoring to reduce cost, since real-time processing is not required in this application.

Keywords: 3D human skeleton; posture recognition; patient care; pressure ulcer prevention



Citation: Chiang, J.-C.; Lie, W.-N.; Huang, H.-C.; Chen, K.-T.; Liang, J.-Y.; Lo, Y.-C.; Huang, W.-H. Posture Monitoring for Health Care of Bedridden Elderly Patients Using 3D Human Skeleton Analysis via Machine Learning Approach. *Appl. Sci.* **2022**, *12*, 3087. <https://doi.org/10.3390/app12063087>

Academic Editor: Simone Morais

Received: 19 January 2022

Accepted: 15 March 2022

Published: 17 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With a growing number of elderly people with disability and chronic diseases [1], long-term care of bedridden elderly people remains a big challenge due to various complications, such as muscular wasting, joint stiffness, nutritional imbalance, and skin problems [2]. One of the most serious complications is pressure ulcers, which are injuries to skin and underlying tissue resulting from prolonged pressure on the skin [3]. The care burden of pressure ulcers is tremendous, and, thus, prevention of pressure ulcers is the most important measure [4]. A systematic review [5] discussed many methods of pressure ulcer prevention and suggested that a non-contact monitoring method is preferred considering the comfort of the patient.

To reduce the occurrence of pressure ulcers, frequent redistribution of the pressure on bedridden patients by regular repositioning is essential [6]. It is commonly recommended to set a repositioning schedule for changing the position of patients every two hours [7]. However, due to a shortage of long-term care workers, repositioning might not

be implemented as often as required. Posture monitoring by using modern health/medical caring technology can potentially solve this problem. In this paper, we present a novel approach based on a remote RGB-D camera to continuously detect and monitor the posture of bedridden elders. The main goal is to notify the caregivers to reposition the patients if they have been staying in the same posture for a long time. In this way, we achieve the following goals: (1) evaluation of the care quality (enough release of pressure) and (2) long-term record of patient's posture variations.

In our system design, a RGB-D (depth) camera is used to overcome the challenge of posture recognition using incomplete skeletons, since elderly people in bed are usually covered with a blanket and only the head and partial shoulders are visible. Recognition of lying postures (left-side, right-side, and supine) solely from traditional RGB information might be difficult and inaccurate. The adoption of depth information from the RGB-D camera will be helpful in determining the 3D skeleton and, hence, the lying postures.

2. Related Works

Conventional research of in-bed behavior analysis includes body movement during sleep, sleep apnea, and sleeping postures. For sleeping posture recognition or behavior analysis, wearable devices [8,9] were attached to the human body or a mattress. However, they often make the patients uncomfortable and affect sleeping quality. Another way is to install pressure sensors on the mattress for sleep monitoring [10] (such as detecting events of normal breathing, apnea, and body motion).

To ensure less inconvenience or interference, image-based approaches were often developed. The RGB, depth [11], and near- or far-infrared (NIR or FIR) images [12] can be used for different purposes. Among them, thermo imaging [13] is free of privacy issues but often expensive. Depth cameras based on ToF (time of flight) principle also suffer from higher cost and cluttered noise from environmental lighting. In a compromise between the performance, cost, and privacy for elderly patients, we choose to use a RGB-D camera which is capable of capturing color and depth information (e.g., the Microsoft Kinect, Intel RealSense, etc.).

Recently, machine learning (ML) approaches for disease diagnosis (e.g., Parkinson's [14] and ovarian cancers [15]) have gained increasing attention. They often used hand-crafted features sophisticatedly defined by a system designer or deep-learning (DL) features automatically trained from the dataset. For health care purposes, electronic health records (EHR) are also used to predict the risk of fall for the elderly using a ML approach [16].

In sleep monitoring, Chang et al. [17] used a Kinect depth sensor to detect and recognize patient movements, motion patterns, and pose positions by gross body motion estimation [18] and in-bed pose analysis. Spatial and temporal features from depth image sequence were aggregated for analysis on a condition of no blanket covering on the body, which seems unreasonable for immobile and elderly patients. Grimm et al. [19] also used Kinect V2 depth image to form a bed-aligned map (BAM) [20] (each cell covering 10 cm × 10 cm) and realized sleeping position classification through convolutional neural networks (CNNs). Li et al. [21] proposed to build a vertical distance map from the Kinect depth image to classify among ten sleeping postures, where a multi-stream CNN architecture considering different resolutions was used. In [21,22], though the subjects covered with blankets were also considered, their goal was sleep quality assessment.

In contrast to image-based methods, some other studies used skeleton-based algorithms for posture/action recognition. Lee et al. [23] proposed to use 3D skeleton from Kinect V2 for sleep pattern monitoring. However, the subjects for the test were not covered with blankets, so that 3D skeletons could be constructed from the depth images. Recently, deep-learning algorithms were developed to estimate human skeleton (a set of articulated 3D joints) from a single RGB image [24–27] to avoid the influence of inaccurate depths and for outdoor use. This kind of 2D skeleton and 3D skeleton was used in [28,29] for fall detection and gesture recognition, respectively.

In view of the prior works, Table 1 compares research on different goals for posture classification to better reflect the goal and condition considered in the proposed work.

Table 1. A comparison between sleep quality assessment and our pressure ulcer prevention.

Goals	Subjects	Blanket Covering	Sleeping Poses
Sleep quality assessment	Healthy	Might be	Several and complicated (like fetus/log/yearner positions in [21,27])
Pressure ulcer prevention	Usually immobile	Usually yes, in air-conditioned environment	Usually 3 (supine, left-, right-oriented) which are hardly differentiable

A collaboration between National Chung Cheng University (CCU) and Ditmanson Medical Foundation Chiayi Christian Hospital was conducted to use a RGB-D camera for capturing the color plus depth images and carry out recognition into three sleep postures for immobile patients. A warning message will be sent to the caregivers to remind them about repositioning the patients if they stay in unchanged postures for a long time (e.g., two hours) or lose a specified posture due to body motions. By referring to [30,31], three lying postures are defined: supine, left, and right. Left or right side lying is performed by inserting pillows under the right or left side back. Since the three postures after blanket covering tend to be visually similar to the supine position, this fine-grained pose classification problem seems to be challenging. In this paper, we propose a 3D skeleton-based system to classify the three postures considering blanket covering (maybe up to the shoulders), with two legs occluded and two hands possibly visible.

3. Proposed Method

Considering the above two observations (blanket covering and tiny difference between postures), a partial 3D skeleton which contains four joints only (i.e., nose, neck, left shoulder, and right shoulder) is to be estimated, as illustrated in Figure 1. The 3D skeleton is not estimated directly from the depth image. Instead, first a 2D skeleton is estimated from RGB information and then combined with the depth cues to form a 3D skeleton. For posture classification, a machine learning approach that accepts various kinds of information (including 2D skeleton, 3D skeleton, depth information, etc.) as hand-crafted features will be used.

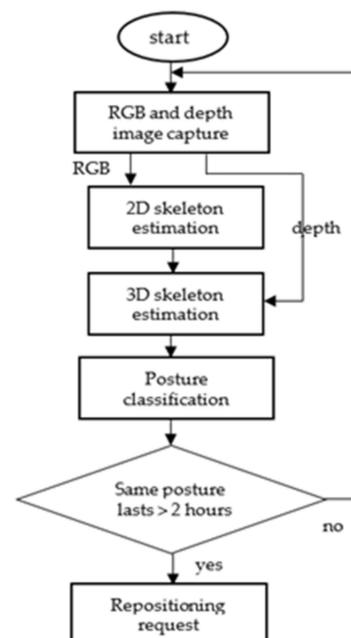


Figure 1. The framework of the proposed scheme.

3.1. System Setup

Figure 2 shows our system setup, where an Intel Realsense RGB-D camera was mounted above the bed at a height of 174 cm with a viewing angle of 10 degrees with respect to the nadir point. Practically, the camera can be mounted on a track above the bed, with a suitable height and tilting angle, so as not to interrupt the working of the caregivers.

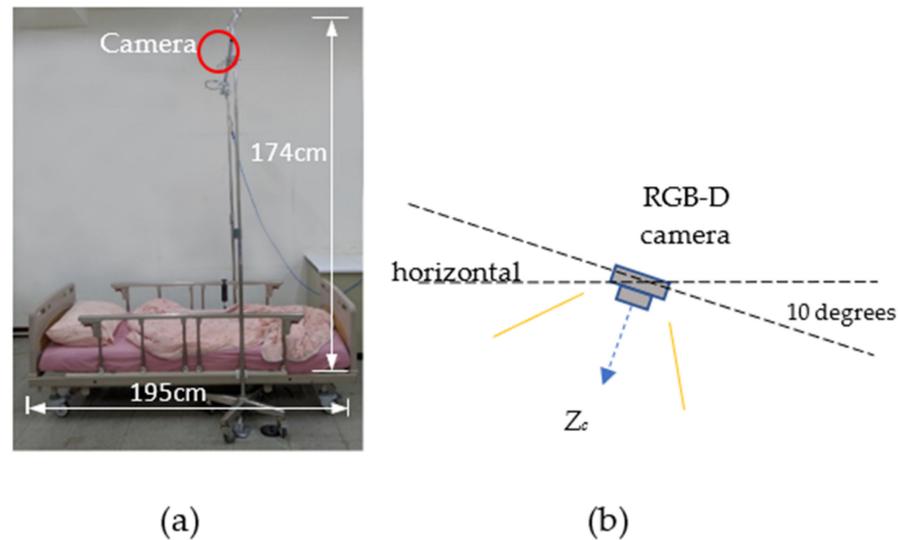


Figure 2. (a) The system setup; (b) the camera tilting angle.

3.2. 3D Skeleton Estimation

Dissimilar from Microsoft Kinect, the 2D skeleton is estimated from the RGB image and then fused with corresponding depth information to form a 3D skeleton. This two-stage method is preferable for RGB-D cameras other than Microsoft Kinect which did not provide a toolbox for skeleton estimation, and, thus, it has wider applications. This was contributed by recent developments on deep-learning neural networks which make the estimation of 2D/3D human skeleton joints from complex backgrounds a reality [32].

We use the OpenPose [25] tool, which is advantageous in speed performance and robustness to joint occlusions. OpenPose is not restricted to the detection of a full skeleton, but all visible or perhaps some occluded joints can be detected. However, only four joints (nose, neck, left shoulder, right shoulder) will be used for posture classification in the proposed system. Figure 3 illustrates for OpenPose (a) the complete skeleton model and (b) an example of 2D skeleton extracted for a sleeping person (with blanket) by using OpenPose.

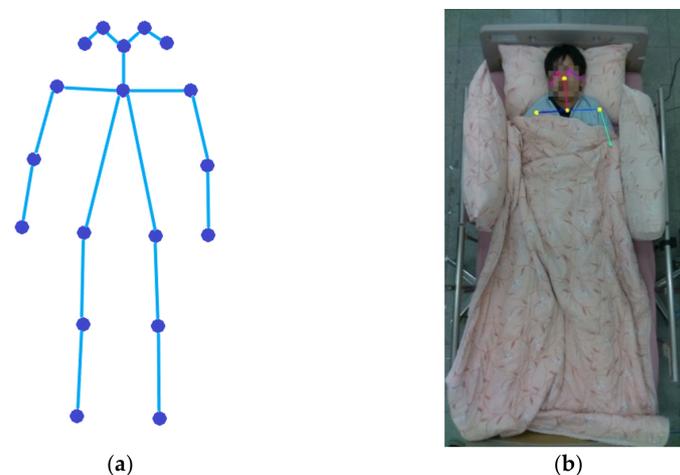


Figure 3. (a) A complete skeleton model used by OpenPose which contains 18 joints. (b) An example of 2D skeleton extracted for sleeping person (with blanket) by using OpenPose.

After aligning the RGB and depth images via parameters provided by the manufacturer, depth information for joints detected by OpenPose can be retrieved to form a 3D skeleton. To stably obtain the depths for the four joints (especially the left and right shoulders), trilateral filtering [33] is performed to the depth values near the detected joints. A trilateral filter consists of three Gaussian kernels as adaptive weights considering depth similarity, luminance similarity, and the distance between the neighboring pixels and the center pixel. It is used to smooth out the variations of depths in a 33×33 window centered at the selected joints. In this way, we obtain the coordinates (u, v) of the 2D joints and their corresponding depths z_c , which is one kind of the feature input used in our proposed classifier and is called “2D skeleton plus depth (u, v, z_c) ”.

With the aid of a calibrated camera, we convert the “2D skeleton plus depth” into a “3D skeleton” via the aid of camera intrinsic parameters, i.e., each joint is represented as the (x_c, y_c, z_c) coordinates in the camera coordinate system with a known (u, v, z_c) .

3.3. Posture Classification

Instead of accepting the RGB/depth images and extracting hand-crafted [17,19,20] or deep-learned [13,21,22] features from them for classification, we adopt a hybrid method for posture classification of bedridden patients. The 3D descriptions (either (x_c, y_c, z_c) or (u, v, z_c)) of the four selected skeleton joints are treated as hand-crafted features for classification. We use Xgboost [34] as our classifier, where the input is a vector containing 12 elements presenting the four joints. The lying postures to be classified include left-side, right-side, and supine.

4. Experimental Results

The Intel Realsense RGB-D D435 camera was chosen for experiments. The same resolution $1280 \times 720 @ 30$ Hz for both the RGB and the depth parts are output for alignment and processing.

4.1. Dataset Collection

Considering the bedridden elderly were weak in health and not suitable for collecting the diverse postures, a total of 20 volunteers (including 15 males and five females) participated in the first-stage experiment for dataset collection and initial training. The volunteers were requested to perform the demanded lying postures, which were supervised by the nursing officer. On-site RGB-D video recording of a limited number of elderly patients was then conducted for cross-testing and classifier refinement.

Since the elderly patients were mobility-compromised and suffered from muscle atrophy, the repositioning for them was realized by placing pillows on specified locations which are very different from those defined in fetus/log/yearner positions in [21,23]. To meet practical situations, we made some arrangements in collecting the dataset which were considered as the changing factors and are listed in Table 2. As a result, 162 samples were recorded (54 samples for each posture with blanket covering) for each volunteer subject. Figure 4 illustrates some captured samples for left-side posture in the dataset.

Table 2. Changing factors in collecting the posture dataset.

Kind of Body Posture	Right-Side, Left-Side, or Supine (3 Cases): Achieved with/without the Pillow Under the Back
Variations of each body posture	<ol style="list-style-type: none"> For left- or right-side posture: low, medium, or high level of body sideward orientation (controlled by pillow number and location) (3 cases) For supine posture: slightly-left, -right, or -centered displacement of the body (3 cases)
Inclined angle of the bedhead	0-degree, 20-degree, and 40-degree (3 cases)
Head orientations	Left-side, right-side, or supine (3 cases)
Bed-side support	Pillow support: upper or lower position (2 cases)



Figure 4. Some samples of left-side lying at different levels of sideward orientation; (a) two pillows under the right back (without covering), (b) two pillows under the right back (with covering), (c) one pillow under the right back (without covering), (d) one pillow under the right back (with covering), (e) the head orientation is not consistent with the body posture, (f) the bedhead is inclined at 40 degrees.

We also collected clinic data of two bedridden elderly patients, including one man and one woman. For the man, his postures include left, right, and supine with 431, 532, and 562 samples, respectively. However, for the woman, only left and right lying with 504 and 552 samples were collected, since the supine posture was not clinically allowed for her.

Both of the experiments for volunteers and clinic patients were conducted by the Institutional Review Board (IRB) of Ditmanson Medical Foundation Chiayi Christian Hospital, Taiwan.

4.2. Experimental Settings

The performance of the proposed technique is evaluated using settings in Table 3. In Settings 1–5, the selected joints are used as input, while in Settings 6–7, the RGB images are used instead for comparison. Settings 1–5 present the ablation study, where clinical data were gradually added to the initial training and testing sets composed of volunteer data only and fully replaced them all. In Setting 1, the training and testing sets include only the volunteers, and the leave-one-person-out strategy [35] is adopted. In Setting 2, the training set includes all the volunteers, while the testing set includes all the clinical data. In Setting 3, in addition to the volunteer samples, partial clinical data were used during training, and the remaining samples were used for testing. More clinical data were used for training in Setting 4. Similarly, in Setting 5, the volunteer samples were excluded for training. For Settings 3–5, equal and random selections were conducted among the three postures.

Table 3. Different Settings 1–7 for training and testing sets in our ablation experiments.

Skeleton-Based Posture Classification		
	Training Data	Testing Data
Setting 1	19 volunteers	One volunteer (leave-one-person-out [35])
Setting 2	20 volunteer (3154 samples) (left, supine, right) = (1053, 1054, 1047)	All clinical data (2581 samples): male (left, supine, right) = (431, 532, 562) female (left, right) = (504, 552)
Setting 3	1. 20 volunteer (3154 samples):(left, supine, right) = (1053, 1054, 1047) 2. Partial clinical data (108 samples):male (left, supine, right) = (18, 36, 18)female (left, right) = (18, 18)	The remaining clinical data (2473 samples): male (left, supine, right) = (413, 496, 544) female (left, right) = (486, 534)
Setting 4	1. 20 volunteer (3154 samples):(left, supine, right) = (1053, 1054, 1047) 2. Partial clinical data (1116 samples):male (left, supine, right) = (186, 372, 186) female (left, right) = (186, 186)	The remaining clinical data (1465 samples): male (left, supine, right) = (245, 160, 376) female (left, right) = (318, 366)

Table 3. Cont.

Skeleton-Based Posture Classification		
	Training Data	Testing Data
Setting 5	Partial clinical data (1116 samples): (left, supine, right) = (372, 372, 372) male (left, supine, right) = (186, 372, 186) female (left, right) = (186, 186)	The remaining clinical data (1465 samples): male (left, supine, right) = (245, 160, 376) female (left, right) = (318, 366)
RGB image-based posture classification		
	Training set	Testing set
Setting 6	19 volunteers (without blanket)	One volunteer (without blanket)
Setting 7	19 volunteers (with blanket)	One volunteer (with blanket)

To verify the effectiveness of the proposed skeleton-based technique, Settings 6 and 7 were considered, and the input for posture classification is the associated RGB images for 20 volunteers. Setting 6 is for volunteers covered without blankets, and Setting 7 is with blankets. Similar to Setting 1, the leave-one-person-out strategy is adopted. Meanwhile, Alexnet [36] was used as the classifier, a deep learning-based architecture consisting of five convolution and two dense layers.

4.3. Results

4.3.1. Comparison with RGB Image-Based Classification

Table 4 summarizes the accuracy ($TP/(TP+FP)$ [35], TP: True Positive, FP: False Positive) and the confusion matrix (ground truth vs. prediction) for Settings 1, 6, and 7. In Setting 1, two kinds of input are considered, including 2D skeleton plus depth and 3D skeleton, while the input is RGB image in Setting 6 and Setting 7. First, it shows that Setting 6 performs better than Setting 7, implying that body occlusion will degrade the classification accuracy. Second, in our proposed technique, even only four joints were used in consideration of blanket covering, the classification accuracies are 95.4% and 95.62%, respectively, for “2D skeleton plus depth” and “3D skeleton” input, which are higher than those in Setting 6 (RGB image-based, without blanket). To further improve the performances, the nose-centered coordinates for joints were used instead of the absolute image or camera coordinates, which indeed shows improvements in accuracy (up to 96.92% for “3D skeleton” in Table 4). This meets our statistics about the nose-distances (with respect to the camera) that a variation between left/right and supine postures for the volunteers (1399.5 mm for left-sided, 1404.2 mm for supine, and 1399.1 mm for right-sided) exists, and removal of this variation by subtracting nose coordinates from joint coordinates (thus, coordinates of the nose joint will become (0,0,0)) is helpful to the classification accuracy. The average nose-distances for the volunteer, male patient, and female patient are also analyzed as: 1400.9 mm, 1530.4 mm, and 1493.4 mm, respectively. There is an even larger variation in nose-distances between different target persons (obviously between the volunteers and the patients, in a range of 93 mm to 130 mm). The sources of variations are mostly due to difference in environment settings (such as the camera position and tilting angle). Although we try to make the environment in the laboratory and the hospital as similar as possible, they are always different in aspects of illumination, the height, and the orientation of the camera. In subsequent experiments, only the results based on nose-centered joint coordinates are shown so as to save space.

4.3.2. Domain Adaptation

As mentioned earlier, we collected data from 20 volunteers and two clinic patients. Although the volunteers imitated the postures of the bedridden patients, there were some inevitable dissimilarities. In this situation, volunteers and real patients belong to two different domains and the domain adaptation performance will be evaluated here. In Setting 2, the training set contains only the volunteers, while the testing set contains all the clinical data. Table 5 shows the results for individual male and female patients, when nose-

centered coordinates of joints are adopted. It reveals that the training set involving only the volunteer samples is not capable of transfer learning when the testing set contains the clinic samples. This poor performance is obvious for the male patient, due to the caregiver’s improper arrangement or small depth difference between left and right shoulders (thus incurring a reverse ordinal relationship between them). Table 6 shows the average depth differences between the left and the right shoulder joints ($diff = depth_R - depth_L$) of three postures for the volunteer, the male patient, and the female patient. For left-sided, the pillow is put under the right side of the back and the right shoulder is closer to the camera, and, hence, the right shoulder depth is smaller than that of the left shoulder. Therefore, the $diff$ value is negative for left-sided and positive for right-sided. Table 6 shows that $diffs$ for the volunteer are -131 and 118 mm for left-sided and right-sided, respectively. However, the $diffs$ for the male patient are -16 and 29 mm for left-sided and right-sided, respectively. The $diffs$ for the male patient are much smaller than those of the volunteers. For the female patient, $diffs$ are -86 and 115 mm for left-sided and right-sided, respectively. The $diffs$ for the female patient are closer to those of the volunteers. This explains why the recognition accuracy of the female patient is better than that of the male patient.

To increase the robustness and the adaptation of the proposed technique, in Setting 3, not only the volunteer samples but also partial clinic samples were used during the training stage. The accuracy and the associated confusion matrix are summarized in Table 7. Compared to Table 5, it indicates that the performance is improved by adding partial clinical samples for training. For the male patient and the female patient, the accuracy is 66.71% and 86.02%, respectively, for “2D skeleton plus depth”. In Setting 4, the training set includes more clinical data, and the performance is listed in Table 8. The accuracy is improved, as expected, since more characteristics of the clinical data are learned during training. It means that the impact due to the gap between the clinical data and the volunteers will be eliminated once the system is trained with more and more clinical data. Specifically, the accuracy for the male patient was boosted from 66.71% to 94.52% for “2D skeleton plus depth”. Besides, the accuracies for the female patient are 97.34% and 96.55% for “2D skeleton plus depth” and “3D skeleton”, respectively. For the input of 2D skeleton plus depth, the accuracy is higher than the highest accuracy (i.e., 96.92%) achievable in Table 4, where only volunteer samples are used. This reflects that the proposed technique is able to overcome the domain-diversities situation.

Table 4. Accuracy and confusion matrix (all numbers in terms of %) for different settings.

Setting 1 (2D Skeleton Plus Depth)				Setting 1 (3D Skeleton)				Setting 6			
Pred.				Pred.				Pred.			
GT	left	supine	right	GT	left	supine	right	GT	left	supine	right
left	95.82	3.99	0.19	left	95.63	4.27	0.09	left	88.52	11.48	0
supine	1.52	94.78	3.70	supine	1.90	94.97	3.13	supine	0.83	90.19	8.98
right	0	4.39	95.61	right	1.0	3.63	96.27	right	0.09	16.67	83.24
Accuracy: 95.40				Accuracy: 95.62				Accuracy: 87.31			
Setting 1 (2D skeleton plus depth, nose-centered joints)				Setting 1 (3D skeleton, nose-centered joints)				Setting 7			
Pred.				Pred.				Pred.			
GT	left	supine	right	GT	left	supine	right	GT	left	supine	right
left	95.73	4.18	0.09	left	96.39	3.51	0.10	left	86.48	13.52	0
supine	2.09	95.63	2.28	supine	1.71	96.30	1.99	supine	5.28	84.44	10.28
right	0.09	2.30	97.61	right	0	1.91	98.09	right	0	21.94	78.06
Accuracy: 96.32				Accuracy: 96.92				Accuracy: 82.99			

Table 5. Accuracy and confusion matrix (all numbers in terms of %) for Setting 2.

Male Patient—Nose-Centered Joints (2D Skeleton Plus Depth)				Male Patient—Nose-Centered Joints (3D Skeleton)			
	Pred.				Pred.		
GT	left	supine	right	GT	left	supine	right
left	29.00	69.61	1.39	left	31.78	64.27	3.95
supine	11.84	75.94	12.22	supine	7.90	78.00	41.10
right	16.73	51.07	32.20	right	16.73	63.88	19.39
Accuracy: 46.55				Accuracy: 43.34			
Female patient—nose-centered joints (2D skeleton plus depth)				Female patient—nose-centered joints (3D skeleton)			
	Pred.				Pred.		
GT	left	supine	right	GT	left	supine	right
left	78.37	12.30	9.33	left	78.37	12.30	9.33
right	17.75	3.99	78.26	right	17.75	3.99	78.26
Accuracy: 78.31				Accuracy: 73.39			

Table 6. The depth difference between the right shoulder and the left shoulder (in mm).

<i>diff</i>	Left-Sided	Supine	Right-Sided
Volunteer	−131.16	−7.21	118.82
Male patient	−16.92	11.134	29.90
Female patient	−86.17	X	115.40

Table 7. Accuracy and confusion matrix (all numbers in terms of %) for Setting 3.

Male Patient—Nose-Centered Joints (2D Skeleton Plus Depth)				Male Patient—Nose-Centered Joints (3D Skeleton)			
	Pred.				Pred.		
GT	left	supine	right	GT	left	supine	right
left	53.51	45.76	0.73	left	57.38	40.44	2.18
supine	9.68	84.27	6.05	supine	9.48	83.67	6.85
right	8.27	24.45	67.28	right	5.70	33.27	61.03
Accuracy: 66.71				Accuracy: 67.05			
Female patient—nose-centered joints (2D skeleton plus depth)				Female patient—nose-centered joints (3D skeleton)			
	Pred.				Pred.		
GT	left	supine	right	GT	left	supine	right
left	88.68	7.20	4.12	left	86.42	7.82	5.76
right	12.17	4.31	83.52	right	13.48	4.31	82.21
Accuracy: 86.02				Accuracy: 82.47			

Table 8. Accuracy and confusion matrix (all numbers in terms of %) for Setting 4.

Male Patient—Nose-Centered Joints (2D Skeleton Plus Depth)				Male Patient—Nose-Centered Joints (3D Skeleton)			
	Pred.				Pred.		
GT	left	supine	right	GT	left	supine	right
left	90.20	9.39	0.41	left	90.20	9.39	0.41
supine	0.063	0.975	0.062	supine	1.25	98.13	0.62
right	0.82	1.64	97.54	right	2.13	6.12	91.75
Accuracy: 94.52				Accuracy: 93.89			

Table 8. Cont.

Female patient—nose-centered joints (2D skeleton plus depth)				Female patient—nose-centered joints (3D skeleton)			
	Pred.				Pred.		
GT	left	supine	right	GT	left	supine	right
left	99.06	0.94	0	left	95.91	2.83	1.26
right	0.82	1.64	97.54	right	0.82	1.91	97.27
	Accuracy: 97.34				Accuracy: 96.55		

Table 9 shows the performance for Setting 5, where both the training and testing data are from the clinic samples. The accuracies for both patients are higher than 96% and competitive to the highest accuracy (i.e., 96.92%) in Table 4. Particularly, the accuracy for the female patient is 99.71% for both “2D skeleton plus depth” and “3D skeleton”.

Table 9. Accuracy and confusion matrix (all numbers in terms of %) for Setting 5.

Male Patient—Nose-Centered Joints (2D Skeleton Plus Depth)				Male Patient—Nose-Centered Joints (3D Skeleton)			
	Pred.				Pred.		
GT	left	supine	right	GT	left	supine	right
left	95.10	4.90	0	left	95.51	4.08	0.41
supine	0.62	98.76	0.62	supine	2.5	97.5	0
right	1.86	0	98.14	right	2.39	1.33	96.28
	Accuracy: 97.31				Accuracy: 96.29		

Female patient—nose-centered joints (2D skeleton plus depth)				Female patient—nose-centered joints (3D skeleton)			
	Pred.				Pred.		
GT	left	supine	right	GT	left	supine	right
left	100	0	0	left	99.37	0.63	0
right	0.27	0.27	99.46	right	0	0	100
	Accuracy: 99.71				Accuracy: 99.71		

4.4. The Performance of 2D Skeleton Estimation

Our proposed technique relies on the success of 2D skeleton extraction and alignment of corresponding joint depths to form 3D skeletons for classification. Though the OpenPose tool is featured for its capability in incomplete skeleton extraction in the cases where some joints are occluded by blankets, some failures are still possible and need to be excluded, for example: (c1) fewer than four specified joints are detected; and (c2) the detected four specified joints do not conform to reasonable bone-length relationships.

For the volunteers, in total 3154 out of 3280 individually captured frames (96.16%) were successful in 3D skeleton extraction, since they follow the guidance of the professional caregivers to perform the demanded sleep postures. By contrast, lying postures of the clinic elderly patients were captured continuously at a 3–6 Hz rate. Table 10 summarizes the collection of the clinic samples in several visits, where only 3401 out of 7548 samples (45.06%) were successfully processed due to long burst failures.

Figure 5 shows some typical failure examples in skeleton generation due to: (a) and (b), the blanket fully covers the shoulders; (c), the scarf has a similar color to that of the pillow; (d) the patient was woken and scratched her face for a while. All of the above situations result in a long burst failure as mentioned earlier. On the contrary, Figure 6 shows examples where 3D skeletons are extracted successfully.

Table 10. Statistics of our clinic video data collection.

	Patient	Time Periods for Video Capturing (Mins)	No. of Captured Frames	No. of Successful 3D Skeleton Extraction	Success Rate (%)
visit 1	male	423	2534	813	32.08
visit 1	female	337	2020	280	13.86
visit 2	female	211	633	493	77.88
visit 3	male	369	1107	661	48.24
visit 4	male	255	763	759	99.48
visit 4	female	164	491	395	80.45
	Total	1759	7548	3401	45.06

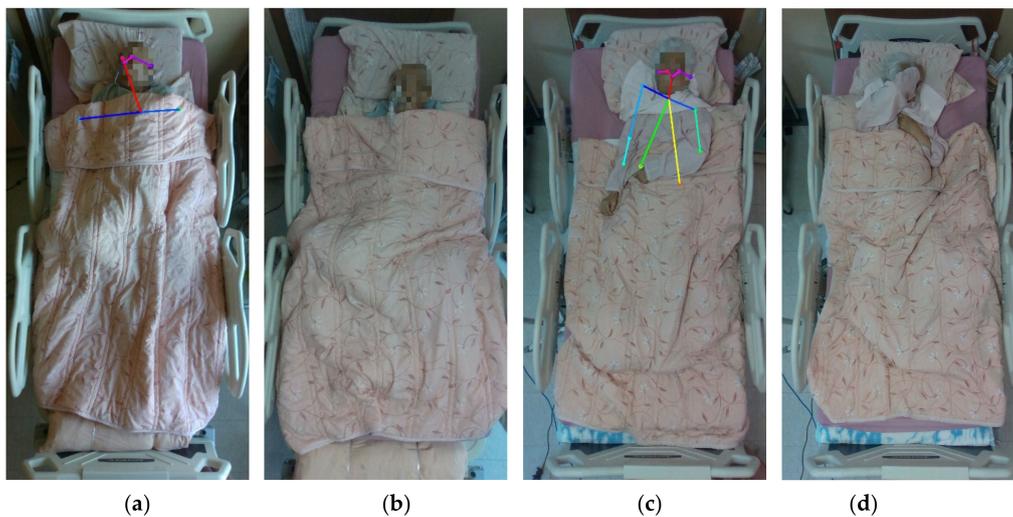


Figure 5. Some failed cases in clinics: incorrect skeleton in (a,c); no skeleton in (b,d).

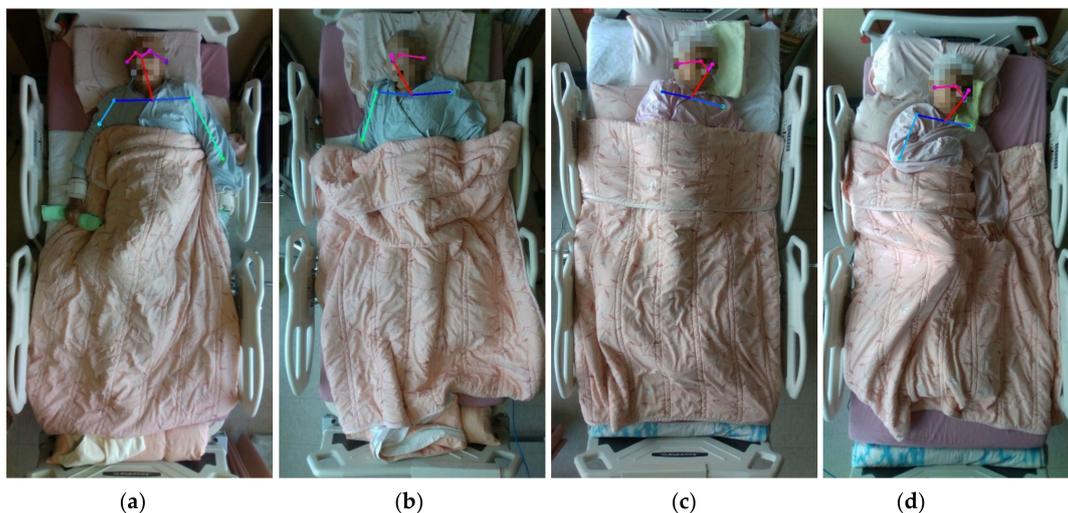


Figure 6. Successful cases in clinics: (a,b) cases 1 and 2 for the male patient, (c,d) cases 1 and 2 for the female patient.

5. Remarks and Conclusions

In this paper, a 3D skeleton-based lying posture classification technique for bedridden patients with blanket covering was proposed and shown to have high recognition rates (from 94.52% to 99.71% for Settings 4 and 5) for applications, outperforming another scheme based on direct feature extraction from RGB images. The blanket covering problem has been overcome successfully. However, there are some interfering factors:

- (1) caregivers' ways of bed sheets covering (e.g., a scarf/towel with a similar color which prevents correct OpenPose operation, nearly full covering of the patients' shoulders by blankets);
- (2) little differences in body orientations of the left-side, right-side, and supine postures.

We still need to avoid them so as to have a robust performance in clinics.

For clinic considerations, the monitoring system can be designed to send an alert for caregivers when any of the following conditions arises:

- (1) The patient's sleep postures have been properly recognized, and the same postures have remained for a long time;
- (2) The patient's 3D skeleton cannot be properly extracted for a long time;
- (3) The patient's recognized posture changes earlier than the schedule time (patients turns by themselves to the posture that was the most comfortable but not allowed yet);
- (4) The patient's recognized postures oscillate between several states (this might be due to patient's body motion or small orientations of the left- or right-side postures).

After designing a RGB-D camera system which was installed with the avoidance of interfering with the caregiver's work, the proposed scheme can be undertaken in an intensive test in hospitals. A desktop computer can also be responsible for multi-point monitoring to reduce the average cost, since real-time processing is not required in this application (less than 1 s per RGB-D frame was spent for OpenPose).

Author Contributions: Conceptualization, writing (original draft), and data curation, J.-C.C., W.-N.L. and H.-C.H.; methodology, J.-C.C. and W.-N.L.; software and validation, K.-T.C., J.-Y.L., Y.-C.L. and W.-H.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan and under a contract number RCN0012 within a collaboration project between National Chung Cheng University and Ditmanson Medical Foundation Chiayi Christian Hospital.

Institutional Review Board Statement: All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Ditmanson Medical Foundation Chiayi Christian Hospital (CYCH-IRB No. 2018082).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study, and written informed consent has been obtained from the patient(s) to publish this paper.

Data Availability Statement: Not applicable.

Acknowledgments: This work was financially supported by the Center for Innovative Research on Aging Society (CIRAS), Advanced Institute of Manufacturing with High-tech Innovations (AIM-HI) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan and under a contract number RCN0012 within a collaboration project between National Chung Cheng University and Ditmanson Medical Foundation Chiayi Christian Hospital (2019–2021).

Conflicts of Interest: Wen-Nung Lie and Jui-Chiu Chiang have received research grants from Ditmanson Medical Foundation Chiayi Christian Hospital.

References

1. Feng, Z.; Glinskaya, E. Aiming Higher: Advancing Public Social Insurance for Longterm Care to Meet the Global Aging Challenge Comment on "Financing Long-term Care: Lessons from Japan". *Int. J. Health Policy Manag.* **2019**, *9*, 356–359. [[CrossRef](#)] [[PubMed](#)]
2. Eckman, K.L. The prevalence of dermal ulcers among persons in the U.S. who have died. *Decubitus* **1989**, *2*, 36–40. [[PubMed](#)]
3. Medeiros, A.B.; Lopes, C.H.; Jorge, M.S. Analysis of prevention and treatment of the pres-sure ulcers proposed by nurses. *Rev. Esc. Enferm. USP* **2009**, *43*, 223–228. [[CrossRef](#)] [[PubMed](#)]
4. Boyko, T.V.; Longaker, M.T.; Yang, G.P. Review of the Current Management of Pressure Ulcers. *Adv. Wound Care* **2018**, *7*, 57–67. [[CrossRef](#)] [[PubMed](#)]
5. Marchione, F.G.; Araújo, L.M.Q.; Araújo, L.V. Approaches that use software to support the prevention of pressure ulcer: A systematic review. *Int. J. Med. Inform.* **2015**, *84*, 725–736. [[CrossRef](#)] [[PubMed](#)]

6. Jocelyn, C.H.S.; Thiara, E.; Lopez, V.; Shorey, S. Turning frequency in adult bedridden patients to present hospital-acquired pressure ulcer: A scoping review. *Int. Wound J.* **2018**, *15*, 225–236. [[CrossRef](#)] [[PubMed](#)]
7. Lyder, C.H.; Ayello, E.A. Chapter 12 Pressure Ulcers: A Patient Safety Issue. In *Patient Safety and Quality: An Evidence-Based Handbook for Nurses*; Hughes, R.G., Ed.; Agency for Healthcare Research and Quality (US): Rockville, MD, USA, 2008.
8. Yoon, H.N.; Hwang, S.; Jung, D.W.; Choi, S.; Joo, K.; Choi, J.; Lee, Y.; Jeong, D.; Park, K.S. Estimation of sleep posture using a patch-type accelerometer based device. In Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society, Milano, Italy, 25–29 August 2015; pp. 4942–4945.
9. Borazio, M.; van Laerhoven, K. Combining wearable and environmental sensing into an un-obtrusive tool for long-term sleep studies. In Proceedings of the ACM SIGHIT International Health Informatics Symposium, Miami, FL, USA, 28–30 January 2012; pp. 71–80.
10. Malakuti, K.; Albu, A.B. Towards an Intelligent Bed Sensor: Non-intrusive Monitoring of Sleep Irregularities with Computer Vision Techniques. In Proceedings of the 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; pp. 4004–4007.
11. Sarsfield, J.; Brown, D.; Sherkat, N.; Langensiepen, C.; Lewis, J.; Taheri, M.; McCollin, C.; Barnett, C.; Selwood, L.; Standen, P.; et al. Clinical assessment of depth sensor based pose estimation algorithms for technology supervised rehabilitation applications. *Int. J. Med. Inform.* **2018**, *121*, 30–38. [[CrossRef](#)] [[PubMed](#)]
12. Faessler, M.; Mueggler, E.; Schwabe, K.; Scaramuzza, D. A monocular pose estimation system based on infrared LEDs. In Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, China, 31 May–7 June 2014; pp. 907–913.
13. Morawski, I.; Lie, W.-N. Two-stream deep learning architecture for action recognition by using extremely low-resolution infrared thermopile arrays. In Proceedings of the International Workshop on Advanced Imaging Technology (IWAIT), Yogyakarta, Indonesia, 5–7 January 2020; Volume 11515, p. 115150Y.
14. Xu, S.; Pan, Z. A novel ensemble of random forest for assisting diagnosis of Parkinson’s disease on small handwritten dynamics dataset. *Int. J. Med. Inform.* **2020**, *144*, 104283. [[CrossRef](#)] [[PubMed](#)]
15. Lu, M.; Fan, Z.; Xu, B.; Chen, L.; Zheng, X.; Li, J.; Znati, T.; Mi, Q.; Jiang, J. Using machine learning to predict ovarian cancer. *Int. J. Med. Inform.* **2020**, *141*, 104195. [[CrossRef](#)] [[PubMed](#)]
16. Ye, C.; Li, J.; Hao, S.; Liu, M.; Jin, H.; Zheng, L.; Xia, M.; Jin, B.; Zhu, C.; Alfreds, S.T.; et al. Identification of elders at higher risk for fall with statewide electronic health records and a machine learning algorithm. *Int. J. Med. Inform.* **2020**, *137*, 104105. [[CrossRef](#)] [[PubMed](#)]
17. Chang, M.-C.; Yi, T.; Duan, K.; Luo, J.; Tu, P.; Priebe, M.; Wood, E.; Stachura, M. In-bed patient motion and pose analysis using depth videos for pressure ulcer prevention. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 4118–4122.
18. Lie, W.-N.; Hsu, F.-Y.; Hsu, Y. Fall-down event detection for elderly based on motion history images and deep learning. In Proceedings of the International Workshop on Advanced Image Technology (IWAIT) 2019, Singapore, 6–9 January 2019; Volume 11049, p. 110493Z.
19. Grimm, T.; Martinez, M.; Benz, A.; Stiefelhagen, R. Sleep position classification from a depth camera using Bed Aligned Maps. In Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, 4–8 December 2016; pp. 319–324. [[CrossRef](#)]
20. Martinez, M.; Schauerte, B.; Stiefelhagen, R. BAM! Depth-based body analysis in critical care. In Proceedings of the 15th International Conference on Computer Analysis of Images and Patterns (CAIP), York, UK, 27–29 August 2013; pp. 465–472.
21. Li, Y.Y.; Lei, Y.J.; Chen, L.C.L.; Hung, Y.P. Sleep posture classification with multi-stream CNN using vertical distance map. In Proceedings of the International Workshop on Advanced Image Technology (IWAIT), Chiang Mai, Thailand, 7–10 January 2018.
22. Mohammadi, S.M.; Kouchaki, S.; Khan, S.; Dijk, D.-J.; Hilton, A.; Wells, K. Two-Step Deep Learning for Estimating Human Sleep Pose Occluded by Bed Covers. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; Volume 2019, pp. 3115–3118.
23. Lee, J.; Hong, M.; Ryu, S. Sleep Monitoring System Using Kinect Sensor. *Int. J. Distrib. Sens. Netw.* **2015**, *2015*, 1–9. [[CrossRef](#)]
24. Toshev, A.; Szegedy, C. DeepPose: Human Pose Estimation via Deep Neural Networks. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 1653–1660. [[CrossRef](#)]
25. Wang, K.; Lin, L.; Jiang, C.; Qian, C.; Wei, P. 3D Human Pose Machines with Self-supervised Learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 1069–1082. [[CrossRef](#)] [[PubMed](#)]
26. Cao, Z.; Hidalgo, G.; Simon, T.; Wei, S.E.; Sheikh, Y. OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *43*, 172–186. [[CrossRef](#)] [[PubMed](#)]
27. Sun, X.; Xiao, B.; Wei, F.; Liang, S.; Wei, Y. Integral Human Pose Regression. In Proceedings of the 13th International Conference on Practice and Theory in Public Key Cryptography, Paris, France, 26–28 May 2010; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2010; pp. 536–553.
28. Lin, C.-B.; Dong, Z.; Kuan, W.-K.; Huang, Y.-F. A Framework for Fall Detection Based on OpenPose Skeleton and LSTM/GRU Models. *Appl. Sci.* **2020**, *11*, 329. [[CrossRef](#)]

29. Nguyen, N.-H.; Phan, T.-D.-T.; Lee, G.-S.; Kim, S.-H.; Yang, H.-J. Gesture Recognition Based on 3D Human Pose Estimation and Body Part Segmentation for RGB Data Input. *Appl. Sci.* **2020**, *10*, 6188. [CrossRef]
30. Doyle, G.R.; McCutcheon, J.A. *Clinical Procedure for Safer Patient Care*; BCcampus Open Textbook Library: Victoria, BC, USA, 2015.
31. Proper Positioning for the Prevention of Pressure Sores and Muscle Contracture. Available online: https://www.elderly.gov.hk/english/carers_corner/positioning/prevention_of_pressure_sores.html (accessed on 18 January 2022).
32. Lie, W.-N.; Lin, G.-H.; Shih, L.-S.; Hsu, Y.; Nguyen, T.H.; Nhu, Q.N.Q. Fully Convolutional Network for 3D Human Skeleton Estimation from a Single View for Action Analysis. In Proceedings of the 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Shanghai, China, 8–12 July 2019; pp. 1–6.
33. Lin, G.-S.; Chen, C.-Y.; Kuo, C.-T.; Lie, W.-N. A Computing Framework of Adaptive Support-Window Multi-Lateral Filter for Image and Depth Processing. *IEEE Trans. Broadcast.* **2014**, *60*, 452–463. [CrossRef]
34. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. *arXiv* **2016**, arXiv:1603.02754v3.
35. Duda, R.O.; Hart, P.E.; Stork, D.G. *Pattern Classification*, 2nd ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2001.
36. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet Classification with Deep Convolutional Neural Networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]