https://doi.org/10.1016/j.autcon.2021.103725

Automated Postural Ergonomic Risk Assessment Using Vision-based Posture Classification

JoonOh Seo[1] and SangHyun Lee[2]

[1] Department of Building and Real Estate, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

[2] Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, United States

ABSTRACT

Because they perform physically demanding manual-handling tasks in awkward postures, construction workers are at high risk of work-related musculoskeletal disorders (WMSDs). Previous studies have developed observational postural ergonomic assessment methods to identify WMSD risks. Although inexpensive and easy to use, these methods are seldom used in construction because they are time-consuming, subject to observer bias, and require well-trained analysts. To address these drawbacks, this paper propose a vision-based method to automatically classify workers' postures for ergonomic assessment. Specifically, it proposes a vision-based method that eliminates the need to collect extensive training-image datasets by employing classification algorithms to learn diverse postures from virtual images, and then identifies those postures in real-world images. In addition, the proposed method extracts features from body silhouettes to lessen the confusion caused by differences between virtual and real-world images, as well as across different lighting conditions and colors of workers' clothing. To assess its feasibility, we conducted laboratory-based tests with varied physical attributes of subjects and image viewpoints. These tests showed that the method had 88.6% classification accuracy, confirming the usefulness of virtual training images for posture classification. Thus, the proposed method has potential for automated and real-time ergonomic risk analysis, and could help to prevent WMSDs not only in the construction industry but in diverse other occupations and tasks.

Keywords: Work-related Musculoskeletal Disorders, Vision-based Posture Classification, Ergonomic Risk Assessment

1. INTRODUCTION

Construction workers are at high risk of work-related musculoskeletal disorders (WMSDs) because they perform physically demanding manual-handling tasks in awkward postures (Everett 1999; Boschman et al. 2012). WMSDs are a leading cause of non-fatal injuries in construction, accounting for 27.5% of such injuries in the United States (CPWR 2018) and 18% of industrial accidents in Hong Kong (OSH 2017). WMSDs are also associated with high costs to employers through absenteeism, lost productivity, increased health care and workers' compensation (NIOSH 2007). Moreover, WMSDs are substantially under-reported, so the problems associated with them are likely to be more severe than indicated by the statistics (Pransky et al. 1999; Punnett & Wegman 2004). Thus, identifying physical exposure to WMSD risks is crucial to implementing workplace ergonomic interventions that will help prevent them (Li & Buckle 1999).

Among the existing methods for measuring exposure to WMSD risks, posture-based ergonomic assessment is one of the most commonly adopted (Janowitz et al. 2006). One of the commonly used approaches widely used in many industries including construction is an ergonomic assessment by simulating tasks at the laboratory environments (Li & Buckle 1999; Seo et al. 2015; Antwi-Afari et al. 2017; Antwi-Afari et al. 2018a). However, the need for technologically sophisticated measurement (e.g., whole-body motion-capture systems) or analysis (e.g., biomechanical analysis) methods may hinder the applicability of this approach in practice. Instead, observation-based postural ergonomic assessment methods quantify risks by systematically classifying the postures of different body parts and scoring them by experienced experts (Li & Buckle 1999). As they are quick and easy to use, they have been widely used for initial screening of specific activities or tasks with higher ergonomic risks at job sites, after which more sophisticated methods can be applied to identify the sources of such risks (Janowitz et al. 2006).

Recently, many researchers in construction have tried to improve these ergonomic assessment methods at construction sites by replacing human observers with wearable sensors or image processing techniques, with the aim of automatically classifying risky postures while tasks are ongoing (Wang et al. 2015). While wearable sensor-based approaches have focused on specific body joints (Yan et al. 2017a; Yan et al. 2018) or risky postures (Antwi-Afari et al. 2018b; Antwi-Afari et al. 2020), the vision-based approaches aim to assess ergonomic risks based on whole body configuration. Such research efforts have ranged from risky posture classification using 2D images from a monocular camera (Seo et al. 2016) to ergonomic posture analysis using 2D or 3D skeleton-based motion data from monocular or depth cameras (Ray & Teizer 2012; Seo et al. 2015; Liu et al. 2016; Dzeng et al. 2017; Yan et al. 2017b; Zhang et al. 2018). Though the use of skeleton data enables estimation of body angles, and is thus a reliable means of detecting awkward postures, 2D image-based posture classification has several comparative advantages. For example, extracting 2D or 3D skeletons from RGB or RGB-D images (e.g., Microsoft KinectTM) requires additional processing time after image collection, whereas 2D image-based posture classification can detect awkward postures directly from the images. Additionally, 2D image-based approaches are less demanding computationally, because 1) they use a selection of feature descriptors from raw images rather than the whole images, and 2) low image resolution does not significantly affect their performance, as their classifications rely only on body silhouettes (Seo et al. 2016). These comparative advantages imply that it might be feasible to develop a stand-alone smartphone application for quick screening of ergonomically risky tasks that may need more sophisticated ergonomic analysis to identify effective intervention methods.

However, a key challenge of the 2D image-based approach is its requirement for large, comprehensive training datasets as a prerequisite for machine learning-based classification (Poppe 2010; Golparvar-Fard et al. 2013). One way to address this issue is through ‘virtual training data’ that allows extraction of extensive training images from a wide range of viewpoints, and which has been successfully utilized for object, face, and gesture recognition (Chiu et al. 2007; Ke et al. 2018; Nikolaev et al. 2018; Tain et al. 2018). Also, in the specific case of 2D image-based ergonomic assessment, the use of virtually created training images to lighten the burden of collecting training data from real-world images has been tested, and deemed appropriate as a means of classifying ergonomically risky postures from a specific viewpoint. To further validate the applicability of the use of virtual training datasets for automated posture classification, it is necessary to address both intra- and inter-class variation attributable to dynamic environments (e.g., changing viewpoints) and also human variability (e.g., clothing and physical attributes).

To this end, the present study proposes and tests a new form of 2D image-based posture classification, in which awkward postures are identified by machine-learning algorithms trained using virtual image datasets to minimize the efforts to collect training images from a real world. In particular, considering varying viewpoints of cameras and workers' different physical attributes (e.g., height and weight), the proposed approach enables to create customized virtual images for the targeted workers to be assessed, which would be challenging when collecting real-world images. To test the proposed approach, we conducted laboratory experiments by collecting diverse views of eight male subjects who were chosen to reflect the range of physical attributes such as height and body mass. Next, virtual training images were created by adjusting our virtual human model to match real conditions (i.e., viewpoints and individual differences). Image features from body silhouettes were then extracted in a manner that sought 1) to minimize color and texture differences between virtual and real-world images, and 2) to capture local variation in trunk and limb movements. Machine-learning algorithms for posture classification were then applied to a set of video images illustrating postures simulated by the eight subjects.

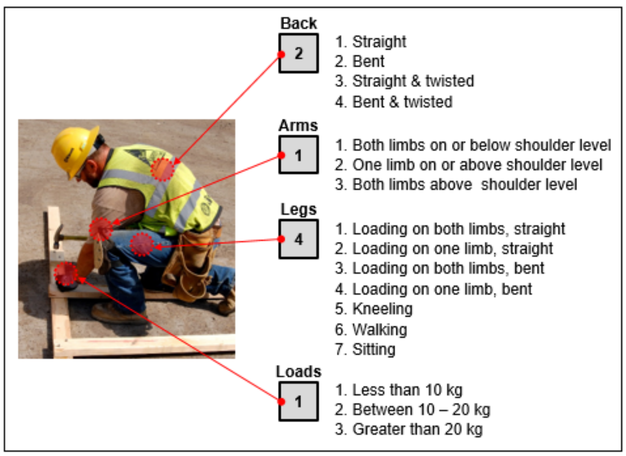
The remaining manuscript is organized as follows. In Chapter 2, existing postural ergonomic analysis methods are introduced. Then, in Chapter 3, machine learning algorithms using virtually created training images for awkward posture classification are presented. Chapter 4 describes laboratory testing and results to validate the proposed approach. Finally, in Chapter 5, potential difficulties and directions for future research are discussed on the basis of the findings.

2. POSTURAL ERGONOMIC RISK ASSESSMENT

Previous epidemiological studies have linked WMSDs to exposure to job-related risk factors such as repetitive motion, heavy lifting, forceful manual exertion and awkward postures (Kumar 2001; Punnett & Wegman 2004). As the level of potential WMSD risk can vary according to the intensity, frequency and duration of such exposure during task performance, quantitative evaluation enables practitioners to identify risky tasks and workers needing immediate ergonomic intervention to prevent WMSDs (David 2005). The ergonomic methods that have been introduced to assess WMSD risks include 1) self-reports, 2) observation, and 3) instrumental or direct measurement (Li & Buckle 1999; David 2005). Of these methods, observation is the most widely used because of its low cost, ease of use, and non-interference with ongoing activity (Genaidy et al. 1994). Observational methods include but are not limited to the Ovako Working Posture Analysing System (OWAS; Karhu et al. 1977; Karhu et al. 1981); Rapid Upper Limb Assessment (RULA; McAtamney & Corlett 1993); Posture, Activity, Tools and Handling (PATH; Buchholz et al. 1966); and Rapid Entire Body Assessment (REBA; Hignett & McAtamney 2000). All are designed to identify potential WMSD risks by recording working postures on proformas and scoring them according to predetermined criteria. Even though each method has its own categories for posture classification, with varying degrees of detail, the level of ergonomic risk in each one is determined via human observation of postural combinations of trunk and limbs.

Of the four methods mentioned above, OWAS and PATH have been deemed the most suitable for use in the construction industry, on the grounds that they are based on *work sampling*: the observation of workers at fixed time intervals to allow estimates of the proportion of time taken in risky postures (Buchholz et al. 1996). Since PATH is based on OWAS, both methods feature similar postural categories and risk-ranking systems. OWAS identifies four working postures for the back, three for the arms, and seven for the legs, and has three weight categories for the load being handled; and the combination of these four variables into a four-digit code (Figure 1) summarizes the whole-body posture. For example, the worker in Figure 1A is hammering a nail while kneeling on one knee. According to the OWAS postural codes, his posture is ‘2’ for the back, ‘1’ for the arms, and ‘4’ for the legs; and because the weight of his hand load is less than 10kg, the load is coded as ‘1’. Thus, the postural code for this posture is ‘2-1-4-1’, which falls into Action Category (AC) 3 (Figure 1B). AC 3 indicates that this posture is distinctly harmful, and thus corrective action should be taken as soon as possible. The action categories for postural ergonomic risk assessment were determined by physicians, work analysts, and workers, and subsequently validated by an international group of ergonomic experts (Karhu et al. 1977).

A. Posture Codes



B. Action Categories (AC)

Back	Arms	1			2			3			4			5			6			7			Legs
		1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	Loads
1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	1	1	1	1	1	1	
	2	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	1	1	1	1	1	1	
	3	1	1	1	1	1	1	1	1	1	2	2	3	2	2	3	1	1	1	1	1	2	
2	1	2	2	3	2	2	3	2	2	3	**3**	3	3	3	3	3	2	2	2	2	3	3	
	2	2	2	3	2	2	3	2	3	3	3	4	4	3	4	4	3	3	4	2	3	4	
	3	3	3	4	2	2	3	3	3	3	3	4	4	4	4	4	4	4	4	2	3	4	
3	1	1	1	1	1	1	1	1	1	2	3	3	3	4	4	4	1	1	1	1	1	1	
	2	2	2	3	1	1	1	1	1	2	3	3	3	4	4	4	1	1	1	1	1	1	
	3	2	2	3	1	1	1	2	3	3	4	4	4	4	4	4	3	3	3	1	1	1	
4	1	2	3	3	2	2	3	2	2	3	4	4	4	4	4	4	4	4	4	2	3	4	
	2	3	3	4	2	3	4	3	3	4	4	4	4	4	4	4	4	4	4	2	3	4	
	3	4	4	4	2	3	4	3	3	4	4	4	4	4	4	4	4	4	4	2	3	4	

C. Stress Ranking System

- AC 1: Normal Posture (no intervention required)
- AC 2: Slightly Harmful (corrective action should be taken during next regular review of work methods)
- **AC 3: Distinctly Harmful** (corrective action should be taken as soon as possible)
- AC 4: Extremely harmful (corrective action should be taken immediately)

Figure 1. Postural Ergonomic Assessment in OWAS

Influential research studies have applied OWAS or PATH to identify potential ergonomic risks during construction tasks. Paquet et al. (2001) found that, in contrast to manufacturing tasks, construction ones exhibit significant variations in exposure to WMSD risk factors, not only across tasks but also between individuals performing the same task. Thus, systematic objective assessment using an observational work-sampling approach such as PATH enables practitioners to see which tasks require immediate intervention and to implement appropriate controls. For example, a pioneering study by Kivi and Mattila (1991) applied OWAS in the building construction industry, and recommended the development of work-redesign measures that would minimize awkward postures on construction sites. Subsequent studies have been conducted on hammering tasks (Mattila et al. 1993), concrete work such as formwork installation, rebar tying and concrete pouring (Li & Lee 1999; Buchholz et al. 2003), iron work (Forde & Buchholz 2004), scaffolding (Saurin & de Macedo Guimarães 2008), and highway-tunnel construction (Tak et al. 2011). Despite the validity and usefulness of these methods, however, they are very time-consuming, as they rely either on direct human observation or human review of video-recordings (Dzeng et al. 2018), with about 30 minutes required for analyzing a single task (Lowe 2004). Some studies have used computer programs to record and analyze postures, but their classification processes has remained manual (Kivi & Mattila 1991; Li & Lee 1999). Also, to help ensure reliable recording of postures, these studies employed trained ergonomists or provided ergonomics training to observers, either of which approaches would imply additional costs if these methods are applied in construction where safety managers would not have appropriate expertise for ergonomic assessments.

To address these issues, many construction researchers have employed advanced sensing technologies to replace human observers with these technologies, with varying degrees of success. These technologies can be divided into two categories, 1) body-attached sensor-based approaches and 2) vision-based approaches. One promising type of body-attached sensor is the Inertial Measurement Unit (IMU; Wang et al. 2015), which consists of an accelerometer, a gyroscope and a magnetometer that jointly measure movements of specific body parts to which the IMU is attached. Thus, if multiple IMUs are attached to the same person, the body angles that are required to identify awkward postures can be calculated based on the relative movements of multiple body parts: e.g., a head and a trunk in the case of back-bending angles (Yan et al. 2017a). Moreover, IMU-based whole-body motion-capture systems can provide 3D skeleton models that can be used to directly calculate specific joint angles (Seo et al. 2017), or to detect ergonomically risky postures through machine-learning techniques (Chen et al. 2017). Recently, a wearable insole pressure sensor has been also used to identify ergonomically hazardous postures by detecting abnormal foot pressure patterns due to over-exertion during construction activities (Antwi-Afari et al. 2018b; Antwi-Afari et al. 2020). Additionally, significant research efforts have been devoted to leveraging the benefits of computer vision techniques for ergonomic assessments. For example, Seo et al. (2016) proposed a 2D image-based posture-classification algorithm to differentiate awkward postures based on body silhouettes. Given that images can provide richer information on body postures than body-attached sensors can, many researchers have proposed ergonomic assessment based on skeletons extracted either from 2D images from an ordinary camera, or from 3D images produced by RGB-D sensors (e.g., Ray & Teizer 2012; Seo et al. 2015; Liu et al. 2016; Yan et al. 2017b; Zhang et al. 2018; Yu et al. 2019b). For example, Dzeng et al. (2017) proposed a novel approach that automatically records postures based on OWAS, and analyzes ergonomic risks using skeleton-based motion data extracted from Microsoft Kinect™. Thanks to the achievement in deep learning algorithms, more reliable and accurate classification of awkward postures on images has been enabled, extending the applicability of vision-based ergonomic assessments (Yu et al. 2019b; Yang et al. 2020; Chu et al. 2020). Also, recent research efforts have tried to combine the data from both computer vision approaches and wearable sensors to not only improve the accuracy of ergonomic risk detection, but also provide additional information (e.g., locations of workers) that would be needed for effective intervention (Cheng et al. 2013; Yu et al. 2019a).

However, it remains unclear which approach would be best suited to the context of construction work, as body-attached sensor-based approaches and vision-based approaches have different limitations. One frequently mentioned limitation of body-attached sensors is the discomfort they cause to workers, which in some cases interferes with ongoing work. Vision-based approaches, meanwhile, have been criticized due to limited site coverage by cameras and the high likelihood of occlusions. Given that the purpose of postural ergonomic assessment is initial screening for risky tasks based on work sampling, data collection at different positions for diverse tasks should be required. When using vision-based approaches based on 2D images, the safety personnel can easily record videos using a hand-held camera or a smartphone, without having to interfere with ongoing work to attach sensors to workers, and quickly move on to other individuals and tasks. For these reasons, vision-based approaches to postural ergonomic assessment appear to be more promising than their sensor-based counterparts.

To sum up, observation-based postural ergonomic assessment methods have been widely used to identify ergonomic risks during occupational tasks including construction. To address the limitation of time-consuming and error-prone manual observations, the recent research efforts have proposed automated posture classification approaches using both wearable sensors and computer vision techniques. Even though both approaches have shown promising results in terms of awkward posture identification, the vision-based approach would be more suitable for quick screening of risk tasks without any interference with on-going work. However, from a technical point of views, the vision-based approaches have suffered from the need for significant manual efforts to collect training images from a real-world (Golparvar-Fard et al. 2013). The performance of vision-based posture classification would heavily rely on the quantity of training images that is large enough to include any possible variations on images. This issue would be more significant for vision-based posture classification of construction workers as the human body consists of the head, the torso and the limbs, creating a variety of postures depending on different tasks unlike other project entities (e.g., equipment, materials) in construction. This has been well-recognized as a significant research challenge in not only the computer science domain, but also the construction domain (Yu et al. 2010; Golparvar-Fard et al. 2013).

3. METHODS

The purpose of the computer vision-based postural classification algorithms proposed in this study is to identify and classify different bodily postures, as defined by existing postural ergonomic evaluation methods, from video or time-lapse images. Figure 2 shows the present study's overall research procedure. Its key basis is the insight that different postures within images create distinguishable pixel patterns (i.e., image features), allowing classification algorithms to learn patterns from training images and differentiate among postures in test images. As discussed above, one of the key challenges of vision-based approaches is the creation of comprehensive training images that reflect variations in real-world conditions such as viewpoints or workers' physical attributes. To address it, training-image datasets obtained through virtual human modeling were used. In addition, the current study applied an algorithm for vision-based posture classification using image features from body silhouettes obtained by background subtraction. These postural classification algorithms lead a classifier to learn diverse postures from virtual training images, and then to classify the postures in real-world images. Additionally, the present research combines newly proposed image features (e.g., shape-based features) with well-established ones (e.g., radial histograms of silhouettes) to better reflect morphological variations in the body silhouettes of people in different postures.

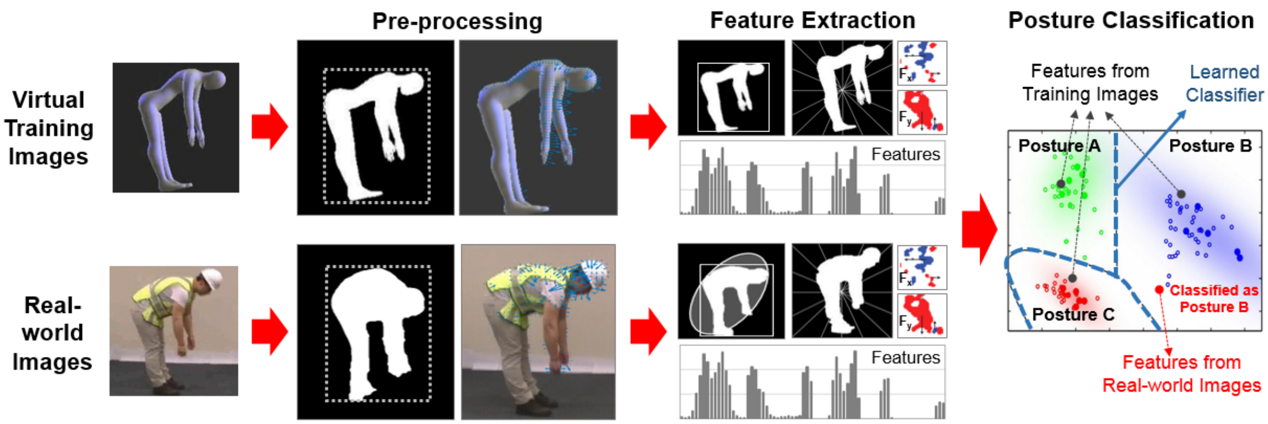


Figure 2. Overall Procedure for Vision-based Posture Classification

3.1. Creating Virtual Training Datasets

Training images for diverse postures were obtained by using virtual human modeling, an emerging technology for motion simulation in a virtual environment (VE; Demirel & Duffy 2007). A virtual human model with specific physical attributes such as height and weight is inserted and animated according to human motion-capture data in a 3D virtual space, thus generating virtual training image datasets of the type illustrated in Figure 3.

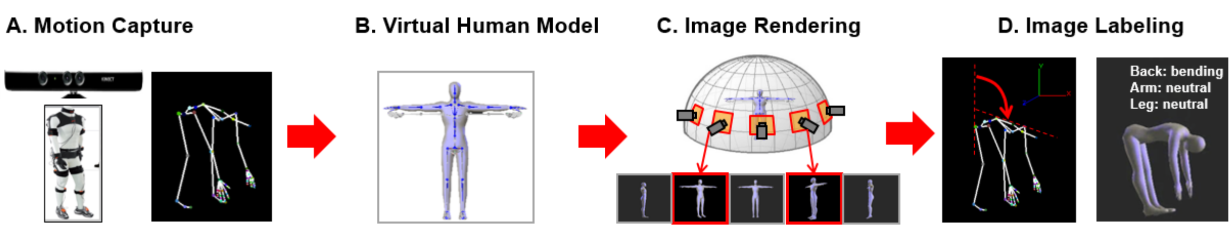


Figure 3. Procedure for Virtual Training Datasets

First, motion data from workers at real-world construction sites was collected using motion-capture systems such as an RGB-D sensor and an IMU-based motion capture system (Figure 3A). Then, the virtual human model could be manipulated into any posture from any viewpoint, with substantial variability in physical attributes, without requiring new motion data from actual workers. Nevertheless, actual workers' motion data was collected because, given the unstandardized and complex nature of construction work, manual manipulation of postures might have missed important configurations, and at this stage, the researchers wished to encompass as diverse a range of workers' postures as possible. That being said, however, if any posture was noted as missing, a virtual human model allowed the creation of that posture simply by modifying existing postures without further real-life observation, making it relatively easy to update the training datasets.

Once the motion data were obtained, the virtual human model was constructed and simulated in the VE on the basis of such data (Figure 3B). Specifically, as body profiles—i.e., silhouettes in images—are significantly influenced by individual differences in height and weight, virtual human models representing diverse distributions of such characteristics within a specified population had to be created by creating multiple variants of the virtual human model. Then, the human model was projected onto an image sphere to create a sequence of images depicting human postures. By changing the positions of the virtual camera, virtual video sequences could be created from all of the possible viewpoints that would exist under real-world conditions (Figure 3C).

Each video image was then labeled according to the postures to be used in training datasets. As described above, current methods of postural ergonomic assessment define whole-body postures by combining the observed postures of specific body parts (Karwowski & Marras 1998). Such combinations in video images can be automatically identified using corresponding motion data where 3D limb positions and body angles are available (Figure 3D). The use of 3D skeleton data permits both accurate and instant postural labeling for training datasets, and can be extended to any type of postural ergonomic assessment simply by varying the criteria used to define postures of interest.

3.2. Obtaining Body Silhouettes through Background Subtraction

Because variation in body shape provides enough information to identify distinct postures and motions in visual data, body silhouettes have been widely used for spatial representation of actions by the human body (Weinland et al. 2011). In the present study, the use of body silhouettes had the additional advantage of being unaffected by variations in color, texture and contrast, thus eliminating confusion that might otherwise have been caused by variations in lighting conditions or in the colors of workers' clothing. Moreover, body silhouettes are unaffected by differences between virtual and real-world images, thereby enabling algorithms trained using only virtual training images to accurately identify distinct postures in real-world images.

Nevertheless, accurate posture classification is dependent on effective techniques for deriving clear body silhouettes from images. The present study employed background-subtraction and noise-removal algorithms to obtain clear body silhouettes, with a bounding box serving as a Region of Interest (ROI) to extract image features at the next step (Figure 4). Background-subtraction algorithms define the current pixel as foreground when the difference between its value and those of pixels in the background model exceeds a threshold value (Piccardi 2004). This study utilized a state-of-the-art background-subtraction algorithm, ViBe, which is robust to lighting changes and the appearance of new objects within the scene as it updates the background model over time (Barnich & Van Droogenbroeck 2001).

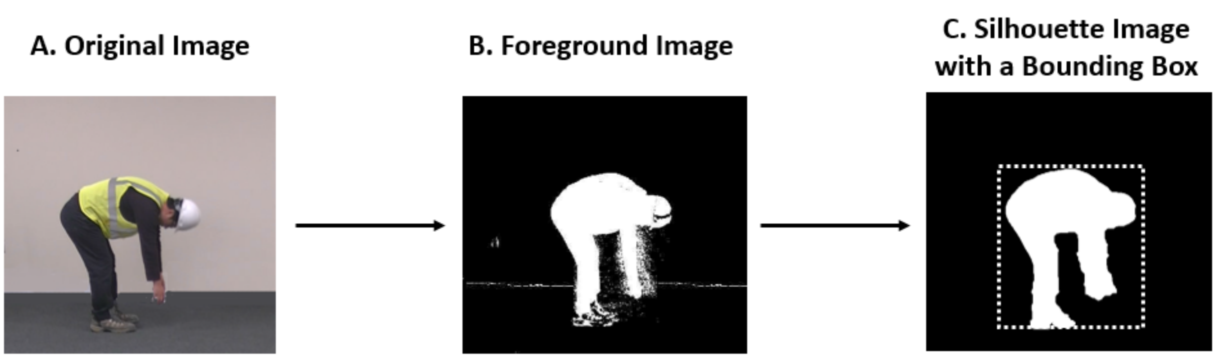


Figure 4. Background Subtraction and Detection of Bounding Box

ViBe deals with the problem of background subtraction as one of classification (Barnich & Van Droogenbroeck 2011). It classifies a current pixel value (*v(x)*, the value of the pixel located at point x in the image) by comparing that value against its corresponding background model at the pixel location x, *M(x)*, which is modeled by v_i, background sample pixel values having been established in the N previous frames (Eq. 1). Specifically, if the number of background pixel samples close to the new pixel value in a Euclidean color space is higher than a given threshold, the current pixel is classified as background. Additionally, the background model is continually modified to adapt to lighting changes or to new objects appearing in a scene via a conservative update method, in which the background model is updated using the value of the current pixel value after the latter has been defined as background.

$$\mathcal{M}(x) = \{v_1, v_2, \dots, v_N\} \text{ where } v_i \text{ is a background pixel sample} \tag{1}$$

After subtraction, the foreground images might nevertheless have noisy pixels in the background. As shown in Figure 4B, for example, shadows on walls and high-contrast edges can result in the false detection of background regions (Elgammal et al. 2000). To remove noisy pixels from the background, several noise-removal algorithms were deployed, and removed objects containing fewer than 50 pixels; then, a median-based filter replaced the noisy pixels (*f(x, y)*, a pixel value at the position of (x, y)) with median values (g(x, y), a median filtered pixel value) in a 5×5 pixel window, as shown in Equation 2 (Dong & Xu 2007). In addition, a morphological closing operation was performed to fill in the narrow gaps and small holes in body silhouettes. Once a clear silhouette was generated, a bounding box was placed around it, and this box served as an ROI for feature extraction (Figure 4C).

$$g(x, y) = median\left(\sum_{i=-2}^{2} \sum_{j=-2}^{2} f(x-i, y-j)\right) \tag{2}$$

3.3. Extracting Image Features from Body Silhouettes

At the next stage, image features representing body postures were extracted from body silhouettes. Because of their power to capture complex body movements, silhouette-based techniques have been widely used for action recognition (Poppe 2010). Generally, temporal variations in body silhouettes are key characteristics of image features when the aim is robust action recognition. However, for purposes of the present study, posture classification had to rely upon the limited information that could be gleaned from a single body silhouette, yet extract posture-specific image features containing information rich enough not only to classify diverse postures, but also to generalize across small variations in workers' appearance.

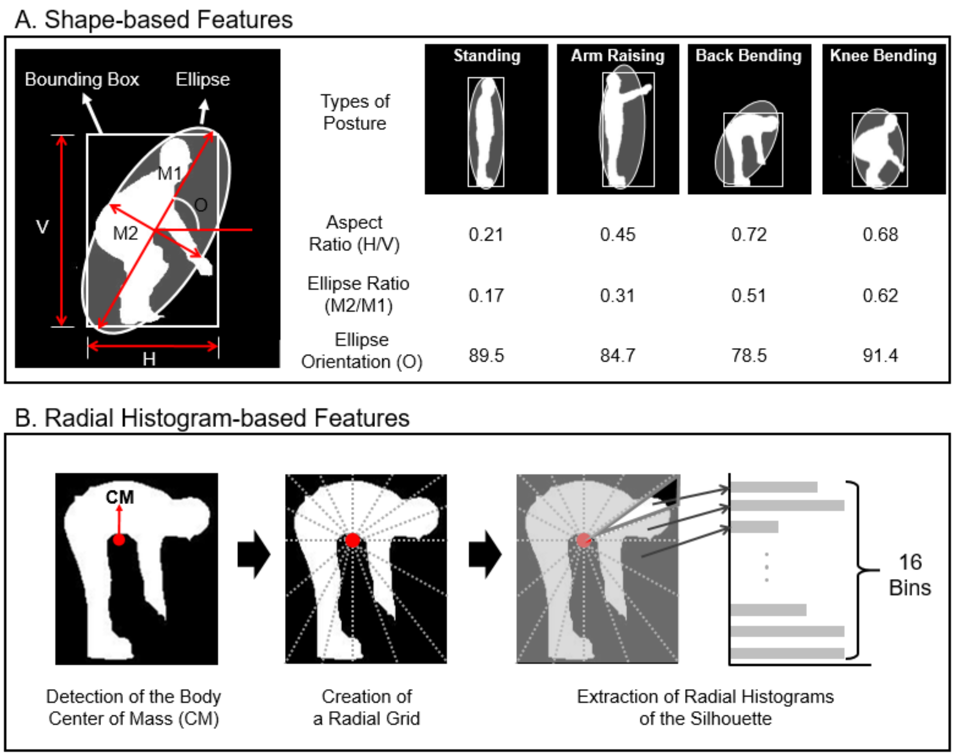


Figure 5. Silhouette-based Image Features

In the current study, attributes of body shapes derived from body silhouettes were used as image features. Those attributes included: 1) the aspect ratio of the bounding box; 2) the ratio of the minor to the major axis of the ellipse that could be fitted to the silhouette; and 3) the orientation of the ellipse itself (Figure 5A). They are intended to capture morphological variations in body silhouettes according to different postures. Aspect ratio is a measurement that provides an intuitive cue about the size of an object, and thus can serve as a morphological feature for human detection or gait analysis (Garcia & Tziritas 1999; Broggi et al. 2000;

Wang et al. 2003). The aspect ratio of the bounding box also provides a unique cue for recognizing bodily postures, especially when differentiating standing from other postures, as shown in Figure 5. In addition, a body silhouette can be defined by an ellipse fitted to it (i.e., by the ratio between the major and minor axes of the ellipse and the slope of the major axis), which helps to classify postures in the bounding box that have similar aspect ratios, such as back-bending and knee-bending. To extract more detailed shape-based features, the bounding box is further divided into 2×1 (two subsets) and 2×2 (four subsets) sub-windows, and features extracted from each sub-window.

As a local descriptor for capturing details in postures, a radial histogram of the silhouette—the center of which is defined as the silhouette's center of mass (CM)—is also extracted from the image (Figure 5B). The position of CM was calculated using the formula

$$\bar{x} = \frac{\sum_{i=0}^{n-1}\sum_{j=0}^{m-1} j\,B[i,j]}{A}, \quad \bar{y} = \frac{-\sum_{i=0}^{n-1}\sum_{j=0}^{m-1} i\,B[i,j]}{A} \tag{3}$$

where B$[i,j]$ is a binary image of the body silhouette; A is the area of the image; and $\bar{x}$ and $\bar{y}$ are the coordinates of the CM with respect to the top left pixel. Then, the bounding box was divided into 16 slices, and the ratio of black to white pixels in each slice histogrammed with 16 bins.

3.4. Classification Algorithm

Once the image features were constructed, a classifier had to learn them from virtual training datasets if it was to accurately classify postures in new testing images. The classifier the researchers selected was a Support Vector Machine (SVM), which is widely used in action recognition (Poppe 2010). In brief, standard SVM classification aims to separate a set of training vectors belonging to two classes, but it can be extended by combining a number of two-class classification SVMs to form a multi-class classifier (Hsu & Lin 2002). The present study implemented the one-against-one method, which constructs k(k-1)/2 classifiers (k being the number of classes), each of which is trained on data from two classes and predicts the class of a testing vector based on majority voting (Kreßel 1999).

3.5. Post-processing for Noise Removal

The proposed method performs posture classification based on frame-by-frame processing, which means that each frame is classified independently. Under real-world conditions, workers perform tasks by varying their postures, but specific postures are generally maintained for a certain period (e.g., several seconds). As such, if classification results show variation across a brief sequence of consecutive frames, it is likely that the postures involved might be incorrectly classified. To eliminate such noise, any classified posture that did not persist for more than 10 consecutive image frames was re-labeled as whatever posture dominated the adjacent frames, as shown in Figure 6.

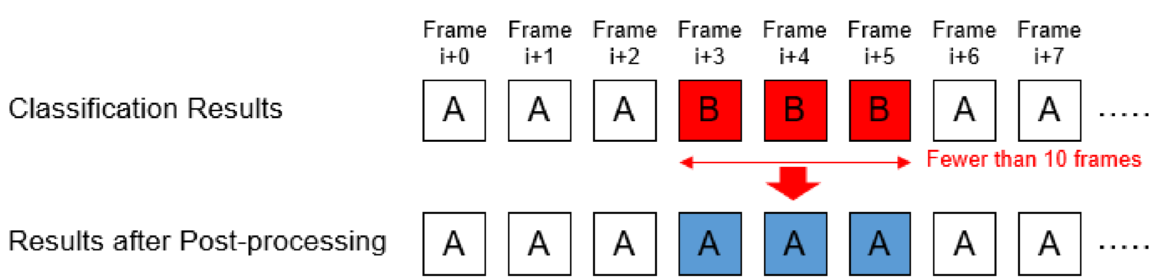


Figure 6. Post-processing of Classification Results

4. LABORATORY TESTING

Laboratory tests of the proposed method were conducted under a variety of conditions, i.e., multiple viewpoints (intraclass variation) and individual differences in physical attributes (interclass variation). Their primary purpose was to establish whether training images from a virtual environment, independent of real-world testing conditions, were applicable to the posture classification of images with both intra- and interclass variation, and if so, whether such classification was more accurate than than that obtained via human observation. To obtain reliable training and testing images, data collection procedures were carefully controlled in the laboratory environment, following a pre-designed protocol (data available from the corresponding author upon request).

4.1. Testing Postures

For purposes of the above-mentioned laboratory test, the researchers selected three representative postures involving different body parts, i.e., back-bending for back posture, arm-raising for arm posture, and knee-bending for leg posture, as shown in Figure 7. Each resulting combination of three postures was defined by reference to the OWAS codes for ergonomic risk: that is, as Category 1, 2, 3, or 4, with a higher number indicating a more risky posture. For example, in OWAS, when the upper body is bent forward or backward by 20 degrees or more, the posture is classified as 'back-bending'. An 'arm-raising' entails both arms being at or above shoulder level; and a 'squat' is when both knees are bent at an angle of 150 degrees or less.

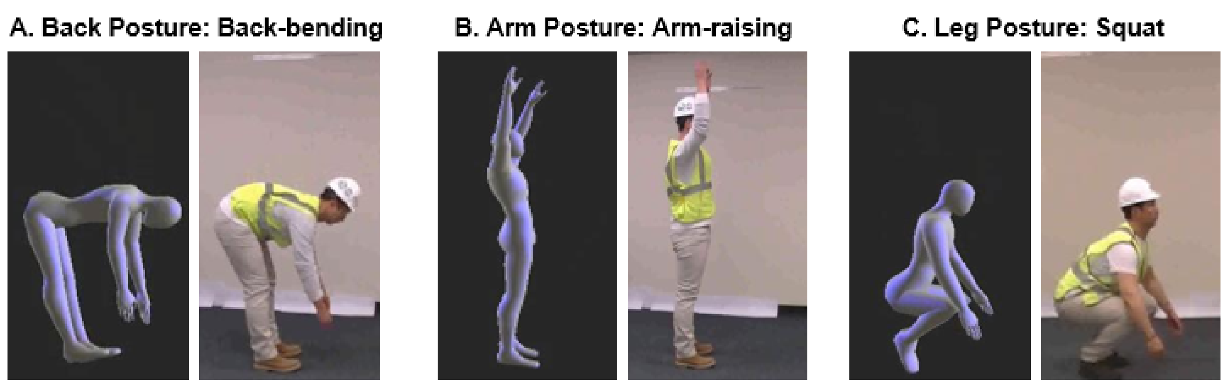


Figure 7. Examples of Postures in Training and Testing Images

4.2. Data Collection

The tests were carried out in the Construction Laboratory at the University of Michigan. As shown in Table 1, male subjects with divergent physical attributes such as height and Body Mass Index (BMI) were recruited to represent human variability in silhouettes. Only male subjects were recruited because the overwhelming majority of construction workers at high ergonomic risk are men; in the U.S., for example, only 2.4% of production workers in construction are women (CPWR 2018). To ensure the representation of typical physical characteristics among males aged 20 and over, six subjects with heights and BMIs between the 25th and 75th percentiles were the main focus of such recruitment (CDC 2012; Flegal et al. 2012). However, to test extreme samples, two additional subjects, one in the 4th and the other in the 96th percentile for height, were also recruited. Thus, testing images including various postures were collected from eight subjects, reflecting possible variations on body silhouettes due to height (ranging from the 4th to the 96th percentile) and BMI (ranging from the 5th to the 75th percentile).

Table 1. Subjects' Heights and BMIs

	#1	#2	#3	#4	#5	#6	#7	#8	Average
Height	163cm (4.1 %tile)	173cm (34.4 %tile)	175cm (44.7 %tile)	175cm (44.7 %tile)	180cm (70.4 %tile)	180cm (70.4 %tile)	181cm (74.8 %tile)	189cm (95.9 %tile)	177cm
BMI	25.6 (normal)	18.5 (under)	24.4 (normal)	25.2 (normal)	23.5 (normal)	29.3 (over)	26.3 (over)	25.8 (over)	24.9

Note: under (underweight, < 18.5 BMI), normal (normal weight, 18.5 to 24.9 BMI) and over (overweight, 25.0 to 29.9 BMI)

The subjects were asked to simulate each posture 10 times. They began by standing up straight, and then either bent their backs up to 90 degrees; raised their arms as high as they comfortably can; or bent their knees as if they were squatting, followed by standing up straight once more. They were given enough time to practice so that they were able to simulate the postures identically.

While simulating each posture, videos were recorded from three viewpoints (left, rear left diagonal, and rear), as shown in Figure 8. These perspectives were chosen in consideration of the fact that front views are not easily obtainable on construction sites, where workers generally face workspaces such as walls.

Figure 8. Testing Images from Three Viewpoints (Left, Left-diagonal and Rear)

From videos recorded at 30 frames per second, a total of 60,091 image frames (or around 7,500 per subject) were extracted as testing images, as shown in Table 2. Additionally, motion data for each subject were collected using a RGB-D sensor (i.e., Microsoft Kinect^{TM}) to identify postures in the corresponding images as ground truth. For the training data, nine different virtual human models representing, respectively, the 15th, 50th and 85th percentiles for male height and BMI were created, and then simulated in a VE using one subject's motion data. Next, image sequences were extracted from the same viewpoints with testing images. Both training and testing images were processed to obtain body silhouettes, after which image features were extracted using MATLAB software.

Table 2. Numbers of Testing Images of Each Posture, by Viewpoint

Postures	Viewpoints			
	Left	Left-diagonal	Rear	Sub-total
Standing	9,616	7,676	9,906	27,198
Back-bending	3,829	3,070	3,937	10,836
Arm-raising	4,823	3,355	3,862	12,040
Knee-bending	3,640	3,122	3,255	10,017
Subtotal	21,908	17,223	20,960	**Total: 60,091**

4.3. Testing Conditions and Measures

Laboratory testing of the proposed method served three purposes: 1) assessment of the overall classification performance by viewpoint, without taking the selection of optimal training images into account (Test #1); 2) establishing the effect of viewpoint-choice on classification performance (Test #2); and 3) measuring the effect of the virtual models' physical attributes on classification performance (Test #3). Below, a comparison between the results of Test #1 and those of Tests #2 and #3 will be followed by a discussion of the advantages of virtual training images that are readily adjustable to real-world conditions.

In Test #1, the proposed method was assessed from each viewpoint and for all subjects, while the SVM classifier learned all the virtual training images from nine virtual models and three viewpoints, without adjusting either for viewpoints or for anatomical variation in the virtual models. As its measures of classification performance, Test #1 utilized *accuracy*, i.e., the ratio of true positives to the total number of images; *precision*, the ratio of true positives to the combined total of true and false positives; and *recall*, the ratio of true positives to the combined total of true positives and false negatives. These three measures were calculated in a confusion matrix to define the performance of a given classification model.

Test #2 examined the hypothesis that, if training images and testing images shared the same viewpoint, classification performance would be improved. For this test, two sets of training images were selected, one from a left view and the other from a left-diagonal view, on the grounds that these only slightly different viewpoints could cause confusion when classifying postures, and thus be more challenging for the SVM classifier. Then, each subject's posture was classified using both the left and left-diagonal sets of training images. Using a paired *t*-test, the classification accuracy when the same viewpoint was used in both the training and testing images was compared against such accuracy when the viewpoints did not match.

Test #3 was designed to test the hypothesis that the selection of training images from a virtual model with similar physical attributes to the subject would enhance classification accuracy. For this test, nine sets of training images were sorted according to the physical attributes of the virtual models, i.e., 3 (short, average, tall) ×3 (underweight, medium, overweight), and used to classify the subjects' postures. Classification accuracy when selecting a virtual model with similar physical attributes to those of each subject was then compared against the accuracy achieved when dissimilar virtual models were selected, again via a paired *t*-test.

4.4. Testing Results

Table 3 presents the proposed system's posture-classification results by viewpoint. The results indicate that the proposed algorithms performed better when testing images from a left view (88.6%) than when testing those taken from a left-diagonal perspective (85.6%) or from the rear (78.4%). More specifically, there was no confusion between back-bending, arm-raising and knee-bending postures in a left view, whereas confusion between these postures increased as the viewpoint shifted from the left toward the rear. This implies that a side view will likely yield the best classification results. Most classification errors occurred as a consequence of confusion between standing and other postures (i.e., back-bending, arm-raising and knee-bending). From all three viewpoints, significant numbers of standing-posture images were wrongly classified as depicting one of the other three postures, and conversely, the other three postures tended to be wrongly recognized as standing. Further investigation revealed that the cause of such errors was similarity in the subjects' transitions from standing to other postures.

Table 3. Posture-classification Results by Testing-image Viewpoint

		Predicted Postures				**Recall*** **
		1	**2**	**3**	**4**	
Actual postures of testing images from **a left view** Accuracy** = **88.6%**	**1***	**8,075**	819	719	3	84.0%
	2	265	**3,564**	0	0	93.0%
	3	12	0	**4,811**	0	99.7%
	4	669	0	0	**2,971**	81.5%
	Precision** **	89.5%	81.3%	87.1%	99.8%	-
Actual postures of testing images from	**1**	**6,558**	538	509	71	85.4%
	2	221	**2,849**	0	0	92.7%
	3	8	0	**3,323**	24	99.0%

a left-diagonal view	**4**	1,104	0	0	**2,018**	64.6%
Accuracy = **85.6%**	**Precision**	83.1%	84.1%	86.6%	95.3%	-
Actual postures of testing images from **a rear view**	**1**	**8,305**	1,402	75	124	83.8%
	2	862	**3,075**	0	0	78.1%
	3	904	197	**2,692**	69	69.7%
	4	895	0	0	**2,360**	72.4%
Accuracy = **78.4%**	**Precision**	75.7%	65.8%	97.2%	92.3%	-

* 1. Standing, 2. Back-bending, 3. Arm-raising, and 4. Knee-bending
** Accuracy: Ratio of true positives to total number of images
*** Recall: Ratio of true positives to combined total of true positives and false negatives
**** Precision: Ratio of true positives to combined total of true and false positives

The findings of Test #2, which investigated postural variability by viewpoint, underline the importance of training-image viewpoint selection. As shown in Table 4, mean accuracy was 88.7% when the same views were used for both training and testing images, but it fell to 80.8% when the alternative view was used for training images ($p = 0.004$). The classification errors would increase as the mismatch between training- and testing-image viewpoints became more marked (e.g., if training images from the rear were combined with testing images from the left). This result also indicates that differences in the viewpoints from which images were captured can produce strong variations in body silhouettes, and thus, that the perspectives from which training and testing images are taken should match each other to enhance classification accuracy.

Table 4. Classification Accuracy by Training-image Viewpoint Similarity or Dissimilarity

Subject	Classification Accuracy	
	Training: Left View Testing: Left View	Training: Left-diagonal View Testing: Left View
#1	90.7%	87.5%
#2	81.6%	83.0%
#3	84.1%	81.6%
#4	89.0%	82.6%
#5	85.5%	83.7%
#6	89.3%	72.4%
#7	94.7%	70.0%
#8	94.2%	85.8%
Mean	**88.7%**	**80.8%**
Standard Deviation	**4.3%**	**5.9%**

Note: $p = 0.004$, paired t-test

Table 5 presents the proposed system's classification accuracy according to which virtual model was selected as the source of the training images (Test #3). Training images taken from virtual models with similar physical attributes to those of the real-world subjects were identified with greater accuracy (mean: 85.8%, standard deviation: 2.4%) than ones that used other virtual models (mean: 83.0%, standard deviation: 3.7%; $p = 0.004$). These results indicate that variations in silhouette shape linked to individuals' physical attributes affected classification performance, and thus, that posture-classification algorithms must take individual differences in height and body mass into account.

Table 5. Classification Accuracy by Training-model Similarity or Dissimilarity to Subjects

Subject	Classification Accuracy	
	Similar Physical Attributes*	Different Physical Attributes
#1	85.6%	81.7%
#2	84.2%	82.0%
#3	83.9%	77.4%
#4	89.4%	89.1%
#5	81.9%	79.6%
#6	85.1%	81.3%
#7	87.3%	85.9%
#8	89.0%	86.6%
Mean	**85.8%**	**83.0%**
Standard Deviation	**2.4%**	**3.7%**

Note: p-value = 0.004, paired t-test
* Height and BMI.

The mean accuracy the proposed system attained when using training images from a virtual model that matched the individual subject, 85.8%, was slightly lower than the overall accuracy of 88.6% obtained when all nine virtual models were used to reflect individual subjects' key physical differences. This implies that, in each subject's body silhouette, there were some variations that could not be adequately reflected by virtual human models, even those of similar height and BMI to the real person. In turn, this suggests that, rather than adjusting the virtual model's height and BMI for each subject, it might be preferable to create virtual models that reflect population variability in body silhouettes, and require the classifier learn such variability via training images of all possible virtual human models.

5. DISCUSSION

The results of experimental testing of the proposed system showed that its overall classification accuracy ranged from 78.4% to 88.6%, depending on image viewpoint. It was also found that the use of training images similar to testing images (e.g., images captured from the same viewpoint or from a virtual human models with similar physical attributes to the human subject) significantly increased overall classification performance. Despite significant differences in color and texture between virtual training images and real-world testing images, the proposed method yielded acceptable performance (88.6%), implying the strong potential for practical applications of automated postural ergonomic risk assessment that minimizes the need to collect real-world training images. In addition, the above testing results imply that the use of virtual human models that reflect population variability, and the selection of very similar viewpoints to those in real-world views, will be essential if accuracy is to be improved. Taken together, these results strongly support the utility of virtually created training datasets with high adjustability for these factors.

On examining incorrectly classified image frames, it was found that most errors were associated with transitional postures, i.e., postures near the frame at which a subject transitioned between standing and another posture such as back-bending, arm-raising, or knee-bending (Figure 9). As this study defined each posture on the basis of body angles (e.g., 20 degrees for back-bending, 150 degrees for knee-bending), postures with body angles close to but not actually meeting these criteria had similar body silhouettes to those that did meet them, and this was the main cause of classification errors. It should also be noted here that the tests used in this study were based on all image frames from the relevant videos, which were shot at 30 frames per second, and that each posture was simulated repetitively within a short cycle time (e.g., 2-3 seconds), and this meant that almost 20% of all images were of transitional postures. However, it is expected that postural transitions would occur less frequently in practice, and that this would reduce the number of errors.

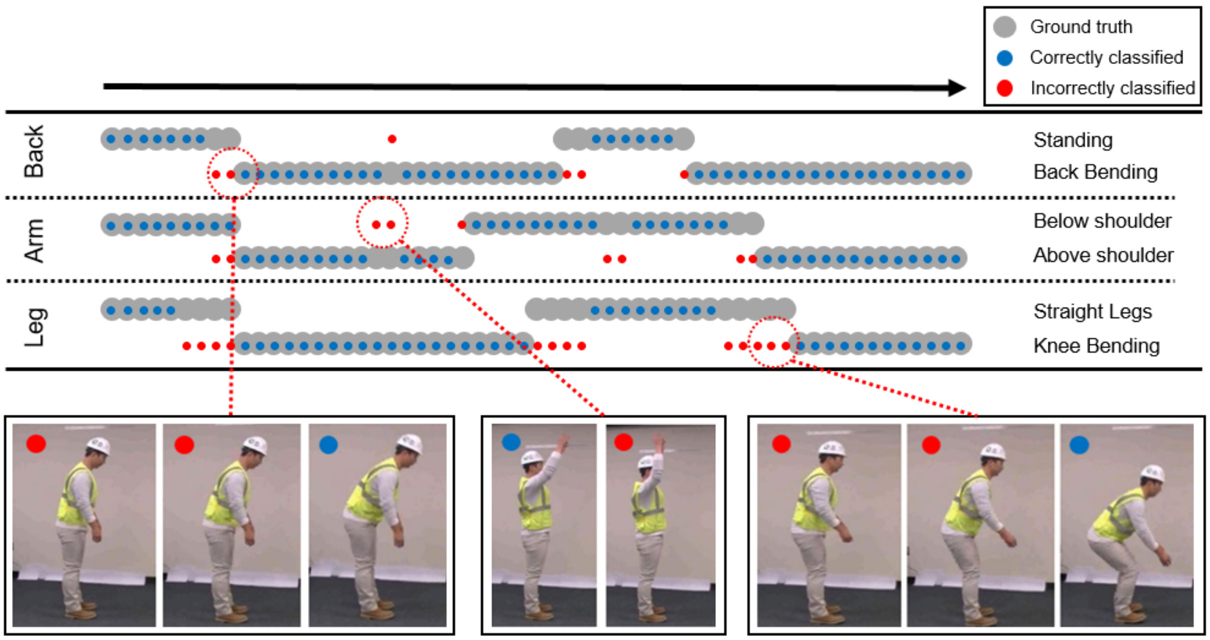


Figure 9. Classification Error Analysis

Further improvement of the proposed system will require algorithmic solutions capable of dealing with transitional postures. Chalamala and Kumar (2016) have found that probabilistic classifiers are preferable to deterministic methods such as SVMs for continuous actions with random changes of class. They proposed a probabilistic model based on transition probabilities (between walking and running) and occurrence probabilities (whether walking or running), and obtained superior performance to that achievable by deterministic methods. Though further studies are required to test the feasibility of a probabilistic approach to postural ergonomic risk assessment, it could represent a viable means of reducing errors resulting from transitional postures.

While further improvement to classification performance would be desirable and may well be possible, the existing system's overall accuracy of nearly 90% compares favorably to observational posture recording. To achieve better performance, recording videos from a left view would be required, but as it is assumed that images will be collected using a hand-held camera or a smartphone camera, the camera angle can be easily adjusted to suit job conditions. Previous studies that investigated human observers' posture-classification accuracy have found that even experienced ergonomic analysts could make errors more than 10% of the time (Burt & Punnett 1999; Paquet et al. 2001; Spielholz et al. 2001; Lowe 2004; Weir et al. 2011). Importantly, postural ergonomic analysis based on human observations takes about 30 minutes per task (Lowe 2004), whereas its 2D image-based equivalent is much more rapid and requires less ergonomics expertise on the part of those performing it, e.g., construction safety managers. Even though further in-depth investigation by human observers may be needed to understand the individual factors such as working habits or the environmental factors such as poorly designed workspaces that contribute to awkward postures, the proposed method will be a useful means of quickly identifying risky activities that need immediate intervention.

Notwithstanding the potential of the proposed vision-based posture classification algorithm, several obstacles still need to be overcome before it can be applied in real-world conditions. First, the test for the algorithm was made only for single postures according to bodily attributes. In reality, however, combinations of postures (e.g., back-bending + knee-bending, or knee-bending + arm-raising) are frequently observed. As such, the proposed method needs to be further validated with more complex postures involving various such combinations. Also, large hand-tools and other objects held by workers can significantly affect the shape of their silhouettes as obtained via background subtraction, and this might lead to classification errors. To address this problem, more sophisticated post-processing algorithms may be required to detect objects and recover clear silhouettes. In addition, since workers are always on the move in the workplace, the views captured by a video camera will also continually be changing. Thus, given that a strong similarity of viewpoint between training images and testing images was found in this study to be key to classification accuracy, it will be essential to determine how best to capture training images of workers who are always on the move. Automated object-orientation detection to identify an object's rotation angles, based on statistical pattern-recognition techniques, could be a solution for determining how target workers are oriented in camera images (Vailaya et al. 2002).

6. CONCLUSIONS

This study has proposed a 2D image-based posture-classification method based on machine-learning algorithms, with the wider aim of automating current postural ergonomic evaluation methods. To assess the proposed method under diverse conditions, laboratory-based tests were conducted with eight male subjects possessing different physical attributes. These tests established that the proposed algorithm is capable of robust posture-classification accuracy, comparable to that attained by human observers. In particular, considering that the use of customized training datasets created and manipulated in a VE showed better classification performance, the proposed approach has great potential for dealing with high variability of human postures that has been one of the challenges in the computer vision domain.

The proposed approach creates a range of opportunities for both research and practice. On the research side, it could be used to tackle several persistent challenges to vision-based posture classification, including 1) insufficient numbers of training images for machine learning, 2) how to deal with changes in color, texture and contrast in images, and 3) the complexity of representing postures across a range of different observational viewpoints and workers' anatomical differences. First, a novel method of using training datasets for diverse postures from a VE was suggested to minimize the effort associated with collecting training images from a real world. Also, the above test results show that the use of body silhouettes can address potential errors caused by differences between virtual and real-world images and/or by variations in image quality. In addition, the proposed approach is robust to intra- and inter-class variability consequent upon changing viewpoints and individual differences in workers' physical attributes that could lead to variations in the imaged postures. Using virtual human models, this approach can create training images in which body mass and viewpoint can be adjusted as required. Beyond these contributions of the proposed approach,

In terms of practice, although the proposed approach needs to be further validated with images containing the more complex postures often adopted in construction, it has the potential to automate ergonomic assessment methods that are currently time-consuming and error-prone because of their manual procedures.

In construction especially, those responsible for on-site health and safety are few in number and generally lack the expertise needed to perform manual ergonomic assessments when identifying WMSD risks. The proposed approach can help even those practitioners who lack sufficient ergonomic knowledge to perform postural ergonomic risk assessments simply by making videos of workers. In short, computer vision-based ergonomic assessment could open the door to proactive control of WMSDs among construction workers by quickly evaluating all tasks, identifying potential risks, and taking timely action to eliminate those risks.

ACKNOWLEDGEMENTS

This study was supported by an Early Career Scheme (PolyU 25210917) from Research Grants Council, Hong Kong, a grant (19CTAP-C151784-01) from Technology Advancement Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government, and a National Science Foundation Award (No. CMMI-1161123), United States.

REFERENCES

Antwi-Afari, M. F., Li, H., Edwards, D. J., Pärn, E. A., Owusu-Manu, D. G., Seo, J., & Wong, A. Y. L. (2018a). Identification of potential biomechanical risk factors for low back disorders during repetitive rebar lifting. *Construction Innovation*, 18(2), pp.206-226, DOI: 10.1108/CI-05-2017-0048

Antwi-Afari, M. F., Li, H., Edwards, D. J., Pärn, E. A., Seo, J., & Wong, A. Y. L. (2017). Biomechanical analysis of risk factors for work-related musculoskeletal disorders during repetitive lifting task in construction workers. *Automation in Construction*, 83, pp. 41-47, DOI: 10.1016/j.autcon.2017.07.007

Antwi-Afari, M. F., Li, H., Umer, W., Yu, Y., & Xing, X. (2020). Construction activity recognition and ergonomic risk assessment using a wearable insole pressure system. *Journal of Construction Engineering and Management*, 146(7), DOI: 10.1061/(ASCE)CO.1943-7862.0001849

Antwi-Afari, M. F., Li, H., Yu, Y., & Kong, L. (2018b). Wearable insole pressure system for automated detection and classification of awkward working postures in construction workers. *Automation in Construction*, 96, pp. 433-441, DOI: 10.1016/j.autcon.2018.10.004

Barnich, O. & Van Droogenbroeck, M. (2011). ViBe: A universal background subtraction algorithm for video sequences, *IEEE Transactions on Image Processing*, 20(6), pp.1709-1724, DOI: 10.1109/TIP.2010.2101613

Boschman, J. S., van der Molen, H. F., Sluiter, J. K., & Frings-Dresen, M. H. (2012). Musculoskeletal disorders among construction workers: a one-year follow-up study. *BMC Musculoskeletal Disorders*, 13(1), http://www.biomedcentral.com/1471-2474/13/19, Accessed 10th Oct 2016 (ISSN: 1471-2474)

Broggi, A., Bertozzi, M., Fascioli, A., & Sechi, M. (2000, October). Shape-based pedestrian detection. *IEEE Intelligent Vehicles Symposium*, pp. 215-220, DOI: 10.1109/IVS.2002.1187920

Buchholz, B., Paquet, V., Punnett, L., Lee, D., & Moir, S. (1996). PATH: a work sampling-based approach to ergonomic job analysis for construction and other non-repetitive work. *Applied Ergonomics*, 27(3), pp. 177-187, DOI: 10.1016/0003-6870(95)00078-X

Buchholz, B., Paquet, V., Wellman, H., & Forde, M. (2003). Quantification of ergonomic hazards for ironworkers performing concrete reinforcement tasks during heavy highway construction. *Journal of Occupational and Environmental Hygiene*, 64(2), pp. 243-250, DOI: 10.1080/15428110308984814

Burt, S., & Punnett, L. (1999). Evaluation of interrater reliability for posture observations in a field study. *Applied Ergonomics*, 30(2), pp. 121-135, DOI: 10.1016/S0003-6870(98)00007-6

Centers for Disease Control and Prevention (CDC) (2012) Anthropometric Reference Data for Children and Adults: United States, 2007–2010, http://www.cdc.gov/nchs/data/series/sr_11/sr11_252.pdf. Accessed 30th April 2016

Center for Construction Research and Training (CPWR) (2018). The Construction Chart Book: The U.S. Construction Industry and Its Workers (Sixth Edition).

https://www.cpwr.com/sites/default/files/publications/The_6th_Edition_Construction_eChart_Book.pdf Accessed 1st August 2019

Chalamala, S. R., & Kumar, P. (2016). A probabilistic approach for human action recognition using motion trajectories. *7th International Conference on Intelligent Systems, Modelling and Simulation (ISMS)*, IEEE, pp. 185-190, DOI: 10.1109/ISMS.2016.39

Chen, J., Qiu, J., & Ahn, C. (2017). Construction worker's awkward posture recognition through supervised motion tensor decomposition. *Automation in Construction*, 77, pp. 67-81, DOI: 10.1016/j.autcon.2017.01.020

Cheng, T., Teizer, J., Migliaccio, G. C., & Gatti, U. C. (2013). Automated task-level activity analysis through fusion of real time location sensors and worker's thoracic posture data. *Automation in Construction*, 29, pp. 24-39, DOI: 10.1016/j.autcon.2012.08.003

Chu, W., Han, S., Luo, X., & Zhu, Z. (2020). Monocular vision–based framework for biomechanical analysis or ergonomic posture assessment in modular construction. *Journal of Computing in Civil Engineering*, 34(4), DOI: 10.1061/(ASCE) CP.1943-5487.0000897

David, G. C. (2005). Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders. *Occupational Medicine*, 55(3), pp. 190-199, DOI: 10.1093/occmed/kqi082

Demirel, H. O., & Duffy, V. G. (2007). Applications of digital human modeling in industry. *International Conference on Digital Human Modeling (ICDHM 2007)*, pp. 824-832, DOI: 10.1007/978-3-540-73321-8_93

Dong, Y., & Xu, S. (2007). A new directional weighted median filter for removal of random-valued impulse noise. *IEEE Signal Processing Letters*, 14(3), pp. 193-196, DOI: 10.1109/LSP.2006.884014

Dzeng, R. J., & Watanabe, K. (2018). Applications of ICTs and action recognition for construction workers. *Trends in Civil Engineering and its Architecture*, 1(3), pp. 48-54, DOI: 10.32474/TCEIA.2018.01.000113

Dzeng, R. J., Hsueh, H. H., & Ho, C. W. (2017). Automated posture assessment for construction workers. *40th IEEE International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 1027-1031, DOI: 10.23919/MIPRO.2017.7973575

Elgammal, A., Harwood, D., & Davis, L. (2000). Non-parametric model for background subtraction. *European Conference on Computer Vision (ECCV 2000)*, pp. 751-767, DOI: 10.1007/3-540-45053-X_48

Everett, J. G. (1999). Overexertion injuries in construction. *Journal of Construction Engineering and Management*, 125(2), pp. 109-114, DOI: 10.1061/(ASCE)0733-9364(1999)125:2(109)

Flegal, K. M., Carroll, M. D., Kit, B. K., & Ogden, C. L. (2012). Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. *Journal of the American Medical Association (JAMA)*, 307(5), pp. 491-497, DOI: 10.1001/jama.2012.39

Forde, M. S., & Buchholz, B. (2004). Task content and physical ergonomic risk factors in construction ironwork. *International Journal of Industrial Ergonomics*, 34(4), pp. 319-333, DOI: 10.1016/j.ergon.2004.04.011

Garcia, C., & Tziritas, G. (1999). Face detection using quantized skin color regions merging and wavelet packet analysis. *IEEE Transactions on Multimedia*, 1(3), pp. 264-277, https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=784465, Accessed 23th July 2016

Genaidy, A. M., Al-Shedi, A. A., & Karwowski, W. (1994). Postural stress analysis in industry. *Applied Ergonomics*, 25(2), pp. 77-87, DOI: 10.1016/0003-6870(94)90068-X

Golparvar-Fard, M., Heydarian, A., & Niebles, J. C. (2013). Vision-based action recognition of earthmoving equipment using spatio-temporal features and support vector machine classifiers. *Advanced Engineering Informatics*, 27(4), pp. 652-663, DOI: 10.1016/j.aei.2013.09.001

Chiu, H. P., Kaelbling, L. P., & Lozano-Pérez, T. (2007, June). Virtual training for multi-view object class recognition. *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, DOI: 10.1109/CVPR.2007.383044

Hignett, S., & McAtamney, L. (2000). Rapid entire body assessment (REBA). *Applied Ergonomics*, 31(2), pp. 201-205, DOI: 10.1016/S0003-6870(99)00039-3

Hsu, C. W., & Lin, C. J. (2002). A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks*, 13(2), pp. 415-425, https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=991427, Accessed 22th July 2016

Janowitz, I. L., Gillen, M., Ryan, G., Rempel, D., Trupin, L., Swig, L., ... & Blanc, P. D. (2006). Measuring the physical demands of work in hospital settings: Design and implementation of an ergonomics assessment. *Applied Ergonomics*, 37(5), pp. 641-658, DOI: 10.1016/j.apergo.2005.08.004

Karhu, O., Härkönen, R., Sorvali, P., & Vepsäläinen, P. (1981). Observing working postures in industry: Examples of OWAS application. *Applied Ergonomics*, 12(1), pp. 13-17, DOI: 10.1016/0003-6870(81)90088-0

Karhu, O., Kansi, P., & Kuorinka, I. (1977). Correcting working postures in industry: a practical method for analysis. *Applied Ergonomics*, 8(4), pp. 199-201, DOI: 10.1016/0003-6870(77)90164-8

Karwowski, W. & Marras, W. S. (1998). The occupational ergonomics handbook. CRC Press LLC, DOI: 10.1002/(SICI)1520-6564(200021)10:2<223::AID-HFM6>3.0.CO;2-I (ISBN 0–8493–2641–9)

Ke, J., Peng, Y., Liu, S., Li, J., & Pei, Z. (2018). Face recognition based on symmetrical virtual image and original training image. *Journal of Modern Optics*, 65(4), pp. 367-380, DOI: 10.1080/09500340.2017.1380854

Kivi, P., & Mattila, M. (1991). Analysis and improvement of work postures in the building industry: application of the computerised OWAS method. *Applied Ergonomics*, 22(1), pp. 43-48, DOI: 10.1016/0003-6870(91)90009-7

Kreßel, U (1999) Pairwise classification and support vector machines. *Advances in Kernel Methods: Support Vector Learning*, MIT Press, pp. 255–268, DOI: 10.5555/299094.299108 (ISBN:0-262-19416-3)

Kumar, S. (2001). Theories of musculoskeletal injury causation. *Ergonomics*, 44(1), pp. 17-47, DOI: 10.1080/00140130120716

Li, G., & Buckle, P. (1999). Current techniques for assessing physical exposure to work-related musculoskeletal risks, with emphasis on posture-based methods. *Ergonomics*, 42(5), pp. 674-695, DOI: 10.1080/001401399185388

Li, K. W., & Lee, C. L. (1999). Postural analysis of four jobs on two building construction sites: an experience of using the OWAS method in Taiwan. *Journal of Occupational Health*, 41(3), pp. 183-190, DOI: 10.1539/joh.41.183

Liu, M., Han, S., & Lee, S. (2016). Tracking-based 3D human skeleton extraction from stereo video camera toward an on-site safety and ergonomic analysis. *Construction Innovation*, 16(3), 348-367, DOI: 10.1108/CI-10-2015-0054

Lowe, B. D. (2004). Accuracy and validity of observational estimates of wrist and forearm posture. *Ergonomics*, 47(5), pp. 527-554, DOI: 10.1080/00140130310001653057

Mattila, M., Karwowski, W., & Vilkki, M. (1993). Analysis of working postures in hammering tasks on building construction sites using the computerized OWAS method. *Applied Ergonomics*, 24(6), pp. 405-412, DOI: 10.1016/0003-6870(93)90172-6

McAtamney, L., & Corlett, E. N. (1993). RULA: a survey method for the investigation of work-related upper limb disorders. *Applied Ergonomics*, 24(2), pp. 91-99, DOI: 10.1016/0003-6870(93)90080-S

National Institute for Occupational Safety and Health (NIOSH) (2007). Simple Solutions: Ergonomics for Construction Workers, DHHS (NIOSH) Publication No. 2007–122, https://stacks.cdc.gov/view/cdc/6912 Accessed 21th July 2016

Nikolaev, E. I., Dvoryaninov, P. V., Lensky, Y. Y., & Drozdovsky, N. S. (2018). Using virtual data for training deep model for hand gesture recognition. *Journal of Physics: Conference Series*, 1015(4), DOI: 10.1088/1742-6596/1015/4/042045

Occupational Safety and Health Branch (OSH) (2017) Occupational Safety and Health Statistics 2017, Labour Department, The Government of the Hong Kong Special Administrative Region, https://www.labour.gov.hk/eng/osh/pdf/OSH_Statistics_2017_EN.pdf Accessed 24th August 2019

Paquet, V. L., Punnett, L., & Buchholz, B. (2001). Validity of fixed-interval observations for postural assessment in construction work. *Applied Ergonomics*, 32(3), pp. 215-224, DOI: 10.1016/S0003-6870(01)00002-3

Piccardi, M. (2004). Background subtraction techniques: a review. *2004 IEEE International Conference on Systems, Man and Cybernetics*, 4, pp. 3099-3104, DOI: 10.1109/ICSMC.2004.1400815

Poppe, R. (2010). A survey on vision-based human action recognition. *Image and Vision Computing*, 28(6), pp. 976-990, DOI: 10.1016/j.imavis.2009.11.014

Pransky, G., Snyder, T., Dembe, A., & Himmelstein, J. (1999). Under-reporting of work-related disorders in the workplace: a case study and review of the literature. *Ergonomics*, 42(1), pp. 171-182, DOI: 10.1080/001401399185874

Punnett, L., & Wegman, D. H. (2004). Work-related musculoskeletal disorders: the epidemiologic evidence and the debate. *Journal of Electromyography and Kinesiology*, 14(1), 13-23, DOI: 10.1016/j.jelekin.2003.09.015

Ray, S. J., & Teizer, J. (2012). Real-time construction worker posture analysis for ergonomics training. *Advanced Engineering Informatics*, 26(2), 439-455, DOI: 10.1016/j.aei.2012.02.011

Saurin, T. A., & de Macedo Guimarães, L. B. (2006). Ergonomic assessment of suspended scaffolds. *International Journal of Industrial Ergonomics*, 36(3), pp. 229-237, DOI: 10.1016/j.ergon.2005.11.006

Seo, J., Alwasel, A., Lee, S., Abdel-Rahman, E. M., & Haas, C. (2017). A comparative study of in-field motion capture approaches for body kinematics measurement in construction. *Robotica*, 1-19, DOI: 10.1017/S0263574717000571

Seo, J., Starbuck, R., Han, S., Lee, S., & Armstrong, T. (2015). Motion-Data–driven Biomechanical Analysis during Construction Tasks on Sites. *Journal of Computing in Civil Engineering*, ASCE, 29(4), DOI: 10.1061/(ASCE)CP.1943-5487.0000400

Seo, J., Yin, K., & Lee, S. (2016). Automated postural ergonomic assessment using a computer vision-based posture classification. *Construction Research Congress 2016*, pp. 809-818, DOI: 10.1061/9780784479827.082

Spielholz, P., Silverstein, B., Morgan, M., Checkoway, H., & Kaufman, J. (2001). Comparison of self-report, video observation and direct measurement methods for upper extremity musculoskeletal disorder physical risk factors. *Ergonomics*, 44(6), pp. 588-613, DOI: 10.1080/00140130118050

Tak, S., Buchholz, B., Punnett, L., Moir, S., Paquet, V., Fulmer, S., ... & Wegman, D. (2011). Physical ergonomic hazards in highway tunnel construction: overview from the Construction Occupational Health Program. *Applied Ergonomics*, 42(5), pp. 665-671, DOI: 10.1016/j.apergo.2010.10.001

Tian, Y., Li, X., Wang, K., & Wang, F. Y. (2018). Training and testing object detectors with virtual images. *IEEE/CAA Journal of Automatica Sinica*, 5(2), pp. 539-546, DOI: 10.1109/JAS.2017.7510841

Vailaya, A., Zhang, H., Yang, C., Liu, F. I., & Jain, A. K. (2002). Automatic image orientation detection. *IEEE Transactions on Image Processing*, 11(7), pp. 746-755, DOI: 10.1109/TIP.2002.801590

Wang, D., Dai, F., & Ning, X. (2015). Risk assessment of work-related musculoskeletal disorders in construction: state-of-the-art review. *Journal of Construction Engineering and Management*, 141(6), DOI: 10.1061/(ASCE)CO.1943-7862.0000979

Wang, L., Tan, T., Ning, H., & Hu, W. (2003). Silhouette analysis-based gait recognition for human identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(12), pp. 1505-1518, DOI: 10.1109/TPAMI.2003.1251144

Weinland, D., Ronfard, R., & Boyer, E. (2011). A survey of vision-based methods for action representation, segmentation and recognition. *Computer Vision and Image Understanding*, 115(2), pp. 224-241, DOI:

10.1016/j.cviu.2010.10.002

Weir, P. L., Andrews, D. M., van Wyk, P. M., & Callaghan, J. P. (2011). The influence of training on decision times and errors associated with classifying trunk postures using video-based posture assessment methods. *Ergonomics*, 54(2), pp. 197-205, DOI: 10.1080/00140139.2010.547603

Yan, X., Li, H., Li, A. R., & Zhang, H. (2017a). Wearable IMU-based real-time motion warning system for construction workers' musculoskeletal disorders prevention. *Automation in Construction*, 74, pp. 2-11, DOI: 10.1016/j.autcon.2016.11.007

Yan, X., Li, H., Wang, C., Seo, J., Zhang, H., & Wang, H. (2017b). Development of ergonomic posture recognition technique based on 2D ordinary camera for construction hazard prevention through view-invariant features in 2D skeleton motion. *Advanced Engineering Informatics*, 34, pp. 152-163, DOI: 10.1016/j.aei.2017.11.001

Yan, X., Li, H., Zhang, H., & Rose, T. M. (2018). Personalized method for self-management of trunk postural ergonomic hazards in construction rebar ironwork. *Advanced Engineering Informatics*, 37, pp. 31-41, DOI: 10.1016/j.aei.2018.04.013

Yu, J., Farin, D., Krüger, C., & Schiele, B., (2010). Improving person detection using synthetic training data. *2010 IEEE International Conference on Image Processing,* IEEE, pp. 3477-3480, DOI: 10.1109/ICIP.2010.5650143

Yu, Y., Li, H., Umer, W., Dong, C., Yang, X., Skitmore, M., & Wong, A. Y. (2019a). Automatic biomechanical workload estimation for construction workers by computer vision and smart insoles. *Journal of Computing in Civil Engineering*, 33(3), DOI: 10.1061/(ASCE)CP.1943-5487.0000827

Yu, Y., Yang, X., Li, H., Luo, X., Guo, H., & Fang, Q. (2019b). Joint-level vision-based ergonomic assessment tool for construction workers. *Journal of Construction Engineering and Management*, 145(5), DOI: 10.1061/(ASCE)CO.1943-7862.0001647

Zhang, H., Yan, X., & Li, H. (2018). Ergonomic posture recognition using 3D view-invariant features from single ordinary camera. *Automation in Construction*, 94, pp. 1-10, DOI: 10.1016/j.autcon.2018.05.033