REBAPose -A Computer vision based Musculoskeletal Disorder Risk Assessment Framework

Sivakumar K S[1], Nikhil Bugalia[2] and Benny Raphael [3]

[1,2,3] Department of Civil Engineering, Indian Institute of Technology Madras, Chennai, India
ce22s018@smail.iitm.ac.in , nikhilbugalia@gmail.com , benny@civil.iitm.ac.in

Abstract –

Work-related Musculoskeletal Disorders (WMSDs) are one of the prominent challenges facing the construction industry, and their effective management requires a risk-quantification strategy. A comprehensive method to assess WMSDs for the construction sector is the Rapid Entire Body Assessment (REBA). The conventional risk-quantification frameworks relying on manual analysis of postures and images are resource-intensive and not scalable. Hence, automated approaches using Computer Vision (CV) are gaining attention. Current CV-based methods to automatically estimate REBA scores do not track all the key points necessary for accurate computation; they rely on heuristics. The present study is part of a wider research effort to develop a CV-based, fully automatic REBA risk calculator. An important task in this effort is to develop and validate a key-point annotation strategy for accurately estimating various body parts and joints necessary for REBA score calculation. This study re-annotated 149,813 human images present in the open-access COCO dataset. About 0.5 million key points were added for three body parts: the midpoint of head, neck, and center hip. Then, this data was used to train the state-of-the-art CV algorithms, MMPose and Alphapose. A comparison of the ML models developed on the newly annotated data with the pre-existing heuristics-based approaches demonstrates significant performance gains in precision and accuracy. The proposed model has the potential to be widely adopted for the precise and quick estimation of the REBA WMSD risk-assessment framework in different industries.

Keywords –

Ergonomics; Computer vision; Pose analytics; REBA Score; COCO; Construction.

1 Introduction

The construction industry worldwide continues to perform poorly regarding Occupational Health and Safety (OHS) issues [1][2]. Among these, Work-related Musculoskeletal Disorders (WMSD), preventable disorders affecting nerves, tendons, muscles, and supporting structures, are receiving increasing academic and practitioner attention due to their significant impact on the construction industry. For example, 41% of registered workers in Hong Kong reported facing WMSD [3]. In the USA, workday loss due to WMSD has increased from 8 days in 1992 to 13 days in 2014 [4]. Hence, developing an appropriate WMSD risk-assessment framework is an important task.

Observation-based methods have been prominently used for WMSD risk quantification and management in construction. In these methods, the worker's postures are observed through work-sampling approaches and classified into risk categories. Several risk-assessment frameworks, such as OVAKO, REBA, and RULA, have been extensively used on construction sites for WMSD risk assessment [5][6]. An important step in this is to classify postures in images and videos into risk categories. However, these conventional observation-based approaches rely on extensive manual inputs. Hence, they are not scalable to the needs of the rapidly increasing construction sector. To overcome this problem, Computer Vision (CV) has been used to identify postures in images and quickly analyze large quantities of images and videos. A highly accurate CV-based risk assessment framework for OVAKO exists [7]. However, to the authors' knowledge, an accurate REBA score estimator has not yet been developed, even though few previous studies have attempted heuristic approaches [8]. Compared to OVAKO, REBA is a comprehensive and accurate risk-assessment framework that is more widely applicable for construction. Therefore, a CV-based approach for accurate REBA estimation may significantly benefit many practitioners.

The present study is part of a wider research effort to develop a CV-based, fully automatic REBA risk calculator called REBAPose. An important task in this effort is to accurately locate various body parts and joints in the images. This work is focused on the task mentioned above and aims to develop and validate a strategy for key-point annotation. The performance of the proposed strategy is compared with the currently adopted heuristics approach to estimate the key points necessary for REBA score estimation. The scope of the present study is limited to developing an accurate training dataset by adding 3 more annotation points to the existing COCO dataset of 149813 images and evaluating the performance of ML algorithms on detecting these newly annotated key points. Such an assessment is expected to make the REBA scoring assessment more accurate.

The study is structured as follows. Section 2 provides an overview of the literature and identifies the essential gaps. Section 3 describes the essentials of CV-based WMSD assessment and the analytical methodology adopted in the current study. Results have been summarized in section 4, followed by discussions in section 5. Conclusions have been outlined in section 6.

2 Literature Review

2.1 Overview of WMSD assessment frameworks

Several methods of WMSD risk assessment are used in construction and other industries. Popular direct observation methods are the OVAKO working posture assessment system (OWAS) [9], Rapid Entire Body Assessment (REBA) [10], and Rapid Upper Limb Assessment (RULA) [11]. REBA is a common assessment method with high accuracy. Recent extensive literature on the topic suggests that REBA is likely to be a comprehensive and accurate analysis method for the construction industry [6]. However, a generic challenge with all observation-based methods is their extensive reliance on manual intervention for analyzing human postures (identifying body parts joints and angles between them) from large quantities of images collected at the site. Hence, a CV-based approach for automatic REBA risk estimation is a great tool of practical relevance that could aid the work of safety practitioners.

2.2 Computer vision frameworks for posture analysis

In recent years, extensive research has been carried out in automatic human pose estimation using images for various applications. Several advanced Machine Learning (ML) algorithms have been developed, such as MMPose [12], and Alphapose [13]. These have been reported to have achieved more than 90% accuracy in identifying various joints and body parts on large-scale standard datasets such as the open-access COCO validation dataset. The current study leverages these standardized algorithms to develop REBAPose.

2.3 Available Posture Datasets

A fundamental need for any CV-based algorithm is an annotated dataset for training the ML algorithms. Over the years, several large-scale standardized publicly available human-annotated datasets have been created, allowing researchers from several areas to quickly develop and validate their own ML approaches. Popular datasets include MS COCO [14], MPII [15], Human36M [16], and AIC[17]. Using pre-existing datasets has several advantages compared to developing a new pose estimation model from scratch. Manual efforts can be reduced, and comparative assessments become possible using published results.

For a reliable estimate of the REBA score, 18 key points need to be estimated. However, none of the existing datasets can track all these 18 key points necessary for REBA score estimation (see Table 1).

Table 1. Body parts and Datasets

Body Part	COCO	BlazePose	MPII	AIC	REBA
Midpoint of Head	X	X	X	X	Yes
Nose	✓	✓	X	X	Yes
Neck	X	X	✓	✓	Yes
Right Shoulder	✓	✓	✓	✓	Yes
Left shoulder	✓	✓	✓	✓	Yes
Right Elbow	✓	✓	✓	✓	Yes
Left Elbow	✓	✓	✓	✓	Yes
Right Wrist	✓	✓	✓	✓	Yes
Left Wrist	✓	✓	✓	✓	Yes
R Index Finger	X*	✓	X	X	Yes
L Index Finger	X*	✓	X	X	Yes
Right Hip	✓	✓	✓	✓	Yes
Left Hip	✓	✓	✓	✓	Yes
Center Hip	X	X	X	X	Yes
R Knee	✓	✓	✓	✓	Yes
L Knee	✓	✓	✓	✓	Yes
R Ankle	✓	✓	✓	✓	Yes
L Ankle	✓	✓	✓	✓	Yes
Total KPs	13	15	13	13	18

X* - Not available in the original COCO dataset but is available in COCO Whole body dataset

Due to the limitations of the dataset, many previous studies on automatic REBA score estimators use heuristics to estimate the 3 key points necessary for REBA score calculation. For example, previous studies have relied on the point “Nose,” available in the COCO dataset, instead of the midpoint of head for angle calculations. Section 3.4 of the current study presents more examples of such heuristics.

Overall, it is evident that none of the existing datasets are sufficient to be readily used for accurate REBA posture analytics, and annotation on additional key points may be necessary for training and testing purposes of the CV model. The current study adopts to improve the existing annotations of the COCO dataset, as this dataset covers huge variations in the type of images collected and has been extensively used in previous studies, allowing for easy benchmarking.

3 Methodology

3.1 Methodology Overview

A three-step process is adopted in the current study and is summarized in Figure 1. The steps include annotation, training, and testing. The detail of each step is presented in the subsequent sections. As indicated in Table 1, key points at the midpoint of HEAD, NECK, and Center HIP are required in addition to the information already available in the pre-annotated COCO dataset. These three points were annotated in the current study to prepare an extensive new dataset for accurate REBA score estimation. It is important to note that the information on the two index fingers is also required for REBA score estimation. While such information was unavailable in the original COCO dataset, it was available from the COCO Whole Body dataset implemented in another study [18]. Hence, these readily available key points have been assimilated in the current study without the need to re-annotate them.

3.2 Annotation Framework

Crowdsourcing is extensively used to annotate large quantities of datasets. However, the quality of crowdsourced work is difficult to assess and control. Hence, we deployed 3rd party vendor to annotate the images for 3 new pose points. Employees of this organization are specialized in annotation services. The current study sought the help of these paid professional data annotators with extensive efforts to quality control. For this purpose, a toolset compatible with the JSON files in the COCO dataset was developed specifically to annotate pose key points. The tool can use the existing annotated key points from datasets such as COCO and build on top of it to annotate additional points.

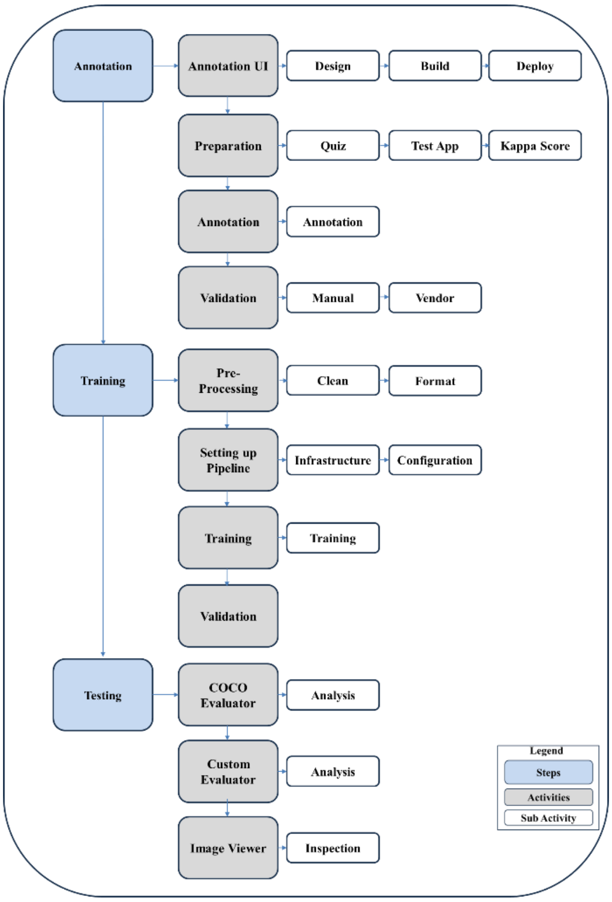

Figure 1. Methodology overview

3.2.1 Annotation UI details

Figure 2 shows the annotation toolset's intuitive user interface. The central area showcases the bounding box of the human being annotated. The pixel coordinates of each key point, along with their human interpretable names, are shown on the left side. Various annotation support functions are on the right side. To help in faster annotation, the framework enabled quick shortcut keys, such as Zoom functionality for the image and changing the color palette for better visibility. The tool was web-based so that multiple annotators could work simultaneously.

Figure 2. Annotation toolset UI Screen

3.2.2 Methods to minimize the bias in annotation

The current study devised a rigorous approach to ensure the annotation process's quality across the 20 different professional annotators. Before the annotation process, each annotator had to complete a ten-question quiz. The questions provided specific directions for certain tasks and could be answered based on information available only for successful completion. Annotators who scored above 90% are made eligible to carry out an annotation. Inter-rater reliability of all the annotators through kappa statistics [19] was also assessed on 100 images. Kappa statistics compares the agreement between two or more annotators by calculating the degree of agreement without considering the probability of consistency due to chance. The coefficient ranges between 0 (likely to be coincidence chance) and 1 (perfect agreement). A kappa value above 0.6 represents an adequate annotator agreement. The current study obtained a Kappa score of 0.67 for 20 annotators.

Once the annotation task was complete, the third-party annotation company was asked to re-check the quality of annotations using different personnel. The research team also randomly checked the annotation for 1000 persons and ensured the quality of the work done. Through this extensive process, an entire training set of 149813 persons and a validation set of 6352 persons were annotated for each of the three new key points. This annotated dataset is used for training and evaluation during testing and can be made accessible to all upon a reasonable request.

3.3 Training of ML Models

Training is a process where machine learning models learn from labeled data. The key steps in this process are preprocessing, setting up ML architecture, training, and validation.

3.3.1 Pre-processing of data

In this pre-processing step, data is cleaned and arranged so that models can learn efficiently. A single image in the COCO dataset may contain multiple persons; hence, a top-down approach is used to train the ML algorithms. The bounding boxes are first generated for every person in the image. Each bounding box is then taken as a single JSON file and shown to annotators for ease of key point annotation. Hence, multiple JSON files could contain annotation information for a single image in the COCO dataset. However, for training purposes, the algorithms require all information in a single file; hence, all individual annotated files are clubbed into a single file using the bounding box information and image ID.

3.3.2 Setting up the ML Model Architecture

For retraining purposes using the modified key points, the current study leverages the ML algorithm architectures of the two state-of-the-art ML models known to have good accuracy on the COCO dataset. These two ML algorithms are MMPose and Alphapose. Both these architectures adopt the most accurate HRNet V2 [20] as the backbone for training.

3.3.3 Training

The two ML models were initialized based on the optimal hyperparameters reported in the previous literature [12][13]. The training validation is calculated at every 5th epoch, and a decision is made whether to stop the training or continue. Percentage Correct Key Point (PCK) is widely used as an accuracy parameter in pose analytics [22]. It is defined as the percentage of key points detected correctly compared to the total number of key points available in the annotated dataset of ground truth. The key points are predicted correctly if the prediction lies within a circle of threshold distance radius, i.e., 20% of the diagonal length of the bounding box [22]. In the current study, the training was stopped at the 100th epoch as PCK at 20% threshold value reached an accuracy of 94.34% and loss (Mean Squared Error - MSE) saturated at a low value of 0.0012 (See Figure 3). Overall, training took around 3 days for the total 149813 images with 100 epochs.

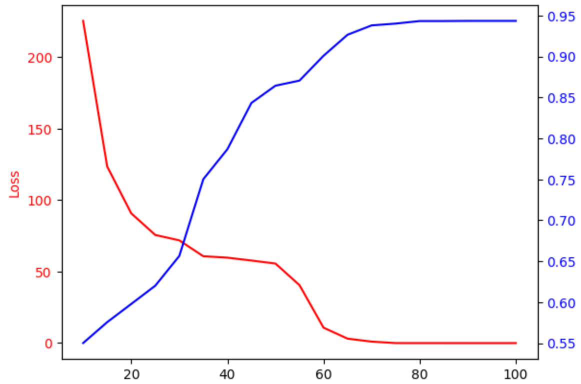

Figure 3. Training accuracy and Loss values

3.4 Testing

To confirm the quality of the newly trained ML model with 18 key points, evaluation is carried out by comparing the key points predicted by the ML algorithm to the validation data annotated by professional humans. Consistent with the state-of-the-art literature on the topic, several metrics were used to evaluate the performance of the ML model, such as a.) The difference in pixel distances between ML predicted and annotated key points b.) commonly used metrics such as Average precision (AP) and Average Recall (AR) based on Object Key point Similarity (OKS) [21] criteria, and c.) Computational performance in inference time, i.e., time taken to detect the key points from the given image.

Further, multi-stage experiments have been designed to evaluate the proposed REBAPose's comparative performance compared to several ML algorithms. In the first stage, the performance of Alphapose and MMPose algorithms is compared with other commonly adopted ML algorithms, such as the Detectron, Mediapipe, YOLO, Movenet, and PoseNet, for the original COCO dataset. Such a step is necessary to evaluate if state-of-the-art algorithms such as Alphapose and MMPose are the best choice, even for the added number of key points.

In the second stage, the predictions of the newly trained models and heuristic methods are compared. One heuristics method [8] computes the five key points as follows: The neck point is estimated as the midpoint of the ML algorithm's right and left shoulder prediction. The hip point is estimated by dividing the distance between the right and left hip points by 2. The midpoint of the head is estimated as the midpoint of the right and left eye points. Index finger values are derived by adding 0.02 times the diagonal length of the bounding box to the wrist point.

Testing is performed on unseen images that are not used for training the algorithm. The COCO dataset contains a set of images for evaluation and a default evaluator algorithm. However, a custom evaluation set and code were necessary for the current study for several reasons. First, the COCO validation dataset image contains many attributes, such as object detection, segmentation, and key points about human joints. Hence, not all images in the validation dataset can be evaluated; only images with humans become relevant. Out of 5000 images, only 2346 with 6352 persons could be used.

The default COCO validation dataset does not contain information about all the persons and the corresponding bounding boxes and key points. This is due to low-precision annotation done by a few annotators in the earlier annotation work. Few images in the COCO dataset have multiple humans. However, the COCO validation dataset has only a few persons annotated with key points and not all. Many modern algorithms can detect more humans than what is annotated in the default COCO validation dataset. The number of persons detected varies depending on the ML algorithm. For example, the *detectron* algorithm could detect about 11,000 persons, whereas the media pipe algorithm could detect only 2000 persons in the same dataset. Hence, only the images of 245 persons commonly detected by all the algorithms are used for comparative evaluation in the current study.

4 Results

4.1 Visual Comparison

Figure 4 visually compares the results obtained from the newly trained REBAPose models and the heuristic approaches. The three essential key points for REBA, i.e., the neck, midpoint of head, and center hip, as estimated by the heuristic approaches, are shown using the blue dots in Figure 4. The same key points estimated based on the newly annotated data are shown using the yellow dots.

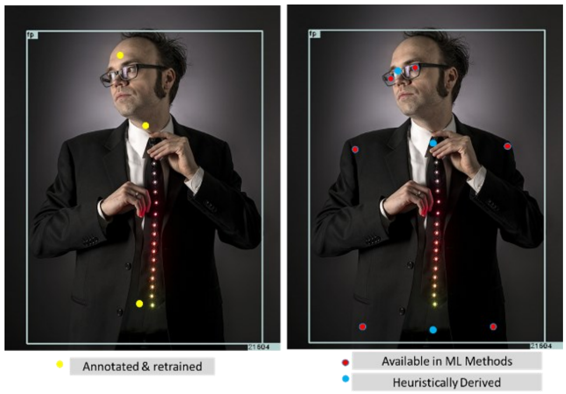

Figure 4. Visual representation of study results

4.2 Pixel Distance Comparison

A comprehensive heatmap of differences in pixel distances between ML predictions and annotated data for various algorithms for each of the body joints is shown in Figure 5. The data in Figure 5 shows the average results obtained for 245 persons (148 single-person images, 97 persons from images containing more than 1 person) from the COCO validation dataset. In Figure 5, except for the results of the two REBAPose architectures, i.e., REBAPose (Alphapose) and REBAPose (MMPose), the results for body parts (midpoint of head, neck, and center hip) from the remaining algorithms are estimated based on heuristics approaches, as described in section 3.4. For REBAPose (Alphapose) and REBAPose (MMPose), the results are obtained using retraining of the algorithm based on the newly annotated key points. Several important inferences can be made from these results.

First, state-of-the-art methods such as MMPose and Alphapose perform better than commonly adopted frameworks, even when only heuristics-based methods are considered. Such results reaffirm the technical superiority of these ML architectures.

The results for the heuristics method for the REBA essential key points, such as neck, midpoint of head, and center hip, further demonstrate that the heuristics-based methods for estimating the three-body parts result in large errors compared to the annotated ground truth, irrespective of the ML methodology applied.

Finally, the results demonstrate that Alphaspose and MMPose architectures, when applied to the newly annotated data, greatly improved the prediction accuracy for all three key points that are newly annotated and retrained, while there is a slight improvement on two index finger points which are retrained and derived COCO whole body dataset.

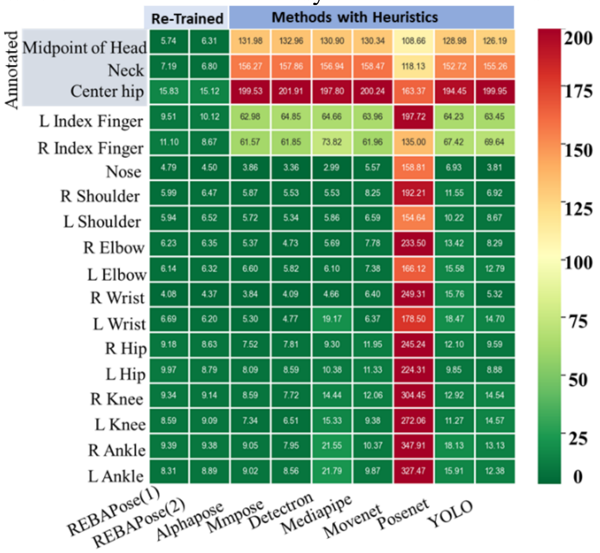

Figure 5. Pixel variation of detection and ground truth

4.3 OKS Analysis

OKS-based criteria are often used to assess how far or close the ML predictions are compared to the ground truth to estimate the performance of CV algorithms for pose estimations. OKS considers both the body part area and the image scale and is defined as a threshold value like 0.5 or 0.75. If the CV model predicted key point pixel is within the threshold distance governed by OKS criteria of the annotated pixel, the prediction is deemed *True Positive*. Similarly, information about false positives, false negatives, etc., can be estimated, and such matrices then help evaluate the necessary metrics, such as the Average Precision (AP) and Average Recall (AR), to understand the efficiency of pose estimation.

A higher OKS threshold is less permissive and requires key points to be more precisely aligned to be considered correct. The results from accuracy parameters at OKS thresholds 0.5 (loose metric), and 0.75 (strict metric), along with their Mean AP (Average of values at 0.5 and 0.95) are given in Tables 2 and 3.

Table 2. Average Precision Results

Methodology	mAP	AP @0.5	AP @0.75
REBAPose(A)	0.588	0.833	0.677
REBAPose(M)	0.530	0.800	0.584
Alphapose	0.224	0.733	0.015
MMPose	0.240	0.767	0.019

REBAPose (A) is based on Alphapose architecture, and REBAPose (M) is based on MMpose architecture.

Table 3. Average Recall Results

Methodology	mAR	AR@0.5	AR @0.75
REBAPose(A)	0.686	0.877	0.785
REBAPose(M)	0.642	0.858	0.723
Alphapose	0.316	0.817	0.094
MMPose	0.329	0.834	0.108

Results indicate that the extensive annotation and retraining on 5 out of 18 REBA key points done in the current study (for REBAPose (A) and (M) algorithms) has resulted in multi-fold accuracy improvement, which is expected for the critical WMSD assessment framework.

4.4 Computational Performance

For CV-based WMSD assessment, the inference time is important, as real-time estimation is often desirable. The quicker models can potentially be used even on mobile phones. Table 4 compares the speed of models.

Table 4. Inference Results

Methodology	Inference time in seconds
REBAPose(M)	0.2025
REBAPose(A)	0.0388
Detectron	0.1138
YOLO	0.0898
Mediapipe	0.0252
Movenet	0.0161
Posenet	0.0169

5 Discussion

5.1.1 Technical advantages of REBAPose

Accurate estimation of the key points (or the body parts) is one of the first steps toward an accurate REBA score calculation. if the body parts are not correctly predicted, the angles between them will have errors directly affecting the REBA score estimations. The REBAPose estimator developed in this study has significantly higher accuracy compared to conventionally adopted heuristics-based methods (see results in Figure 5, Tables 2 and 3). Due to annotation and retraining, there was improvement towards precise detection in pixel difference between detected and annotated ground truth for the neck, center hip, and midpoint of the head. For example, for the midpoint of head, REBAPose variation from ground truth was just about 5 pixels (Figure 5), while for other heuristic approaches such a variation was a few orders of magnitude apart at about 130 pixels (see Figure 5). Similar improvements can also be seen for the two other newly trained points as compared to heuristics approaches (Figure 5). Due to such a significant improvement in the accuracies of each of the three newly annotated points, the overall accuracy of REBAPose is also improved by about 100 % (see Tables 2 and 3).

Another important parameter to measure the performance of ML models is inference time. Results in Table 4, indicate the inference time of all models. Results indicate that the REBAPose (Alphapose), despite the high accuracy does not suffer from longer inference time and is the best model regarding both improved accuracy and inference time.

5.2 Key academic contribution

One of the essential academic contributions of the current study is the creation of a unified and comprehensive dataset that can be readily used for WMSD risk-assessment tasks, especially using the REBA method. Such a task has been achieved by developing an integrated annotated dataset comprising 13 key points from the original COCO dataset (Table 1), 2 key points annotated from the COCO whole body dataset (Left and Right Index figures), and 3 newly annotated key points (neck, midpoint of head and center hip) for about 1,50,000 images available in the COCO dataset. Previous studies have not carried out such an extensive annotation. Previous studies have retrained the datasets with only a limited number of images.

REBAPose essentially removes the heuristic approach on three body parts neck, midpoint of head, and center hip, and hence provides accurate angle measurements forming the basis for REBA scoring. Because of annotation and retraining, the accuracy of all points improved.

The extensively annotated dataset can be used for training and testing future CV-based algorithms for higher technical performance. Moreover, the REBAPose estimator developed as part of the current study can be used as a pre-trained model for accurately estimating the human postures for REBA scores in various fields where there is a concern about the OHS performance of their workers.

The newly annotated information about the center hip and the neck can enable a more accurate estimation of the spine, which is necessary for correcting yoga postures or for medical researchers treating spondylitis.

5.3 Limitations and Future Work

While the currently developed REBAPose model is industry agnostic, its accuracies can be further evaluated for specific trades. For example, future studies could test the applicability of the REBApose models for construction-specific datasets such as the MOCS datasets [23]. The present study focussed only on 2D pose estimation and did not address 3D pose improvements which may be covered in future work. As the REBA score depends on angles between body parts, 3D pose estimation is expected to get more precise angles than 2D points.

In the current study, human-annotated data was considered as the ground truth to estimate the accuracy of the work. However, the accuracy of these human annotations must be validated. The data from actual safety practitioners, such as industrial hygienists, could also be collected and compared against the predictions made by REBAPose. REBA scoring mechanism involves computing three scores A, B and C and these scores rely on angles formed between body parts. Since the present study can accurately detect body parts (key joints), it is likely to find precise angles that are formed between body parts. So present study is likely to have better REBA score calculations to reflect real ground conditions. However, the impact of an improved algorithm for enhanced key point detection performance on the enhancement of the accuracy of the REBA score needs to be validated in future studies.

The human pose estimation is just one of the components in REBA score calculation. Several other factors cannot be determined using a CV-based pose estimation technique. Examples include the weight of the objects held by personnel and the coupling factors. In future studies, this aspect will be considered.

6 Conclusions

Existing approaches for WMSD risk assessment using REBA rely on readily available pre-annotated datasets. While such approaches using heuristics for estimating REBA scores are easy and quick, they lack precision. This study aims to address this gap by developing new REBAPose estimators. More than 150,000 images selected from the COCO dataset were re-annotated to add 3 essential key points for calculating the REBA score. The REBAPose estimators in this study use state-of-the-art ML architectures, including Alphapose and MMPose. The proposed approach results in a 200% improvement in mAP and mAR values compared to predictions made using the heuristics approaches (see Tables 2 and 3). The inference time of the REBAPose estimators is also comparable, making them a fast and accurate method for pose estimations in future studies.

Acknowledgments

We acknowledge the contributions of Hriddhi Mondal, Antony Samuel, and Likhith Kumara of Tata Consulting Services (TCS) for building and deploying annotation UI. We are also thankful for the exploratory research grant provided by IIT Madras (Project Number

RF22230544CERFER008976).

Data Availability

The dataset used in the current study can be made available for academic use upon written request to authors.

References

[1] K. Bhattacharjee, N. Bugalia, and A. Mahalingam, "An analysis of safety practices for small, medium, and large construction projects: A resilience engineering perspective," Saf Sci, vol. 169, Jan. 2024, doi: 10.1016/j.ssci.2023.106330.

[2] N. Bugalia, V. Tarani, J. Kedia, and H. Gadekar, "Machine Learning-Based Automated Classification Of Worker-Reported Safety Reports In Construction," Journal of Information Technology in Construction, vol. 27, pp. 926–950, 2022, doi: 10.36680/j.itcon.2022.045.

[3] W. Yi and A. Chan, "Health profile of construction workers in Hong Kong," Int J Environ Res Public Health, vol. 13, no. 12, Dec. 2016, doi: 10.3390/ijerph13121232.

[4] X. Wang, X. S. Dong, S. D. Choi, and J. Dement, "Work-related musculoskeletal disorders among construction workers in the United States from 1992 to 2014," Occup Environ Med, vol. 74, no. 5, pp. 374–380, May 2017, doi: 10.1136/oemed-2016-103943.

[5] D. Kee, "Systematic Comparison of OWAS, RULA, and REBA Based on a Literature Review," International Journal of Environmental Research and Public Health, vol. 19, no. 1. MDPI, Jan. 01, 2022. doi: 10.3390/ijerph19010595.

[6] T. Jones and S. Kumar, "Comparison of ergonomic risk assessment output in four sawmill jobs," International Journal of Occupational Safety and Ergonomics, vol. 16, no. 1, pp. 105–111, 2010, doi: 10.1080/10803548.2010.11076834.

[7] S. Pinzke and L. Kopp, "Marker-less systems for tracking working posturesFresults from two experiments," 2001.

[8] S. O. Jeong and J. Kook, "CREBAS: Computer-Based REBA Evaluation System for Wood Manufacturers Using MediaPipe," Applied Sciences (Switzerland), vol. 13, no. 2, Jan. 2023, doi: 10.3390/app13020938.

[9] O. Karhu, P. Kansi, and I. Kuorinka, "Correcting working postures in industry: A practical method for analysis," 1977.

[10] S. Hignett and L. M. Ergonomist, "Rapid Entire Body Assessment (REBA)," 2000.

[11] L. Mcatamney and E. N. Corlett, "RULA: a survey method for the investigation of work-related upper limb disorders," 1993.

[12] T. Jiang et al., "RTMPose: Real-Time Multi-Person Pose Estimation based on MMPose," Mar. 2023, [Online]. Available: http://arxiv.org/abs/2303.07399

[13] H.-S. Fang et al., "AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time," Nov. 2022, [Online]. Available: http://arxiv.org/abs/2211.03375

[14] T.-Y. Lin et al., "Microsoft COCO: Common Objects in Context," May 2014, [Online]. Available: http://arxiv.org/abs/1405.0312

[15] M. Andriluka, L. Pishchulin, P. Gehler, and B. Schiele, "2D human pose estimation: New benchmark and state of the art analysis," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, Sep. 2014, pp. 3686–3693. doi: 10.1109/CVPR.2014.471.

[16] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu, "Human3.6M: Large scale datasets and predictive methods for 3D human sensing in natural environments," IEEE Trans Pattern Anal Mach Intell, vol. 36, no. 7, pp. 1325–1339, 2014, doi: 10.1109/TPAMI.2013.248.

[17] J. Wu et al., "AI Challenger : A Large-scale Dataset for Going Deeper in Image Understanding," Nov. 2017, doi: 10.1109/ICME.2019.00256.

[18] S. Jin *et al.*, "Whole-Body Human Pose Estimation in the Wild," Jul. 2020, [Online]. Available: http://arxiv.org/abs/2007.11858

[19] J. Cohen, "A coefficient of agreement for nominal scales," Educ Psychol Meas, vol. 20, no. 1, pp. 37–46, 1960.

[20] Z. Zhao, A. Song, S. Zheng, Q. Xiong, and J. Guo, "DSC-HRNet: a lightweight teaching pose estimation model with depthwise separable convolution and deep high-resolution representation learning in computer-aided education," International Journal of Information Technology (Singapore), vol. 15, no. 5, pp. 2373–2385, Jun. 2023, doi: 10.1007/s41870-023-01297-5.

[21] "COCO evaluation." Accessed: Dec. 14, 2023. [Online]. Available: https://cocodataset.org/#keypoints-eval

[22] Y. Yang and D. Ramanan, "Articulated human detection with flexible mixtures of parts," *IEEE Trans Pattern Anal Mach Intell*, vol. 35, no. 12, pp. 2878–2890, 2013, doi: 10.1109/TPAMI.2012.261.

[23] A. Xuehui, Z. Li, L. Zuguang, W. Chengzhi, L. Pengfei, and L. Zhiwei, "Dataset and benchmark for detecting moving objects in construction sites," Autom Constr, vol. 122, Feb. 2021, doi: 10.1016/j.autcon.2020.103482.