Post-Lift Analysis of Thermal Imprint for Weight and Effort Detection

Austin Dykeman
dykemaa@clarkson.edu
Clarkson University
Potsdam, NY, USA

Joseph Judge
judgejc@clarkson.edu
Clarkson University
Potsdam, NY, USA

Priyo Ranjan Kundu Prosun
prosunp@clarkson.edu
Clarkson University
Potsdam, NY, USA

Gurpreet Kaur
gukaur@clarkson.edu
Clarkson University
Potsdam, NY, USA

Owen Talmage
otalmage@delsys.com
Delsys
Natick, MA, USA

Sean Banerjee
sbanerje@clarkson.edu
Clarkson University
Potsdam, NY, USA

Natasha Kholgade Banerjee
nbanerje@clarkson.edu
Clarkson University
Potsdam, NY, USA

ABSTRACT

Warehousing and manufacturing environments are seeing a proliferation of various body-mounted and environment biosensing mechanisms to assess elements of worker activity indicative of signs of fatigue. In this work, we provide one of the first investigations of thermal imaging technologies to perform automated detection of object weight and subject perceptions of weight using thermal imprints left behind upon interactions with objects of varying masses. We use data recorded from multi-viewpoint thermal cameras when subjects lift and lower weighted cartons with awareness of the weight, as well as blind, i.e., without prior weight knowledge. We demonstrate statistically significant differences across weight classes in mean thermal intensity of thermal imprints on regions of interest extracted around packages in post-lift thermal images for 24 subjects performing 10 randomized interactions each with 0 lb, 15 lb, 30 lb, and 45 lb weights. Using convolutional neu-ral networks trained to be subject independent, we show average accuracies of 75.83% and 70.25% for weight and subject-reported strenuousness detection for weight-blind lifts, and accuracies of 73.84% and 68.92% for weight-aware lifts. Our work plays an impor-tant role in strengthening multi-modal algorithms for monitoring workers for signs of fatigue.

CCS CONCEPTS

• Applied computing → Consumer health; • Computing methodologies → Neural networks.

KEYWORDS

healthcare artificial intelligence, post-lift, thermal imprint, fatigue, deep networks

ACM Reference Format:
Austin Dykeman, Joseph Judge, Priyo Ranjan Kundu Prosun, Gurpreet Kaur, Owen Talmage, Sean Banerjee, and Natasha Kholgade Banerjee. 2022. Post-Lift Analysis of Thermal Imprint for Weight and Effort Detection. In ACM/IEEE International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE' 22), November 17–19, 2022, Washington, DC, USA. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3551455.3559606

CHASE' 22, November 17–19, 2022, Washington, DC, USA

ACM ISBN 978-1-4503-9476-5/22/11…$15.00
https://doi.org/10.1145/3551455.3559606

1 INTRODUCTION

Workers in blue collar environments such as manufacturing and warehousing are at risk for work-related musculoskeletal disorders (WMSDs) due to engagement in repetitive lifting and lowering interactions with heavy objects [5, 6, 11, 21]. Current manufacturing and warehousing environments are seeing an increase in a variety of sensing technologies for tasks such as detection of sub-optimal posture and fatigue, including wearable biosensors to measure, e.g., heart rate or electromyography (EMG) signals [1, 13, 14, 19, 20], and at-a-distance monitoring through off-the-shelf cameras [24, 25].

Much existing work on fatigue monitoring largely tends to rely on fatigue predictors extracted from physiological or biomechanical measurements, and use subject perceptions as a validation tool, rather than incorporating subject perceptions directly in fatigue monitoring. Subject perceptions of fatigue, especially in long-term activity, may not necessarily align with short-range parameters drawn from physiological measurements. Additionally, the existing work performs an overall detection of fatigue, rather than learning-driven detection of finer parameters of the task, e.g., weight lifted. Given the potentially complex dependency of subject perceptions of the task on the task type, task parameters, and subject demographics, task parameters and subject task perceptions should be taken into account in decision-making for fatigue management.

In this paper, we investigate the potential of thermal cameras to facilitate worker activity monitoring by using them to perform

automated non-intrusive detection of weight lifted and subject perceptions of effort performed during lifting of weight as subjects interact with objects of varying masses. Thermal cameras with sensitivity in the human temperature range offer considerable benefits in person activity tracking [8], due to well-defined intensity differences between human targets and background environments. However, clear benefit of thermal sensors is manifested if multiple tasks can be performed via a single sensor, e.g., if thermal-based human tracking in a warehouse environment may be coupled with detection of signs for fatigue.

We leverage the capability of thermal sensors to record thermal imprints on objects in the environment as an indirect mechanism to monitor the potential for fatigue by automatically detecting the weight lifted and the subject's perceived effort in lifting the weight. By performing assessment after the interaction, rather than during the process, our approach enables providing information on the extent to which an individual should continue performing a par-ticular or similar activity based on how strenuous they perceive the activity just performed. Through automated continuous moni-toring of weight lifted in warehouse environments, our work can be incorporated into holistic artificial intelligence (AI) algorithms that detect the potential for workers to be fatigued. Our work can also be used to detect need for assistance through imprints after cursory attempts on lift.

Our work analyzes the thermal imprint left behind after interactions in the form of lifting, lowering, and releasing identical packages in four weight classes—0 lb, 15 lb, 30 lb, and 45 lb. We assess the potential for post-lift data to inform on weight and subject-reported effort by analyzing thermal images of hand-prints left behind after box release due to heat transferred from the hands of the subject to the box. Our hypothesis is that thermal images show higher intensities when individuals grasp and lift heavier objects based on higher heat transferred from greater pressure application and longer duration of interaction. We evaluate the hypothesis by analyzing the relationship between object weight and mean thermal imprint intensity on the box surface, and conclude that inter-class intensity differences are statistically significant. Based on our findings, i.e., that a positive relationship exists between weight and thermal intensity, we provide convolutional neural networks (CNNs) that predict weight and effort in the form of subject-reported strenuousness during lifting from the input thermal imprint images. We provide an average accuracy of 75.83% and 70.25% on detection of weight and effort when the subject lacks weight knowledge, and ac-curacy of 73.84% and 68.92% when the subject has prior knowledge of weight being lifted.

2 RELATED WORK

2.1 Fatigue Detection in Industrial Settings

A number of approaches use wearable sensors to perform fatigue detection; a systematic review may be found in Ahn et al. [1]. Maman et al. [20] use a chest-strapped heart rate monitor and four inertial measurement unit sensors attached on the right ankle, right wrist, hip, and torso for physiological data collection. Similar to our experiments, they request participants to pick and load 10 kg, 18 kg, 26 kg cartons onto a dolly and transport and palletize the cartons at a destination. They hand-define features indicative of fatigue, e.g., heart rate reserve (HRR), acceleration, and jerk, and use logistic and linear regression to respectively detect fatigue and predict its level. Maman et al. [19] builds upon this work to make recommendations on using wearable sensors to detect fatigue. Unlike our work that predicts subject ratings, Maman et al. use subject ratings only to validate results.

Jebelli and Lee [14] perform a similar comparison of subject perceptions to features from electromyography (EMG) sensors. A challenge remains that while hand-crafted features from independent sensors may align with subject perceptions of fatigue under lab-specific settings, they may deviate under complex tasks or long duration activity. Jebelli et al. [13] use a wearable wrist biosensor that acquires multiple physiological signals, and provide 87% accuracy in distinguishing low-intensity (standing and talking), medium-intensity (cleaning, finding tools, and cutting sheets), and high-intensity (drywall install and heavy material movement) activities. The method focuses on activities that involve significant postural changes, rather than activities involving similar posture.

Papoutsakis et al. [24, 25] present the sustAGE system that provides 70% accuracy in classifying ergonomically sub-optimal postures from OpenPose [4] skeletons estimated from RGB videos of line workers using graph convolutional networks. They also use heart rate sensors to demonstrate a correlation between estimated risk for strain and increase in heart rate. Their method focuses on differences in posture, rather than variations in weight lifted or in predictions of subject perception of effort in situations where whole-body posture is unlikely to demonstrate separability, e.g., lifting and lowering in the same manner.

2.2 Using Thermal Imprints on Object Surfaces

Traditionally, thermal sensors have been used for applications such as person and animal identification and tracking, and monitoring of buildings and food quality, where the target displays obvious differences in temperature from the environment [8]. A number of approaches have investigated the benefit of using the thermal imprint that are left behind through human interactions on the environment. Work exists to enable natural surface interaction by tracking left-behind thermal imprint, with the work spanning tracking of finger motion across the surface and classification of swipe pressure [7, 12, 16, 17, 22, 23, 27]. Work has also been performed on detecting material characteristics based on thermal imprint properties, such as strength of imprint and imprint dissipation delay [9] and frictional force [28]. Recent work has used thermal imprints on object surfaces to detect human hand hold on cups [15], and to capture and predict grasp profiles and hand pose on objects [2, 3].

3 DATA COLLECTION

3.1 Capture Setup

To capture thermal information on subject interaction with objects of different weights, we use a set of three Sierra Olympic Viento-G Gig-E Vision thermal cameras placed along the east, west, and north directions of a circular space of diameter 22 feet. Each camera has a spatial resolution of 640×480, and captures images at a frame rate of 30 frames per second (FPS). To ensure continuous capture with minimal frame drops, we disable fast field correction on the thermal cameras. We use a 9mm lens with each camera, and store

16-bit PNG images. The cameras record auto-white-balanced non-radiometric thermal images, with typical thermal intensity ranges between 4,900 to 5,600.

We connect each thermal sensor to a custom-built computer with an AMD Ryzen 2700X processor, 16GB of RAM, 4.5TB of SSD storage, and an NVIDIA GTX 1080Ti GPU. The three computers act as subordinate capture computers, and are controlled via a central control machine via a C# command-line application developed at our lab. The command-line application enables user-driven real-time capture and storage of 16-bit PNG images. We use individual computers per camera to minimize latency and frame drops that are likely to arise as more cameras are connected to the same computer.

Each thermal camera is paired with an RGB-D sensor on the same computer, namely a Microsoft Azure Kinect. We use a DHCP server to perform initial synchronization of each computer's internal clock. Within each computer, we synchronize the thermal and RGB-D sensors by recording millisecond epoch timestamps for images captured at each time instant by the two cameras. Across the color cameras of the RGB-D sensors, we program an overhead green light to automatically flash on and off twice prior to the start of a subject's interaction. We use inter-frame subtraction, thresholding, and averaging to obtain the rising and falling edges of light flash, and detect the synchronization frame as that for the second falling edge. While the thermal cameras have an external trigger, relying on a hardware trigger mechanism may prevent proliferation to commercial off-the-shelf cameras, e.g., StereoLabs ZED which lacks a hardware trigger and instead uses Precision Time Protocol (PTP) to synchronize across multiple sensors.

3.2 Subject Recruitment

To recruit subjects, we sent out a general email to students, faculty, and staff at Clarkson University, United States, under IRB number 21-03. We requested each subject to visit the capture environment in two sessions to perform lifting, lowering, and release interactions with cartons of varying weights. We requested that subjects perform the sessions be separated by at least 24 hours. In one of the sessions, prior to each interaction with a carton, the subject was provided knowledge of the weight lifted. In the other day, the subject received no knowledge of the weight lifted prior to the interaction. We refer to the sessions with and without weight knowledge as 'non-blind' and 'blind' sessions respectively. We randomized the order of non-blind and blind sessions across subjects. Each session was completed within 90 minutes. 28 subjects visited the lab for capture. We acquired informed consent from each subject before the study. Upon completing both sessions, each subject was compensated in the monetary amount of $25.

3.3 Study Protocol

During each session, the experimenter informed each subject that they would interact with four weights—0 lb, 15 lb, 30 lb, and 45 lb—, with 10 presentations per weight, yielding 40 weight presentations. All 40 weight presentations were randomized across the subject. We varied randomizations across subjects and across sessions. Each weight was presented as a carton of size 18"×12"×4" containing enough dumbbells to sum up to the corresponding weight class. The experimenter used a dolly to facilitate movement of cartons between presentations. To ensure that visual, auditory, vibrational, or temporal aspects of motion did not provide clues to the subject on the weight, the experimenter requested the subject to step away from the interaction location, and moved the dolly with consistent pressure and movement rate, determined by the maximum weight. The experimenter positioned the dolly at the center of the capture point, such that it was nearly equidistant from the three thermal cameras, locked the casters to restrict motion, and requested the subject to stand behind the dolly facing the north camera. In the non-blind session, the experimenter provided prior knowledge of the weight to the subject.

Subject directions for interaction performance consisted of lifting the weight, holding it for around 2 seconds or as comfortable, lowering the weight back on the dolly, and releasing after lowering. No other instructions were provided, e.g., the subject was free to use any stance, grip on the object, and time duration of lift, hold, lowering, and release. The experimenter also provided recommendations on safe posture during the interaction. All but one interaction was performed using a two-handed grasp, and the exception was performed using a one-handed grasp. After the interaction, the experimenter requested the subject to rate how strenuous they found the interaction on the Likert scale from 1 to 5, with 1 being least strenuous and 5 being most strenuous. In the blind session, the subject was requested to provide a post-guess of weight. The experimenter ensured that adjacent captures were spaced by no less than 90 seconds to permit heat diedown and return of the box to the ambient temperature. We obtained the heat diedown duration through empirical testing prior to the start of our data collection.

We eliminated data from 4 subjects due to issues with capture, storage, and data loss, and retained data from the remaining 24 subjects as our thermal dataset. For each interaction, we manually marked the start and end times of the lift as the timestamps of the color images in the north-facing camera corresponding to first hand touch and first hand release. For each color image, we use the color-thermal synchronization and cross-color synchronization to propagate start and end times to all thermal images. With 24 subjects, 2 sessions, and 40 captures per subject per session, we obtained data for 1,920 interactions. At 3 views per interaction, we obtained 5,760 thermal videos. Figure 1 shows sample thermal images from the three cardinal directions as a subject performs a lifting, lowering, and release action with a box. Time instant '0 seconds' corresponds to the manually marked location of start at first hand release. The last three frames demonstrate the hand imprint left behind on the carton after release, especially in the left and right views of the subject.

4 EXTRACTING POST-LIFT THERMAL DATA

We obtain post-release thermal images by searching past the manually labeled hand-release location for a thermal frame where the subject has moved their hand away from the box. For most captures, we find that after 1 second, the subject hand shows no overlap with the box pixels in the thermal image. For a few captures, we find the overlap-free point to be at 1.5 second after first-hand-release. Typically, in these captures, the subject is still struggling with releasing the box with both hands, e.g., if the box is too heavy and placement on the dolly is challenging.

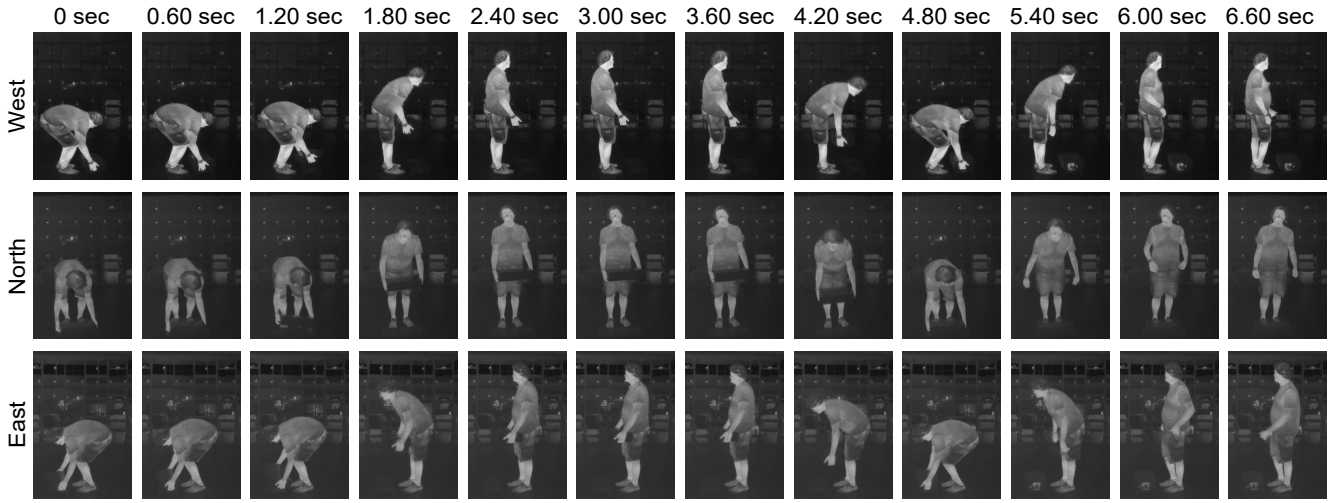

Figure 1: Frames from west, north, and east viewpoints showing Subject 7 performing lift, lowering, and release of a 30 lb carton.

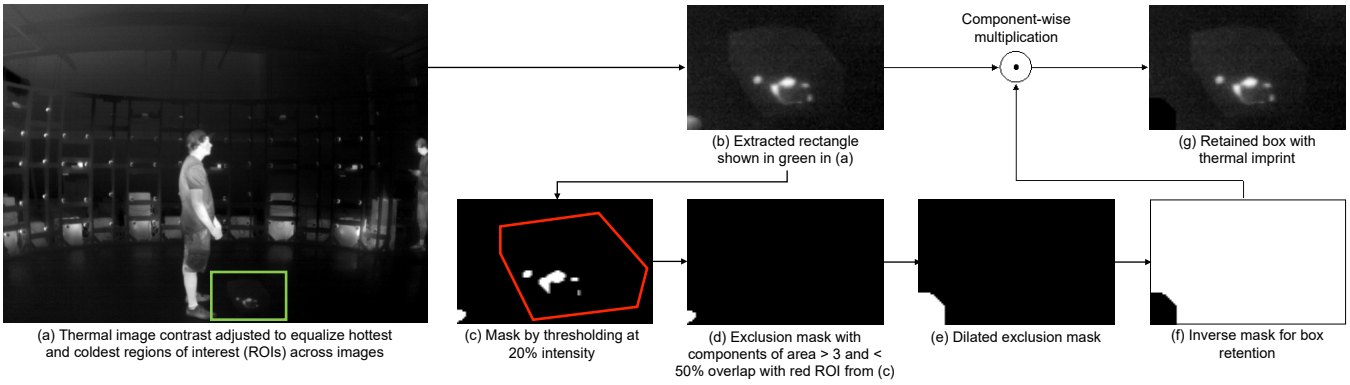

Figure 2: Flowchart demonstrating approach to extract thermal frames with hand imprint for the right viewpoint. The same approach is used for the left viewpoint.

The Sierra Olympic Viento-G cameras perform automated white balancing, a common setting for lower-end thermal imagers. For each thermal image, we correct white balancing by using a reference image to mark the hottest and coldest regions, and scaling the thermal image in each capture linearly to equalize the hottest and coldest regions across all captures. In our capture environment, the hottest region tends to be a thermal camera belonging to the opposite view, while the coldest region tends to be the floor. Figure 2(a) shows a thermal frame after performing linear contrast adjustment.

Given the thermal image in each capture, we perform a close crop of the image around the region of the box by using the reference image to define the crop region, and cropping thermal frames in all captures to match the reference region. Figure 2(b) shows an example of a close crop. Since subject feet are often close to the box, we use the reference frame to define a mask region around the box shown in red in Figure 2(c). To ensure high-fidelity retention of hand prints without cut-off, we perform a binary thresholding of the image as shown in Figure 2(c), and obtain all connected components, which largely tend to be imprints and feet. We keep those components that have an area larger than 3 square pixels, and have a low overlap with the reference mask of less than 50% the size of the component. The process largely eliminates subject legs and feet and retains imprints within the vicinity of the reference mask as shown in Figure 2(d). We perform a final morphological dilation operation to remove mid-range components on the body. We use the retained components to mask the thermal image.

5 EVALUATING RELATIONSHIP BETWEEN THERMAL INTENSITY AND WEIGHT

Figure 3 shows example thermal images for various weights on the blind and non-blind days for Subject 7 as generated by our approach. Subject hand prints belonging to lower weights, especially 0 lb, typically show lower intensities. The increase in thermal intensity of imprint with increasing weight may be attributed to

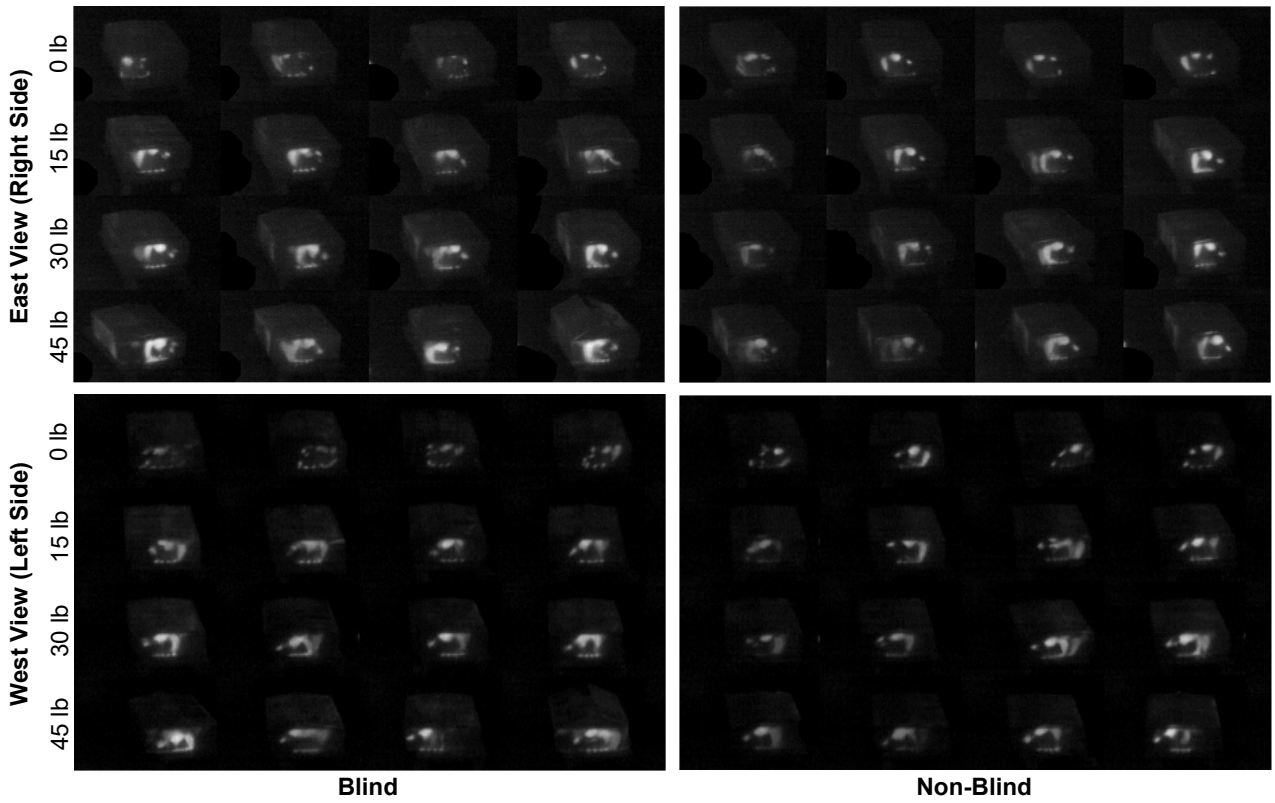

Figure 3: Thermal images upon box release for 4 lifts for the 4 weight classes from the east (right) and west (left) viewpoints from blind and non-blind days for Subject 7.

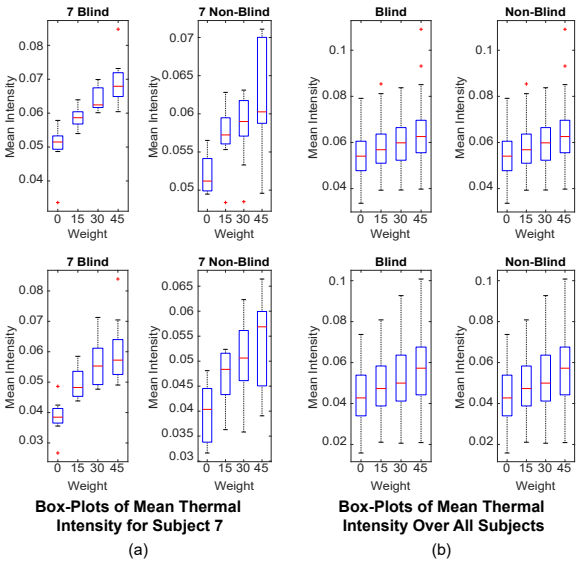

Figure 4: Box plots of mean thermal intensity versus weight for (a) Subject 7 and (b) all subjects.

increased heat transferred to the object due to longer interaction times during lifting and lowering, and to higher pressure. In Figure 4, we show box plots of mean thermal intensities for Subject 7 and for all subjects The box plots indicate positive dependence of mean thermal intensity with increase in weights.

We use Friedman's Test with Nemenyi post-hoc test to determine if there is a statistically significant difference between intensity for each weight class. The null and alternative hypothesis are:

H_0: There is no difference in mean thermal intensity for pairs of weight classes.

H_a: One or more pairs of weight classes have differences in mean thermal intensity.

We use the R programming language to conduct the statistical test and we reject the null hypothesis overall and in pairwise tests. The Friedman test provides a p-value < 2.2e-16, and the Nemenyi test provides p-values of 1.4e-13, < 2.2e-16, < 2.2e-16, 2.8e-5, 4.3e-14, and 0.0011 for 0-15, 0-30, 0-45, 15-30, 15-45, and 30-45 comparisons using mean thermal intensity from blind lifts. For mean thermal intensity from non-blind lifts, the Friedman test provides a p-value < 2.2e-16, and the Nemenyi test provides p-values of 1.6e-13, < 2.2e-16, < 2.2e-16, 7.6e-6, 3.2e-14, and 0.00058 for 0-15, 0-30, 0-45, 15-30, 15-45, and 30-45 comparisons.

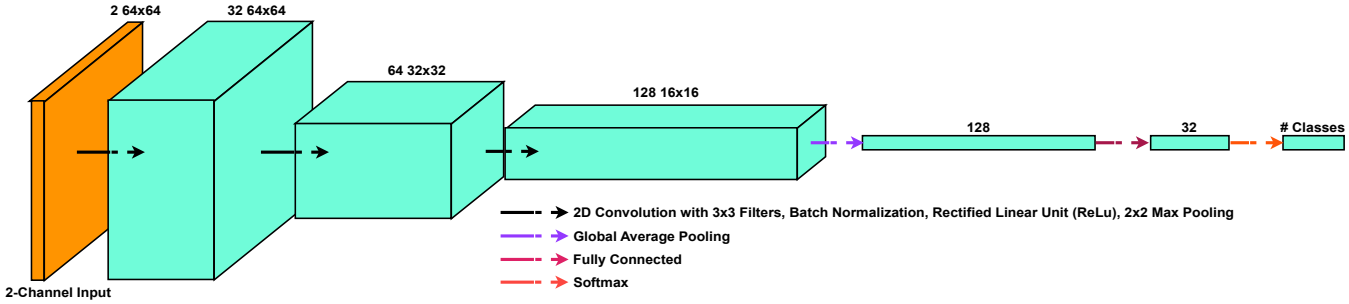

Figure 5: 2D CNN used to classify thermal images. The 2-channel input takes in thermal images from the east and west views concatenated together and resized to 64×64. The outputs are two or more classes for weight, strenuousness, or knowledge of weight.

6 USING NEURAL NETWORKS TO DETECT WEIGHT AND EFFORT FROM IMPRINTS

Based on our findings of statistically significant differences between mean intensities of different weight classes, we investigate the ability of learning algorithms to leverage the intensity differences for prediction of weight. We use 2D convolutional neural networks (CNNs) to enable self-learning of relevant features from raw data. We use the 2D CNNs to predict 4-class weight, i.e., the 4 weight categories studied in this work of 0 lb, 15 lb, 30 lb, and 45 lb. We found that subject-provided guesses on weight in the blind session were confounded in 34 instances over the higher two classes. We expect CNNs to similarly confound adjacent weight classes. We evaluate the predictive ability of coarser CNNs in two forms

- 3-class weight. We train CNNs to predict a 'Low' weight class assigned 0 lb weights, a 'Mid' weight class assigned 15 lb weights, and a 'High' weight class assigned 30 lb and 45 lb weights.
- 2-class weight. We train CNNs to predict a 'Low' weight class assigned 0 lb and 15 lb weights, and a 'High' weight class assigned 30 lb and 45 lb weights.

Our data collection reveals that subjects tend to predict high strenuousness for higher weight classes, and lower strenuousness for lower weight classes. In addition to weight prediction, we investigate the capability of the neural networks to predict effort invested as subject-reported strenuousness. We observe a class imbalance in strenuousness ratings on the 1 to 5 scale, with subjects largely providing ratings toward the lower end of the scale. We handle the imbalance by predicting 2-class strenuousness, where ratings of 2 and below were assigned to the 'Low' class, and ratings of 3 and above were assigned to the 'High' class.

We were also interested in identifying whether thermal imprints may reveal if the subject had prior knowledge of weight, e.g., based upon the location or smearing of the imprint if an adjustment had to be performed to account for a compensation of interaction without weight knowledge. We show results of training CNNs to predict the subject's prior weight knowledge, where ground truth for knowledge comes from the non-blind sessions, and ground truth for lack of knowledge comes from blind sessions.

Figure 5 shows the neural network architecture that we use in this work. The 2-channel input to the network takes in images from the east and west views processed according to Section 5 and resized to be of size 64×64. The architecture consists of three blocks that each perform 3× convolution with stride 1 and padding to retain image size, batch normalization, rectified linear unit (ReLU) activation, and 2×2 max pooling. The three blocks consist of 32, 64, and 128 convolutional filters respectively. Outputs from the third block are accumulated via global average pooling to form a 128-dimensional feature vector that is subjected to a dense layer that generates a 32-dimensional vector, followed by a softmax layer with as many nodes as the predicted number of classes. The predicted class corresponds to the one with the highest softmax output.

We provide user-agnostic CNNs where the subjects from the test set are not present in the training set, and user-aware CNNs where data from a subject is distributed across the training and test set. The purpose of the user-agnostic CNNs is to enable generalization to novel subjects without need for retraining. The user-aware CNNs are useful in situations where prior knowledge of the user is available to an AI monitoring system, e.g., in a warehouse environment where employees may go through a registration phase and provide input data about lifting patterns.

Our primary goal is to investigate predictive ability when the user lack awareness of the weight lifted, as packages in warehousing environments may lack information on the mass of the internal contents. As such we conduct a set of predictions that use training and test data solely from the blind session. We are also interested in analyzing predictive ability if a subject has access to information on the weight lifted, e.g., if package contents can be seen, or if package labels and/or warehouse organization provide clues to weight. For this, we conduct a second set of predictions that use training and test data solely from the non-blind session. For baseline comparison, we also provide CNNs that use training and test data mixed from the two sessions. We do not provide weight knowledge detection for blind only and non-blind only CNNs.

We perform 4-fold classification for user-agnostic CNNs and 2-fold classification for user-aware CNNs. With thermal data from 24 subjects, we have 6 subjects per fold for user-agnostic CNNs and 12 per fold for user-aware CNNs. We assign subjects to folds at random. We perform ablation testing over batch sizes of 4, 8, and 16, learning rates of 1e-4 and 1e-5, and number of epochs of 25, 50, 75, and 125, and obtain results at the maximum accuracies for each prediction output. The accuracy is obtained by counting the number of true positives within each class and averaging over all classes.

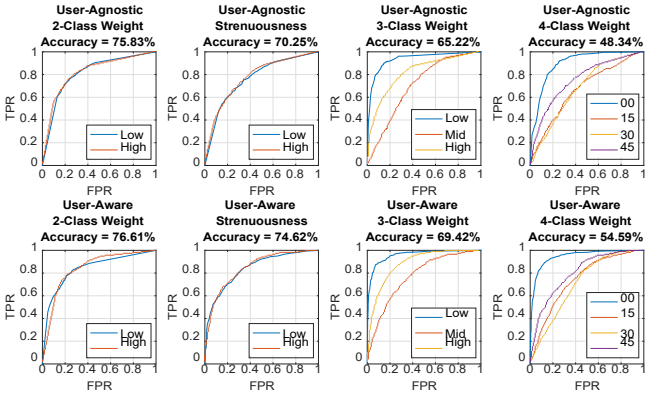

Figure 6: ROC curves for user-agnostic and user-aware prediction of weight classes, strenuousness, and awareness of weight using images on post-lift thermal signature from blind data only.

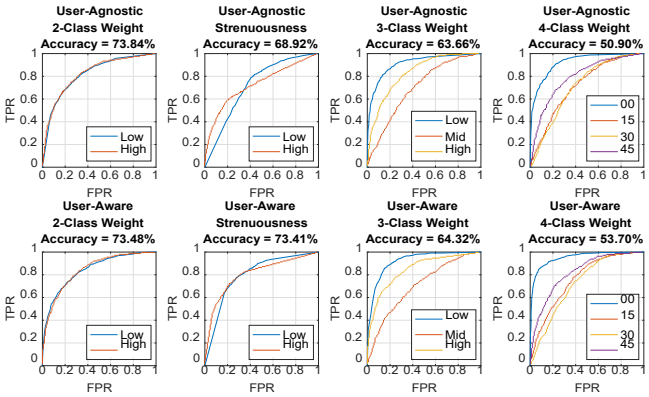

Figure 7: ROC curves for user-agnostic and user-aware prediction of weight classes, strenuousness, and awareness of weight using images on post-lift thermal signature from non-blind data only.

We generate an ROC curve for each class by varying the threshold above which samples for a class are labeled as belonging to the class, computing the true positive rate (TPR) for the class, obtaining the false positive rate (FPR) as the number of times that samples from other classes are classified as belonging to the class in question, and plotting the TPR against the FPR at varying thresholds.

7 RESULTS

We show ROC curves for each type of output and class using blind data in Figure 6, non-blind data in 7, and mixed data in Figure 8. We list the maximum accuracy in the title for each plot. Using user-agnostic CNNs trained and tested with blind data, we observe accuracies of 75.83% and 70.25% for 2-class weight and strenuousness prediction, and 65.22% and 48.34% for 3- and 4-class weight prediction. As shown by the confusion matrices for blind data in Figure 9, 2-class weight prediction shows accuracies of 77.3% and 74.3% for the low and high weight classes. Strenuousness predictions are weighted more toward the lower class, where for more examples, samples from the higher class are classified as belonging to the lower class. We see this phenomenon recur in the 3- and 4-class weight prediction, where the lowest weight class is predicted with higher accuracy rising to 92% for 0 lb weight in 4-class prediction. We observe that the greatest misclassification occurs for the 30 lb weight class in 4-class weight, with most samples misclassifying as belonging to adjacent weight categories. Building in user-awareness improves classification accuracy, with greatest improvements seen for the finer weight CNNs, at 69.42% and 54.59% for 3- and 4-class weight prediction.

Comparing classifiers trained on blind and non-blind data separately, we observe that average accuracies using blind data are generally higher than using non-blind data, with the exception being 4-class weight prediction. Comparing the confusion matrices for blind and non-blind data across Figures 9 and 10, we observe that prediction for 2-class networks is biased against the higher class, whereas predictions are more balanced for the networks that use blind data. With higher weight categories, we observe a reverse trend for user-agnostic prediction. Accuracies for most higher weight classes with non-blind data are higher at the risk of the lower weight class. User-aware predictions are biased toward the lower class, with lower predictions for most higher weight classes in comparison to networks trained and tested with blind data.

A possible explanation for higher general performance using blind data may be attributable to the subject being unbiased by prior knowledge of weight, where effort performed depends largely on performing the lift optimally based on the subject's immediate perception of the box weight on grasping. This behavior may manifest as separable thermal signatures on the box. On the non-blind day the subject may be biased by the prior knowledge and their own perceived ability, and perform a sub-optimal grasp if they believe they can easily lift a certain weight. The sub-optimal grasp may generate weaker thermal signatures if, due to mis-perceived ability, the subject fails to perform a firm grasp on a heavier box. Further investigation will benefit from the use of surface electromyography sensors to measure muscle activation during lifting and lowering, and analysis of timing in movement during the lifting phase. Mixing blind and non-blind data provides classifiers with the highest accuracy, suggesting that higher quantities of data may benefit the networks in learning class separation boundaries.

From the perspective of assessment of potential for injury, it is more important for classifiers to detect heavier weights. Incorrect detection of lighter weights may compromise usability, however, may guarantee safety. Analyzing the ROC curves for higher weight categories using mixed data in Figure‘8, we observe that at an FPR of 0.3, TPRs of 0.8475, 0.8780, and 0.7206 are obtained for the highest weight classes in 2-, 3-, and 4-class prediction.

We find detection of prior knowledge of weight to perform close to chance for user-agnostic weight knowledge detection as seen in Figure 8, indicating that differences in interaction patterns with and without knowledge are too subtle for detection from post-lift signatures. The accuracy for knowledge of weight for user-aware prediction is higher, with a value of 69.55%. However, data being collected on different days is a threat to validity for the results on user-aware prediction of prior weight knowledge. While attempts

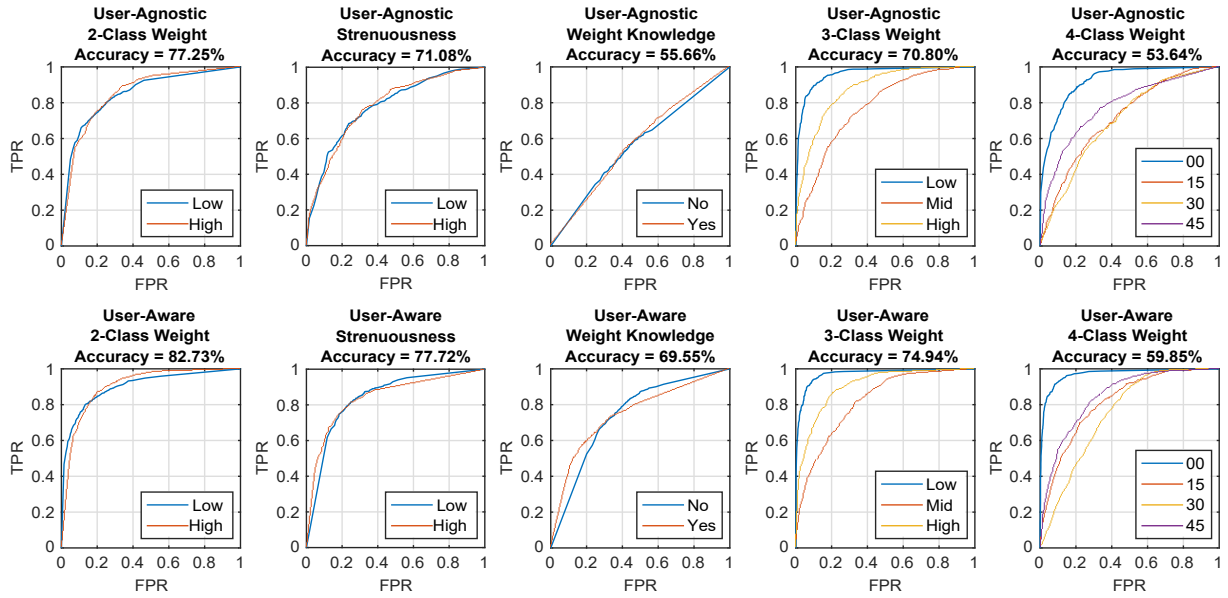

Figure 8: ROC curves for user-agnostic and user-aware prediction of weight classes, strenuousness, and awareness of weight using images on post-lift thermal signature.

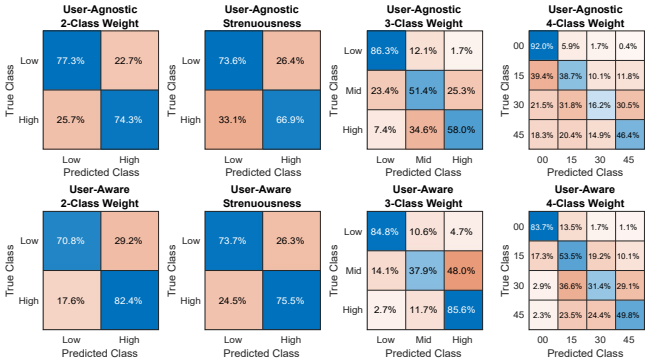

Figure 9: Confusion matrices to predict weight, strenuousness, and, weight knowledge using blind thermal images.

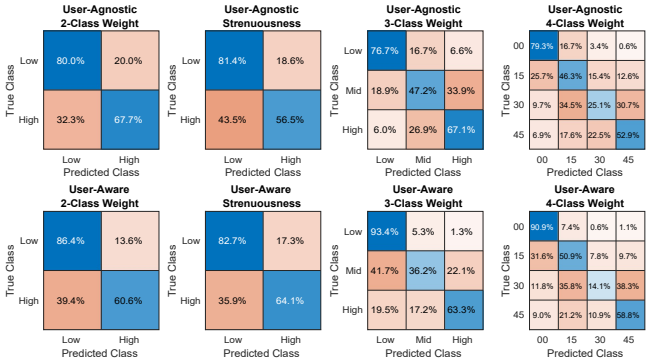

Figure 10: Confusion matrices to predict weight, strenuousness, and, weight knowledge using non-blind data.

were made to keep the orientation of the dolly consistent across days and to randomize within-day standing patterns, chance variation in body temperature and hold location across days may generate separation between blind and non-blind data.

We summarize per-subject individual class accuracies in Tables 1 and 2 for agnostic and aware predictions respectively. Focusing on 2-class weight classification, we observe that separability is higher. For instance, in the case of Subject 14, as shown in Figure 12, thermal maps for boxes from the lower classes are of visually lower intensity than those from the higher classes. For subjects with lower average classification accuracy, properties of the extracted thermal image may influence separability. For instance, for Subject 25, as shown in Figure 13, the box plots on the right demonstrate that mean thermal intensities for the lower class on blind days are spread over a larger range, influencing separability of the lower class from higher classes.

One factor influencing the separability may be that despite efforts to perform normalization of the image with respect to hottest and coldest regions as per Section 5, images from the blind day for Subject 25 demonstrate a wide variability in normalized intensities for the non-imprint regions of the box, especially on the left side. We find that for some images of Subject 25, the box intensity in the original non-preprocssed images approaches that of the ground intensity, whereas in others, the ground appears cooler than the box. The reason for the differences is likely due to non-linear trun-cation of the box and ground intensities during gamma correction by the camera at the time of white balancing, rather than a true difference between the temperature of the box and the ground. In the absence of control over white balancing, a likely concern in

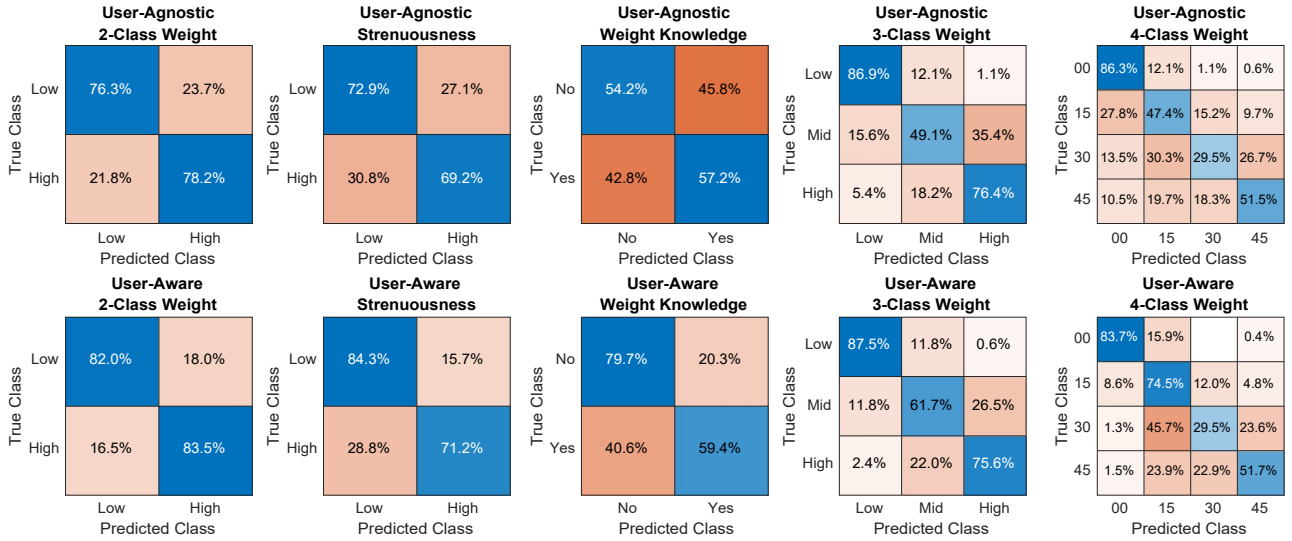

Figure 11: Confusion matrices to predict weight, strenuousness, and, weight knowledge using mixed blind and non-blind thermal images.

ID	2-Class Weight		2-Class Stren.		Weight Knowl.		4-Class Weight				2-Class Weight		2-Class Stren.		4-Class Weight				2-Class Weight		2-Class Stren.		4-Class Weight			
	Lo	Hi	Lo	Hi	No	Yes	0	15	30	45	Lo	Hi	Lo	Hi	0	15	30	45	Lo	Hi	Lo	Hi	0	15	30	45
	Mixed										Blind								Non-Blind							
6	0.88	0.68	0.97	0.57	0.00	1.00	0.60	0.60	0.40	0.75	0.93	0.55	0.95	0.42	1.00	0.30	0.00	0.10	0.78	0.75	0.80	0.78	0.75	0.60	0.40	0.35
7	0.84	0.93	0.40	1.00	0.49	0.74	1.00	0.60	0.70	0.50	0.68	0.90	0.53	0.97	1.00	0.10	0.05	0.45	0.89	0.75	0.63	0.94	0.94	0.65	0.15	0.80
8	0.93	0.42	0.75	NaN	0.00	1.00	0.45	0.85	0.35	0.25	0.97	0.33	0.86	NaN	1.00	0.20	0.00	0.00	0.95	0.28	0.65	NaN	0.70	0.55	0.45	0.10
10	0.75	0.95	0.95	0.55	0.50	0.40	1.00	0.25	0.05	0.10	0.70	0.82	0.75	0.78	0.90	0.85	0.35	0.45	0.93	0.55	1.00	0.10	1.00	0.25	0.20	0.05
11	0.88	0.72	0.80	0.97	0.65	0.54	1.00	0.60	0.60	0.16	0.88	0.64	0.90	0.47	1.00	0.05	0.00	0.05	0.82	0.87	0.93	0.58	1.00	0.55	0.30	0.63
12	0.97	0.68	0.86	0.79	0.97	0.35	1.00	0.05	0.40	0.10	0.97	0.57	0.95	0.35	1.00	0.00	0.00	0.05	1.00	0.17	0.92	0.60	1.00	0.25	0.05	0.50
13	0.80	0.90	0.72	0.78	0.00	1.00	0.85	0.40	0.37	0.60	0.78	0.92	0.87	0.44	1.00	0.80	0.47	0.75	0.85	0.74	0.57	0.89	1.00	0.50	0.26	0.35
14	0.82	0.95	0.96	0.34	0.95	0.38	1.00	0.35	0.25	0.68	0.82	0.77	0.83	0.81	1.00	0.45	0.10	0.95	0.90	0.90	0.87	0.53	1.00	0.65	0.50	0.84
15	0.78	0.68	1.00	0.00	0.38	0.62	0.95	0.85	0.25	0.40	0.88	0.68	0.82	0.67	1.00	0.75	0.30	0.20	0.85	0.47	0.95	0.33	0.50	0.85	0.05	0.15
16	0.80	0.74	0.95	0.73	0.95	0.23	1.00	0.50	0.11	0.30	0.85	0.72	0.84	0.80	0.85	0.90	0.21	0.10	0.85	0.72	1.00	0.00	0.90	0.40	0.21	0.15
17	0.46	0.95	0.81	0.35	0.95	0.15	0.89	0.20	0.25	0.70	0.79	0.40	0.66	0.60	0.95	0.00	0.05	0.80	0.67	0.55	0.81	0.55	0.47	0.10	0.30	0.60
18	0.82	0.75	0.41	0.97	0.82	0.17	1.00	0.63	0.60	0.65	0.54	0.90	0.67	0.85	1.00	0.37	0.00	0.50	0.72	0.88	0.74	0.68	0.90	0.58	0.50	0.70
19	0.79	0.60	1.00	0.07	0.70	0.56	0.80	0.74	0.15	0.30	0.67	0.80	0.90	0.33	0.70	0.16	0.45	0.60	0.87	0.35	1.00	0.12	0.60	0.53	0.40	0.55
20	0.53	1.00	0.78	0.80	0.80	0.53	0.85	0.70	0.30	0.90	0.72	0.95	0.36	0.96	0.70	0.10	0.30	0.90	0.75	0.90	0.80	0.80	0.15	0.45	0.35	0.85
21	0.90	0.88	0.85	0.90	0.00	1.00	0.90	0.65	0.55	0.95	0.80	0.97	0.93	0.72	0.70	0.55	0.15	1.00	0.93	0.72	0.90	0.79	0.85	0.70	0.25	0.90
22	0.39	1.00	0.25	1.00	0.00	1.00	0.21	0.21	0.00	1.00	0.71	0.97	0.42	1.00	0.63	0.32	0.00	1.00	0.29	1.00	0.31	1.00	0.32	0.16	0.10	1.00
24	0.67	0.85	0.21	1.00	0.68	0.41	0.89	0.60	0.30	0.30	0.54	0.72	0.40	0.81	1.00	0.00	0.00	0.15	0.92	0.23	0.57	1.00	0.95	0.30	0.00	0.40
25	0.90	0.33	0.98	0.15	0.72	0.75	1.00	0.10	0.00	0.00	0.95	0.12	0.93	0.24	1.00	0.50	0.00	0.00	0.90	0.23	1.00	0.00	1.00	0.35	0.05	0.00
26	0.69	0.95	0.39	1.00	0.56	0.64	1.00	0.84	0.53	0.85	0.62	0.92	0.51	1.00	1.00	0.63	0.79	0.50	0.69	0.97	0.95	0.00	1.00	0.63	0.68	0.65
28	0.79	0.82	0.68	0.90	0.00	1.00	0.74	0.30	0.25	0.94	0.90	0.79	0.89	0.75	0.95	0.50	0.20	1.00	0.82	0.79	0.70	0.80	0.89	0.40	0.35	0.67
29	0.45	0.97	0.42	1.00	0.85	0.49	0.65	0.35	0.32	0.65	0.33	0.97	0.50	1.00	0.80	0.20	0.05	0.60	0.55	0.97	0.62	0.94	0.55	0.15	0.00	0.95
30	0.92	0.33	0.68	0.32	1.00	0.05	1.00	0.00	0.00	0.00	0.95	0.56	0.85	0.28	0.95	0.26	0.05	0.00	0.79	0.59	1.00	0.00	0.90	0.26	0.00	0.10
31	0.64	1.00	1.00	0.38	0.28	0.59	0.89	0.70	0.30	0.95	0.74	1.00	0.74	0.81	0.95	0.55	0.20	1.00	0.77	0.95	0.89	0.69	0.63	0.65	0.35	0.95
32	0.88	0.72	0.81	0.74	0.75	0.12	1.00	0.30	0.05	0.35	0.82	0.85	0.81	0.53	1.00	0.75	0.20	0.05	0.68	0.95	1.00	0.32	1.00	0.60	0.10	0.50

Table 1: Per-subject accuracies for user-agnostic predictions ('Lo'='Low', 'Hi'='High', 'Stren.'='Strenuousness', 'Knowl.'='Knowledge').

ID	2-Class Weight		2-Class Stren.		Weight Knowl.		4-Class Weight				2-Class Weight		2-Class Stren.		4-Class Weight				2-Class Weight		2-Class Stren.		4-Class Weight			
	Lo	Hi	Lo	Hi	No	Yes	0	15	30	45	Lo	Hi	Lo	Hi	0	15	30	45	Lo	Hi	Lo	Hi	0	15	30	45
	Mixed										Blind								Non-Blind							
6	0.85	0.72	0.93	0.72	0.60	0.90	0.85	0.75	0.05	0.30	0.72	0.78	0.95	0.53	0.80	0.60	0.25	0.30	0.97	0.45	0.93	0.55	0.80	0.25	0.10	0.45
7	0.82	0.78	0.81	0.80	0.59	0.67	0.94	0.75	0.35	0.55	0.84	0.80	0.65	0.89	1.00	0.80	0.35	0.45	0.79	0.80	0.67	0.83	0.94	0.60	0.10	0.85
8	0.93	0.62	0.99	NaN	0.95	0.47	0.75	0.75	0.25	0.35	0.78	0.45	0.93	NaN	0.75	0.60	0.15	0.05	0.90	0.55	0.99	NaN	0.90	0.40	0.30	0.15
10	0.85	0.90	0.95	0.60	0.65	0.78	1.00	0.85	0.25	0.50	0.80	0.85	0.85	0.88	1.00	0.55	0.25	0.25	1.00	0.57	0.78	0.70	1.00	0.80	0.15	0.80
11	0.75	0.97	0.90	0.47	0.70	0.18	0.95	0.75	0.35	0.37	0.62	0.90	0.80	0.82	0.95	0.65	0.30	0.37	0.88	0.56	0.90	0.53	1.00	0.60	0.20	0.37
12	0.90	0.85	0.73	0.72	1.00	0.53	1.00	0.65	0.20	0.45	0.75	0.78	0.95	0.44	1.00	0.40	0.35	0.50	0.93	0.40	0.97	0.53	1.00	0.15	0.05	0.55
13	0.85	0.87	0.89	0.56	0.59	0.78	0.95	0.85	0.37	0.40	0.82	0.82	0.84	0.56	1.00	0.65	0.53	0.15	0.93	0.51	0.80	0.50	1.00	0.70	0.16	0.40
14	0.93	0.85	0.79	0.88	0.97	0.51	0.95	0.75	0.30	0.53	0.70	0.92	0.74	0.88	1.00	0.60	0.35	0.79	0.95	0.59	0.81	0.81	1.00	0.20	0.05	0.63
15	0.82	0.85	0.99	0.33	0.75	0.35	0.70	0.85	0.35	0.40	0.80	0.72	0.78	0.33	0.90	0.70	0.05	0.05	0.95	0.23	0.90	0.00	1.00	0.50	0.05	0.55
16	0.85	0.85	0.89	0.78	0.74	0.47	0.75	0.95	0.26	0.35	0.70	0.85	0.92	0.66	0.85	0.80	0.37	0.40	0.85	0.72	0.79	0.73	0.85	0.80	0.05	0.75
17	0.74	0.68	0.54	0.65	0.97	0.53	0.84	0.50	0.35	0.35	0.62	0.78	0.46	0.70	0.42	0.40	0.40	0.55	0.59	0.70	0.69	0.60	0.95	0.25	0.05	0.40
18	0.85	0.90	0.79	0.72	0.97	0.50	0.80	0.74	0.50	0.85	0.74	0.90	0.56	0.93	0.85	0.58	0.25	0.60	0.79	0.60	0.74	0.62	0.95	0.63	0.25	0.55
19	0.69	0.75	0.72	0.82	0.88	0.54	0.70	0.79	0.25	0.40	0.64	0.80	0.74	0.70	0.70	0.74	0.35	0.35	0.77	0.78	0.74	0.45	0.95	0.53	0.10	0.75
20	0.78	0.97	0.78	0.88	0.97	0.68	0.75	0.65	0.45	0.80	0.53	0.93	0.45	1.00	0.80	0.20	0.25	0.90	0.88	0.85	0.80	0.92	0.95	0.60	0.25	0.85
21	0.90	0.95	0.90	0.87	0.70	0.85	0.75	0.80	0.40	0.90	0.80	0.82	0.80	0.95	0.95	0.30	0.25	0.70	0.97	0.62	0.85	0.79	0.90	0.45	0.20	0.85
22	0.82	0.95	0.75	0.67	0.76	0.75	0.58	0.53	0.25	0.90	0.79	0.97	0.54	1.00	0.79	0.53	0.45	0.80	0.76	0.90	0.67	0.70	0.68	0.53	0.00	0.90
24	0.64	0.80	0.86	0.71	0.70	0.62	0.74	0.85	0.35	0.50	0.54	0.85	0.57	0.86	0.68	0.80	0.45	0.40	0.79	0.68	0.78	0.57	0.74	0.60	0.05	0.40
25	0.80	0.62	0.83	0.35	0.93	0.88	0.90	0.75	0.15	0.05	0.93	0.42	0.89	0.21	0.95	0.50	0.00	0.00	0.95	0.12	0.87	0.35	1.00	0.55	0.05	0.15
26	0.87	0.92	0.97	1.00	0.74	0.67	0.90	0.84	0.37	0.70	0.69	0.95	0.70	0.00	1.00	0.53	0.47	0.85	0.90	0.64	0.90	1.00	0.95	0.63	0.37	0.70
28	0.87	0.84	0.81	0.72	0.93	0.57	0.95	0.90	0.35	0.56	0.62	0.97	0.76	0.88	0.68	0.25	0.45	1.00	0.85	0.66	0.89	0.60	0.95	0.60	0.20	0.72
29	0.60	0.95	0.69	1.00	0.85	0.51	0.50	0.50	0.05	0.70	0.45	0.97	0.50	1.00	0.35	0.00	0.16	0.90	0.57	0.79	0.85	0.89	0.55	0.35	0.05	0.75
30	0.85	0.74	0.81	0.56	0.77	0.72	0.85	0.74	0.16	0.15	0.77	0.79	0.74	0.32	0.80	0.42	0.58	0.15	0.95	0.38	0.77	0.32	0.85	0.32	0.16	0.10
31	0.85	0.97	0.83	0.50	0.60	0.41	1.00	0.70	0.40	0.85	0.69	0.93	0.77	0.91	0.89	0.75	0.20	0.95	0.95	0.80	0.89	0.59	0.89	0.65	0.25	1.00
32	0.88	0.72	0.79	0.68	0.80	0.40	1.00	0.70	0.30	0.50	0.65	0.85	0.81	0.68	0.95	0.50	0.35	0.55	0.85	0.62	0.74	0.71	1.00	0.55	0.20	0.50

Table 2: Per-subject accuracies for user-aware predictions ('Lo'='Low', 'Hi'='High', 'Stren.'='Strenuousness', 'Knowl.'='Knowledge').

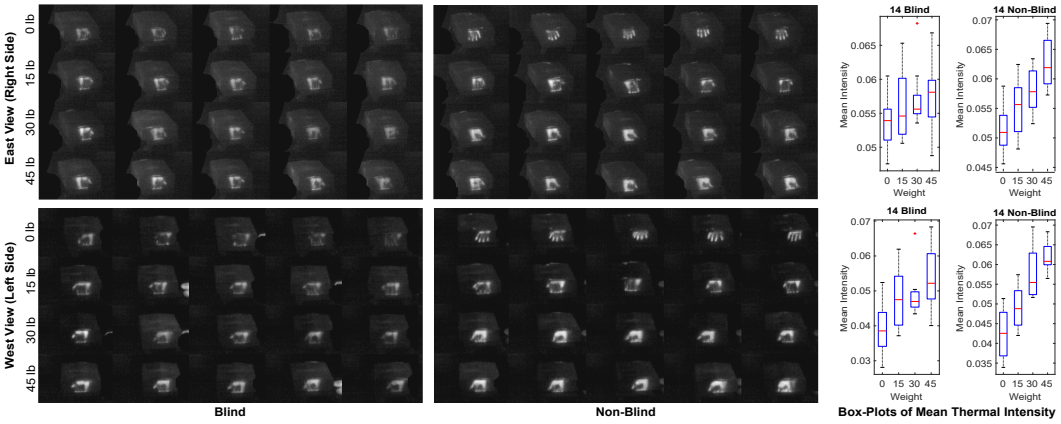

Figure 12: Left and right hand thermal hand prints for Subject 14 on blind and non-blind days. As the subject interacts with heaver weight classes, the thermal intensity increases. The boxplot on the right shows a positive trend in mean intensity.

consumer thermal cameras, a potential corrective approach may be to segment out the imprint region, rather than perform detection using the entire set of thermal intensities in the processed image.

The 4-class weight classification demonstrates challenges in separability for user-agnostic classification, where samples from most users are classified in the lowermost class. These results conform the summaries in the confusion matrices in 9, 10, and Figures 11.

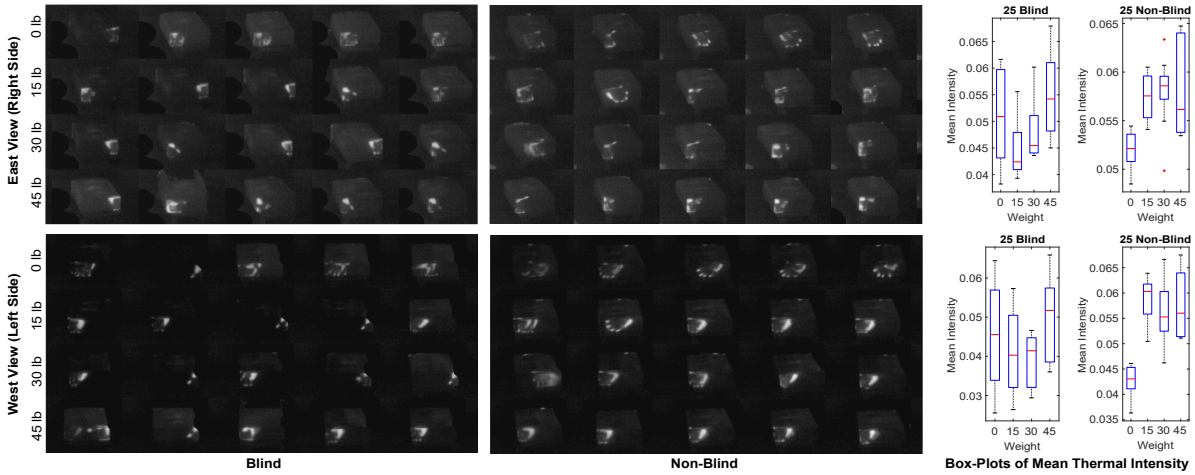

Figure 13: Left and right hand thermal hand prints for Subject 25 on blind and non-blind days, together with box plots of mean thermal intensity. Some samples from the lower-weight object classes show higher mean thermal intensities on blind days causing intensity distributions to span a larger range, and influencing separability from higher weight classes.

Figure 13 demonstrates that separability amongst the higher three classes is lost for Subject 25 on the non-blind day. Mean average intensities as per the box plots show similar within class mean, with a lower mean and highest spread for the highest class. A potential approach to address this issue may be to use a multi-step technique that identifies coarse features of interest such as grab location prior to extracting finer details on thermal patterns and intensity.

8 DISCUSSION

In this work, we have presented an approach that uses thermal images after lift, lowering, and release interactions to perform detection of weight and effort as subject-reported strenuousness. Our work demonstrates that hand imprints as recorded by thermal cameras provide promise in performing coarse separation of weight and strenuousness by showing accuracies upwards of 73% and 68%. Post-lift analysis enables informing monitoring systems on the effort perceived by the individual in performing a task, and whether a similar task should be continued by the individual. Such monitoring may be factored into higher-level AI algorithms that perform continuous personnel management, by re-allocating individuals from tasks that induce more fatigue and lower productivity to less fatiguing tasks, and by moving persons who demonstrate lower levels of fatigue to higher performance tasks, thereby ensuring continuous optimal output.

While we demonstrate the promise of thermal imprints, our work also reveals challenges in using off-the-shelf thermal cameras that are worth future study, given the benefits of thermal cameras in ensuring non-invasive at-a-distance monitoring. As discussed in Section 7, we find that the non-linear white balancing of thermal cameras influences normalization of input images, especially when the cooler non-imprint portions of the box approach surrounding temperatures. While the current approach uses a linear normalization, estimating the parameters of a sigmoid gamma correction function may improve accuracies. To focus on the features of the thermal imprint rather than the box, it may also be beneficial to segment out the region around the imprint using alpha matting. As part of future work, we are interested in conducting studies of the relationship of imprint temperature and weight by using radiometric cameras to eliminate the confounding factor of non-linear white balancing. Another aspect of importance is that core and skin temperatures vary across subjects, and even within a single subject, continuous aerobic activity over time increases core and skin temperature [10, 26] . While a single session was completed in an hour and a half, for some subjects, the session may have resulted in a rise in body temperature. The accuracies of imprint-based detection may be improved if variabilities in body temperature can be factored out, e.g., by performing a body-specific normalization of the imprint using thermal pixels on the hand or the upper body.

Another confounding factor worth study is the influence of force versus time of lift on the higher intensity of the thermal imprint for heavier weights. Lifting activities with the heavier weights typically take a longer duration in our study, than with the lighter weights. Since the heat flow equation [29] is time-dependent, the longer interaction duration may result in more heat being transferred to the box, though the analysis may be complexified by diffusion to the remaining carton and heat loss to the ambient environment. Studies on swipe pressure on plane surfaces provides some qualitative evidence that application of a higher force on a surface provides a higher intensity of imprint than a lower force applied for the same period of time, however, detailed studies are necessary to examine the effect for interactions such as grab, lift, hold, lower, and release. Future work should investigate the interplay of time, force applied as measured by force sensors, object weight, and interaction type, e.g., grab and release without movement, lifting and lowering, dragging.

In this work, we focus on using the raw thermal imprint on the package. However, work on using thermal cameras for behavior monitoring can benefit from a more fine-grained analysis on the

grab and movement patterns employed by the user based on the weight of the carton. E.g., we notice that for higher weights, users may grab the carton along the diagonal, while for lower weights they may grab the carton symmetrically at the two sides. Users also showed signs of struggle in lifting higher weights, and spent more time at the lowering phase in careful stabilization of the box prior to release, presumably also performing synergistic posture stabilization. During lift, asymmetrical patterns may be observed for individuals employing a higher effort level. To study spatial grab location and box motion, future work may benefit from extracting the 3D pose of the box, either using geometric techniques such as perspective n-point [18] , or by tying the thermal cameras with depth sensors, performing 3D cuboid extraction from depth images, and texturing extracted cuboids with the thermal imprint. This approach also provides the benefit of being view-independent enabling propagation to unstructured spaces.

This work focuses solely on the post-lift thermal imprint in images. However, the thermal data has the potential to provide additional parameters of information, such as variability in skin temperature, duration of interaction, and information on subject posture during performance by enabling extraction of subject silhouettes. Future work should focus on extraction of multi-modal parameters from thermal data to enable continuous assessment for signs of fatigue. Future work can also investigate estimation of large-scale parameters in industrial environments, e.g., ambient temperature, sources of heat, and presence of other humans to signal for assistance, to enable holistic monitoring and intervention for personnel in industrial settings.

ACKNOWLEDGMENTS

This work was funded by National Science Foundation grant IIS-2026559.

REFERENCES

[1] Changbum R Ahn, SangHyun Lee, Cenfei Sun, Houtan Jebelli, Kanghyeok Yang, and Byungjoo Choi. 2019. Wearable sensing technology applications in construction safety and health. Journal of Construction Engineering and Management 145, 11 (2019), 03119007.

[2] Samarth Brahmbhatt, Cusuh Ham, Charles C Kemp, and James Hays. 2019. Contactdb: Analyzing and predicting grasp contact via thermal imaging. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. IEEE, Piscataway, NJ, 8709–8719.

[3] Samarth Brahmbhatt, Ankur Handa, James Hays, and Dieter Fox. 2019. Contactgrasp: Functional multi-finger grasp synthesis from contact. In 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, Piscataway, NJ, 2386–2393.

[4] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2019. OpenPose: realtime multi-person 2D pose estimation using Part Afinity Fields. IEEE transactions on pattern analysis and machine intelligence 43, 1 (2019), 172–186.

[5] Lora Cavuoto and Fadel Megahed. 2016. Understanding fatigue and the implications for worker safety. In ASSE Professional Development Conference and Exposition. ASSP, Park Ridge, IL, 16–19.

[6] Viviane de Freitas Cardoso, Claudia Aparecida Stefane, Fernanda Cabegi de Barros, Josiane Sotrate Gonçalves, Leandro Corrêa Figueiredo, and Tatiana de Oliveira Sato. 2022. Influence of gender and age on musculoskeletal symptoms in white-collar and blue-collar workers: a cross-sectional study. International Journal of Occupational Safety and Ergonomics 1, 1 (2022), 1–10.

[7] Tim Dunn, Sean Banerjee, and Natasha Kholgade Banerjee. 2018. User-independent detection of swipe pressure using a thermal camera for natural surface interaction. In 2018 IEEE 20th International Workshop on Multimedia Signal Processing (MMSP). IEEE, Piscataway, NJ, 1–6.

[8] Rikke Gade and Thomas B Moeslund. 2014. Thermal cameras and applications: a survey. Machine vision and applications 25, 1 (2014), 245–262.

[9] Jacob Gately, Ying Liang, Matthew Kolessar Wright, Natasha Kholgade Banerjee, Sean Banerjee, and Soumyabrata Dey. 2020. Automatic Material Classification Using Thermal Finger Impression. In International Conference on Multimedia Modeling. Springer, Berlin, Germany, 239–250.

[10] M Gleeson. 1998. Temperature regulation during exercise. International Journal of Sports Medicine 19, S2 (1998), S96–S99.

[11] Renée Govaerts, Bruno Tassignon, Jo Ghillebert, Ben Serrien, Sander De Bock, Toon Ampe, Ilias El Makrini, Bram Vanderborght, Romain Meeusen, and Kevin De Pauw. 2021. Prevalence and incidence of work-related musculoskeletal disorders in secondary industries of 21st century Europe: a systematic review and meta-analysis. BMC musculoskeletal disorders 22, 1 (2021), 1–30.

[12] Daisuke Iwai and Kosuke Sato. 2005. Heat sensation in image creation with thermal vision. In Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology. ACM, New York, NY, 213–216.

[13] Houtan Jebelli, Byungjoo Choi, and SangHyun Lee. 2019. Application of wearable biosensors to construction sites. II: Assessing workers' physical demand. Journal of Construction Engineering and Management 145, 12 (2019), 04019080.

[14] Houtan Jebelli and SangHyun Lee. 2019. Feasibility of wearable electromyography (EMG) to assess construction workers' muscle fatigue. In Advances in informatics and computing in civil and construction engineering. Springer, Berlin, Germany, 181–187.

[15] Yijun Jiang, Sean Banerjee, and Natasha Kholgade Banerjee. 2019. Predicting Human Grasp Locations on Cup Handles by Using Deep Neural Networks to Infer Heat Signatures from Depth Data. In 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). IEEE, Piscataway, NJ, 25–30.

[16] Daniel Kurz. 2014. Thermal touch: Thermography-enabled everywhere touch interfaces for mobile augmented reality applications. In Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on. IEEE, Piscataway, NJ, 9–16.

[17] Eric Larson, Gabe Cohn, Sidhant Gupta, Xiaofeng Ren, Beverly Harrison, Dieter Fox, and Shwetak Patel. 2011. HeatWave: thermal imaging for surface user interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, New York, NY, 2565–2574.

[18] Xiao Xin Lu. 2018. A review of solutions for perspective-n-point problem in camera pose estimation. In Journal of Physics: Conference Series, Vol. 1087. IOP Publishing, Bristol, UK, 052009.

[19] Zahra Sedighi Maman, Ying-Ju Chen, Amir Baghdadi, Seamus Lombardo, Lora A Cavuoto, and Fadel M Megahed. 2020. A data analytic framework for physical fatigue management using wearable sensors. Expert Systems with Applications 155 (2020), 113405.

[20] Zahra Sedighi Maman, Mohammad Ali Alamdar Yazdi, Lora A Cavuoto, and Fadel M Megahed. 2017. A data-driven approach to modeling physical fatigue in the workplace using wearable sensors. Applied ergonomics 65 (2017), 515–529.

[21] Nurhayati Mohd Nur, Siti Zawiah Dawal, and Mahidzal Dahari. 2014. The prevalence of work related musculoskeletal disorders among workers performing industrial repetitive tasks in the automotive manufacturing companies. In Proceedings of the 2014 international conference on industrial engineering and operations management Bali, Indonesia. IEOM Society International, Southfield, MI, 1–8.

[22] Kenji Oka, Yoichi Sato, and Hideki Koike. 2002. Real-time tracking of multiple fingertips and gesture recognition for augmented desk interface systems. In Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on. IEEE, Piscataway, NJ, 429–434.

[23] Karri Palovuori and Ismo Rakkolainen. 2015. The heat is on: thermal input for immaterial interaction. In Proceedings of the 19th International Academic Mindtrek Conference. ACM, New York, NY, 152–154.

[24] Konstantinos Papoutsakis, George Papadopoulos, Michail Maniadakis, Thodoris Papadopoulos, Manolis Lourakis, Maria Pateraki, and Iraklis Varlamis. 2022. Detection of Physical Strain and Fatigue in Industrial Environments Using Visual and Non-Visual Low-Cost Sensors. Technologies 10, 2 (2022), 42.

[25] Konstantinos Papoutsakis, Thodoris Papadopoulos, Michalis Maniadakis, Manolis Lourakis, Maria Pateraki, and Iraklis Varlamis. 2021. Detection of physical strain and fatigue in industrial environments using visual and non-visual sensors. In The 14th PErvasive Technologies Related to Assistive Environments Conference. ACM, New York, NY, 270–271.

[26] MICHAEL F Roberts and C BRUCE Wenger. 1979. Control of skin circulation during exercise and heat stress. Medicine and science in sports 11, 1 (1979), 36–41.

[27] Elliot N Saba, Eric C Larson, and Shwetak N Patel. 2012. Dante vision: In-air and touch gesture sensing for natural surface interaction with combined depth and thermal cameras. In Emerging Signal Processing Applications (ESPA), 2012 IEEE International Conference on. IEEE, Piscataway, NJ, 167–170.

[28] Mitsuhiko Shimomura, Masahiro Fujiwara, Yasutoshi Makino, and Hiroyuki Shinoda. 2022. Estimation of Frictional Force Using the Thermal Images of Target Surface During Stroking. In International Conference on Human Haptic Sensing and Touch Enabled Computer Applications. Springer, Berlin, Germany, 234–242.

[29] David Vernon Widder. 1976. The heat equation. Vol. 67. Academic Press, Cambridge, MA.