A Computer Vision Algorithm to Identify High Force Exertions from Facial Expressions

Hamed Asadi, Guoyang Zhou, Jae Joong Lee, Vaneet Aggarwal, Denny Yu

Abstract

Cumulative exposure to repetitive and forceful activities may lead to musculoskeletal injuries which not only reduce workers' efficiency and productivity, but also affect their quality of life. Thus, widely accessible techniques for reliable detection of unsafe muscle force exertion levels for human activity is necessary for their well-being. However, measurement of force exertion levels is challenging and the existing techniques pose a great challenge as they are either intrusive, interfere with human-machine interface, and/or subjective in the nature, thus are not scalable for all workers. In this work, face videos and machine learning techniques are used to detect the high force exertion levels of over 65%MVC (Maximum Voluntary Contraction), representing the discomfort feeling in force exertion; thus providing a non-intrusive and scalable approach. Efficient feature extraction approaches have been investigated, including the movement of different landmarks of the face. Based on the data collected from 18 subjects, features extracted from the face videos give 92% overall accuracy in prediction the discomfort force levels. Further 0.78 recall and 0.88 AUC (Area Under Curve) values indicates the model's performance in detecting the forceful exertions. The approach is also shown to be robust to the correctly identify force level when the person is talking, even though such datasets are not included in the training.

Keywords: Computer Vision, High Force Exertions, Facial Expressions, Machine Learning

1. Introduction

Musculoskeletal disorders (MSDs), such as sprains or strains resulting from overexertion, accounts for 349,050 cases for all workers [1] annually. This means that 33 workers in every 10,000 suffer an injury severe enough that they must take time away from work. [1] Although overall percentage of the workforce getting hurt is small, these injuries are preventable. Furthermore, they not only impact individual worker's health and quality of life[2], they also result in significant cost employees and society (e.g., workers compensation, medical care, loss productivity, training temporary workers). The annual cost of the injuries in the United States are nearly $60 billion in direct workers compensation costs [3]. Due to high direct and indirect cost of MSDs, there is a strong motivation for all stakeholders (e.g., employers, workers, and researchers) to identify factors that lead to MSDs and actively monitor and eliminate worker exposure to these factors.

A comprehensive report by the National Institute for Occupational Safety and Health (NIOSH) lists high/sustained force, repetitive movements, and poor biomechanical postures as contributors to MSDs, with conclusion that evidence exists linking force to musculoskeletal injuries [4]. Similarly, in other studies, risk factors for MSDs include repetition, posture, vibration, and forceful exertions [5, 6]. High force exertion levels are reported as the most common contributing factors with sufficient evidence to suggest a causal relationship for work-related musculoskeletal disorders (MSDs) [5, 6, 7, 8, 9].

Several key physiological and biomechanical mechanisms are proposed for how force exertions lead to injuries. Different tissue can be damages by various injury mechanisms including acute and prolonged exertions [2]. As an example, chronic low back pain can be a result of tears in the soft tissues [10]. For instance, high and/or frequent force exertions initiates lumbar disc damage and degeneration [11]. As another example, prolonged force exertions could lead to wrist injuries where frequent force exertions by the hand (e.g., pinching and griping) lead to and exacerbate inflammation of the carpal tunnel cumulative tissue stress can eventually lead to injuries [12].

Force is one of the hardest to measure because it is difficult to observe and depend on individual's effort. For example, changes in expressions are subtle unless high forces and strong efforts are needed. Many methods are currently available to measure the high force exertion level. However, each method vary in reliability and feasibility as they are either 1) intrusive (e.g., disrupts the worker while they are performing their job), 2) interfere with human machine interface (e.g., need to install force gauges on tool-handles and machine controls), 3) subjective, and most importantly 4) not widely scalable across all workers, jobs, and workplaces as trained ergonomics and safety professionals are needed to implement these methods. Kong et. all [13] evaluated comfortable or uncomfortable feelings for the grip force level and reported 65% Maximum Voluntary Contraction (MVC) as the transition level to change from comfort to discomfort. To better understand the comfort and discomfort concept and better the need for a new evaluation model was emphasized [14].

This paper proposes objective and automated predictions high force exertion levels which has minimum distractions on workers and could be used in wide variety of workplaces by using the videos of the person to predict the high level of force exertions. Innovations in computer vision techniques can address many of the deficiencies in the current approaches. This paper proposes a new objective approach, which can be widely accessible and is not intrusive to workers.

2. Methods

The overall algorithm proposed in this paper is using video recording of participant's face video as an input and classifies the comfort/discomfort level of a grip force.

2.1. Participants

Eighteen healthy volunteers participated in this study. The participants were recruited from a university population through email including a description of the study. This study was reviewed by the university's Institutional Review Board and exclusion criteria were current musculoskeletal impairments that prevented participants from performing force exertions. Sixteen males and 4 females participated in the study, all were right hand dominant, and their ages ranged from 18 to 24 years. The details of all the subjects that participated in the study is given in Table 1.

2.2. Study Setup

The power grip dynamometer (Lafayette Hydraulic Hand Dynamometer, Lafayette Instrument Company, IN, USA) was used to measure the grip force of each subject. This devise helps in measuring the maximum isometric strength of the hand and forearm muscles and hence helps us collecting the ground truth of force exertion level for each subject.

A GoPro HERO4 camera (GoPro, San Mateo, CA, USA) was used to capture the video of subjects while they were performing different kind of activities. The GoPro is placed in front of the subject, 0.5 meter away from face, and video recorded the subject during the entire experiment. Video recordings are done at 50 frames per second.

Figure 1: The experimental setup with a subject holding a grip dynamometer and pulse oximeter attached to the earlobe. The PPG signals were recorded using pulse Shimmer3 GSR+ (Shimmer, Dublin, Ireland), but the resulting PPG were not used in this study

	Female (n=10)		
	Mean ± SD	Min	Max
Age (years)	20.8±2.0	19	24
Weight (lb)	125.2±23.7	100	188
Grip Force (lb)	63.0±23.8	30	105
	Male (n=8)		
	Mean ± SD	Min	Max
Age (years)	21.4±1.3	20	23
Weight (lb)	143.8±18.9	120	170
Grip MVC (lb)	92.3±13.9	64	120

Table 1: Data for 18 subjects in our experiment

2.3. Data Collection

At the beginning of the data collection session, participants were provided a description of the study, and written consent was collected. Subjects were seated in front of the white background to minimize the noise in video processing in detecting the face. The handheld dynamometer was adjusted by hand size to ensure standardized and comfortable gripping postures for each subject.

Participants received a 5-minutes practice period to familiarize with grip device and exertion force at varying levels. The overall study involved exertion at six levels of force and one activity. In the first trial, each participant performed a grip exertion at maximum force exertion. The subjects were instructed to maintain the maximum force for 9 seconds (note that although the magnitude of the force may decrease during the 9-seconds, participants continued to exert their maximum effort). The recordings were stopped after 9 seconds.The second exertion trial was 0% grip force. In this trial, subjects were asked to hold the grip dynamometer without exerting any grip force. The subjects rested for 2 minutes between each force exertion levels to prevent fatigue effects from carrying over to the next force exertion trial. Then the 15, 33, 50, and 75%MVC colelcted and the videos of face. In each trial, subjects were asked to exert exactly the percentage of their maximum grip contraction and if the subject exerted force $\pm 10\%$ of the specified force, the trial restarted (after 2 minutes gap). The distribution of the grip force for different subjects is reported in Table 1.

2.4. Video Processing

The videos of several subjects are recorded under different force exertion levels as explained in section 2.3. Each video is processed using the DeepFace algorithm proposed in [15]. This is a state-of-the-art algorithm developed by researchers at Facebook. DeepFace is a face recognition algorithm that consists of four main stages: 1. Detect 2. Align 3. Represent, and 4. Classify.

There have been other work in developing algorithm for facial recognition [16, 17, 18, 19, 20, 21], but DeepFace [15] reached an accuracy of 97.35% in Labeled Faces in the Wild (LFW) dataset and reduced the error in face recognition of current state-of-the art by more than 27%. The high accuracy in DeepFace is achieved by revisiting both alignment and representation step. 3D face alignment has been done using piecewise affine transformation and face representation is derived using 9-layer neural network which is a key for the high performance. Therefore, we utilized DeepFace for recognizing faces in our approach.

The 9 seconds video of each subject is trimmed to 7 seconds before passing it to DeepFace. The first 2 seconds of videos are removed because each subject requires initial 1 to 2 seconds to reach to the required force level. Each video is recorded at 50 frames per second and hence, consists of 350 frames We process all these frames using DeepFace that recognizes and aligns the face of each subject across the frames using 68 landmark points on the face. Figure 2 shows how DeepFace is used to extract faces from the each frame in the video. Figure 2 (a) is an example of an actual frame in the video. DeepFace recognizes the face of the person in each and crops the face out of it as shown in Figure 2 (b). This algorithm helps identify 68 landmark points on the face as depicted in Figure 2 (c) and track these 68 landmark points over the whole video The 68 landmark points represents the contour of the face, eyebrows, eyes, lips, and nose (Figure 3). Detecting and aligning the face in each frame of the video is one of the most critical step in our overall methodology, because relevant features to train a neural network will be extracted from the output of DeepFace.

Figure 2: The steps followed for feature extraction from each frame of the video. (a) The actual image (one of the many frame) from the video captured during the experiment. (b) The detected and aligned face using DeepFace. (c) The face along with the 68 landmarks on it. These 68 landmark points are used by DeepFace in face recognition.

2.5. Feature selection

The extraction of “right” features is important as it plays significant role in training a neural network. The choice of relevant features leads to the simplification of the models which in turn requires shorter training time

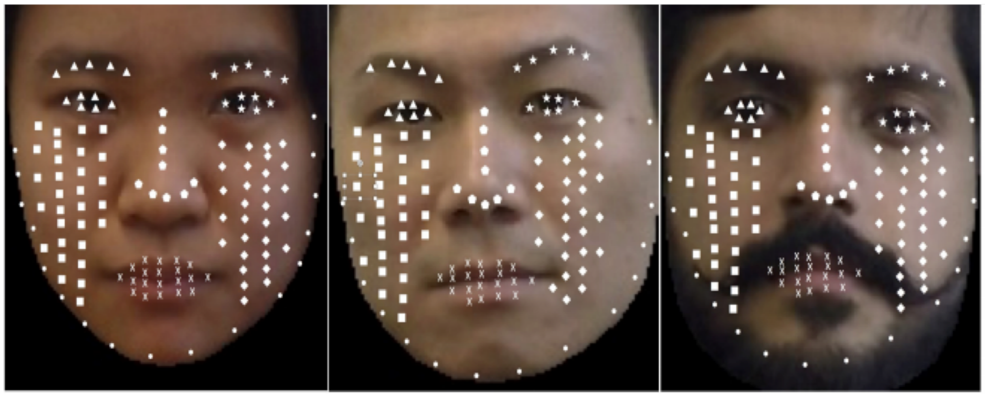

Figure 3: The location of 128 landmark points on the face for different subjects. Additional 60 landmarks have been identified on the face for efficient model training.

[22]. “Right” set of features helps in avoiding the curse of dimensionality and leads to generalization of the model by reducing the variance in the model [23]. Choosing the subset of features from the available data reduces redundancy in the input to the neural networks and subsequently improving the performance. We will extract relevant features from the Frames that has been processed by DeepFace.

2.5.1. Facial feature

Deepface utilizes the information of 68 landmark points on the face. Our proposed method uses 128 landmark points on the face as shown in Figure 3. Based on 68 landmark points given by Deepface, we located 60 more landmarks on the face that lies on the left and right cheeks. 30 landmarks on each cheek is located based on the location of landmarks on the contour of the face and eyes. Different landmark points can be grouped together based on the location on the face as: 1: Contour of Face (17 landmarks), 2: Eyes (12 landmarks) 3: Eyebrows (10 landmarks), 4: Nose (9 landmarks), 5: Lips (20 landmarks), 6: Cheek (60 landmarks). All the 128 landmark points were tracked in 350 frames for each video. The location (x and y co-ordinate values) of each landmark was extracted and based on the location, the movements of each landmark with respect to its location in the first frame were calculated over the entire video. The movements of the distances between landmarks were calculated using Euclidean distance equation, which is presented in equation 1.

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad (1)$$

The average and standard deviation of facial landmarks' movement within each group are chosen as the domain features (Figure 4). The total of twelve feature were used in training the final model. As the subject increases the effort level, facial expression tends to change and there are clear differences in the movement of the movement of facial landmarks for different force exertion. For all of the groups, the average movements for discomfort levels were higher than discomfort, however they are some overlapping areas in each group. It is also interesting to see that movement of noses landmarks are always the least while contours movement tend to be the highest.

Specifically, as hand force level increase to discomfort level, facial landmarks tend to move apparently and more variably, which can potentially indicate that facial landmarks movement between frames can be a good feature for the prediction model.

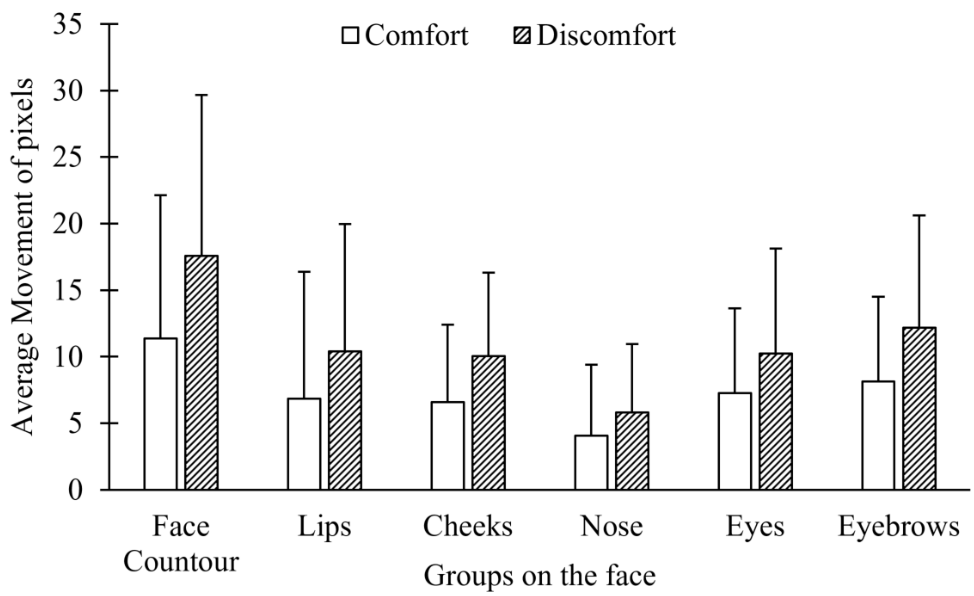

Figure 4: Movement of facial landmarks

2.5.2. Comfort and Discomfort Groups

Based on the previous study stating the change of the feeling from comfort to discomfort in 65%MVC [13], the current study used it as the threshold.

Thus, 0, 15, 33, and 50%MVC and considered in comfort group and 75 and 100%MVC are considered as discomfort group in grip force exertion. Based of our **Research Hypothesis**, since we are proposing an automated method for predicting the high force exertions, the discomfort group was chosen to be the true prediction of the proposed model. The confusion matrix is shows in Table 2.

	Actual Discomfort	Actual Comfort
Predicted Discomfort	True Positive (**TP**)	False Positive (**FP**)
Predicted Comfort	False Negative (**FN**)	True Negative (**TN**)

Table 2: Confusion matrix

2.6. Machine Learning Models

2.6.1. Data re-sampling

The dataset is unbalanced between two levels since it has four comfort levels and two discomfort levels. Unbalanced dataset is an issue that can lead to unreliable learning model. Specifically, when dealing with detecting abnormal situations, such as fraud detection, because the unbalanced dataset can bring biased toward the majority group[24]. Current state-of-art techniques for solving this issue are the followings: 1. Sampling techniques: Down-sampling the majority, Over-sampling the minority and re-sampling both. 2. Cost sensitive learning: Assigning a heavier cost to wrong classification of the minority and focused on reducing the overall cost [25]. Randomly down-sampling the majority is chosen as the strategy in this study. In addition, multiple classifier evaluation indexes, other than overall accuracy, are used to further evaluate our models and overcome the unbalanced dataset problem

2.6.2. Evaluation Methods

In many binary classification cases, dataset will be divided into positive and negative samples, and confusion matrix shown in table 2 will be applied to describe the model's performance. To prevent injuries, detecting the discomfort levels over comfort seems to be more important in actual workplaces. For this reason, this study focus more on detecting discomfort hand force level to prevent injuries, discomfort levels are considered as positive samples and comfort groups as negative samples. Precision, Recall and F1 score were used to evaluate our models' performance in classifying the positive samples. The receiver operating characteristic (ROC) curves and the values of area under curve (AUC) were used to evaluate our models' overall performance. False alarm rate was also addressed in order to better introduce ROC and AUC techniques [26].

Sensitivity, commonly known as recall and true positive rate, represents the ratio of number of correctly classified positive samples to total number of positive samples. In this study, this index is the representation of the probability of true prediction when the subject is exerting force in discomfort level. Recall rate is calculated by using the equation 2.

$$Recall = \frac{True\ positive}{True\ positive + False\ negative} = \frac{True\ positive}{positive} \tag{2}$$

Precision represents the ratio of number of correctly classified positive samples to total number of predicted positive samples output from the model. In this study, precision means that when the model gives a positive sample prediction, what is the probability that this prediction is correct. The precision is calculated using equation 3.

$$Precision = \frac{True\ positive}{True\ positive + False\ positive} \tag{3}$$

F1 score is generally considered as the comprehensive measurement of precision and recall, and the equation used to calculate it is shown in equation 4.

$$F1\ Score = 2 * \frac{precision * recall}{precision + recall} \tag{4}$$

False alarm rate, or false positive rate, is used to measuring how many comfort level samples are wrongly classified as discomfort level in our studies and it is calculated as equation 5.

$$False\ alarm\ rate = \frac{False\ positive}{False\ positive + True\ negative} \tag{5}$$

Through combining true positive and false negative rate, ROC curve is able to make a balance between benefits and thus become common in evaluate the overall performance of binary classification with imbalanced data set. An example of ROC curve is shown in Figure 5. However, since comparing differences between curves is difficult, a single scalar value, known as area under curve (AUC) is introduced. AUC is the area between the diagonal line and the Rcurve. The relative classifier's performance to AUC score is proposed in table 3 [27].

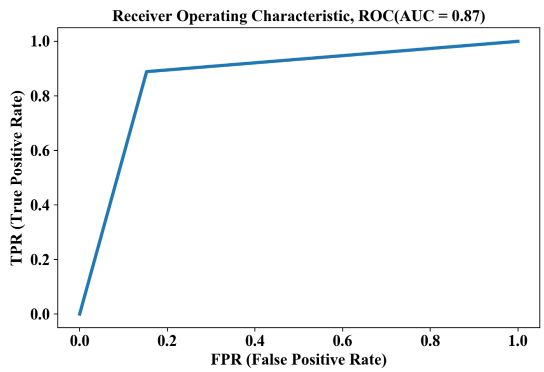

Figure 5: Example ROC Curve

AUC Range	Performance
1-0.9	Excellent
0.9-0.8	Good
0.8-0.7	Fair
0.7-0.6	Poor
0.6-0.5	Fail

Table 3: A general guide to evaluate classifier using AUC

2.6.3. Various Learning Models and Choosing Best Performance

learning models: After all features are extracted and dataset is balanced, the following supervised learning models are trained:

1.Random Forest
2.Support Vector Machine (SVM)
3.Bagging K-Nearest Neighbor
4.Neural Network

Models training: Each model was trained using re-sample balanced dataset and pre-assigned labels. Best parameters of each model was found using grid search method, and specified callback methods for Neural Network.

Testing: Each model was tested using K-fold cross validation approach. Specifically, leave one subject out, leave two subjects out and leave three subjects approaches were applied.

Choosing The Best Model: When testing each model we trained using leave one out cross validation approach at the earlier stage, we found that SVM, Random Forest and Bagging KNN models did not perform well and the Neural Network model outperform these models in significant degree. Therefore, the Neural Network chosen and used it for further analysis. Details information about the performance of each model will be shown in result section.

Details of Neural Network Model: A neural network with 1 input, 3 hidden and 2 output layers as shown in Figure 6 was built. For each hidden layer, 128 neurons was used. The activation function used in the training of network was exponential linear units (ELUs) [28] as defined in equation 2. Batch normalization was used in each hidden layer [29]. In the output layer, two neurons were used for discomfort hand force level and comfort hand force level. The best performance of the network was achieved with using SGD as an optimizer along with binary cross-entropy as a loss function. In addition, Model Check Point and Early Stop techniques were implemented to find the best parameters and stop the training.

$$f(x) = \begin{cases} x & if\ x \geq 0 \\ \alpha(e^x - 1) & if\ x \leq 0 \end{cases} \tag{6}$$

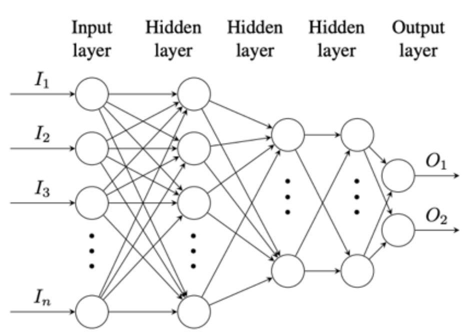

Figure 6: The architecture of a fully connected neural network with three hidden layers.

2.7. Multi-level force exertion and activity tasks classification

Other than two-level classification (Comfort and Discomfort), three-level (i) and activity tasks (ii) classifications were also performed using the same methodology.

(i) For *three-level* force exertions, Neural Network model was trained to classify three levels. No/Low (0%), medium (50%), and high (100%) force exertion levels considered as the classification groups. The facial feature are given in figure 7.

The model was tested using one leave out cross validation approach. Through setting each label as positive sample in alternate order, multi-label classification model can be evaluated through the precision, recall and F1 score of each label.

(ii) For *activity classification*, Talking task was considered as it is frequent in actual workplaces. For testing the model's robustness in classifying the activity tasks from force exertions, Talking video were recorded. The activity data only used for testing in the best model, and no training was performed for activity classifications.

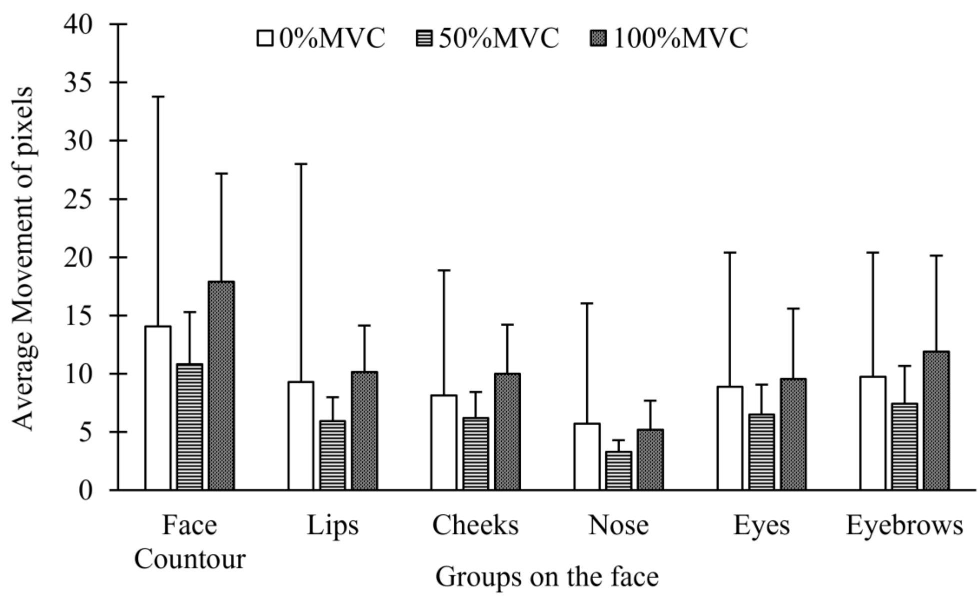

Figure 7: Movement of facial landmarks for three-level classification

3. Results

3.1. Results of different classifications

The performance of various models tested with leave one out approach is given in Table 4. The Neural Network Model shows higher accuracy in comparison to rest of the models.

Model	Precision	Recall	F1 Score	AUC	Overall Accuracy
SVM	0.52	0.61	0.56	0.67	0.69
Random Forest	0.51	0.81	0.62	0.71	0.68
Bagging KNN	0.51	0.75	0.61	0.70	0.68
Neural Network	**0.97**	**0.78**	**0.86**	**0.88**	**0.92**

Table 4: Performance of each model

3.2. Classification of Comfort and Discomfort Force Exertions

Neural Network Model was tested individually using leave one out, leave two out and leave three out cross validation approaches. For leave three out approach, there were 816 combinations to select 3 subjects from 18 total subjects. Due to heavy process of training 816 models, 153 combinations were randomly chosen for testing. The testing results of each approach are shown in Table 5.

Tesing Approach	Precision	Recall	F1 Score	AUC	Overall Accuracy
Leave 1 out	0.97	0.78	0.86	0.88	0.92
Leave 2 out	0.85	0.81	0.83	0.87	0.89
Leave 3 out	0.82	0.75	0.78	0.83	0.86

Table 5: Results of each testing approach

The overall accuracy decreased from 92% to 86% as expected. Even with leave three subjects out, the model still showed satisfactory performance with AUC larger than 0.8, which can demonstrate the effectiveness of the selected features and the neural network model.

The model's performance at each force level was reported in Table 6. Accuracy was applied to evaluate the classification performance at each force level. As expected, lower accuracy was observed in 50% and 75% levels.

Force Level (%MVC)	0	15	33	50	75	100
Accuracy (%)	100	100	100	94	67	89

Table 6: Prediction accuracy at each force level

3.3. Various Force exertion classification

Table 7 demonstrate the result of each label.

Label	Precision	Recall	F1 Score
No/Low	0.79	0.61	0.69
Medium	0.67	0.78	0.72
High	0.79	0.83	0.81

Table 7: Results of three level classification

The overall accuracy of the three level classification model is 74%.

3.4. Test Model Robustness in Activity Tasks

Since the talking was performed without any force exertions in this studies, the talking situation was considered as comfort level. The best model was used for analyzing the talking situation. Model's accuracy was 77.8% when tested on talking data.

4. Discussion

Computer vision and machine learning can predict the force exertion level using extracted facial features and provides a novel approach for such estimation. Understanding force exertion levels has important implications across domains and applications, and in this work, we demonstrate the approach in the context of workplace injuries. Specifically, varying levels of force and duration/frequency of these forces are predictive of musculoskeletal injuries. This section provides more discussion on using machine learning in prediction of force exertion level and provides more insights on the feature selection that is introduced in this work.

4.1. Deep-face and Neural Networks in Classifying High Force Exposures

There are various methodologies [30, 17, 18, 16, 21, 19, 20, 31] proposed that can achieve facial recognition but the methodology proposed in [15] outperforms other methods and results in the accuracy of 97.35% in Labeled Faces in the Wild (LFW) dataset, reducing the error in face recognition of current state-of-the-art by more than 27%. This method is more robust and the explanation on DeepFace is discussed in section 2.4. The 9 layer neural network used in Deepface makes it more robust to detect faces in the video for our study and henceforth extract relevant features from the video frames. These facial features represent a key component for force classification.

The neural networks are known to be universal approximators [32] and hence are used to identify the underlying function explaining the relationship between the features and response variable. This approach is used extracted features to classify the force exertions and add additional novelty by leveraging the underlying physiological mechanisms of generating muscle forces to improve force classification accuracy.

4.2. Non-contact Exposure Assessment

The force exertions has been considered as one of the main contributing factors in current risk assessment tools [33, 34, 35]. The high variability of the identified risk score with respect to the estimated force exertion parameters is reported in current assessment tools. For example, the Strain Index Assessment [35] score will double if the intensity of the exertion changes from 20% to 40% [9]. In addition, Bao et al. reported weak correlation values between the ergonomists estimates and the worker's self-reports for pinch and grip force. Further exploration suggested among relationships of worker's self-reports, the ergonomist's estimates and the directly measured hand forces [36]. The proposed non-contact assessment method for classifying force levels can provide an objective automated estimations of hand forces.

4.3. Classification of Comfort/Discomfort Force Exertions

The performance of the best model for predicting the comfort (0, 15, 33, 50%MVC) and discomfort (75 and 100%MVC) is given at table 5. The proposed new objective approach can identify the high force exertions with over 90% accuracy. This model and approach could be translated into widely accessible tool in workplace which is not intrusive to workers.

"U" pattern was observed in reporting each level's accuracy. At levels near the considered threshhold (65%MVC) the accuracy of classification would decrease as expected.

4.4. Classification of Low/Medium/High Force Exertions

The performance of the model for predicting the three-level force groups was reported to be 74%. The lower performance could be due to the fact that facial features of minimum hand force exertion and medium hand force exertion do not have significant difference as we can observe in figure. Results from table 7 demonstrates that the model perform well in identifying maximum hand force exertion while has shortage in identifying medium and minimum hand force exertion. The average movement of 50% was calculated to be less than the resting level (0%). This seems to be unlikely, but one potential reason could be the attention of the subject in holding the 50% force. The subject's face seems to move more in resting while they are not concentrating on one task (more variation in 0% than 50%).

4.5. Analysis of Activity Tasks with Current Model

The potential reason for unsatisfied performance of the best model when tested on talking data is that facial movement are more significant when people are talking, especially the movement of lip and face contour, which can confuse the model. Walking situation was also being concerned at the early stage of our studies. However, significant loss of facial data when subjects are walking caused our attention. Subjects' face were not being clearly detected in some of the frames, which caused the failure of landmarks plotting. The applicability of DeepFace algorithm when people are under intense body movement like walking is under suspicion, although DeepFace algorithm is proven to be capable of detecting and centralizing the face under various scenes.

References

[1] U. B. of Labor Statistics, "Survey of occupational injuries and illnesses, in cooperation with participating state agencies," *U.S Department of Labor*, 2016.

[2] R. Wells, "Why have we not solved the msd problem?" *Work*, vol. 34, no. 1, pp. 117–121, 2009.

[3] Liberty Mutual Research Institute for Safety, "2017 Liberty Mutual Workplace Safety Index," *Liberty Mutual Research Institute for Safety*, pp. 39–40, 2017. [Online]. Available: https://www.libertymutualgroup.com/about-liberty-mutual-site/news-site/Pages/2017-Liberty-Mutual-Workplace-Safety-Index.aspx

[4] National Institute for Occupational Safety and Health (NIOSH), "Musculoskeletal disorders and workplace factors: a critical review of epidemiologic evidence for WMSDs of the neck, upper extremity and low back," 1997. [Online]. Available: https://stacks.cdc.gov/view/cdc/21745

[5] P. W. Buckle and J. Jason Devereux, "The nature of work-related neck and upper limb musculoskeletal disorders," *Applied Ergonomics*, vol. 33, no. 3, pp. 207–217, 2002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0003687002000145

[6] W. M. Keyserling, "Workplace Risk Factors and Occupational Musculoskeletal Disorders, Part 2: A Review of Biomechanical and Psychophysical Research on Risk Factors Associated with Upper Extremity Disorders," *AIHAJ - American Industrial Hygiene Association*, vol. 61, no. 2, pp. 231–243, mar 2000. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/15298660008984532

[7] S. P. Schneider, "Musculoskeletal injuries in construction: A review of the literature," pp. 1056–1064, nov 2001. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/104732201753214161

[8] K. G. Hauret, B. H. Jones, S. H. Bullock, M. Canham-Chervak, and S. Canada, "Musculoskeletal injuries: Description of an under-recognized injury problem among military personnel," *American Journal of Preventive Medicine*, vol. 38, no. 1 SUPPL., pp. S61–S70, jan 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0749379709006746

[9] E. Koppelaar and R. Wells, "Comparison of measurement methods for quantifying hand force," *Ergonomics*, vol. 48, no. 8, pp. 983–1007, jun 2005. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/00140130500120841

[10] A. Schwarzer, "The prevalence and clinical features of internal disc disruption in patients with chronic lbp," *Spine*, vol. 20, pp. 1878–1883, 1995.

[11] M. A. Adams, B. J. Freeman, H. P. Morrison, I. W. Nelson, and P. Dolan, "Mechanical initiation of intervertebral disc degeneration," *Spine*, vol. 25, no. 13, pp. 1625–1636, 2000.

[12] B. Fung, K. Chan, L. Lam, S. Cheung, N. Choy, K. Chu, L. Chung, W. Liu, K. Tai, S. Yung *et al.*, "Study of wrist posture, loading and repetitive motion as risk factors for developing carpal tunnel syndrome," *Hand surgery*, vol. 12, no. 01, pp. 13–18, 2007.

[13] Y.-K. Kong, D.-M. Kim, K.-S. Lee, and M.-C. Jung, "Comparison of comfort, discomfort, and continuum ratings of force levels and hand regions during gripping exertions," *Applied Ergonomics*, vol. 43, no. 2, pp. 283–289, 2012.

[14] P. Vink and S. Hallbeck, "Comfort and discomfort studies demonstrate the need for a new model," 2012.

[15] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1701–1708.

[16] O. Barkan, J. Weill, L. Wolf, and H. Aronowitz, "Fast high dimensional vector multiplication face recognition," in *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 1960–1967.

[17] X. Cao, D. Wipf, F. Wen, G. Duan, and J. Sun, "A practical transfer learning algorithm for face verification," in *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 3208–3215.

[18] D. Chen, X. Cao, F. Wen, and J. Sun, "Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 3025–3032.

[19] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a similarity metric discriminatively, with application to face verification," in *Computer Vision*

and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 1. IEEE, 2005, pp. 539–546.

[20] G. B. Huang, H. Lee, and E. Learned-Miller, "Learning hierarchical representations for face verification with convolutional deep belief networks," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 2518–2525.

[21] Y. Sun, X. Wang, and X. Tang, "Deep convolutional network cascade for facial point detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2013, pp. 3476–3483.

[22] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to statistical learning*. Springer, 2013, vol. 112.

[23] M. L. Bermingham, R. Pong-Wong, A. Spiliopoulou, C. Hayward, I. Rudan, H. Campbell, A. F. Wright, J. F. Wilson, F. Agakov, P. Navarro *et al.*, "Application of high-dimensional feature selection: evaluation for genomic prediction in man," *Scientific reports*, vol. 5, p. 10312, 2015.

[24] B. Krawczyk, "Learning from imbalanced data: open challenges and future directions," *Progress in Artificial Intelligence*, vol. 5, no. 4, pp. 221–232, Nov 2016. [Online]. Available: https://doi.org/10.1007/s13748-016-0094-0

[25] C. Chen and L. Breiman, "Using random forest to learn imbalanced data," *University of California, Berkeley*, 01 2004.

[26] A. Tharwat, "Classification assessment methods," *Applied Computing and Informatics*, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2210832718301546

[27] F. Li and H. He, "Assessing the accuracy of diagnostic tests," *Shanghai archives of psychiatry*, vol. 30, no. 3, pp. 207–212, 06 2018. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/30858674

[28] D.-A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (elus)," 01 2016.

[29] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 02 2015.

[30] P. J. Phillips, J. R. Beveridge, B. A. Draper, G. Givens, A. J. O'Toole, D. S. Bolme, J. Dunlop, Y. M. Lui, H. Sahibzada, and S. Weimer, "An introduction to the good, the bad, amp; the ugly face recognition challenge problem," in *Face and Gesture 2011*, March 2011, pp. 346–353.

[31] M. Osadchy, "Synergistic face detection and pose estimation with energy-based models," *Journal of machine learning research : JMLR.*, 2007.

[32] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.

[33] S. Hignett and L. McAtamney, "Rapid entire body assessment," in *Handbook of Human Factors and Ergonomics Methods*. CRC Press, 2004, pp. 97–108.

[34] L. McAtamney and E. N. Corlett, "Rula: a survey method for the investigation of work-related upper limb disorders," *Applied ergonomics*, vol. 24, no. 2, pp. 91–99, 1993.

[35] J. Steven Moore and A. Garg, "The strain index: a proposed method to analyze jobs for risk of distal upper extremity disorders," *American Industrial Hygiene Association Journal*, vol. 56, no. 5, pp. 443–458, 1995.

[36] S. Bao, N. Howard, P. Spielholz, and B. Silverstein, "Quantifying repetitive hand activity for epidemiological research on musculoskeletal disorders–part ii: comparison of different methods of measuring force level and repetitiveness," *Ergonomics*, vol. 49, no. 4, pp. 381–392, 2006.