Predicting Standing Reach Postures using Deep Neural Networks

Matthew P. Reed[1], Sheila M. Ebert[1], Clive D'Souza[2], Monica L.H. Jones[1]
[1]University of Michigan [2]University of Pittsburgh

Proactive ergonomic analysis of occupational tasks using digital human figure models requires accurate prediction of worker postures. A wide range of methods have been proposed and used, including posture libraries, statistical methods including regression, and optimization approaches that incorporate hypothesized criteria such as strength maximization. A common challenge in the implementation of any method is ensuring that the resulting postures are consistent with the kinematic linkage of the figure model, the boundary constraints are satisfied, including those relevant to the task, and that the figure remains in balance after taking into account external forces. Neural network methods have been applied to human posture prediction for more than 25 years, but successful implementation for human posture prediction requires careful consideration of the relevant constraints. This paper describes the implementation of DNN methods within the Human Motion Simulation Framework, which provides a hierarchical structure for posture and motion prediction applicable to any human figure model.

INTRODUCTION

Posture prediction for ergonomics analysis using human figure models has been a topic of research for several decades (Chaffin 2001, Duffy 2013). A wide range of methods have been proposed, including posture libraries (Park et al. 2005), regression and other statistical methods (Faraway and Reed 2007) and optimization based on hypothesized biomechanical criteria (Yang et al. 2011). That none of these methods has become ubiquitous in ergonomics analysis highlights the fact that all are substantially limited in generalizability. Efforts at implementation also uncover difficult issues that may not be immediately apparent during the original development of the methods.

In addition to having documented quantitative performance relative to human subject data, an accurate method also must:

- produce results that are consistent with human figure model linkages scaled over the wide range of human body dimension variability,
- comply with the boundary constraints, such as the feet on the floor, the hand(s) on the task targets, and body bracing on external surfaces,
- not result in interpenetration of any obstacles, while at the same time taking advantage of bracing opportunities where appropriate, and
- remain in balance while taking into account any external forces, for example those associated with task exertions.

As a consequence of these challenges, all current tools use hybrid approaches that combine aspects of kinematic optimization (including data-guided inverse kinematics), biomechanical optimization (for example, limiting shoulder moments), and knowledge-based heuristics (such as determining the appropriate trade-off between stoop and squat in lifting by reference to data). The knowledge obtained from human subject data may be embedded in weights that are applied during the optimization process.

Progress in posture prediction for ergonomic analysis has been slow in part because every human figure model uses a different kinematic linkage, with different joint angle definitions and degrees of freedom. Consequently, any method formulation that is dependent on a specific model is unlikely to be readily implementable in other tools.

Artificial Neural Networks and Deep Learning

The past ten years have seen an explosion of research and application of deep neural networks (DNN) and their variants (Aggarwal 2018). The availability of highly capable, freely available software tools such as PyTorch and TensorFlow have made DNN methods available to researchers in every discipline (Abadi et al. 2016). At a basic level, DNN methods enable the fitting of networks of interconnected elements with very high numbers of parameters in a robust and rapid manner. These networks capture complex non-linearities in training data that are difficult to discover and exploit in hand-crafted statistical models. Although most applications of DNN methods are focused on classification problems, neural network methods are also widely applied to regression problems that have vectors of continuous variables as outputs.

DNN methods have been applied to human posture prediction for at least 25 years (Jung and Park 1994, Perez and Nussbaum 2008, Zhang et al. 2010). Some recent papers have used DNN and related methods for human motion prediction in the context of collaborative robotics (Cheng et al. 2019, Kratzer et al. 2020, Zhang et al. 2020). A recent study with direct applicability to ergonomics analysis used DNN methods to model lifting postures and focused on landmark locations rather than joint angles (Aghazadeh et al. 2020). The current paper focuses on the implementation of DNN landmark predictions applicable to any human figure model.

For posture prediction, a large set of training data spanning the relevant posture is obtained, typically from optical motion capture systems. The parameterization of the data is critical. In most cases, the posture is parameterized by joint angles or quaternions, but as noted above this ties the predictions to the particular linkage definition. Prediction of Cartesian landmark locations (including joint center locations) has improved the performance of DNN models (Li et al. 2021). Equally critical is the parameterization of the task specification. In many posture-prediction applications this can be very general ("walk"), but for ergonomics analysis a detailed parameterization of the task is required (hand locations, stance, required force), as well as the characteristics of the subject, are needed.

This paper describes the development of a DNN posture-prediction method incorporated into the Human Motion Simulation (Humosim Framework), a hierarchical, model-agnostic approach for predicting task-oriented postures and motions (Reed et al. 2006). The overall goal of the paper is to present and discuss the challenges and potential solutions for using this powerful data analysis tool more widely for posture prediction in ergonomics applications.

METHODS

Data Source

As part of a larger laboratory study of task-oriented postures, 25 men and women with a wide range of age and body size performed one-hand, push-button reaches to a wide range of target locations scaled by their stature. Figure 1 shows the range of target locations, which were placed on azimuths from 45 degrees to the left of the subject to 90 degrees to the right. The reach distances were scaled relative to the subject's maximum reach capability in each direction (see Figure 2 for an example), and the heights were set to 110%, 63%, and 41% of stature (approximately overhead, elbow height, and mid-thigh height). Foot placements were prescribed as either side-by-side or with the right foot in front of the left, with subject-selected lateral spacing in all cases.

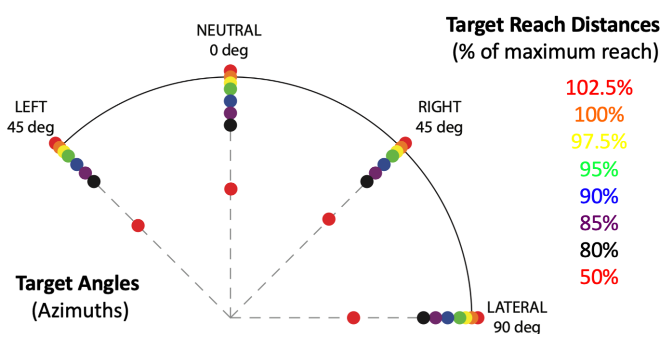

Figure 1. Top-view schematic of target locations.

The participant's body dimensions were recorded using standard anthropometric techniques and an optical motion capture system (Vicon) was used to track their postures. Figure 2 shows the retroreflective markers attached to the participants. Custom software was used to estimate joint center locations based on the surface landmarks. The participant stood on a force platform with self-selected side-by-side or tandem (one foot further forward) foot placements.

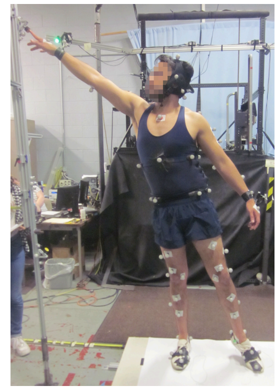

Figure 2. Retroreflective motion-capture targets on a subject completing a maximum lateral reach.

The current results are based on predicting the locations of a set of joint centers and landmarks shown in Figure 3 in Cartesian space. This approach provides the advantages noted above for applying the results to any figure linkage having these joints and landmarks. Prior to prediction, the location data were expressed relative to the centroid of the initial base of support and scaled by participant stature. After processing and verification, a total of 3103 trials were available for analysis. The mean (sd) of stature was 1741 (85) mm; BMI 25.6 (4.8) kg/m^2; age 38 (17) years.

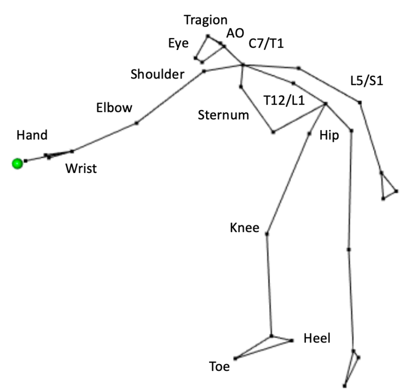

Figure 3. Extracted posture, showing modeled joints and landmarks. A large number of additional landmarks are available in the underlying data, but this constitutes approximately a minimal set necessary to obtain a complete posture sufficient for whole-body ergonomics analysis.

DNN Modeling

The task input was described by the azimuth, elevation, and distance of the target (i.e., spherical coordinates) scaled by stature. The participant's stature, BMI, sitting height, sex, and age were also used as predictors, along with the fore-aft distance between the participant's foot placement (tandem offset). The output was the 3D stature-scaled coordinates of 29 landmarks and joints.

A DNN model structure was determined by typical methods that involve searching for values of various hyperparameters that yield the best predictions. The model was fit in R using the caret package to interact with the Keras library interface to TensorFlow (Gulli and Pal 2017). The ADAM optimizer was used with a mean-square-error (MSE) criterion. The data were randomly split 80/20 into training and test sets. Hyperparameter tuning was conducted on a grid using random sampling of 80% of the training data. At each training epoch, 30% of the training data were randomly withheld for validation. Figure 4 shows the selected model structure, which was trained for 120 epochs.

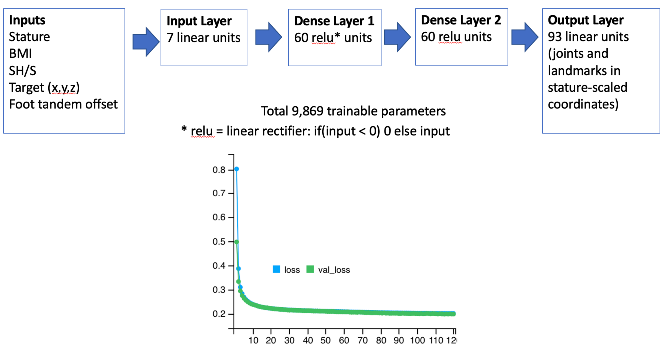

Figure 4. DNN model structure and plot of training and validation loss function (MSE) across 120 epochs

Implementation

The DNN model predicts a set of landmarks and joints that then must be interpreted through a kinematic linkage, considering the requirements and constraints mentioned earlier, to obtain a complete posture prediction. For the current demonstration, a partial implementation of the Humosim Framework in the R language was used.

Figure 5 shows the structure of the prediction process within the Framework. Because the results are not guaranteed to be consistent with any particular kinematic linkage, certain degrees of freedom are prioritized. The model first sets the pelvis position and orientation according to the prediction, adjusting the height to ensure that the inverse kinematics (IK) algorithms can place the feet on the predicted locations. The IK matches the predicted plane of each lower extremity but may produce a different knee location from the prediction depending on the particular figure linkage.

The lumbar spine degrees of freedom are then adjusted to best match the predicted thorax orientation and the clavicle segments set to place the glenohumeral joints as close as possible to the predicted locations. IK is then used to position the upper extremities to respect the task constraints while matching the predicted elbow locations as closely as possible. The head and neck posture is generated similarly, with the head orientation prioritized over location. Balance is then checked, based on the particular segment mass distribution of the figure. If the projected center of pressure lies outside of the base of support the pelvis location is adjusted and the rest of the segments iterated in the same sequence. Also, if the upper-extremity IK is not capable of achieving the required hand position, the torso iterates to improve the shoulder location. For reaches similar to those in the training set, the method usually converges in a single iteration, although a balance correction is needed, or, if the posture is extreme, then several iterations are sometimes needed.

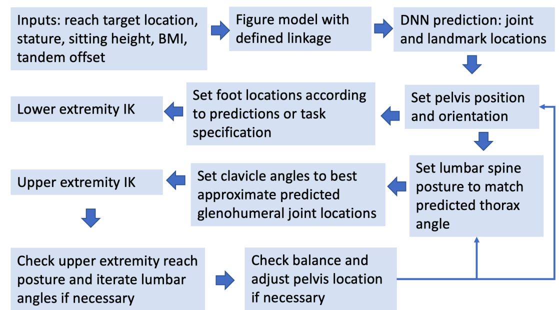

Figure 5. Schematic of simplified Humosim Framework prediction process for standing reach postures using DNN.

RESULTS

DNN Predictions

The DNN predictions were evaluated against the 20% of data that were withheld. Figure 6 shows comparisons of observed and predicted postures for a few reach locations. Neglecting the left upper extremity, which was not involved in the task, MSE for joint locations across the test set increased with the distance along the linkage from the feet: 17 mm at the right ankle joint, 38 mm at the right hip joint, 43 mm at C7/T1, and 61 mm at the right wrist joint. We note that the MSE is most useful for comparing between alternative models, as large variability in the underlying data that is not related to the predictors. As a reference dimension, the mean stature in the data set was 1693 mm, so the largest of these MSEs is about 4% of stature.

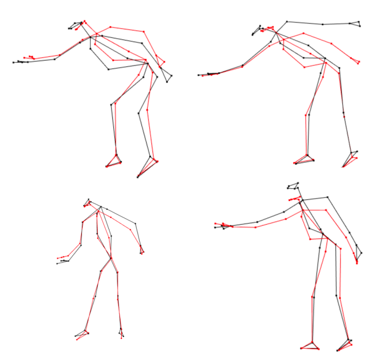

Figure 6. Comparison observed (black) postures from the test set with predicted postures (red) from the DNN model.

The model can also be used to explore the effects of subject covariates on posture. Figure 7 depicts two reach cases that demonstrate the effects of varying stature when reaching to the same target. Sex and age were not important predictors after accounting for overall body size.

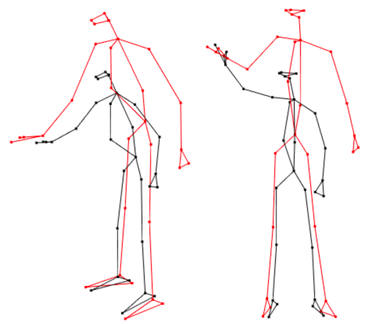

Figure 7. Effects of stature (black = 1520 mm, red = 1870 mm) on predictions of standing reach postures.

Implementation

Figure 8 shows some examples of the Framework implementation of the DNN along with the corresponding DNN predictions. The figures show the blue Framework prediction deviating from the green unadjusted DNN prediction as needed to reach the right-hand target and maintain balance. The bottom two reaches are outside of the range of the posture in the training data, showing the extrapolation capability of both the DNN and the framework.

The DNN prediction is essentially instantaneous, providing support for real-time interaction. The DNN is capable of interpolating and extrapolating very smoothly, although anomalies such as unrealistic segment proportions are observed in more-extreme conditions. However, the Framework implementation is not meaningfully affected by these discrepancies and continues to provide plausible results well beyond the training data.

DISCUSSION

The DNN approach generates a "black box" model with model coefficients that are not amenable to interpretation. Instead, the performance of the model is evaluated qualitatively by manipulating the input parameters and quantitatively by comparing with withheld data. The Framework implementation compensates for the limitations of the predictions, including the inevitable mismatches between the figure segment lengths and those generated by the predictions. Conceptually, all segment lengths could be included as predictors, but the results would still not perfectly conform with the linkage and so minimal benefit would be obtained. The posture could be parameterized as joint angles, rather than as Cartesian coordinates, but implementation would still require conversion to Cartesian space via the figure linkage and the other adjustments needed to meet the task constraints. In our experience, joint-angle-based methods require larger adjustments due to non-linearities inherent in the segment motions. For example, small discrepancies in proximal joint angles produce magnified effects at the end effector (hand, head, foot) locations. With the current approach, these discrepancies do not accumulate across the linkage.

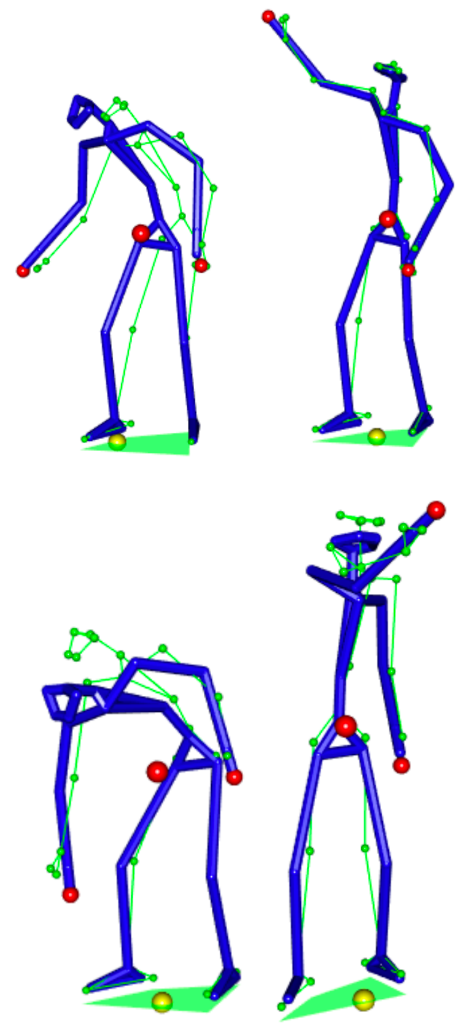

Figure 8. Examples of one-hand standing reach predictions from the Humosim Framework implementation developed to demonstrate the DNN application (blue lines). The DNN predictions are shown in green with thin lines. The small red spheres are hand targets. The large red sphere is the whole-body center of mass, computed based on the posture and segment masses. The yellow sphere is the projected center of mass relative to the green base of support.

The Humosim Framework approach explicitly prioritizes some degrees of freedom based on biomechanical understanding. For example, shoulder location is more important than elbow location for most ergonomic analyses, and pelvis position and torso orientation are very important for determining balance and lower-back loading. In our experience, this prioritization produces better results for ergonomic analyses than would be obtained by, for example, optimizing the posture to minimize the discrepancy across all predicted joints and landmarks. Recent comparisons of model-predicted and measured postures have highlighted the importance of evaluating the effects of model discrepancies on the ergonomic analysis (for example, low-back loading) rather than focusing exclusively on measures of posture error (Potash et al. 2022). That is, what ultimately matters for ergonomics analysis using human figure models is that the analysis correctly identifies potentially injurious or stressful scenarios.

The choice of hyperparameters (for example, the number of nodes in each layer) may not be optimal in a global sense. That is, the MSE was calculated across all posture degrees of freedom without consideration of the importance for ergonomics analysis. For the one-hand reaches, the non-task arm posture is of minimal importance, and hence it would be reasonable to exclude those degrees of freedom or reduce the weight in the analysis. Ultimately, though, the model performance on withheld from withheld subjects should be used to understand prediction accuracy and precision.

As with other laboratory studies, the current results may not be representative of postures that an experienced worker would use for similar tasks. The current work is limited by the scope of postures in the training set and the stick figure used to represent the outcomes in this paper is primitive. Implementation with a full 3D surface manikin could reveal additional limitations, such as with shoulder postures.

We plan to extend this work in several ways. We have access to more than 50,000 task-oriented motions that provide an opportunity to expand the range of postures the model can predict. We also anticipate that expanding the model to predict motions will be straightforward: instead of predicting single locations for each joint, we instead predict the control points of splines fitted to motion data (Faraway and Reed 2007). Preliminary investigation of this approach is promising. As the model becomes more capable and mature, we plan to make it publicly available to facilitate implementation.

ACKNOWLEDGMENTS

This work was supported by the Partners of the Human Motion Simulation Laboratory and by the USCAR Ergonomics Task Force. We thank the subjects for their many hours of patient participaton and the many students and staff who contributed to the data collection and analysis, including students Yuhao Chen, Yue Luo, Nicholas Sandhu, Justin Haney, Tianke Kirk Wang, Mia Carminita, Baibhav Panda, Claire Stemper, Junyi Cheng, Zhongfu Bai, Sabrina Lin, Austin Gier, Charles Marchand, Sam Harris, Kieran O'brien, Kip Schimmoeller, Kai Schiefer, Rachel Prell, Ashton Russo, and Dylan Plummer.

REFERENCES

Abadi, M., Barham, M., Chen, J. and 18 others (2016). TensorFlow: A System for Large-Scale Machine Learning. *Proc. 12th USENIX Symposium on Operating Systems Design and Implementation.*

Aggarwal, C.C. (2018). *Neural Networks and Deep Learning*. Springer, Berlin.

Aghazadeh, N. Arjmand, A.M. Nasrabadi (2020). Coupled artificial neural networks to estimate 3D whole-body posture, lumbosacral moments, and spinal loads during load-handling activities, *Journal of Biomechanics*, 102:109332

Chaffin, D.B., ed. (2001). *Digital Human Modeling.* SAE International, Warrendale, PA.

Cheng, Y., Zhao, W., Liu, C., and Tomizuka, M. (2019), Human motion prediction using semi-adaptable neural networks, *Proceedings American Controls Conference*, 4884–4890.

Duffy, V.G. (ed.) (2013). *Handbook of Digital Human Modeling Research for Applied Ergonomics and Human Factors Engineering*. CRC Press, Boca Raton, FL.

Faraway, J. and Reed, M.P. (2007). Statistics for digital human motion modeling in ergonomics. *Technometrics*, 49(3):277-290.

Gulli, A, and Pal, S. (2017). *Deep learning with Keras*. Packt Publishing Ltd.

Jung, E.S. and Park, S. (1994). Prediction of human reach posture using a neural network for ergonomic man models. *Computers in Industrial Engineering*, 27(1-4):369-372.

Kratzer, P, Toussaint, M., and Mainprice, J. (2020). Prediction of human full-body movements with motion optimization and recurrent neural networks, Proc. IEEE International Conference on Robotic Automation, 1792–1798.

Li, L., S. Prabhu, Z. Xie, H. Wang, L. Lu and X. Xu (2021). Lifting posture prediction with generative models for improving occupational safety, *IEEE Transactions on Human-Machine Systems*, vol. 51, no. 5, pp. 494-503.

Park, W., Martin, B.J., Choe, S., Chaffin, D.B., and Reed, M.P. (2005). Representing and identifying alternative movement techniques for goal-directed manual tasks. *Journal of Biomechanics*, 38:519-527.

Perez, M.A. Nussbaum, M.A. (2008) A neural network model for predicting postures during non-repetitive manual materials handling tasks, *Ergonomics*, 51:10, 1549-1564

Potash, Z, Harari, Y., and Riemer, R. (2022). Effect of using real motion versus predicted motion as input for digital human modeling of back and shoulder loads during manual material handling. *Applied Ergonomics*, 101:103675

Reed, M.P., Faraway, J., Chaffin, D.B., and Martin, B.J. (2006). The HUMOSIM Ergonomics Framework: A new approach to digital human simulation for ergonomic analysis. *Proc. 2006 SAE Digital Human Modeling for Design and Engineering Conference.* SAE International.

Yang, Y., Marler, T., and Rahmatalla, S. (2011). Multi-objective optimization-based method for kinematic posture prediction: development and validation. *Robotica*, 29:245-253.

Zhang, B., Horvath, J.F.M., Molenbroek, J.F.M., and Snijders, C. (2010). Using artificial neural networks for human body posture prediction. *International Journal of Industrial Ergonomics*, 40 :414e424

Zhang, J., Liu, H., Chang, Q., Wang, L., and Gao, R. X. (2020). Recurrent neural network for motion trajectory prediction in human-robot collaborative assembly, CIRP Ann., vol. 69, no. 1, pp. 9–12, 2020.