hazards (i.e., human factors). The literature also discusses the issue of offering training in a language that is preferable to the workers.

**Table 2. The Shortcomings of Traditional Construction Safety Training**

| Limitation | Example References |
|---|---|
| 1. Poorly designed training material and resources | (Lin et al. 2018; Namian et al. 2016; Wilkins 2011) |
| 2. Training does not encourage training transfer (i.e., application of learned concepts in the workplace) | (Bahn and Barratt-Pugh 2012; Namian et al. 2016; Schoenfisch et al. 2017) |
| 3. Training is not designed upon on a clear understanding of why workers fail to recognize and manage safety hazards | (Albert et al. 2014a; Jeelani et al. 2018b; Namian et al. 2016) |
| 4. Training disregards the difference between explicit and tacit knowledge which may require district training strategies | (Albert et al. 2014a; Hallowell 2012; Lingard et al. 2015) |
| 5. Training does not promote active learning (e.g., lecture-based) | (Bhandari et al. 2019; Namian et al. 2016; Wilkins 2011) |
| 6. Use of low-engagement training approaches (e.g., trainer centric with limited discussion) | (Demirkesen and Arditi 2015; Haslam et al. 2005; Namian et al. 2016) |
| 7. Training is delivered using pedagogical approaches as opposed to andragogical approaches (i.e., optimized for adult learners) | (Bhandari and Hallowell 2017; Bhandari et al. 2019; Wilkins 2011) |
| 8. Training does not facilitate inductive and analytical learning | (Bhandari and Hallowell 2017; Namian et al. 2016) |
| 9. Training is generic and lacks context-based learning experiences | (Bahn and Barratt-Pugh 2012; Bhandari and Hallowell 2017) |
| 10. Training that is not delivered in the worker's preferred language | (Canales et al. 2009; Cunningham et al. 2018; Lin et al. 2018) |
| 11. Training that is not designed to maximize knowledge retention and skill acquisition | (Haslam et al. 2005; Jeelani et al. 2018b; Namian et al. 2016) |
| 12. Training is not tailored/personalized to the training needs of individual workers | (Jeelani et al. 2017; Jeelani et al. 2018b) |
| 13. Training without opportunities for practice, feedback, self-reflection, and retention assessment | (Albert et al. 2014a; Jeelani et al. 2018b; Pereira et al. 2018) |
| 14. Training minimally targets emotions and situational interest | (Bhandari and Hallowell 2017; Bhandari et al. 2019) |
| 15. Trainers unaware of the theory of learning and skill acquisition (e.g., visual, kinesthetic learning) | (Haslam et al. 2005; Li et al. 2012; Wilkins 2011) |

Finally, Table 3 presents some of the non-traditional safety training methods that have been proposed in the literature to overcome several of the discussed shortcomings. These include several visualizations, activity-based, and technology-driven safety training solutions.

**Table 3. Selected Nontraditional Construction Safety Training Delivery Methods**

| Delivery Method | Description | Example References |
|---|---|---|
| 1. E-learning tools | Use of the internet or a storage media (e.g., CD) to deliver health and safety training (video lectures, readings, and interactive tools) | (Acar et al. 2008; Ho and Dzeng 2010) |
| 2. Peer-led Training | Workers receive safety training from their experienced peers | (Sinyai et al. 2013; Williams Jr et al. 2010) |
| 3. Game Technology-based Safety Training | Game engine technology is used in creating a 3D virtual environment for executing certain construction operations safely or scenario-based simulations. The 3D environment can be displayed on monitor or VR goggles | (Guo et al. 2012; Li et al. 2012; Lin et al. 2018; Mo et al. 2018; Zhao and Ye 2012) |
| 4. Immersive Virtual Environment (IVE) | A 3D Virtual construction site environment is displayed on Immersive VR power-wall and can be seen through active glasses | (Sacks et al. 2013) |
| 5. System for Augmented Virtuality Environment Safety (SAVES) | A high-fidelity 3D environment was developed to immerse workers in different work scenarios and assess their hazard recognition skill | (Albert et al. 2014a) |
| 6. Mixed Reality (MR) Safety Training System | Integrating tracking system, game engine, and VR goggles in creating virtual construction environment on the googles | (Bosché et al. 2015) |
| 7. Participatory Videos (PV) Intervention Training | Workers are filmed acting while performing certain operations. Then, the video is displayed and discussed during the training session (bottom-up approach) | (Lingard et al. 2015) |
| 8. Naturalistic Injury Simulation (NIS) | Live safety demos are demonstrated resembling actual construction injuries using artificial body parts. The demos target the workers' emotions | (Bhandari and Hallowell 2017) |
| 9. 360-Degree Panorama Safety Training | Actual 360-Degree images of the construction site are taken, and safety-related layers are augmented for the trainee to detect the site's hazards | (Jeelani et al. 2017; Pereira et al. 2018) |
| 10. Personalized safety Training using Eye-Tracking | Leveraging computer vision and eye-tracking technologies in developing personalized hazard recognition training | (Jeelani et al. 2018a; Jeelani et al. 2018b) |

## DISCUSSIONS AND CONCLUSIONS

Despite the adoption of training interventions, an unacceptable number of injuries continue to be reported. (Albert et al. 2014b). Therefore, it is important to understand why training efforts are failing. Therefore, the current research effort focused on understanding barriers to effective training and outlining shortcomings associated with traditionally adopted training methods. A content analysis of existing research revealed 12 barriers to effective training and 15 shortcomings associated with traditional training practices. The research also identified several newly proposed training interventions that seek to address some of the limitations associated with traditional training methods.

The findings of the current research can be used to set the agenda for future construction safety research. For example, efforts must focus on tackling some of the widespread barriers to effective safety training. Moreover, training interventions that address the identified limitations of existing training methods must be developed. Although the research identified a few safety training interventions that have been proposed to address some of the identified limitations, several of the limitations continue to be unaddressed. Future research may focus on these areas to promote safety in the construction industry.

## REFERENCES


Acar, E., Wall, J., McNamee, F., Carney, M., and Oney-Yazici, E. (2008). "Innovative Safety Management Training Through e-Learning." *Arch. Eng. and Design Manag.t*, 4(3-4), 239-250.

Albert, A., Hallowell, M. R., Kleiner, B., Chen, A., and Golparvar-Fard, M. (2014a). "Enhancing construction hazard recognition with high-fidelity augmented virtuality." J. *Constr. Eng. Manage.,* 140(7).

Albert, A., Hallowell, M. R., and Kleiner, B. M. (2014b). "Enhancing Construction Hazard Recognition and Communication with Energy-Based Cognitive Mnemonics And Safety Meeting Maturity Model: Multiple Baseline Study." *J. Constr. Eng. Manage.,* 140(2), 04013042.

Bahn, S., and Barratt-Pugh, L. (2012). "Emerging Issues of Health and Safety Training Delivery in Australia: Quality and Transferability." *Procedia - Social and Behavioral Sciences*, 62, 213-222.

Bahn, S., and Barratt-Pugh, L. (2012). "Evaluation of the mandatory construction induction training program in Western Australia: Unanticipated consequences." *Eval. Program Plan.*, 35(3), 337-343.

Bhandari, S., and Hallowell, M. R. (2017). "Emotional Engagement in Safety Training: Impact of Naturalistic Injury Simulations on the Emotional State of Construction Workers." *J. Constr. Eng. Manage.,* 143(12), 04017090.

Bhandari, S., Hallowell, M. R., Correll, J. (2019). "Making Construction Safety Training Interesting: A Field-Based Quasi-Experiment to Test the Relationship between Emotional Arousal and Situational Interest among Adult Learners." *Saf. Sci.,* 117, 58-70.

Bosché, F., Abdel-Wahab, M., and Carozza, L. (2015). "Towards a mixed reality system for construction trade training." *J. Comput. Civil. Eng.*, 30(2), 04015016.

Canales, A. R., Arbelaez, M., Vasquez, E., Aveiga, F., Strong, K., Walters, R., Jaselskis, E. J., and Jahren, C. T. (2009). "Exploring training needs and development of construction language courses for American supervisors and Hispanic craft workers." *J. Constr. Eng. Manage.,*, 135(5), 387-396.

Cunningham, T. R., Guerin, R. J., Keller, B. M., Flynn, M. A., Salgado, C., and Hudson, D. (2018). "Differences in safety training among smaller and larger construction firms with non-

native workers: Evidence of overlapping vulnerabilities." *Safety Science*, 103, 62-69.
Demirkesen, S., and Arditi, D. (2015). "Construction safety personnel's perceptions of safety training practices." *Intern. J. of Project Management*, 33(5), 1160-1169.
Goldenhar, L. M., Moran, S. K., and Colligan, M. (2001). "Health and safety training in a sample of open-shop construction companies." *J. of Safety Research*, 32(2), 237-252.
Guo, H., Li, H., Chan, G., and Skitmore, M. (2012). "Using game technologies to improve the safety of construction plant operations." *Acc. Anal. & Prevention*, 48, 204-213.
Hallowell, M. R. (2012). "Safety-Knowledge Management in American Construction Organizations." *Journal of Management in Engineering*, 28(2), 203-211.
Hallowell, M. R., and Gambatese, J. A. (2009). "Construction Safety Risk Mitigation." *J. Constr. Eng. Manage.,* 135(12), 1316-1323.
Haslam, R. A., Hide, S. A., Gibb, A. G. F., Gyi, D. E., Pavitt, T., Atkinson, S., and Duff, A. R. (2005). "Contributing factors in construction accidents." *Applied Ergonomics*, 36(4), 401-415.
Hinze, J. (2006). *Construction safety*, Prentice Hall.
Ho, C.-L., and Dzeng, R.-J. (2010). "Construction safety training via e-Learning: Learning effectiveness and user satisfaction." *Comp.s & Educ.*, 55(2), 858-867.
Hung, Y.-H., Winchester, W. W., Smith-Jackson, T. L., Kleiner, B. M., Babski-Reeves, K. L., and Mills, T. H. (2013). "Identifying fall-protection training needs for residential roofing subcontractors." *Applied Ergonomics*, 44(3), 372-380.
Jaselskis, E. J., Anderson, S. D., and Russell, J. S. (1996). "Strategies for achieving excellence in construction safety performance." *J. Constr. Eng. Manage.,* 122(1), 61-70.
Jazayeri, E., Huang, L., and Dadi, G. B. "Creating a Safety Training and Competence Model for Construction Craft Professionals." *Proc., Construction Research Congress 2018, 2-4 April 2018*, ASCE - ASCE, 241-250.
Jeelani, I., Albert, A., Azevedo, R., and Jaselskis, E. J. (2017). "Development and Testing of a Personalized Hazard-Recognition Training Intervention." *J. Constr. Eng. Manage.,* 143(5).
Jeelani, I., Albert, A., Han, K., and Azevedo, R. (2018a). "Are Visual Search Patterns Predictive of Hazard Recognition Performance? Empirical Investigation Using Eye-Tracking Technology." *J. Constr. Eng. Manage.*, 145(1).
Jeelani, I., Han, K., and Albert, A. (2017). "Development of Immersive Personalized Training Environment for Construction Workers." *ASCE International Workshop on Computing in Civil Engineering 2017*, 407-415.
Jeelani, I., Han, K., and Albert, A. (2018b). "Automating and scaling personalized safety training using eye-tracking data." *Autom. in Construction*, 93, 63-77.
Li, H., Chan, G., and Skitmore, M. (2012). "Visualizing safety assessment by integrating the use of game technology." *Autom. in Construction*, 22, 498-505.
Lin, K.-Y., Lee, W., Azari, R., and Migliaccio, G. C. (2018). "Training of Low-Literacy and Low-English-Proficiency Hispanic Workers on Construction Fall Fatality." *Journal of Management in Engineering*, 34(2).
Lingard, H., Pink, S., Harley, J., and Edirisinghe, R. (2015). "Looking and learning: using participatory video to improve health and safety in the construction industry." *Construction Management and Economics*, 33(9), 740-751.
McCabe, M. J., and Snyder, C. M. (2007). "Academic journal prices in a digital age: A two-sided market model." *The BE Journal of Economic Analysis & Policy*, 7(1).
Mo, Y., Zhao, D., Du, J., Liu, W., and Dhara, A. (2018). "Data-driven approach to scenario

determination for VR-based construction safety training." *Construction Research Congress 2018: Safety and Disaster Manag., CRC.*
Namian, M., Albert, A., Zuluaga, C. M., and Jaselskis, E. J. (2016). "Improving Hazard-Recognition Performance and Safety Training Outcomes: Integrating Strategies for Training Transfer." *J. Constr. Eng. Manage.,* 142(10).
Namian, M., Albert, A., Zuluaga, C. M., Behm, M. (2016). "Role of Safety Training: Impact on Hazard Recognition and Safety Risk Perception." *J. Constr. Eng. Manage.,* 142(12), 04016073.
Pereira, R. E., Gheisari, M., and Esmaeili, B. (2018). "Using panoramic augmented reality to develop a virtual safety training environment." *Construction Research Congress 2018: Safety and Disaster Management,* 2018-April, 29-39.
Rodriguez-Garzon, I., Lucas-Ruiz, V., Martinez-Fiestas, M., and Delgado-Padial, A. (2015). "Association between perceived risk and training in the construction industry." *J. Constr. Eng. Manage.,* 141(5).
Sacks, R., Perlman, A., and Barak, R. (2013). "Construction safety training using immersive virtual reality." *Const. Manag.t and Economics*, 31(9), 1005-1017.
Schoenfisch, A. L., Lipscomb, H., Sinyai, C., and Adams, D. (2017). "Effectiveness of OSHA Outreach Training on carpenters' work-related injury rates, Washington State 2000–2008." *Amer. j. of industrial medicine*, 60(1), 45-57.
Sinyai, C., Stafford, P., and Trahan, C. (2013). "Doing It Old School: Peer- Led Occupational Safety Training in The U.S. Construction Industry." *McGill J. Educ.*, 48(3), 605-611.
Wang, Y., Goodrum, P. M., Haas, C., Glover, R., and Vazari, S. (2010). "Analysis of the benefits and costs of construction craft training in the United States based on expert perceptions and industry data." *Construction Management and Economics*, 28(12), 1269-1285.
Wilkins, J. R. (2011). "Construction workers' perceptions of health and safety training programmes." *Construction Management and Economics*, 29(10), 1017-1026.
Williams Jr, Q., Ochsner, M., Marshall, E., Kimmel, L., and Martino, C. (2010). "The impact of a peer-led participatory health and safety training program for Latino day laborers in construction." *Journal of Safety Research*, 41(3), 253-261.
Zhao, D., and Ye, Y. "Using virtual environments simulation to improve construction safety: An application of 3D online-game based training." *Proc., 3rd International Symposium on Artificial Intelligence,* Springer Verlag, 269-27

# Construction Worker Posture Estimation Using OpenPose

Aanuoluwapo Ojelade[1] and Frederick Paige[2]

[1]Occupational Ergonomics and Safety, Dept. of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA. E-mail: aanu@vt.edu
[2]Construction and Engineering Management, Dept. of Civil Engineering, Virginia Tech, Blacksburg VA. E-mail: freddyp@vt.edu

## ABSTRACT

Visual field observations are critical in construction project management. Cost reduction and increases in data usability have made RGB cameras more prevalent on construction sites, expanding capabilities of the modern construction manager. The eyes of a construction manager are irreplaceable tools which monitor and subjectively evaluate construction activities. The variety of "right" and "wrong" ways to complete a construction task make the automation of visual field operations a difficult task. Existing vision-based methods are limited to RGB-depth cameras which are rarely used on construction sites. This paper presents a method for using deep learning techniques to aid the automatic recognition of construction worker activities from RGB camera footage and to what extent can computer vision accurately detect construction worker activity. OpenPose human estimation algorithm was used to create 2D human pose samples of construction activities using RGB cameras. The samples are used to train and test a time distributed feed forward neural network with LSTM Keras model classifier. The results of this work in progress set the stage for improving the usefulness of visual data for project management in a variety of construction scenarios and other labor-intensive sectors. Monitoring the safety and productivity of construction workers will become more efficient allowing much-needed increases in our capacity to develop and maintain infrastructure.

## INTRODUCTION

The advancement of information and technology (computer vision, sensing, good data storage, high internet speed, and high-resolution video cameras) have improved construction project documentations. This is evident in how the presence of surveillance cameras reduced human interference with construction activities by continuously monitoring construction activities. For example, videotaped data collected on the construction site has reduced the cost of manual data collection and subjective analysis of workers output by observers. However, manual analysis of videotaped data is expensive due to the cost and time taken to analyze videotaped data.

Construction tile and marble setters typically install ceramic, interior stone, mosaic tiles and marble tile (Bls.gov 2018). Due to the nature of work done by tile and marble setters, the number of observations needed to report tile and marble setters' activities can grow astronomically, hence increasing the project labor cost. For example, Shehata and El-Gohary (2011) stated that 336 observations are required to report six workers' productivity when tiling 870 square meters. Furthermore, the observers must record each active and inactive time of the tile and marble workers. This could result in over productivity (also known as Hawthorne Effect) (Jones 1992) of the tile and marble workers, hence the need to automate the activity analysis of tile and marble workers.

The authors hypothesis that estimating postures from videotaped data on construction site will set the stage for improving construction workface and project management in a variety of construction scenarios. Therefore, this preliminary work presents a method for using deep learning techniques to aid the automatic recognition of construction worker activities from RGB only camera footage.

## LITERATURE REVIEW

Recently, computer vision and machine learning has evolved in computer science. This emerging field has been applied to human activity recognition. Zhang and Tian (2012) used RGB-D camera to collect human activities under different lighting conditions. The authors used histogram pooling to encode the data and used the features to train linear one-vs-all Support Vector Machine (SVM). Sung et al. (2011) infer three human activities from five different locations. Bag-of-pose features were extracted from RGB-Depth camera. Two-layered Maximum Entropy Markov Model (MEMM) was compared against one-layer MEMM and Naïve Bayes machine learning algorithm. Kepski and Kwolek (2014) used k-Nearest Neighbor algorithm to differentiate between fall and lying down. The data used for the study was collected from a ceiling mounted 3D depth camera. The model gave an accuracy of 100% when tested on real-life datasets. Chen et al. (2014) extracted features from body inertial measurement unit and RGB-Depth camera. The features were fused together, and they later used Sparse Representation Classifier (SRC) to recognize eleven human actions. Dubois and Charpillet (2013) used RGB-Depth camera and Hidden Markov Model (HMM) algorithm to detect falls among ageing people. After processing the data, the authors were able to recognize falls with zero false prediction among seven other activities.

Recently, researchers have applied human activity recognition to the construction field. Rafibakhsh et al. (2012) tested the effectiveness of inexpensive XBOX Kinect camera against high end time-of-flight cameras (D-Imager EKL3104 and MESA SR-3000) for data collection on the construction job site. The result of the study shows that it is possible to use the sensor for spatial sensing on the construction site. Weerasinghe et al. (2012) proposed a method to detect and track construction workers using RBD-Depth images. Escorcia et al. (2012) performed a study to recognize drywall construction activity. The authors extracted color and depth from the construction workers activity using Microsoft Kinect sensor. The data was later encoded into a bag-of-pose histogram to represent the useful features. The actions were predicted using Multiclass SVM classifier. The 80% overall performance of the algorithm shows the promise of activity recognition using RGB-D camera in construction. Khosrowpour et al. (2014) used Microsoft Kinect sensor to recognize drywall activities on construction site. Multi-class one-vs-all SVM classifier was used to evaluate the bag-of-pose features and average accuracy of 76% was reported. Despite the promise that comes with the use of RGB-D camera to recognize construction activities on the jobsite, the time taken to install and calibrate RGB-D camera on jobsite is not trivial, RGB-D camera is not as widely and cheaply acceptable as RGB camera and ferromagnetic interference can cause significant noises in the output (Zhang et al. 2018). Furthermore, activity estimation requires at least four cameras to be installed on the jobsite. Installing at least four RGB-D cameras will cause interference between these cameras (Rafibakhsh et al. 2012).

Some researchers have used RGB cameras to recognize construction activities. Yang et al. (2016) encoded eleven RGB recorded construction actions as bag-of-features using k-means clustering algorithm. The authors used the bag-of-features to train and evaluate non-linear SVM. The overall performance of 59% shows that more work needs to be done to improve the performance of activity recognition using RGB cameras. Therefore, this study will focus on using Recurring Neural Network which is a deep neural network algorithm to estimate construction activities from OpenPose feature generated datasets.

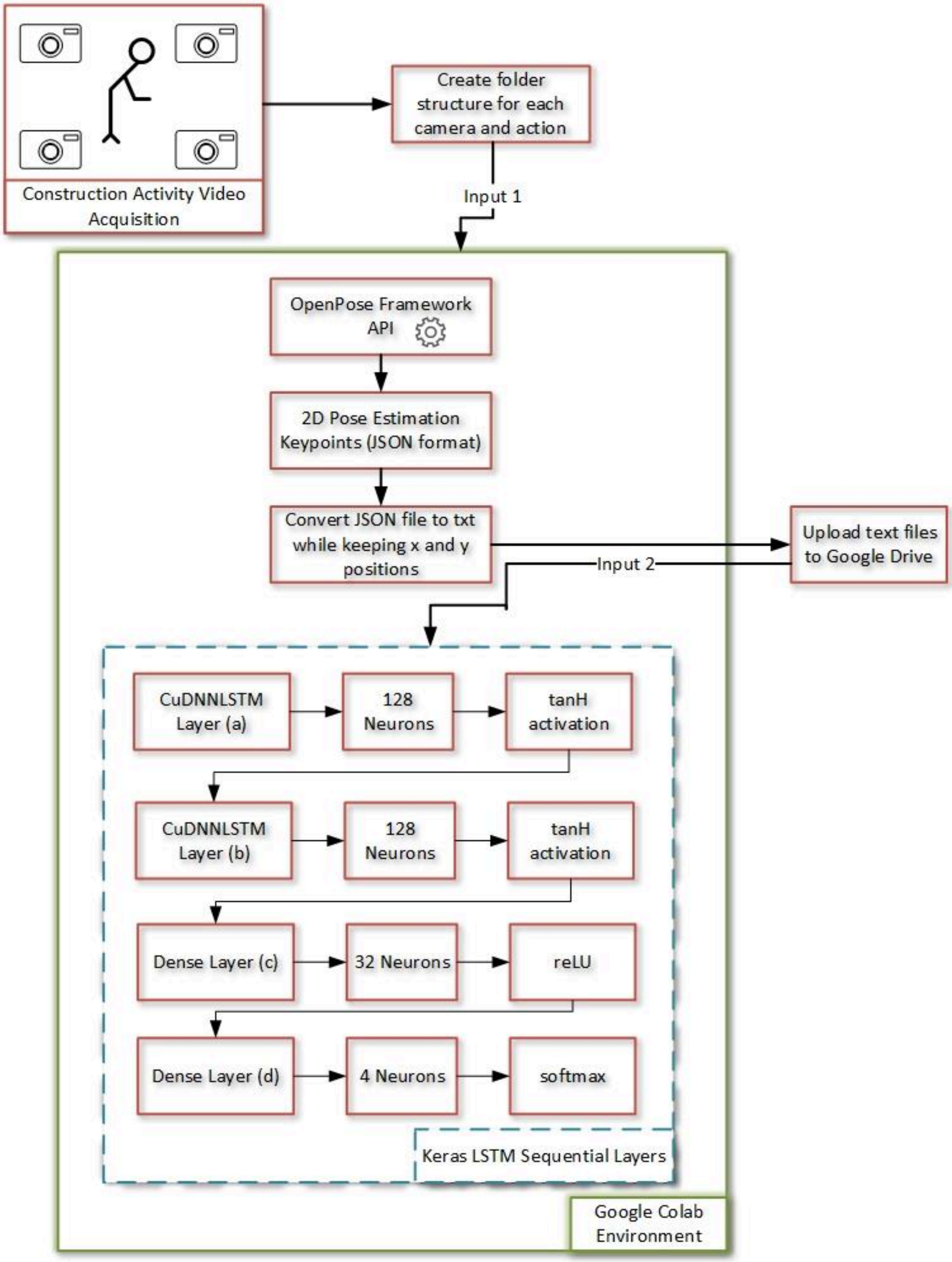


**Figure 1. Overall framework**

## METHODOLOGY

The proposed system architecture (shown in Fig 1) aims to estimate construction activity performed by tile and marble setters. The system can be broken down into three distinct sections which address three different aspects. The first aspect is responsible for collecting construction activity footage on an actual construction site. The recorded footage is converted into worker action datasets in the second stage. Last, the human 2D joint is extracted as JavaScript Object Notation (JSON) files using OpenPose and later converted to text files. The text file serves as input for the Long Short-Term (LSTM) classifier, which is used to estimate construction worker activities.

## Dataset Description

Datasets are critical to training and testing deep neural networks. This study focuses on analyzing construction tile and marble setter activity videos which were recorded with a fixed camera. Four cameras were used to record the actions of a tile and marble setter. The cameras were positioned in strategic location around the worker to reduce occlusion, capture different point of view and capture different scale. After several trial and error attempts, the cameras were fixed on a tripod and were positioned three meters away from the construction worker to avoid self-occlusion. Each camera has a different range of view and field of view, hence the trial and error approach to determine the safe distance to reduce self-occlusion. Four frequently observed flooring actions were selected to form the dataset. These actions are “fixing markers”, “Installing tiles”, “Stacking tiles”, and “Inspecting work. The construction activity recorded was neither scripted nor directed. The workers performed the activity as they would have normally performed the activity under normal conditions.

The resulting video was automatically segmented into four clips for each camera with each clip containing a single action type. The clips were labelled manually with their respective action name. The recorded videos were 51mins and 48 seconds long. The information of the data set is shown in Table 1. Note that VL and FR signifies video length and frame rate respectively.

**Table 1. Recorded Video Statistics**

| | Camera 1 | | Camera 2 | | Camera 3 | | Camera 4 | |
|---|---|---|---|---|---|---|---|---|
| **Action** | **VL** | **FR** | **VL** | **FR** | **VL** | **FR** | **VL** | **FR** |
| fixing markers | 5:01min | 24 | 4:38mins | 24 | 4:38 | 24 | 4:38 | 24 |
| Installing tiles | 2:15min | 24 | 2:20 | 24 | 2:20 | 24 | 2:20 | 24 |
| Stacking tiles | 2:31mins | 24 | 2:04 | 24 | 2:04 | 24 | 2:04 | 24 |
| Inspecting work | 4:01mins | 24 | 3:42mins | 24 | 3:42 | 24 | 3:42 | 24 |

## Camera Hardware

The only camera used in this study was GoPro Hero 7, and it was used to capture floorers activity. This camera consists of an RGB camera module, and it is capable of recording at either 4K, 2.7K, 1080p, or 720p resolution. Frame rate of 60, 30, or 24 fps can be achieved using any of these resolutions mentioned earlier. Since most surveillance cameras use 1080P resolution and 24fps mode for recording, this study recorded floorers activity at 1080P resolution and 24 fps.

## Feature extraction

The first goal of this system was to locate and track workers in the construction site. Then we aimed to jointly detect human body, foot, and hand. OpenPose library was proposed by Cao et al. (2017). OpenPose library was the first real-time system to jointly detect 135 key points on a single RGB image (human body, foot, hand, and facial key points), hence its adoption for this study. OpenPose was developed using OpenCV and Caffe and it can be used on various platforms such as Linux, Windows, and Mac. OpenPose also provides support for CUDA GPUs, OpenCL GPUs, and CPU-only devices. OpenPose takes an RGB image of size $wxh$ as input and outputs 2D locations of anatomical key points for each body detected in the input image. As shown in Fig 2, OpenPose uses a two-branch multi-stage Convolutional Neural Network (CNN) to predict human body key points. First, it uses a feedforward network to simultaneously predicts a set of 2D confidence maps of body part locations and a set of 2D vector fields of part affinities which encode how different body parts are associated. Finally, the confidence maps and the affinity fields are parsed by greedy inference to output the 2D key points for all the people in the image (Cao et al. 2018). OpenPose outputs 2D skeletal data in JSON format. The JSON file consists of body part location coordinates key points. The extracted coordinates key points in the JSON file represent an observation at certain point in time, hence the dataset can be regarded as a sequential data.

**Recurrent Neural Network**

Recurrent Neural Networks (RNN) is a special type of neural network containing deep learning models, simple structures, and its output layer is added back to the next input. The updated layer is fed back to the same layer, i.e. it is good at remembering the analysis that was performed up to a given point by maintaining a state or context (Rumelhart et al. 1986). RNN was introduced to solve the limitation of feedforward neural network to maintain context for sequential data. Although RNN performs well around sequential data, it is known to have the following issues (1) it is computationally expensive to maintain the large number of units, (2) it is sensitive to changes in parameters, (3) increases exponentially or drops down to near zero and stabilize. To solve this problem, Hochreiter and Schmidhuber (1997) developed Long Short-Term Memory (LSTM). LSTM ensures that information is kept over a long period of time and additionally, it solves the oversensitivity to parameter changes. Therefore, the authors adopt RNN – LSTM for this study.

**Implementation**

OpenPose framework and Keras LSTM algorithm were implemented through Google Colaboratory or Colab. Google Colab is a free Python 3 Jupyter notebook environment which offers free Tensorflow with 12 GB GPUs. Custom Python 3 codes will be executed in the Google Colab environment.

The JSON data points returned by OpenPose framework are converted to text files using a custom code written in Python 3. The data in the text file were prepared for Keras LSTM model by splitting the data into 70% training data set and 30% testing data set. The data points labels were coded as numeric values to remove bias from the dataset. Each data point is 1 by 36 array.

The Keras LSTM RNN consist of a sequential model. The model consists of 128 neurons CuDNNLSTM layer (CuDNNLSTM is the GPU optimized version of the CPU LSTM layer), 32 neurons of Dense layer and an additional 4 neurons of Dense layer which signifies the dataset classes. The CuDNNLSTM layer utilizes tanh activation method, the middle dense layer utilizes rectified linear units (Relu) while the last Dense layer utilizes softmax activation method. The model was compiled with “adam optimizer” and categorical crossentropy loss function. The training and test data were fitted together alongside with their labels. The overall accuracy and confusion matrix were printed out and presented in the discussion section.