

(Siebert and Teizer, 2014) and measurement (Wang, X et al., 2017), and 3D model creation (Xie et al., 2012).

However, despite overall growing popularity within the domain of Civil Engineering, the utilization of the UAS and photogrammetry technologies in construction is still at an early stage. Little research has been conducted from a pragmatic perspective to evaluate the effectiveness of this emerging technology due to practical issues such as local regulations, limited resource of test fields, or strict flight conditions. Therefore, this work aims to conduct a quantitative analysis to evaluate the influence of important UAS flight parameters and site conditions on measurement accuracy. According to practical experience and existing literature, flight altitude, image overlapping rate, GCPs, and soil type are key factors during the operation of UAS and modeling quality control (Sibert and Teizer, 2014; Mesas-Carrascosa et al., 2016; Nassar and Jung, 2012). The goal of this project is to compare the positional accuracy of points when applying different factor parameters in order to identify the effectiveness of each factor and interactions among them, thereby providing a practical reference for managers and engineers to allow for efficient application of UAS and photogrammetry in construction projects.

INFLUENTIAL FACTORS

One of the most important flight parameters during UAS operations is the flight altitude. It not only determines the relative size of the pixels of an image, but also the flight durations and the area to be covered (Christiansen et al., 2017). To be more specific, the flight altitude is related to the Ground Sampling Distance (GSD). GSD is the distance between two consecutive pixel centers measured on the ground. The larger the value of GSD, the lower the spatial resolution of the image and the lower visibility of the details. In selecting flight altitude, it is essential to consider the balance between the spatial resolution and area covered. Higher spatial resolution will contribute to image quality but may result in prolonged flight duration.

Another crucial factor is the overlapping rate between images. Photogrammetry is a technology of image processing to interpret the shape and location of an object from one or more photographs of that object. It aims to reconstruct an object from two-dimensional (2D) graphic form to three-dimensional (3D) form. The shape and position of an object are determined by reconstructing bundles of light rays which define the spatial direction of the ray to the corresponding object point. From the intersection of at least two corresponding and separated rays, an object point can be located in 3D space. Therefore, image processing is based on automatically finding thousands of common points between images. Each characteristic point in an image is called a key-point. When two key-points, from two different images captured at different locations, are found to be the same, they will match together. When there is high overlap between images, the camera on the UAS is able to capture a larger common area to generate more matched key-points and thus improve the computational accuracy.

For the use in surveying application, an absolute accuracy test is mandatory. The quality of the 3D model depends on the number of images and manual tie points. The use of Ground Control Points (GCPs) is an effective method to improve accuracy. GCPs are points with known coordinates measured by highly accurate GPS units in the area of interest. The photogrammetry software is able to process projects with or without geo-locations, but accurate GCPs improve the global accuracy of the project. GCPs will give the scales, orientations, and positions to the final results (Wang, J. et al., 2012). Therefore, the number of GCPs and their distribution are important to control the modeling quality and accuracy of measurements.

Lastly, the material of the mapping surface also has great impact on the quality of models

during the image processing. A 3D image is a non-contact measurement method applied to produce a 3D representation of a physical object (Furukawa and Ponce 2010). The point cloud model is the major output of image processing through photogrammetry. A point cloud is composed of a set of vertices used to represent the external surface of objects in a 3D coordinate system. The photogrammetry software generates a point cloud model through measuring a large number of points on the surface of an object (Nassar and Jung 2012). Therefore, the different surface material of an object may affect the modelling quality at various levels.

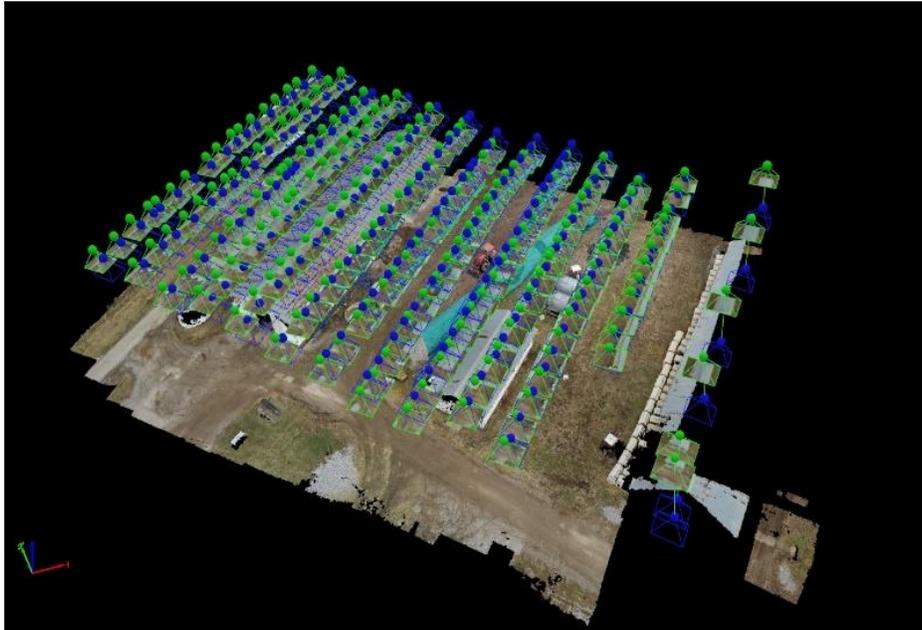


Figure 1. Grid Flight Pattern over the Study Area

METHODOLOGY

The UAS used for image acquisition in this study was the DJI Inspire 1. This UAS is a vertical takeoff and landing aircraft powered by a 22.2V battery (See Figure 1). Its system has a maximum takeoff weight of 7.71lbs and maximum wind resistance of up to 10m/s. The maximum flight duration is approximately 18 minutes. The UAS is equipped with a 1/2.3 inch CMOS sensor with a 20mm lens, and the stock camera has 4096×2160 resolution for still images (DJI 2018). During operation, the UAS autopilot sends a signal to the equipped sensor to capture a photo while, simultaneously, recording the geo-referencing information, such as location and navigation angles, which can be used for post-processing on an SD card. The study area was 163×247 ft in size and located at Coldstream Dairy Research Farm Complex in Lexington, Kentucky.

Multiple flights were conducted following the scheme presented in Figure 1. This flight plan is compatible with different flight altitudes, image overlapping rates, and the use of GCPs. A set of flight missions were performed at altitudes of 60, 90, 120 and 150 ft. Due to the height of wire poles on the farm, it was dangerous to fly the UAS lower than 60ft. For each altitude level, the UAS captured photos based on two different forward and side overlapping rates respectively: 40%-70% and 60%-90%. All the flight missions followed the grid pattern because this study aimed to perform mapping over an area with large size rather than modeling a vertical object (See Figure 1). In addition, flight missions were performed under the same weather conditions,

especially wind speed. In this study, Pix4Dmapper photogrammetry software was selected to process images and generate 3D point cloud and DSM models of the study area. Afterward, images captured by each flight were processed with and without GCPs. The coordinates of GCPs were measured by an EPOCH 50 GNSS Rover. A total of 16 GCPs were measured and spaced evenly across the area of interest to minimize the errors in scale and orientation.

Table 1. RMSE (ft) of Flights Processed by Different Number of GCPs

Flight Altitude (ft)	Image Overlapping Rate (%)	No GCPs	1 GCPs	4 GCPs	8 GCPs	12 GCPs	16 GCPs
60	70%-40%	4.09	4.08	0.92	0.78	0.55	0.52
	90%-60%	5.35	4.25	0.29	0.28	0.28	0.28
90	70%-40%	4.99	4.55	0.81	0.70	0.64	0.50
	90%-60%	3.31	3.31	0.32	0.29	0.28	0.28
120	70%-40%	3.12	3.19	0.60	0.60	0.50	0.43
	90%-60%	6.85	5.74	0.50	0.45	0.31	0.29
150	70%-40%	3.49	3.65	0.64	0.57	0.38	0.38
	90%-60%	2.91	2.89	0.51	0.48	0.30	0.31

RESULTS AND DISCUSSIONS

The major output of image processing was a point cloud model. The accuracy of the position of each point directly contributed to the linear or volumetric measurements. To be more specific, the positional absolute accuracy was the indicator or measure of how a spatial object was accurately positioned on the map with respect to its true position on the ground, within an absolute reference frame – such as UTM coordinate system (Küng et al., 2011). The 16 GCPs perform as checkpoints to be used for measurement of positional accuracy, no matter how many GCPs are used for processing. In this study, the positional accuracy of points was evaluated by Root Mean Square Error (RMSE). (Luhmann, Thomas, et al. 2014, Siebert and Teizer, 2014).

The CPU and memory specifications of the desktop used for analysis are as follows, Intel(R) Core(TM) i7-4790 CPU @ 3.60GHz, with 32GB of RAM. The operating system was Windows 7 Professional, 64-bit, and the photogrammetry platform was Pix4Dmapper Pro.

Table 1 shows the RMSE of each point with varying numbers of GCPs applied at multiple flight altitudes and image overlapping rates. It can be observed that the errors significantly decrease when all GCPs were used for processing because the GCPs provide an accurate orientation of the coordinates reference system. Also, the results show random RMSE behavior when no GCPs were used due to the lack of geometric constraints on the aerial-triangulation computation. This behavior seems to be independent of flight heights and image overlapping rates. With the image overlapping settings, the results indicate higher overlapping rates result in smaller errors when applying all GCPs regardless of flight altitudes. However, as the flight altitude increases, the errors decrease in magnitude if there is a lower overlapping rate.

According to observations of the collected data, lower flight altitude with higher image overlapping rate and the use of GCPs results in better positional accuracy. A multiple regression analysis was used to verify the results based on the observations. According to the results (See Figure 2), the use of GCPs is a statistically significant predictor because its p-value of t-test is smaller than 0.05.

Parameter Estimates				
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	1.3955889	0.464759	3.00	0.0046*
Flight Altitude	-0.003622	0.003968	-0.91	0.3668
Overlapping Rate[90%-60%-70%-40%]	-0.349009	0.188214	-1.85	0.0711
Number of GCPs	-0.136464	0.023004	-5.93	<.0001*
(Flight Altitude-105)*Overlapping Rate[90%-60%-70%-40%]	0.0054048	0.005611	0.96	0.3412
(Flight Altitude-105)*(Number of GCPs-6.83333)	-6.181e-5	0.000686	-0.09	0.9286
Overlapping Rate[90%-60%-70%-40%]*(Number of GCPs-6.83333)	-0.034275	0.032532	-1.05	0.2984
(Flight Altitude-105)*Overlapping Rate[90%-60%-70%-40%]*(Number of GCPs-6.83333)	0.0002693	0.000097	0.28	0.7827

Figure 2. Estimations of the Independent Variables Significance

In this study, the effect of soil type on the volumetric measurement accuracy was tested by modeling four samples composed of different soil types. The four soil types were sand, clay, fine grade gravel, and coarse grade gravel. The actual volumes of samples were based on the standards measured by the manufacture. All samples were piled in similar shapes under the same weather and illumination conditions (See Figure 3).



Figure 3. Sample Piles of Different Soil Types

As seen in Table 2, the results indicated that the measured volume of clay had the smallest error. In addition, as the soil granularity increased and as the color of material became lighter, the accuracy of measurement decreased. The reason may be that coarser surface textures created more visual noise on the surfaces of the models and light-colored and glossy surfaces tended to saturate images leading to difficulties in visual interpretation.

CONCLUSION

This study aimed to investigate how important flight parameters of the UAS and environmental factors impacted measurement accuracy through experimental flights and statistical analysis of positional errors computed through photogrammetry technologies.

After detailed comparisons and analysis for each flight plan, one can derive that the combination of low flight altitudes, high image overlapping rate, the use of a proper number of

GCPs and modeling surface of clay soil type can maximize the measurement accuracy. The positional errors become much smaller when more than 1 GCP are used for processing because GCPs provide an accurate orientation of the coordinate reference system. This behavior is independent of flight heights and image overlapping rates. With the image overlapping and flight altitude settings, however, higher overlapping rates result in larger errors as the flight altitude increases, and the errors decrease if selecting the low overlapping rate. This tendency did not change when a different number of GCPs were applied. Although GCPs were the most influential factor, based on the results of multiple regression analysis, it does not mean an unlimited number of GCPs would be an optimal strategy to guarantee accurate measurements. In the experiment, there were no significant differences in the errors when comparing the results from using 4 GCPs as opposed to 16 GCPs. The selections of parameter values largely depend on the level of accuracy required by users.

Table 2. Impact of Soil Types on the Accuracy of Volumetric Measurements

Soil Type	Number of Calibrated Photos	Actual Volume (ft ³)	Computed Volume (ft ³)	% Error
Clay	11	1.5	1.48	1.33
Sand	11	0.5	0.53	6.00
Gravel	10	0.5	0.47	6.00
Rock	10	0.5	0.45	10.00

The limitations of this study mainly came from the selection of the UAS equipment and photogrammetry software. The UAS, an especially low-cost device, limited the sensor payload in weight and dimension, so low-weight sensors with small-format amateur cameras had to be used. When compared to more expensive UAS with large format cameras, the UAVs acquired a higher number of images in order to obtain the same image coverage and comparable image resolution. Moreover, the low-cost sensors were less stable, which resulted in a lower image quality. When processing the images collected by the UAV, this study did not research the differences caused by varying devices. In the future, more research could be conducted regarding how different devices and other potential environmental factors impact the measurement accuracy when the limitations of UAS technology can be solved – such as inaccurate geo-referencing capability and limited battery capacity.

REFERENCES

- Bendea, H., Chiabrando, F., Tonolo, F. G., & Marenchino, D. (2007, October). Mapping of archaeological areas using a low-cost UAV. The Augusta Bagiennorum test site. In XXI International CIPA Symposium (Vol. 1).
- Chou, T. Y., Yeh, M. L., Chen, Y. C., & Chen, Y. H. (2010). Disaster monitoring and management by the unmanned aerial vehicle technology.
- Christiansen, M. P., Laursen, M. S., Jørgensen, R. N., Skovsen, S., & Gislum, R. (2017). Designing and Testing a UAV Mapping System for Agricultural Field Surveying. *Sensors*, 17(12), 2703.
- d'Oleire-Oltmanns, S., Marzolf, I., Peter, K. D., & Ries, J. B. (2012). Unmanned aerial vehicle

- (UAV) for monitoring soil erosion in Morocco. *Remote Sensing*, 4(11), 3390-3416.
- Ezequiel, C. A. F., Cua, M., Libatique, N. C., Tangonan, G. L., Alampay, R., Labuguen, R. T., ... & Loreto, A. B. (2014, May). UAV aerial imaging applications for post-disaster assessment, environmental management and infrastructure development. In *Unmanned Aircraft Systems (ICUAS), 2014 International Conference on* (pp. 274-283). IEEE.
- Xie, F., Liu, Z., Gui, D., & Liu, H. (2012). Study on construction of 3D building based on UAV images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 469-473.
- Furukawa, Y., & Ponce, J. (2010). Accurate, dense, and robust multiview stereopsis. *IEEE transactions on pattern analysis and machine intelligence*, 32(8), 1362-1376.
- Gómez-Candón, D., De Castro, A. I., & López-Granados, F. (2014). Assessing the accuracy of mosaics from unmanned aerial vehicle (UAV) imagery for precision agriculture purposes in wheat. *Precision Agriculture*, 15(1), 44-56.
- Grenzdörffer, G. J., Engel, A., & Teichert, B. (2008). The photogrammetric potential of low-cost UAVs in forestry and agriculture. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 31(B3), 1207-1214.
- Hugenholtz, C. H., Whitehead, K., Brown, O. W., Barchyn, T. E., Moorman, B. J., LeClair, A., & Hamilton, T. (2013). Geomorphological mapping with a small unmanned aircraft system (sUAS): Feature detection and accuracy assessment of a photogrammetrically-derived digital terrain model. *Geomorphology*, 194, 16-24.
- Luhmann, T., Robson, S., Kyle, S., & Boehm, J. (2013). Close-range photogrammetry and 3D imaging. Walter de Gruyter.
- Metni, N., & Hamel, T. (2007). A UAV for bridge inspection: Visual servoing control law with orientation limits. *Automation in construction*, 17(1), 3-10.
- Mesas-Carrascosa, F. J., Notario García, M. D., Meroño de Larriva, J. E., & García-Ferrer, A. (2016). An analysis of the influence of flight parameters in the generation of unmanned aerial vehicle (UAV) orthomosaics to survey archaeological areas. *Sensors*, 16(11), 1838.
- Nassar, K., & Jung, Y. H. (2012). Structure-From-Motion Approach to the Reconstruction of Surfaces for Earthwork Planning. *Journal of Construction Engineering and Project Management*, 2(3), 1-7.
- Pix4D (2016). <https://support.pix4d.com/hc/en-us/articles/202557459#label>
- Siebert, S., & Teizer, J. (2014). Mobile 3D mapping for surveying earthwork projects using an Unmanned Aerial Vehicle (UAV) system. *Automation in Construction*, 41, 1-14.
- Wang, J., Ge, Y., Heuvelink, G. B., Zhou, C., & Brus, D. (2012). Effect of the sampling design of ground control points on the geometric correction of remotely sensed imagery. *International Journal of Applied Earth Observation and Geoinformation*, 18, 91-100.
- Wang, X., Al-Shabbani, Z., Sturgill, R., Kirk, A., & Dadi, G. B. (2017). Estimating Earthwork Volumes Through Use of Unmanned Aerial Systems. *Transportation Research Record: Journal of the Transportation Research Board*, (2630), 1-8.

Perceptions for Crane Operations

Bo Xiao¹; Keith Yin Kong Lam²; Jieyu Cui³; and Shih-Chung Kang, Ph.D., P.Eng.⁴

¹Graduate Student, Dept. of Civil and Environmental Engineering, Univ. of Alberta, 9211-116 St. NW, Edmonton, AB T6G 1H9, Canada. E-mail: bxiao2@ualberta.ca

²Undergraduate, Dept. of Civil and Environmental Engineering, Univ. of Alberta, 9211-116 St. NW, Edmonton, AB T6G 1H9, Canada. E-mail: kylam@ualberta.ca

³Undergraduate, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong Univ., 800 Dongchuan Rd., Shanghai, China 200240. E-mail: cuijieyurachel@gmail.com

⁴Professor, Dept. of Civil and Environmental Engineering, Univ. of Alberta, 9211-116 St. NW, Edmonton, AB T6G 1H9, Canada. E-mail: sckang@ualberta.ca

ABSTRACT

Sensors and computation increase the precision, efficiency, and agility of crane operations. This paper presents an ongoing work of developing computational methods to enhance the crane operations. This research focuses on three significant challenges for crane operators: (1) identifying construction equipment and their activities, (2) identifying and tracking personnel, and (3) tracking the rigging object. In the preliminary stage, we focus on the first challenge to identify construction equipment and activities. 5,000 images have been collected and manually labeled for training deep learning detection algorithms. In the future steps, we will employ inception-SSD method to locate personnel, trucks, and excavators. After that, we will propose a method that recognizes excavator activities from crane views. Once learning algorithm is the reliable, it will benefit the crane operators to operate the cranes with confidence. It can also reduce the difficulty of crane operation. Training time and safety concerns will be reduced simultaneously.

INTRODUCTION

Sensors and computation increase the precision, efficiency, and agility of crane operations. The computer vision technology has gained great success in construction automation field. This paper presents an ongoing work of developing computational methods to enhance the crane operations. Crane erections are often in the critical paths for construction projects, and efficiency of crane operations directly influences the overall project performance (Neitzel et al., 2001). Meanwhile, crane cableway caused by inertial forces and winds makes it difficult for operators to control safely. Additionally, a dynamic site environment with moving people and construction equipment add challenges for safe lifting. The final objective of this research is overcoming three major challenges for crane operators: (1) identifying construction machines and activities, (2) identifying and tracking personnel and (3) tracking rigging objects.

Crane perception based on the smart sensors and controllers make crane operations more efficient and safer. As illustrated in Figure 1, labels S and C represent smart sensors and controllers respectively. The main idea of crane perception is widely deployed sensors acting like cameras in the environment, which allow the remote users to "know" the working progress. Information collected from the sites will be sent for processing uses artificial intelligence for crane control. The visual perception automatically identifies high risk and high-value works. This allows crane operators and remote users to sense the environment, prepare for the next tasks, and most importantly, prevent potential risks.

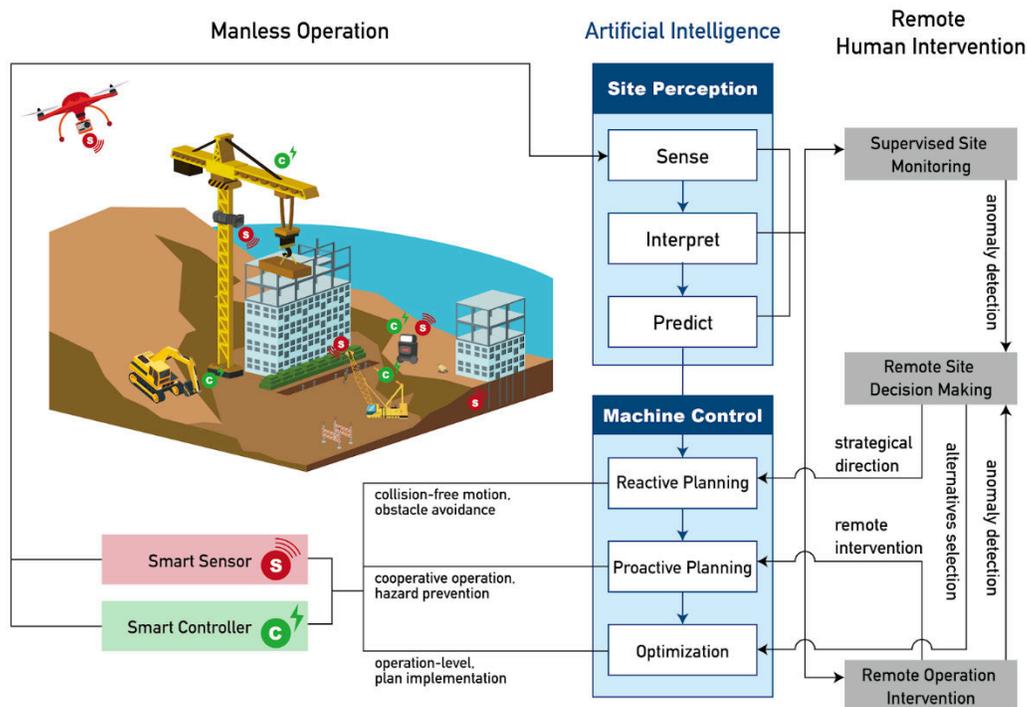


Figure 1. Vision of Crane Perceptions

In the preliminary stage, we have focused on identifying construction equipment and their activities. These construction videos can be used to identify construction equipment which prevent potential collisions during crane operations. Recognizing equipment activities will help remote users to estimate the valuable productivities. Until now, we have manually labeled 5000 images of equipment and tested them with two deep learning algorithms, YOLOV3 and Inception-SSD. The detection results indicate that the Inception-SSD performs better than YOLOV3 in our dataset. After that, we have proposed a method that puts the detected objects to 3D CNN classifiers which recognizes excavator activities from crane view. The proposed method can be extended to other construction equipment such as lifters, bulldozers, and backhoes. In the future, we will focus on developing reliable tracking system for personnel and rigging objects tracking.

LITERATURE REVIEW

The early work of object detection is the cascade detector (Viola and Jones 2001), which consists of multiple stages. Each stage is an ensemble of simple classifiers. The difficulty of detection comes from the huge differences within the same category. To fill this gap, various kinds of deformable template methods (Coughlan et al., 2000; Cootes et al., 2001), and part-based methods (Crandall et al., 2005; Amit and Trouve, 2007) have been conducted in computer vision community. In recent, the convolutional neural networks (CNNs) have been demonstrated in object detection and achieved reliable performance. The CNNs represents images through designed structure of many layers for feature extraction and transformation, which makes the detector understands the images from a higher level (Krizhevsky et al., 2012; Vedaldi and Zisserman, 2015). Girshick (2015) proposed the Fast r-cnn detection model and Ren et al. (2015) proposed the Faster r-cnn model. Redmon et al. (2016) introduced the YOLO darknet into

detection, which can reach real-time performance. Liu et al. (2016) proposed the SSD model to exploit the information of the tiny image area.

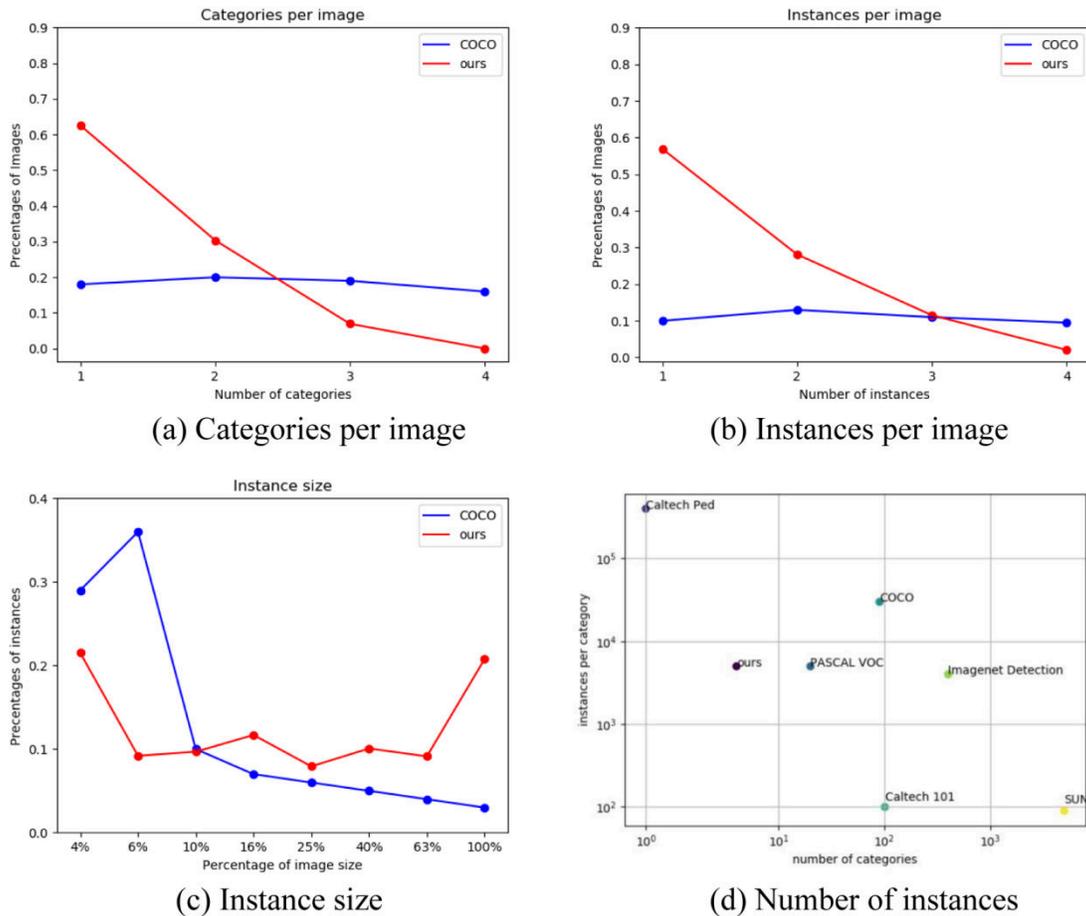


Figure 2. The analysis of construction equipment detection dataset

Crane operators work in difficult environments. It often requires them to complete erection activities without a clear view while being alert to all possible risks. Many researchers have developed methods to use cameras to enhance the perceptions while operating cranes. Gong and Caldas (2009) installed multiple cameras on the crane boom to identify construction activities from the video streams. Weerasinghe and Ruwanpura (2009) tracked construction resources to reduce waste. Rezazadeh and Brenda (2012) have developed automated methods to detect and track trucks to monitor productivity in real time. Han and Lee (2013) developed a method to protect workers from potential collisions by using cameras. Yang et al. (2014) employed Gaussian background subtraction (Wren et al., 1997) to detect crane jibs to analyze crane activities from video streams. Kim and Chi (2017) have developed tracking methods to locate construction equipment. Xiao and Zhu (2018) tested 15 tracking algorithms in construction videos and identified stable trackers in various backgrounds.

CONSTRUCTION EQUIPMENT DETECTION

Object detection is the primary section in this research. We need to identify equipment, personnel and rigging objects by detectors. It is important to choose a reliable detector to conduct

this research. The deep learning detection methods have shown high performance in many applications and we decided to adopt this technology in our research to identify construction objects. In order to evaluate detection algorithms, we have collected and manually labeled 5000 images of equipment and tested them with deep learning detection algorithms YOLOV3 and Inception-SSD.

There are four types of construction equipment labeled in the current dataset, which are truck, excavator, loader, and backhoe. We have analyzed our dataset from different perspectives in Figure 2 and compared it with COCO dataset (Lin et al., 2014), which is a well-known detection dataset in computer vision. In Figure 2(a), it shows that most images in our dataset contain one or two categories, while COCO has a uniform distribution. In Figure 2(b), 58% of images contain only one instance and 30% of the images contain two instances. We can find that our dataset is a specific dataset for construction equipment, while COCO is a general dataset for daily life objects. In Figure 2(c), the instance size of our dataset is larger than COCO. Based on these differences, the detection algorithms perform well in computer vision need to be re-evaluated with the construction equipment dataset. It needs to point out that this research is an on-going research. Comparing with other mature datasets (Figure 2(d)), we have limited number of categories and instances. And we will put more efforts into this construction detection dataset.

YOLOV3 and Inception-SSD have been selected to test on our dataset because of their promising performance on COCO. This dataset has been separated to trainset (90%) and valid-set (10%). In this research, we have used the Mean Average Precision (mAP, LeCun et al., 2015) to evaluate the detection results. mAP is the evaluation criteria decided by Precision and Recall. Precision measures how accurate the algorithm is, but it cannot reflect the performance of finding all positive instances. mAP is able to show the detector performance in both accuracy and robustness. Higher value of mAP means better detection performance. The testing results can be found in Table 1. It shows that both detectors have higher mAP on construction dataset, which means detecting construction categories is simpler than detecting general categories. This result indicates that detecting construction categories by using vision sensors installed on the crane is a reliable option. Inception-SSD methods perform better than YOLOV3 from an overall view and we will employ Inception-SSD for the crane perception.

Table 1. The testing performance of detectors

	AP@0.5 Truck	AP@0.5 Excavator	AP@0.5 Loader	AP@0.5 Backhoe	mAP@0.5 Overall
YOLOV3	0.71	0.93	0.91	0.93	0.87
Inception-SSD	0.80	0.93	0.94	0.95	0.91

EXCAVATOR ACTIVITY RECOGNITION

We have proposed a method to recognize construction activities by 2D-CNN detector and 3D-CNN classifier. 3D-CNN means putting continuous images instead of a single image into the CNN model as the inputs. The 3D-CNN classifier U3D (Tran et al., 2015) has been used in this research. Because excavation is the most common activity in construction sites, we will take excavator as an example to illustrate our methods. The overview of proposed methodology has been shown in Figure 3.

In order to identify excavator activities, multiple continuous images will be the inputs instead of a single image. This kind of input is well-suited for exploiting spatiotemporal features. For each new frame, we will compare it with previous t frames. Inception-SSD will detect excavators