Towards Autonomous Mining via Intelligent Excavators

Hooman Shariati, Anuar Yeraliyev, Burhan Terai, Shahram Tafazoli, Mahdi Ramezani
Motion Metrics International Corp.
Vancouver, BC, Canada
{hooman, Anuar, burhan, shahram, mahdi} @motionmetrics.com

Abstract

In this paper we present our first step solution towards global challenge of safety, productivity, profitability and energy-efficiency in mining. Our solution (intelligent excavator) provides complete monitoring solution for excavators that relies on deep neural networks to produce accurate, actionable data for mine. Our solution helps mines to increase shovel efficiency, reduce unexpected downtime cost, enable planned maintenance. We use a multi-frame convolutional LSTM-based object detection approach to accumulate valuable information across video frames without significant computational overhead. Our experiments on dataset captured in several mines across the world show that we can detect objects of interest with accuracy of more than 90% on 10 FPS. Furthermore, we show that our approach generalizes well to mining sites and equipment types not encountered in our training set. Finally, our work on detecting the types of objects encountered in a mining equipment could be used as a first step in developing a perception module that could provide autonomous excavators with the required knowledge of their environment in order to make optimal decisions.

1. Introduction

The Mine Safety and Health Administration (MSHA) notes that in 2017, nearly 40 percent of the mining fatalities and more than 30 percent of injuries involved mobile equipment [1]. Beyond safety, the need to improve efficiency remains a vital mining industry need. Surprisingly, worldwide mining operations are as much as 28 percent less productive today than a decade ago, according to McKinsey research. The results from McKinsey's new MineLens Productivity Index (MPI) [2], which adjusts for declining ore grades and mine cost inflation, show that the pronounced decline in productivity is evident across different commodities and is seen in most mining companies and geographies. The decline in productivity happens despite significant advances in technology over the last decades, which is a motivation for larger technological innovation in this area. Another primary technology focus is in the area of fuel/power consumption and emissions. The global mining industry consumes nearly 11% of the world's power. Improving energy efficiency and reducing greenhouse gas (GHG) emissions are priorities for the mining industry [3]. All these factors should be considered as well as never-ending increasing global demand for metals and minerals.

Our long-term vision to address concerns about safety, productivity, profitability and energy-efficiency in mining is moving towards automation. Autonomous excavators are expected to play a pivotal role to reach this goal. Shovels and excavators are typically the first kind of equipment to be used in the material handling process and, as a result, any changes in productivity in this first stage can have a domino effect on downstream processes and efficiency. The promise of autonomous shovels is to provide a degree of efficiency so that mines can better handle the productivity demand placed upon them. Autonomous shovels working in conjunction with autonomous trucks, can potentially form a vastly improved milieu.

In this paper, we present intelligent shovels; our first step solution towards autonomous shovels which provides complete monitoring solution for shovels and excavators that relies on deep neural networks to produce accurate, and actionable data. Heavy equipment such as mining shovels and excavators are used to load material ore from a mine into a haul truck or onto a conveyor for transportation to a processing location. Loading operations generally involve at least some element of danger as the payload being transferred may be heavy and could cause severe injury to operators involved in the loading operation. Moreover, an operator has a very limited field of view and is not able to see the contents of the bucket from the cabin. Accordingly, there exists a need to provide efficient monitoring of loading operations by involved operators to ensure that safe loading practices are followed, and that any loading equipment malfunction or damage is quickly identified [4]. Additionally, the payload may include undesired materials such as a detached tooth or large rocks (boulder) that should not be loaded. Such
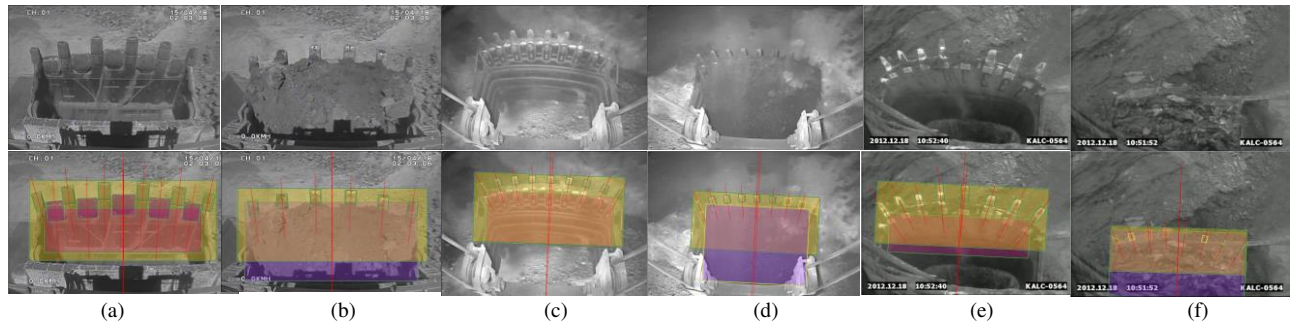
Figure 1. Samples of bucket images and labels for different shovel type (Hydraulic, Bucyrus, and P&H) typically encountered in a mining operation. The first row shows the actual frames, while the second row shows the rectangular bounding boxes labeled by humans to denote different objects. Yellow rectangles show the bucket, purple rectangles show the material inside of the bucket, and orange rectangles show empty buckets. Empty Hydraulic Bucket (a), Full Hydraulic Bucket (b), Empty Bucyrus Bucket (c), Full Bucyrus Bucket (d), Empty P&H Bucket (d), Full P&H Bucket (e).

undesirable materials in the payload may potentially cause equipment damage during loading or during later processing of the payload. There remains a need for systems and methods and for monitoring loading operations to ensure safe and/or efficient operation of the involved equipment [5]. Intelligent shovels solution features a rugged camera that overlooks the shovel bucket, and an in-cab monitor that provides real-time status updates as well as a remote monitoring portal. A set of human-labeled video frames are used as training inputs to train a deep neural network architecture to perform multiple object detection. The neural network when combined with post processing, provides a comprehensive monitoring solution capable of missing tooth detection [6], tooth wear monitoring [7], fragmentation analysis [8] and foreign object detection which are elements of intelligent shovel solution.

2. Materials

Figure 1 shows some of the images captured by our cameras installed on three main types of excavators that are in operation in different mining sites across the world. As it can be seen, in terms of visual features, there are differences between these excavator types. Aside from the difference in aspect ratios of the buckets, in Cable excavators the camera tends to be installed much further from the bucket than in hydraulic shovels, which causes both the bucket and the objects inside of it to look much smaller in images. Consequently, to be practical, any object detection models for the task of bucket monitoring must be able to capture the variation of visual features that is present across shovel types.

There are several objects of interest in frames such as these that must be accurately detected and localized in order to monitor shovel's operation. These can be classified into 3 categories:

A. Sacrificial Ground Engaging Tool (GET) which often includes hardened metal teeth and adapters for digging into the material.

B. Specific areas of the bucket's surface used for differentiating empty vs. full buckets.

C. Type of the material inside of the buckets.

In a shovel or excavator, sacrificial components are the parts that engage the ground causing them to wear down over time during normal operation and, also, partially or completely break on occasion. Bucket's teeth and adaptors are among the most important sacrificial components that can cause severe safety concerns when they break as they end up stuck in the crusher. Aside from causing the entire production line to stop for days and losing millions of dollars, dislodging these broken pieces from the crusher is a dangerous task (due to the inertial forces involved) and has led to injuries in the past. By detecting these events (i.e. broken or missing sacrificial parts) in real-time, we can stop the trucks (containing the broken pieces) before they get to subsequent production line stages and cause damage.

Understanding the real-world sizes of the shovel components and the material inside of it are of high importance to a mine's operation and safety. On one hand, if we know that a component has been worn down it can be replaced before it breaks, thus avoiding the safety concerns associated with broken pieces. In addition, scheduling the replacement of several components together will dramatically reduce shovel down-time. On the other hand, understanding the sizes of the objects inside of the bucket, helps us avoid dropping large boulders into the trucks which damages the trucks and causes safety concerns, as well as, help the mine optimize their blasting operations to yield particles of correct sizes. We compare the detected width of the bucket in pixels with its actual physical width in order to calculate a conversion factor that we use to know the real sizes of all the other objects in the frame,

including the sizes of the rocks inside the bucket.

The detection of the bucket's surface features also facilitates real-time pose estimation (position and orientation) of the bucket with respect to the fixed camera frame. Landmarks on the bucket provide object points (3D points of landmark corners) and axis points (3D points to construct axes) to solve the Perspective N-point (PNP) problem [9] and resolve a transformation matrix (rotation and translation) between the camera and landmark coordinate frames. This transformation not only corrects for distortion artifacts as a result of bucket pose to resolve actual feature geometry for GET components and their wear characteristics with time, but also facilitates an inverse kinematics solution to compute joint variables of the shovel with respect to the bucket's position and orientation [10]. Real-time tracking of the shovel's kinematic configuration is critical for several control functions such as collision avoidance and end-effector wrench monitoring for estimates of payload and cutting resistance between material and teeth - all of which are crucial to the mine's safety and efficiency. Augmenting this technique with sensor fusion from measured joint kinematics using an Extended Kalman Filter (EKF) improves robustness with a state estimator even when certain landmark features are temporarily occluded [11]. In general, vision-based pose estimation is highly advantageous in that sensors are compatible with a wide variety of shovel types and are non-invasive, making them relatively easy to install.

Finally, effective classification of the material inside of the bucket is important for optimizing mine blasting process and significantly saving time and cost. Also, for both safety and efficiency reasons it is important to detect any foreign objects in the bucket such as pieces of broken metal or plastic, etc. and, also, larger than normal boulders. If not detected in time, these objects damage other equipment, stop the production line, and in many cases cause severe safety concerns.

3. Related Work

Despite the large body of research that has been done on developing the prerequisite technologies needed to achieve autonomous vehicles, little work has been done on autonomous excavators specifically. As a result, while a lot of the research done on autonomous driving could be leveraged for autonomous excavators as well, we have focused on the missing technologies that are specific to the task of autonomous excavators.

In terms of high-definition maps, which are among the most crucial components for autonomous vehicles [12], autonomous excavators have lower resolution and accuracy requirements. This is because unlike other vehicles that must travel large distances across different

regions, excavators are usually confined to a mine site. However, when it comes to Perception, autonomous excavators have different requirements than other vehicles. Perception refers to any software and machine learning modules that is responsible for acquiring raw sensor data from on vehicle sensors such as cameras, lidar, and radar, and converting this raw data into complete scene understanding for the autonomous vehicle [13]. Compared to autonomous cars, excavators need to detect a much smaller category of objects (only the categories of objects that are present on a mining site). They can also (in most cases) spend more computational time on detection algorithms, since they operate at lower speeds than an autonomous car. However, autonomous excavators need to differentiate objects that are visually much more like each other (i.e. rocks vs. fine dirt vs. mud, etc.) to know whether they are digging the right type of material. In addition, heavy dust, extreme weather, extremely low contrast that are present in most mining sites, makes detection using visual features more difficult than most driving conditions.

Due to interruptions in internet connectivity, limited bandwidth on mining sites and strict timing requirements for detecting safety critical objects, all mission critical applications need to run on our ruggedized industrial embedded system. The embedded system operates in temperature range of -40 to 55 degrees Celsius, random vibration profile of 3 Grms in frequency range of 5 to 500 Hz, and half a sinusoid shock profile of 30G with 10ms duration. The embedded systems have 32 GB of RAM, a quad core Intel CPU with a base clock cycle of 2.80 GHz, a 4 GB Nvidia GTX1050Ti GPU. Consequently, we not only require a model that is very fast at inference-time but also one with a small memory foot print.

Additionally, single frames contain many artifacts of low contrast, intermittent occlusion due to dust, extremely low contrast, low brightness which all contribute to difficulty in object detection.

4. Methods

Figure 2 shows the overall architecture of our multi-frame object detection model. We propose a multi-frame Long Short-Term Memory (LSTM-) based approach in order to accumulate valuable information across frames, make object detection more robust without any significant computational overhead. We formulate a problem by learning a function $F(I) = F(L(s)) = F(L(Y(I)))$ where L represents LSTM, Y represents YOLO [14], and s and I are sequences of feature maps and images respectively. We use single-shot object detector YOLO-v2 [15] as our main backbone which operates on each frame. YOLO architecture predicts a tensor of shape $SxSx5*BxC$, where S is the number of the grid cells, B is the number of
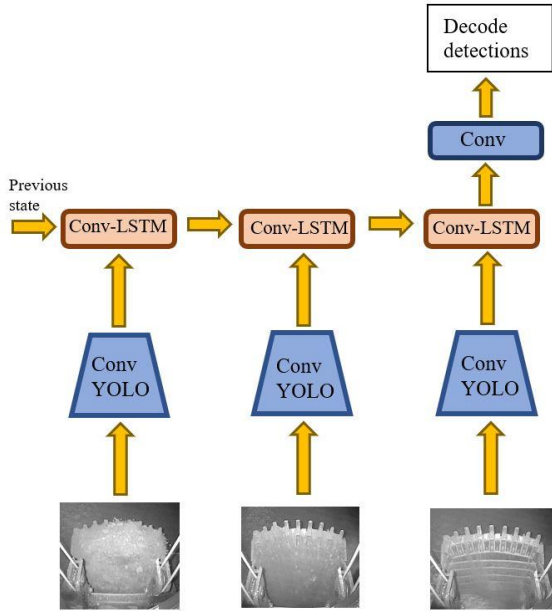
Figure 2 High-level architecture of our multi-frame object detection approach using YOLO backbone.

bounding box predictors and C is the number of classes. Following work on real-time video object detection of Bottleneck-LSTM [16] with an architecture of Convolutional LSTM [17] and SSD [18] which uses MobileNet [19] as a main single-frame backbone for feeding into LSTM layers, we use a similar method but with YOLO backbone, since Bottleneck-LSTM is dedicated to mobile devices running CPUs. In our experiments YOLO showed to be twice faster than SSD on a Nvidia GTX1050Ti GPU and a 640x480 input image. Instead of taking the last prediction layer of YOLO, we feed second last layer before activation into LSTM.

Information from the feature maps of YOLO is propagated and aggregated in Convolutional LSTM which models both spatial and temporal information. Last output layer would predict the output tensor of the same size as YOLO backbone, which would then be decoded into

bounding box representation.

Single-frame YOLO detector is first trained on the whole training set, then we build a full model with Convolutional LSTM and continue training until convergence. For training of YOLO batch size of 16 is used while for training the full model we used batch size of 2 due to limitations in the GPU memory.

Our final model is a LSTM with YOLO which was trained on sequences of 10 steps and can operate online for each incoming frame. Empirically, we learned that a sequence length of 10 (which corresponds to 1 second), at training time, gives us enough temporal information without sacrificing accuracy similar to experiments done in [16]. Furthermore, in the original LSTM paper the same sequence length was used during training as well. Due to the relatively slow movement between individual frames we use every third frame, thus reducing the frame rate from 30 to 10 fps which also allows us to have larger temporal receptive field. We use Adam optimizer [20] with learning rate=0.001 and its standard hyperparameters. The model is implemented in Keras and Tensorflow [21]. We utilize multi-scale training with data augmentation and at test-time our input resolution is 640x480. Anchor boxes are selected with k-means, however since we have imbalance of object instances and significantly varying sizes of bounding boxes, we run k-means twice on smaller and larger size objects, and then combine the anchor boxes.

5. Experiments and Results

Figure 3 shows the results of our multi-frame prediction on different shovels types. For each shovel type, our training dataset contained about 3000 annotated images. We used a fixed validation set of 250 images (not seen during training) for our hyper-parameter optimization. In order to get a complete picture of our model generalization, we have evaluated our approach on 4 different test sets each containing 250 images:

Test Set A: Contains examples, not seen during training or validation, that are from the same mining sites and bucket types present in the training set. We use our

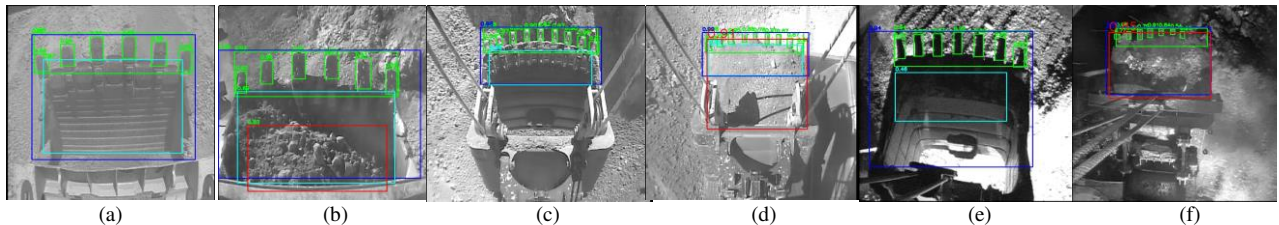| (a) | (b) | (c) | (d) | (e) | (f) |

Figure 3. Predicted bounding boxes for different shovel types using our multi-frame object detection model. Dark blue bounding boxes capture the bucket, green bounding boxes detect teeth and teeth-line objects, red bounding boxes show fine material inside of the bucket, and light blue bounding boxes detect empty buckets. Empty Hydraulic Bucket (a), Full Hydraulic Bucket (b), Empty Bucyrus Bucket (c), Full Bucyrus Bucket (d), Empty P&H Bucket (e), Full P&H Bucket (d).

Test Set	All Classes	Teeth	Buckets	Teeth-line	Surface Patterns	Material Inside
A: new examples	0.926	0.899	0.935	0.957	0.881	0.956
B: new mining sites	0.920	0.890	0.938	0.953	0.883	0.939
C: new bucket types	0.890	0.874	0.908	0.944	0.773	0.950
D: new hard examples	0.856	0.849	0.791	0.915	0.819	0.904

Table 1 Shows the results of our *single-frame* object detection model on 4 different test sets. All reported numbers are Mean Average Precision of the Intersection Ober Unions of predictions at 10 different thresholds.

performance on this test set as a baseline.

Test Set B: Contains examples of mining sites that were not represented in our training or validation sets but include excavators with bucket types that were observed in our training set. We use this test set to quantify the recognizability of our solution across different mining sites. This is important because mines in different locations can have vastly different natural and artificial lighting conditions, camera lens qualities, and dust/ shadow issues.

Test Set C: Contains examples of bucket types that were not represented in our training or validation sets, but from mining sites that we observed during training. This is perhaps one of the most important indicators of the overall performance of our object detection solution because different bucket types with largely different visual features can be installed on the same excavator type. Since it is not desirable to limit mining sites to use a specific bucket type for their excavators, ant practical bucket monitoring solution must generalize well to previously unseen bucket types.

Test Set D: Contains difficult examples that were not included in our training or validation sets. These include frames that capture conditions with lots of dust, different shadows, low lighting, too much lighting (i.e. glare), dirty camera lenses, or low contrast situations. These frames were selected from the ones that our previous solutions performed specially poorly on or are selected by human experts because they contained above conditions. We use this test set to get a measure for our worst-case-scenario performance in terms of the worst imaging conditions that might be encountered in different operations.

In terms of evaluation metrics, we used the mean average precision (mAP) at different intersection over union (IoU) thresholds. We used the following thresholds:(0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95). At each threshold, a true positive happens when there is a single object that matches the ground truth with and IoU above the thresholds; a false positive happens when the predicted object had no associated ground truth objects; while a false negative happens when there is no prediction for a given ground truth object. We then compare each prediction to all ground truth objects in order to calculate a precision value.

Then, for each class of objects and for each image, we take the mean of our precision values at each threshold. The final score returned for each class is then the mean taken over the individual average precisions of all the images in the dataset. Finally, we take the mean of scores for each class in order to report a final score value for the entire model that represents the mean precision values at all intervals, for all classes, and all images.

Table 1 shows the results of our experiments on each of the data sets above using our single frame detection stage. As you can see in the second row (of table 1), our models generalize very well to new mining sites that might, potentially, have different amount of dust, lighting conditions, or camera qualities. As you can see in the third row of Table 1, except for bucket patterns, most of the objects perform well on bucket types not previously seen during training or validation. These results can be attributed to the fact the largest difference between the visual appearances of different bucket types is in the patterns on the bucket's surface. Finally, the last row of Table 1, shows the effects of bad imaging conditions (i.e. dusty, shadowy, images taken in low lighting or low

Test Set	All Classes	Teeth	Buckets	Teeth-line	Surface Patterns	Material Inside
A: new examples	0.953	0.962	0.952	0.980	0.901	0.969
B: new mining sites	0.949	0.962	0.962	0.972	0.972	0.952
C: new bucket types	0.914	0.959	0.959	0.970	0.970	0.944
D: new hard examples	0.883	0.913	0.913	0.940	0.941	0.900

Table 2 Shows the results of our *multi-frame* object detection model on 4 different test sets. All reported numbers are Mean Average Precision of the Intersection Ober Unions of predictions at 10 different thresholds.

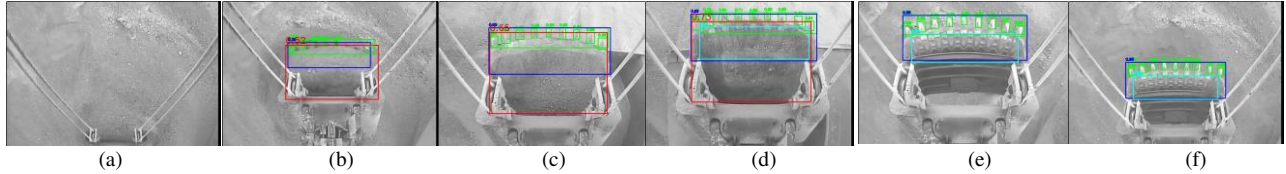|					
(a)	(b)	(c)	(d)	(e)	(f)

Figure 4. A typical Dig-Carry-Dump cycle of operation for mining shovels. Bounding boxes are predictions using our multi-frame object detection model. Dark blue bounding boxes capture the bucket, green bounding boxes detect teeth and teeth-line objects, red bounding boxes show fine material inside of the bucket, and light blue bounding boxes detect empty buckets.

contrast, or with dirty camera lenses, etc.). As expected, all objects perform slightly worse than baseline. It is interesting to note that the adverse effect of imaging conditions on the performance of bucket patterns is less than the effects of observing a new type of object.

Table 2 shows the results of our experiments on each of the data sets above using our *multi-frame* detection stage. In these experiments, we used a model trained with sequences lengths of 10. As previously mentioned, we record our videos at 30fps; so, with a skip rate of 3 (i.e. skipping every 2 frames), each of our sequences contains the information corresponding to 1 second of video. During inference, we continuously append the hidden state of the current element to hidden state used for predicting the next frame. In other words, we do not batch our frames into sequence lengths of 10 during inference time.

Table 2 shows that on all our test sets, majority of our objects perform better using multi-frame detection than they do with single frame detection. We observe the largest performance increase on teeth and teeth-line respectively. This is because these objects exist (and look similar) in most frames in any given sequence. Therefore, they benefit the most from combining the information of several frames. In addition, you can see in Table 2, that the objects that benefit the least form multi-frame detection are the material inside of bucket, and bucket pattern. This is intuitive since the consecutive frames in any given sequence differ in these objects the most.

Figure 4 shows the results of our predictions on a sequence of frames encompassing a typical dig-carry-dump cycle. As you can see, some frames will contain the full bucket and some will show an empty bucket (i.e. with bucket patterns visible), while it is possible that all the frames contain the teeth and the teeth-line.

Finally, it is interesting to note that the only place where using multi-frame prediction does not provide much advantage over single frame prediction is in the case of bucket pattern on our test set containing new bucket types. This could be attributed to the fact that our backbone (i.e. single frame prediction) does so poorly on predicting the patterns on buckets it has never seen before that even having access to multiple frames does not improve the performance. Furthermore, as mentioned above, in a given sequence, bucket patterns are among the objects with the largest variations. The bucket patterns not only go out of view when the bucket is digging or carrying material, but their appearance is also largely influenced by the orientation of the bucket with respect to the camera.

In terms of failure cases, the most common are mistakes in object detection due to bad imaging conditions or camera installations. Some of the false positives and false negatives in our detections could be solved by relying on the fact that we have access to multiple frames to make a reliable decision. However, when faced with poor imaging conditions that last for a relatively long time such as dusty conditions, dirty lens, or bad installation, one cannot rely on future frames to recover from our faulty predictions and could end up with wrong decisions. In future, we plan to develop solutions to detect these poor imaging conditions (i.e. classify an image as dusty or shadowy or taken at wrong camera position) to improve our decision-making logic.

6. Conclusion

We have proposed a multi-frame object detection solution in the scenario of autonomous mining shovels operation. Our approach runs in real-time and achieves near human-level accuracy on all objects. Currently, we are working on testing other categories such as foreign objects, boulders, lip-shrouds, wing-shrouds and produce longer term predictions. In future, we plan to utilize our real-time object detection to create a perception module capable of tracking various objects and incorporating this knowledge into a general understanding about the environment. This will enable us to create a decision-making module that combines a given excavator's perception of its environment with a high-definition map of its surrounding to autonomously guide its operation.

References

[1] M. S. a. H. Administration, "Mine Injury and Worktime, Quarterly," Mine Safety and Health Administration, 2017.
[2] A. Lala, M. Moyo, S. Rehbach, and R. Sellscho "Productivity in mining operations: Reversing the downward trend." McKinsey. https://www.mckinsey.com/industries/metals-and-

mining/our-insights/productivity-in-mining-operations-reversing-the-downward-trend.

[3] P. Bhattar. "Mining industry's energy problem." https://www.wartsila.com/twentyfour7/energy/mining-industry-s-energy-problem.

[4] R. K. Ridley *et al.*, "System and a method for detecting a damaged or missing machine part," 2013.

[5] S. T. Bilandi, N. Nabavi, A. O. Esfahani, I. L. Bell, and M. M. I. Corp, "Method and apparatus for generating an indication of an object within an operating ambit of heavy loading equipment," US9030332B2, 2015.

[6] S. N. Lim, N. Zhou, J. V. B. Soares, and G. E. Co, "System and method for detecting missing tooth in mining shovel," US9886754B2, 2018.

[7] S. T. Bilandi *et al.*, "Method and apparatus for locating a wear part in an image of an operating implement," US20180130222A1, 2018.

[8] S. T. Bilandi *et al.*, "Method and apparatus for identifying fragmented material portions within an image," US20190012768A1, 2019.

[9] A. Penate-Sanchez, J. Andrade-Cetto, and F. Moreno-Noguer, "Exhaustive Linearization for Robust Camera Pose and Focal Length Estimation - IEEE Journals & Magazine," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 35, no. 10, pp. 2387-2400, February 2013.

[10] F. Michel, A. Krull, E. Brachmann, and M. Ying Yang, "Pose Estimation of Kinematic Chain Instances via Object Coordinate Regression," presented at the British Machine Vision Conference, Swansea, 2015.

[11] L.-H. Lin, P. D. Lawrence, and H. Robert, "Stereo vision based swing angle sensor for mining rope shovel," presented at the International Conference on Intelligent Robots and Systems, Taipei, Taiwan, 2010.

[12] S. Bauer *et al.*, "Using High-Definition maps for precise urban vehicle localization - IEEE Conference Publication," presented at the IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), Rio de Janeiro, Brazil, 2016.

[13] A. Petrovskaya and S. Thrun, "Model based vehicle detection and tracking for autonomous urban driving | SpringerLink," *Autonomous Robots* vol. 26, no. 2-3, pp. 123-139, 2019.

[14] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," presented at the Computer Vision and Pattern Recognition, 2015/06/08, 2015.

[15] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," presented at the Computer Vision and Pattern Recognition, 2016/12/25, 2016.

[16] M. Liu and M. Zhu, "Mobile Video Object Detection with Temporally-Aware Feature Maps," presented at the Computer Vision and Pattern Recognition, Salt Lake City, US, 2017/11/17, 2018.

[17] X. SHI, Z. Chen, H. Wang, D.-Y. Yeung, W.-k. Wong, and W.-c. WOO, "Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting," in *Advances in Neural Information Processing Systems*, 2019, pp. 802-810.

[18] W. Liu *et al.*, "SSD: Single Shot MultiBox Detector," presented at the European Conference on Computer Vision, 2016.

[19] A. G. Howard *et al.*, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," 2017/04/17 2017.

[20] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," presented at the International Conference for Learning Representations, San Diego, US, 2014/12/22, 2015.

[21] M. Abadi *et al.*, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," 2016/03/14 2016.