sciendo Organization, Technology and Management in Construction 2021; 13: 2341–2352

Research Article Open Access

Martina Šopić*, Mladen Vukomanović, Diana Car-Pušić and Ivica Završki

Estimation of the excavator actual productivity at the construction site using video analysis

DOI 10.2478/otmcj-2021-0003
Received: May 03, 2020; Accepted: January 12, 2021

Abstract: Current estimates of the actual productivity of heavy construction machinery at a construction site are not supported by an appropriate and widely used methodology. Recently, for the purpose of estimating the actual productivity of heavy construction machinery, vision-based technologies are used. This paper emphasizes the importance of estimating actual productivity and presents a way (i.e. a research framework) to achieve it. Therefore, the aim of this paper is to propose a simple research framework (SRF) for quick and practical estimates of excavator actual productivity and cycle time at a construction site. The excavator actual productivity refers to the maximum possible productivity in real construction site conditions. The SRF includes the use of a video camera and the analysis of recorded videos using an advanced computer program. In cases of continuous application of SRF, a clear and transparent base for monitoring and control of earthworks can be obtained at an observed construction site.

Keywords: earthworks, excavator, actual productivity, cycle time, video analysis, SRF

1 Introduction

Planning earthworks performance is a difficult and demanding task. When planning earthworks, the planner is faced with the problem of having to split up the required work into manageable activities, because earthworks activities are influenced by many factors, including site layout, machinery selection, weather characteristic, and condition of haul road (Mawdesley et al. 2002). A major risk in any civil engineering project is that the contractor may, during the course of construction, encounter physical obstructions or conditions on the site (e.g. ground conditions, sub-surface water, foundation rock) that are unexpected and unforeseeable at the time of tender and may delay his work or cause an increase in costs (Ndekugri and Mcdonnell 1999). Actual productivities achieved in earthworks are considerably different from the theoretical productivity values assumed at the planning stage (Hola and Schabowicz 2010). Construction is generally considered to be conducted in highly challenging site conditions, because of the machineries involved in operation, transfer of materials, moving workers, and changing status (Kim H. et al. 2014). It is difficult to accurately estimate the productivity of an earthwork, because earthwork productivity varies depending on the unique geologic conditions, types of earthwork equipment, and equipment allocation plan (Kim H. et al. 2018a). Therefore, when planning the performance of earthworks, it is quite a challenging task to accurately predict and estimate soil and rock categories, weather impact, productivity of construction machinery, as well as required time and cost. Also, construction site represents a unique, dynamic, and complex environment with possible occurrence of unforeseen circumstances and dangers that can negatively affect the work performance. Hence, when earthworks begin at a construction site, deviations in the plan may be expected.

Productivity performance of heavy construction machinery (e.g. loaders, excavator, hauling trucks) has a significant role in the earthworks operation (Salem et al. 2017). Productivity monitoring of heavy equipment, which is a process of measuring, analyzing, and improving the operational efficiency and performance of the equipment, is a major task for managing and completing earthmoving projects successfully (Kim J. and Chi 2020). Monitoring the earthworks progress and accurate estimation of construction machinery productivity provides a detailed insight into the performance, feedback on the correctness of the

***Corresponding author: Martina Šopić,** Faculty of Civil Engineering, University of Rijeka, Rijeka, Croatia.
E-mail: martina.sopic@uniri.hr
Mladen Vukomanović and Ivica Završki, Faculty of Civil Engineering, University of Zagreb, Zagreb, Croatia
Diana Car-Pušić, Faculty of Civil Engineering, University of Rijeka, Rijeka, Croatia

decisions made, more accurate report of the time and cost necessary for the activity, early detection of low productivity, as well as any other possible defects (Šopić et al. 2018). For proper construction site management and plan revisions during construction, it is necessary to understand a construction site's status in real time (Kim H. et al. 2018b). Earthworks operation needs an automated system to capture job-site data, with the capability of handling a large volume of data in near real time to make a correct productivity estimation, and to calculate the time needed as well as the cost (Montaser and Moselhi 2014). There is a need for low cost, reliable, and automated method for activity analysis of the construction machinery at the construction site that can be widely applied across all construction projects (Golparvar-Fard et al. 2013). Therefore, estimating the actual achieved productivity of construction machinery at a construction site may reduce wrong, inaccurate, and subjective estimates that were assumed at the time of planning the performance of earthworks.

Lately, for the purpose of estimating the actual productivity of construction machinery at a construction site, researchers use a variety of methods and tools of fast-growing wireless technologies, such as (non-vision-based) sensing and tracking technologies (Montaser and Moselhi 2012-2014; Montaser et al. 2012; Ibrahim and Moselhi 2014; Alshibani and Moselhi 2016), and vison-based technologies (Bügler et al. 2017; Chen et al. 2020; Kim J. and Chi 2020). The vision-based methods can visualize the equipment state directly from images and videos, so that it becomes easy to identify false recognition and analyze the reasons behind low productivity, whereas the non-vision-based methods have difficulty in categorizing the activities in detail and they require attaching sensors to monitor the equipment (Chen et al. 2020). Different approaches are available to monitor earthworks equipment using computer vision; the first approach is to develop (or use) software to analyze the visual capture from common electronic tools, such as video cameras, and the second method is to employ other sensing devices together with video cameras to obtain more data from the scene (Rezazadeh Azar et al. 2013). This paper will be oriented toward researches that involve video camera and video analysis.

Visual recording devices, such as video camera, have been broadly used to facilitate work progress or safety monitoring in construction sites (Chi and Caldas 2011). Operation-level visual monitoring is commonly conducted to better understand on-site productivity and safety (Kim J. et al. 2019). The productivity analysis and optimization could be made with the surrounding information provided by the videos (Chen et al. 2020). The use of visual tracking, to track construction equipment and workers, has been recently promoted to facilitate construction automation (Xiao and Zhu 2018).

The aim of this paper is to propose a simple research framework (SRF) for quick and practical estimates of excavator actual productivity and cycle time at a construction site during earthworks. This SRF includes the use of a video camera for the purpose of recording excavator during earthworks at a construction site and the analysis of recorded videos using an advanced computer program.

2 Background

2.1 Artificial intelligence, machine learning, and deep learning

Automatic detection, recognition, and tracking of various objects and activities of construction resources (i.e. equipment and/or workers) in videos can be possible with artificial intelligence (AI) or, more precisely, with machine learning (ML) and deep learning (DL). The meaning of AI, in a simplified and figurative manner, suggests the assignment of "cognitive" ability to machines. ML is a subfield of AI, and the meaning of ML, in a simplified and figurative manner, suggests training the machines to achieve "learning" abilities. ML can be broadly defined as computational method using past information available to the learner (e.g. electronic data collected and made available for analysis, such as digitized human-labeled training sets) to improve performance or to make accurate predictions (Mohri et al. 2018). The emphasis is on "automatic," i.e. ML is concerned about general purpose methodologies that can be applied to many datasets, while producing something that is meaningful (Deisenroth et al. 2020). Some ML algorithms used to build models are Hidden Markov Model (HMM), Gaussian mixture model (GMM), support vector machine (SVM), Histogram of Oriented Gradients (HOG), etc. HOG is an algorithm used for fast and accurate object (or edge) recognition and detection in computer vision or image processing (i.e. an image feature descriptor). SVM is a supervised learning algorithm (i.e. training algorithm to make predictions, by labeling all the data) for classification, nonlinear regression, and outlier's challenges, using a "learned" pattern and data analysis. GMM is unsupervised (i.e. training algorithm to find hidden patterns), probabilistic and modeling algorithm for density estimation and data clustering, with a notable application for object detection and object tracking. HMM is a probabilistic, graphical, and modeling algorithm with prominent application for speech and image processing, handwriting, text and gesture recognition, machine maintenance, and activity analysis in videos.

DL is a subfield of ML. The “deep” in DL comes from the many (hidden) layers that are built into the DL models, which are typically neural networks (Wehle 2017). One of the names that DL has gone by is Artificial Neural Networks (ANNs) (Goodfellow et al. 2016). ANNs are known as universal function approximators because they are able to learn any function, no matter how wiggly, with just a single hidden layer (Maini and Sabri 2017). Therefore, one hidden layer or more hidden layers refer to one or more layers between input and output data in an algorithm. When an ANN has two or more hidden layers, it is called deep neural network (DNN) (Géron 2017). DL, as a new generation of ANN (Ghorbani et al. 2020), is in the intersections among the research areas of neural networks, artificial intelligence, graphical modeling, optimization, pattern recognition, and signal processing (Deng and Yu 2014). Some DL algorithms (i.e. neural network architectures) are Convolutional Neural Network (CNN, ConvNet) and Recurrent Neural Networks (RNN). CNN and RNN have great abilities to represent complex relationships between input data (e.g. images) and output data (e.g. object classes) (Kim J. and Chi 2019). CNN is one of the most popular DL approaches in the field of graphic processing as CNN performs well in image processing and directly deals with raw images (Liu et al. 2019). Therefore, CNN is a powerful algorithm, used for analysis of visual imagery and for performing tasks such as detection, recognition, segmentation, and classification of object or features in an image. RNN is a powerful algorithm used for classification, labeling, and generation. In RNN, the output from the previous task is fed as input to the current task. Long Short Term Memory (LSTM) is a type of RNN, useful in time series prediction.

Building models with ML or DL algorithms is possible in advanced computer programs (i.e. environments), such as MATrix LABoratory (MATLAB) (Cleve Moler, MathWorks), Python (Guido van Rossum), and Visual Studio (Microsoft). This kind of advanced computer programs may provide numerous analysis (of data, video, or graphics), advanced model development or simulation for solving mathematical, engineering, and scientific problems, as well as provide designing visualizations and help in building or training applications.

2.2 Literature review

One of the significant importance of video analysis is reflected in the application of ML and DL. ML and DL allow automatic recognition of various construction operational resources and activities. However, video analysis indicates some weaknesses in implementation, considering estimations of the actual productivity of construction machinery at jobsites. Other limitations, when using a video camera and video analysis include video length, overlapping of resources, camera motion, background cluster (i.e. background data set), bad weather conditions, poor lighting conditions, etc. This paper is dealing with one of the problems in using video camera and video analysis, more precisely, the problem of estimating the actual productivity of an excavator at a construction site.

2.2.1 Classical Computer Vision

Zou and Kim (2007) presented an image processing system for effective measurement of hydraulic excavator idle time. Their system was developed in MATLAB, using Hue Saturation and Value (HSV) color space (i.e. color parameters for image processing and analysis). Gong and Caldas (2011) analyzed 12-minute video record of a bobcat loader (i.e. three cycle time) during earthworks. They developed a prototype program called “Construction Video Analyzer,” written using C++ programming language, which integrated several commercial strength libraries, such as Microsoft Foundation Class, DirectX, and Intel Open Source Computer Vison Library. The purpose of their video interpretation was to automatically monitor the bobcat’s time utilization, production cycle, and abnormal production scenarios.

2.2.2 Machine Learning

Rezazedeh Azar and McCabe (2012) presented two approaches to detect and distinguish off-highway dump trucks from other earthmoving machines in digital construction videos, with HOG detector as the main descriptor. Rezazadeh Azar et al. (2013) presented a vision-based framework that can recognize and estimate dirt loading cycles. In their research, some of the techniques used were the HOG descriptor and SVM classifier, which they implemented using Open Source Computer Vision Library (i.e. OpenCV) in Visual C++ environment. Golparvar-Fard et al. (2013) presented a computer vision-based algorithm for recognizing every single action of earthmoving construction equipment (i.e. for excavator and dump truck). Their proposed method was based on the HOG descriptor and SVM classifier, which they trained using MATLAB. Memarzedeh et al. (2013) presented a computer vision-based algorithm for automated detection of construction workers and equipment from site video streams. Their method was based on using HOG descriptor with Hue-Saturation-Value (HSV) colors (which they denoted as HOG+C) and SVM classifier. Implementation of their proposed algorithm was in MATLAB with several components in C++ for faster computation. Rezazadeh Azar (2016) developed a monitoring system to identify dump trucks and excavators in construction jobsites, using HOG for object detection or recognition and SVM classifier. Implementation of their developed system was in Visual C++ express environment with the use of three open-source libraries. Bügler et al. (2017) proposed a methodology using the combination of two different visual technologies, more precisely, through combination of photogrammetry and video analysis. The purpose of their methodology was to monitor the progress and estimate the productivity of earthworks when deep (large) amounts of excavation is necessary. In their research, target detection was performed using GMM. Kim J. et al. (2018) pointed out that an essential step to measure and monitor the performance of earthworks is the activity identification. For that, they developed a methodology for automatic activity identification of the excavators and dump trucks in videos during earthworks. The automatic activity identification in their paper referred to the "work," "work (move)," "work (stop)," and "idle" activity status of the excavators and dump trucks, according to the types of operational process at a construction site. For their research, they adapted a Tracking-Learning-Detection system (TLD), developed by Kalal et al. (2011), which is the evolving technology of the TLD Vision (an AI company). Also, they used advanced computer programs, more precisely, C++ programming language with Visual Studio and MATLAB. They highlighted that their methodology may be improved using CNN tracking algorithms.

2.2.3 Deep Learning

Kim H. et al. (2018b) proposed a deep convolutional network-based construction object-detection method to accurately recognize construction equipment (i.e. dump truck, excavator, loader, concrete mixer truck, and road roller). In their study, the construction equipment detection model was the region-based fully convolutional network (R-FCN) (i.e. region proposal approach). Their model was trained with a relatively small amount of data by applying transfer learning. Roberts and Golparvar-Fard (2019) presented a vision-based baseline system for detection, tracking, and activity analysis for construction resources, with particular attention to excavators and dump trucks. Their system enabled automatic identification of visually distinctive working activities of excavators and dump trucks from individual frames of video sequence. Working activities for excavator were referring to "idling," "loading bucket," "swinging bucket," "dumping," and "moving" activities, while working activities for dump trucks were referring to "idling," "moving," and "filling" activities. Their system was based on CNN, HMM, GMM, and SVM classifiers, which they trained using MATLAB. Kim J. and Chi (2019) proposed a vision-based action recognition framework (i.e. excavator detection, excavator tracking, and excavator action recognition), considering the excavators' two types of sequential working patterns (visual features and operation cycles), for automated earthworks operation analysis. In their research, they used a Tracking-Learning-Detection (TLD) system to track excavators, while sequential working patterns (visual features and operation cycles) were learned with CNN and Double-Layer Long Short-Term Memory (DLSTM). They highlighted that one of the limitations of their DL model was excessive computational training time and a large amount of training data to build the model. Chen et al. (2020) pointed out some limitations of vision-based methods that were present at the time of writing their paper. More precisely, they pointed out that previous vision-based methods primarily focused on automatic recognition of working activities of construction equipment with little discussion of their practical values, such as using vision-based methods for productivity control of earthworks. Other limitations that they pointed out include difficulties for automatic recognition of working activities if the video sequence is long, or in cases when multiple construction machinery are simultaneously captured. To overcome some of those limitations, Chen et al. (2020) presented a framework that automatically recognized activities and analyzed the productivity of multiple excavators. They used three CNNs to detect, track, and recognize the activities of excavators. They also developed an algorithm to analyze the activity recognition results and calculate the productivity of the excavators. Their activity recognition model was developed in Python environment. Chen et al. (2020) highlighted some limitations in their research, such as overlapping in the detection and tracking results, and light condition of the video (i.e. too bright or too dark). They pointed out that some limitations of their method could be improved using more than one camera, and that the video should be pre-processed by adding a filter. Also, they pointed out that the process of validation of long videos (i.e. about an hour duration) is intensive in terms of time and labor, but they will still test more long videos in the future. Kim J. and Chi (2020) proposed a multi-camera vision-based productivity monitoring methodology that analyzed videos captured from multiple non-overlapping cameras at the jobsites. In their research, earthworks equipment was tracked using TLD system, whereas sequential working pattern of earthworks equipment was learned and analyzed using CNN and DLSTM models. Their model was able to recognize the individual actions of excavators (i.e. "digging," "swinging full," "dumping," "swinging empty," "moving," and "stopping") and dump trucks (i.e. "moving" and "stopping"). Kim J. and Chi (2020) pointed out that their methodology, to their knowledge, is the first attempt in monitoring the productivity of earthworks equipment using multiple (non-overlapping) cameras and that findings of their study can contribute in developing more reliable automated productivity monitoring of earthworks operations. Also, they highlighted that shortcomings of their vision-based monitoring (such as occlusion and tracking errors) can be resolved through the integration with the Internet-of-Things-based (IoT) system, such as Global Positioning System (GPS), Radio Frequency Identification (RFID), and accelerometers.

3 Methodology

This paper proposes an SRF for quick and practical estimation of excavator actual, maximum possible, productivity and cycle time at a construction site during earthworks. This SRF includes the use of a video camera at a construction site and the analysis of recorded videos using an advanced computer program. A particular highlight in using video camera is to record the excavator during excavation and loading of materials into the tipper truck. As the length of the video may present a problem in video analysis, we recommend to record only a few short videos. More precisely, to record, and then analyze, the excavator when performing excavation and loading of materials, from the beginning of loading into the empty box bed of the tipper truck, until the box bed of the tipper truck is fully loaded. The required number of this kind of (short) videos depends on each situation at a construction site. As the application of SRF implies an estimate of the excavator actual, maximum possible, productivity and cycle time of one of the stages of earthworks within good operating conditions, it is not necessary to record more than five to ten videos and analyze two to five of them.

In SRF, the advanced computer program used to analyze recorded videos is MATLAB. MATLAB is, at the same time, a software package for numerical computing and modeling, and also a higher programming language for various scientific and technical applications (Kovačić 2013). Video analysis in SRF includes manual input of appropriate labels of excavator working activities using a specific label automation algorithm that is built in MATLAB. A specific label automation algorithm speeds up the input of labels by interpolating the locations of region of interest (ROI) across the time interval. Otherwise, just about every frame in the video would have to be processed with labels. Labels of excavator working activities imply repetition of excavator working cycles during excavation and loading of materials into the tipper truck. Therefore, the labels of excavator working activities (i.e. cycles) can be named as "excavation," "swingBucketFull," "loadingTipperTruck," and "swingBucketEmpty." After labeling excavator working activities, labels from video can be extracted, enabling some insights into the earthworks. More precisely, by extracting and displaying labeling results, it is possible to notice, calculate, or estimate some earthworks information, such as total duration of loading materials into a tipper truck, the number of excavator buckets used to load a tipper truck, excavator actual productivity, duration of excavator cycle time. Also, the results of labeling can serve as a clear and transparent base for monitoring and control of earthworks if some excavator working activities are (suddenly) justifiably or unjustifiably longer than usual. Figure 1 shows the SRF diagram with goals.

Alshibani (2018) points out that measuring the actual productivity of earthworks operation, which involve heavy machines, can be a complex task and it can be defined as

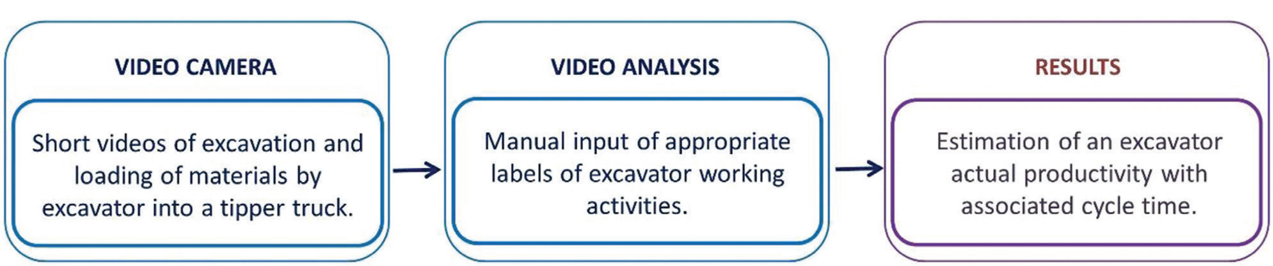

Fig. 1: SRF diagram with goals.

a measure of the output of machines in a certain period of time (i.e. m^3/h). Radujković et al. (2015) provide a similar formula (Eq. 1) for productivity, where the productivity measure (P) is the ratio of production quantity (Q) and time spent (T). Mentioned formula (Eq. 1) and unit measure (m^3/h) are proposed to be used within SRF. However, mentioned formula (Eq. 1) can be transformed into a similar formula (Eq. 2) to better suit SRF. Therefore, in transformed formula (Eq. 2) the actual, maximum possible, productivity (P) is still the ratio of production quantity (Q) and time spent (T), but, the production quantity (Q) (numerator) is the product of the bucket capacity (q) in m^3, with the number of loaded buckets (n_q); and the time spent (T) (denominator) is time, in hours, when excavator is performing earthworks. In formula (Eq. 2), the number of loaded buckets refers to the number of necessary buckets to fully load tipper truck; and time, in hours, when excavator is performing earthworks, refers to excavation and loading materials into the tipper truck, until the tipper truck is fully loaded.

$$P = \frac{Q}{T} \tag{1}$$

$$P = \frac{Q}{T} = \frac{q \times n_q}{T} \tag{2}$$

Similar to the excavator actual productivity (P) estimate, duration of the excavator one cycle time (t) can also be estimated. One repetitive excavator cycle time (t) consists of the excavation, swinging the full excavator bucket to the tipper truck (or near the pit), loading of material into the tipper truck (or disposing near the pit), and swinging back the empty excavator bucket for the excavation. Therefore, to estimate the duration of the excavator one cycle time (t), a formula (Eq. 3) is proposed to be used within SRF. In the proposed formula (Eq. 3), the duration of the excavator one cycle time (t) is the ratio of time (T), in seconds, when excavator is performing earthworks, and number of excavator working cycles (n_{wc}). In the proposed formula (Eq. 3) the time, in seconds, when excavator is performing earthworks, refers to the time required for excavation and loading materials into the tipper truck, until the tipper truck is fully loaded. When an excavator is loading materials into a tipper truck, it is assumed that the number of excavator working cycles (n_{wc}) will be equal to the number of loaded bucket (n_q).

$$t = \frac{T}{n_{wc}} \tag{3}$$

For the practical application of SRF, we selected one construction site, located in Rijeka (Croatia), where earthworks were performed at the time of recording. We recorded five short videos of the excavator during excavation and loading of materials into the tipper truck. At the observed construction site, for performing the earthworks, one excavator and one tipper truck were activated. The excavator and tipper truck were in good shape and well-maintained. Manufacturer of the excavator was Hyundai and the model was 220LC-9, with bucket capacity of 0.9 m^3, whereas manufacturer of the tipper truck was MAN and the model was TGA 35.350 M 8 × 4, with the box bed capacity of 15 m^3. At the observed construction site, the operating conditions of the machines were good, the use of working time was excellent, and the weather was sunny. Before excavation, the material was broken using a hydraulic hammer.

For video analysis, we selected two characteristics and common videos from the observed construction site. The duration of the first video was about 8 minutes, whereas the duration of the second video was about 10 minutes. Each video was showing excavation and loading of materials by excavator, from the beginning of loading into the empty box bed of the tipper truck until the box bed of the tipper truck was fully loaded. In MATLAB, we manually inputted labels, associated with appropriate excavator working activities (i.e. cycles) using label automation algorithm. As written earlier, the labels of excavator working activities (i.e. cycles) were named "excavation," "swingBucketFull," "loadingTipperTruck," and "swingBucketEmpty." After labeling excavator working activities, MATLAB also provides a preview of processed videos, in which we can see appropriate appearance of inputted labels, linked with certain excavator activity in progress. Here, Figure 2 shows some (cropped) pictures from the first video (8 minute video duration) with labels displayed, and Figure 3 shows some (cropped) pictures from the second video (10 minute video duration) with labels displayed.

4 Results

After labeling the excavator working activities, we can extract the labels from both videos to see the results (Figures 4 and 5). By displaying the results of labeling from the first video (Figure 4) we can enumerate that tipper truck was loaded with 11 buckets of the excavator. Also, we can enumerate 11 repeated excavator working cycles, which is, obviously, matching with the number of loaded

Fig. 2: Four (cropped) pictures from the first video with the display of labels.

bucket of the excavator. However, we can notice that there is an exception, or deadlock, in working cycles of the excavator. The exception is caused by the site manager who came to convey some information to the operator of the excavator. Although the deadlock was justified, the duration of the excavation and loading of materials by excavator into the tipper truck was extended.

By displaying the results of labeling from the second video (Figure 5) we can enumerate that the same tipper truck was loaded with 15 buckets of the excavator, which is 4 additional buckets than in the first video. The exact reason for the difference in the number of buckets is unknown. This can happen if the excavator buckets were not completely filled, or if there is a difference in material class and content. However, in some cases, the earthworks are paid by the number of (completely) filled tipper trucks that transport material from the construction site to the dump place. So, if every tipper truck is not completely filled, it may cause higher price for the earthworks!

From the second video, we can also enumerate 15 repeated excavator working cycles. But, as in the first video, there are exceptions in working cycles of the excavator. Those two exceptions were caused by the large stones that were excavated and disposed near the pit. Although these deadlocks, as in the first video, are justified, they again extended the duration of the excavation and loading of materials by excavator into the tipper truck.

With the purpose to estimate actual, maximum possible, productivity of the excavator at the observed construction site, we can use the results of labeling (Figures 4 and 5). In this case, the actual, maximum possible, productivity of the excavator is presenting uninterrupted and continuous work operation by excavator at the observed construction site in Rijeka. Figures 4 and 5 are showing the duration length of works (in seconds) when the excavator is performing earthworks and the duration length (in second) of all exceptions. All the duration lengths are simplified and rounded, in a way that the last number is

Fig. 3: Four (cropped) pictures from the second video with the display of labels.

0 or 5. Therefore, the estimates of the actual, maximum possible, productivity of the excavator at the observed construction site in Rijeka, from the first and the second video are given below. The first calculation is from the first video, and the second one is from the second video.

$$P_1 = \frac{Q_1}{T_1} = \frac{q \times (n_q)_1}{T_1} = \frac{0.9 \times 11}{\frac{(105+225)}{3600}} = 108\ \frac{m^3}{h}$$

$$P_2 = \frac{Q_2}{T_2} = \frac{q \times (n_q)_2}{T_2} = \frac{0.9 \times 15}{\frac{(335+95+20)}{3600}} = 108\ \frac{m^3}{h}$$

The calculations of the estimated average excavator cycle time, using the results of labeling (Figures 4 and 5) are given later. The first calculation is from the first video, and the second is from the second video. The estimated average excavator cycle time refers to the continuous work operation by excavator, without exceptions or deadlocks, in working activities of the excavator.

$$t_1 = \frac{T_1}{(n_{wc})_1} = \frac{105+225}{11} = \frac{330}{11} = 30\ sec$$

$$t_2 = \frac{T_2}{(n_{wc})_2} = \frac{335+95+20}{15} = \frac{450}{15} = 30\ sec$$

From the previous calculations, it can be concluded that the estimated actual, maximum possible, productivity of the excavator, at the observed construction site in Rijeka, is 108 m^3/h. Furthermore, estimated duration of the average excavator cycle time is 30 seconds. Excavator actual, maximum possible, productivity and cycle time

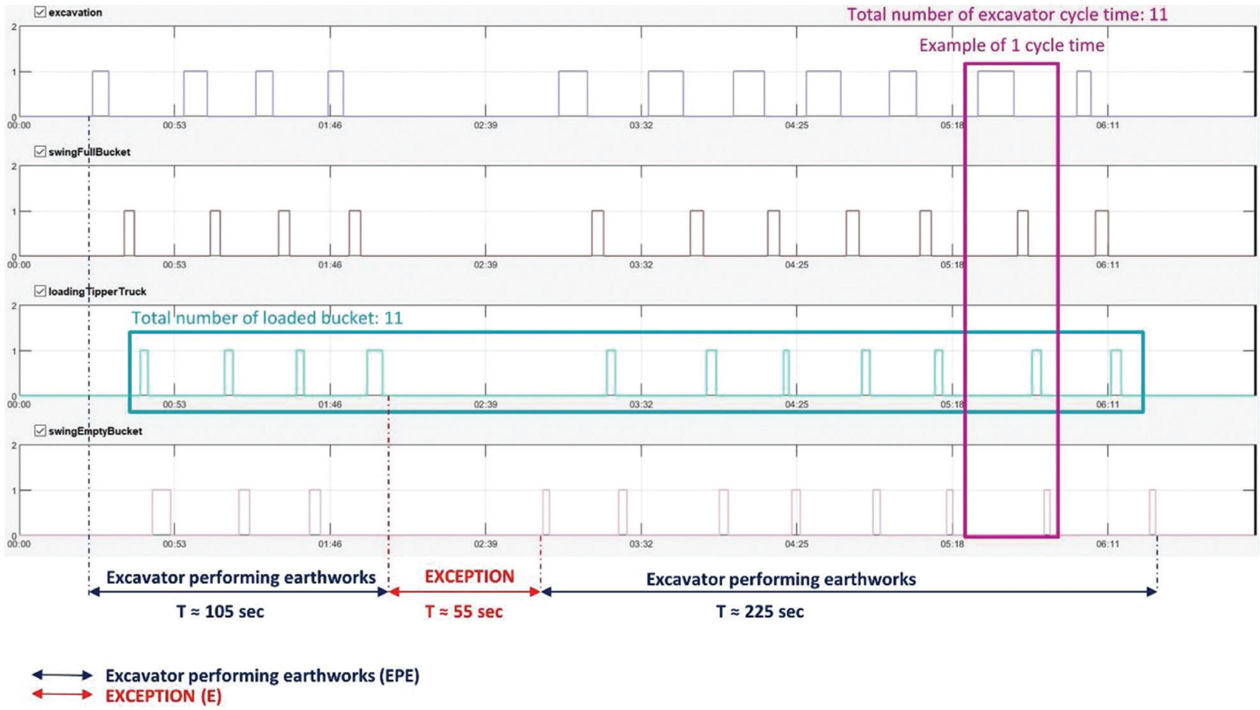

Fig. 4: The results of labeling from the first video.

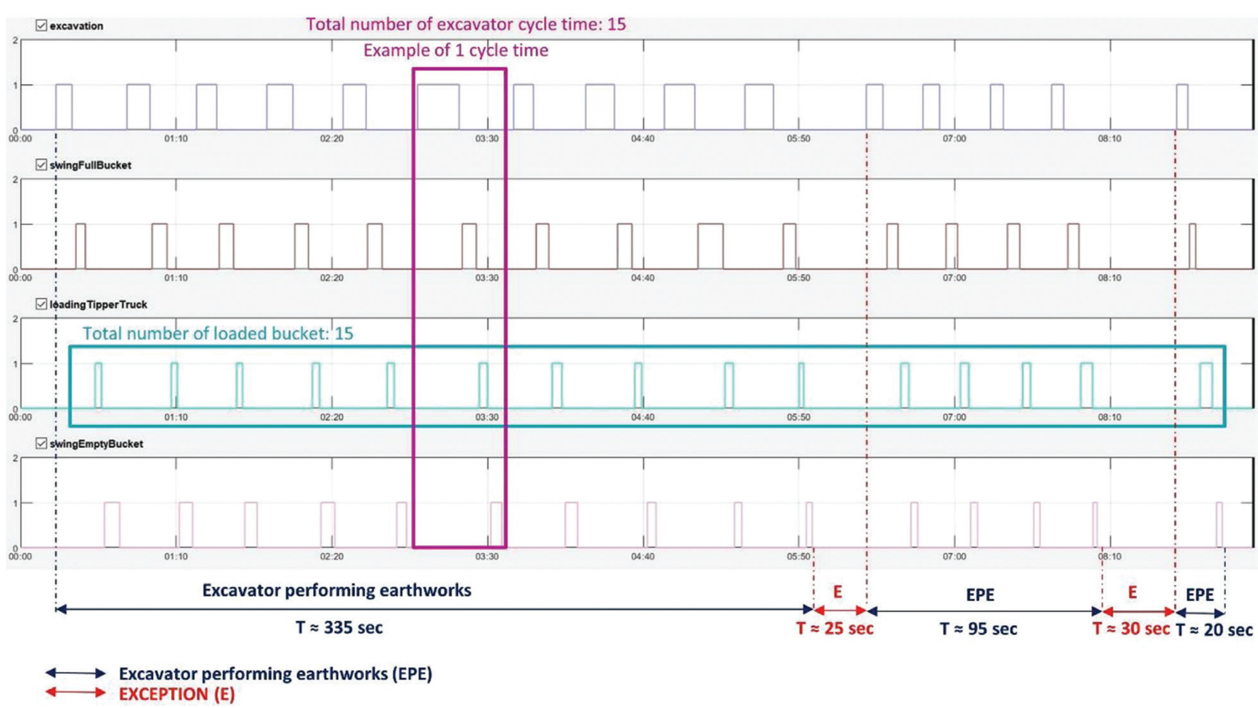

Fig. 5: The results of labeling from the second video.

refer to one of the stages of earthworks. The calculations of the excavator estimated actual productivity and average cycle time are valuable inputs for managing earthworks. However, some exceptions, or deadlocks, while excavator is performing earthworks, as we have seen in both the videos, may be expected. The (un)justified exceptions are extremely important. The way exceptions will be processed or calculated depends on the project manager.

5 Discussion

This paper proposes an SRF for quick and practical estimation of excavator actual, maximum possible, productivity and cycle time at a construction site during earthworks, using video analysis of recorded videos. The benefits of using a video camera to record earthworks are reflected in visual reproductions of the real circumstances in which heavy construction machines operated at a construction site during recording. Furthermore, when a video camera is used, heavy machines can be recorded from a great distance. Also, the use of a video camera does not interfere the normal operation of the heavy machines and does not require installation of some specific tools on the observed heavy machines. By visual insight into real, realistic, and true situations at a construction site, the possibilities of inaccurate interpretation of the collected data is reduced. However, it is necessary to find a position that allows a clear and noticeable representation of earthworks operations. Overlapping of heavy construction machines in videos can cause incorrect data generation in video analysis. Adverse weather conditions, such as rain, fog, snow, or strong wind, as well as conditions with poor natural or artificial lighting, result in videos with reduced sharpness or blurry content. Recorded videos should be with appropriate quality and length, as the computer configuration, as well as advanced computer program for video analysis, may limit the possibilities of video analysis. Therefore, one should be extra careful when selecting the subject to be recorded, as well as during recording.

By applying SRF it is possible to notice, calculate, or estimate some earthworks information, such as total duration of loading materials into a tipper truck, the number of excavator buckets used to load a tipper truck, excavator actual productivity, duration of excavator cycle time. Also, the application of SRF can serve as a clear and transparent base for monitoring and controlling of earthworks if some of the excavator working activities are (suddenly) justifiably or unjustifiably longer than usual. This type of research framework can also be applied for a tipper truck in cases when the dump site is at a same construction site, close to the excavation (Šopić and Vukomanović 2019). Label session can be also used in the form of a training data set (i.e. ground truth label data) to train some kind of object detector. Object detector can be a machine or a DL detector, depending on the training functions (algorithms) used.

In cases of SFR application to only one of the stages of earthworks, the simplicity of SRF is also reflected in the number of videos that needs to be analyzed from the construction site. As previously mentioned, the application of SRF implies an estimate of the excavator actual, maximum possible, productivity and cycle time. Therefore, at the construction site where good operation conditions prevail and with excellent use of working time of good shape and well-maintained machines there is no need to analyze more than a couple of videos. The first one is to estimate the excavator actual, maximum possible, productivity and cycle time, and the second one is to confirm the estimates from the first video. Selected videos should visualize a characteristic and common scene from an observed construction site. In cases where two consecutive measurements show significantly deviating values, it is recommended to analyze more than two videos, using an arithmetic mean or some other statistical data processing. In case of poor working conditions at the construction site, it is recommended to better organize the working conditions, as much as possible, and then to apply SRF.

Applying SRF at the observed construction site in Rijeka, it was calculated that the estimate of the excavator actual, maximum possible, productivity is 108 m^3/h. Using the methodology for calculating the excavator planned work efficiency described in Tijanić et al. (2019) research, it can be calculated that the excavator planned work efficiency for the same conditions would be less than 60 m^3/h. Difference between planned work efficiency and actual productivity indicates the importance of estimating actual productivity.

Although SRF presents practical, simple, and quick research framework for estimates of excavator actual, maximum possible, productivity and cycle time, further research and improvements to SRF are necessary. Improvements should be geared toward integrating SRF with non-vision-based technologies (such as GPS, RFID, accelerometers, sensors). Also, improvements should be directed toward developing a fully automated and autonomous system that will be able to first detect the heavy machine, then recognize the machine's working activities, and finally estimate its actual productivity.

6 Conclusion

This paper proposes an SRF. The main contributions of SRF are explicit values of an excavator actual, maximum possible, productivity and cycle time related to one of the stages of earthworks at a construction site. SRF should be applied at the beginning of earthworks and after any significant change in performance of earthworks. With an insight into the productivity it is possible to estimate the required time and cost of the activity.

In cases of continuous (daily) application of SRF, a clear and transparent base for monitoring and control of earthworks can be obtained. In good operating conditions, continuous application means the daily recording of five to ten videos and the analysis of two to five videos during the performance of earthworks. The main focus of continuous application of SRF is to detect if some excavator working activity is (suddenly) justifiably or unjustifiably longer than usual and exceptions, or deadlocks, in working activities of the excavator. Continuous application of SRF enables enhanced supervision of works. Early detection of an unusual work performance offers the possibility to implement appropriate and corrective measures in time. Therefore, with continuous insight into productivity and deadlocks, it is possible to track the dynamics of performance and timely detect unfavorable and unacceptable actions. However, if SRF is applied continuously, care should be taken to ensure that SRF does not become time-consuming.

Acknowledgments

The authors thank everyone who participated in creating the videos from the observed construction site in Rijeka. This research has been fully supported by the University of Rijeka under the project number uniri-pr-tehnic-19-18 and uniri-tehnic-18-125.

References

Alshibani, A. (2018). Automation of measuring actual productivity of earthwork in urban area, a case study from Montreal. *Buildings, 8*(12), 178.

Alshibani, A., & Moselhi, O. (2016). Productivity based method for forecasting cost & time of earthmoving operations using sampling GPS data. *Journal of Information Technology in Construction (ITcon), 21*(3), pp. 39-56.

Bügler, M., Borrmann, A., Ogunmakin, G., Vela, P. A., & Teizer, J. (2017). Fusion of photogrammetry and video analysis for productivity assessment of earthwork processes. *Computer-Aided Civil and Infrastructure Engineering, 32*(2), pp. 107-123.

Chen, C., Zhu, Z., & Hammad, A. (2020). Automated excavators activity recognition and productivity analysis from construction site surveillance videos. *Automation in Construction, 110*, p. 10304.

Chi, S., & Caldas, C. H. (2011). Automated object identification using optical video cameras on construction sites. *Computer-Aided Civil and Infrastructure Engineering, 26*(5), pp. 368-380.

Deisenroth, M. P., Faisal, A. A., & Ong, C. S. (2020). "Introduction and Motivation" in pre-publication version *Mathematics for Machine Learning*, Cambridge University Press, pp. 11-16.

Deng, L., & Yu, D. (2014). Deep learning: Methods and applications. *Foundations and Trends® in Signal Processing, 7*(3-4), pp. 197-387.

Géron, A. (2017). Multi-Layer Perceptron and Backpropagation. In: Adams, N. (ed.), *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, O'Reilly Media, USA, pp. 326-330.

Ghorbani, M., Bahaghighat, M., Xin, Q., & Özen, F. (2020). ConvLSTMConv network: A deep learning approach for sentiment analysis in cloud computing. *Journal of Cloud Computing, 9*(1), pp. 1-12.

Golparvar-Fard, M., Heydarian, A., & Niebles, J. C. (2013). Vision-based action recognition of earthmoving equipment using spatio-temporal features and support vector machine classifiers. *Advanced Engineering Informatics*, 27(4), pp. 652-663.

Gong, J., & Caldas, C. H. (2011). An object recognition, tracking, and contextual reasoning-based video interpretation method for rapid productivity analysis of construction operations. *Automation in Construction*, 20(8), pp. 1211-1226.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Introduction. In: Dietterich, T. (ed.), *Deep Learning*, The MIT Press, Cambridge, Massachusetts, London, England, pp. 1-28.

Hola, B., & Schabowicz, K. (2010). Estimation of earthworks execution time cost by means of artificial neural networks. *Automation in Construction, 19*(5), pp. 570-579.

Ibrahim, M., & Moselhi, O. (2014). Automated productivity assessment of earthmoving operations. *Journal of Information Technology in Construction (ITcon), 19*(9), pp. 169-184.

Kalal, Z., Mikolajczyk, K., & Matas, J. (2011). Tracking-learning-detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 34*(7), pp. 1409-1422.

Kim, H., Bang, S., Jeong, H., Ham, Y., and Kim, H. (2018a). Integration of imaging and simulation for earthmoving productivity analysis. In: *ISARC, Proceedings of the International Symposium on Automation and Robotics in Construction in Berlin, Germany, 2018,* International Association for Automation and Robotics in Construction (IAARC) Publications, Vol. 1, pp. 704-707.

Kim, H., Elhamim, B., Jeong, H., Kim, C., & Kim, H. (2014). On-site safety management using image processing and fuzzy inference. *In Computing in Civil and Building Engineering, 2014*, pp. 1013-1020.

Kim, H., Kim, H., Hong, Y. W., & Byun, H. (2018b). Detecting construction equipment using a region-based fully convolutional network and transfer learning. *Journal of Computing in Civil Engineering, 32*(2), pp. 04017082.

Kim, J., & Chi, S. (2019). Action recognition of earthmoving excavators based on sequential pattern analysis of visual features and operation cycles. *Automation in Construction, 104*, pp. 255-264.

Kim, J., & Chi, S. (2020). Multi-camera vision-based productivity monitoring of earthmoving operations. *Automation in Construction, 112*, pp. 10312.

Kim, J., Chi, S., & Seo, J. (2018). Interaction analysis for vision-based activity identification of earthmoving excavators and dump trucks. *Automation in Construction, 87*, pp. 297-308.

Kim, J., Ham, Y., Chung, Y., & Chi, S. (2019). Systematic camera placement framework for operation-level visual monitoring on

construction jobsites. *Journal of Construction Engineering and Management, 145*(4), pp. 04019019.
Kovačić, B. (2013). Uvod u Matlab. In: *Matematički alati u elektrotehnici*, udžbenik, Tehničko veleučilište u Zagrebu, pp. 6-16.
Liu, Q., Feng, C., Song, Z., Louis, J., & Zhou, J. (2019). Deep learning model comparison for vision-based classification of full/empty-load trucks in earthmoving operations. *Applied Sciences, 9*(22), p. 4871.
Maini, V., & Sabri, S. (2017). Neural networks & deep learning. In: Maini, S. (ed.), *Machine Learning for Humans*, e-book, pp. 68-80, Available at https://everythingcomputerscience.com/books/Machine%20Learning%20for%20Humans.pdf [accessed 11 March, 2020].
MathWorks. Available at https://www.mathworks.com/products/matlab.html [accessed 11 March, 2020].
Mawdesley, M. J., Al-Jibouri, S. H., Askew, W. H., & Patterson, D. E. (2002). A model for the automated generation of earthwork planning activities. *Construction Innovation, 2*(4), pp. 249-268.
Memarzadeh, M., Golparvar-Fard, M., & Niebles, J. C. (2013). Automated 2D detection of construction equipment and workers from site video streams using histograms of oriented gradients and colors. *Automation in Construction, 32*, pp. 24-37.
Microsoft. Available at https://visualstudio.microsoft.com/ [accessed 11 March, 2020].
Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2018). Introduction. In: Bach, F. (ed.), *Foundations of Machine Learning*, The MIT Press, Cambridge, Massachusetts, London, England, pp. 1-8.
Montaser, A., & Moselhi, O. (2012). RFID+ for tracking earthmoving operations. In: Cai, H. (ed.), *Construction Research Congress 2012: Construction Challenges in a Flat World in West Lafayette, Indiana, 2012*. Americal Society of Civil Engineering (ASCE) Publication, Cdr edition (May 22, 2012), pp. 1011-1020.
Montaser, A., & Moselhi, O. (2013). Tracking scraper-pusher fleet operations using wireless technologies. In*: 4th Construction Specialty Conference Proceedings of the International Conference in Montreal, Quebec, 2013,* Canadian Society for Civil Engineering Annual Conference (CSCE 2013*)*, CON-073-1 – CON-073-10.
Montaser, A., & Moselhi, O. (2014). Truck+ for earthmoving operations. *Journal of Information Technology in Construction (ITcon), 19*(25), pp. 412-433.
Montaser, A., Bakry, I., Alshibani, A., & Moselhi, O. (2012). Estimating productivity of earthmoving operations using spatial technologies. *Canadian Journal of Civil Engineering, 39*(9), pp. 1072-1082.
Ndekugri, I., & Mcdonnell, B. (1999). Differing site conditions risks: A FIDIC/engineering and construction contract comparison. *Engineering Construction and Architectural Management, 6*(2), pp.177-187.
Python Software Foundation. Available at https://www.python.org [accessed 11 March, 2020].
Radujković, M., Burcar Dunović, I., Dolaček Alduk, Z., Nahod, M.M., & Vukomanović, M. (2015). Uvod u organizaciju građenja. In: Radujković, M. (ed.), *Organizacija građenja*, Sveučilište u Zagrebu, Građevinski fakultet, Zagreb, pp. 15-55.
Rezazadeh Azar, E. (2016). Construction equipment identification using marker-based recognition and an active zoom camera. *Journal of Computing in Civil Engineering, 30*(3), p. 04015033.
Rezazadeh Azar, E., & McCabe, B. (2012). Automated visual recognition of dump trucks in construction videos. *Journal of Computing in Civil Engineering, 26*(6), pp. 769-781.
Rezazadeh Azar, E., Dickinson, S., & McCabe, B. (2013). Server-customer interaction tracker: Computer vision–based system to estimate dirt-loading cycles. *Journal of Construction Engineering and Management, 139*(7), pp. 785-794.
Roberts, D., & Golparvar-Fard, M. (2019). End-to-end vision-based detection, tracking and activity analysis of earthmoving equipment filmed at ground level. *Automation in Construction, 105*, p. 102811.
Salem, A., Salah, A., Ibrahim, M., & Moselhi, O. (2017). Study of factors influencing productivity of hauling equipment in earthmoving projects using fuzzy set theory. *International Journal of Innovation, Management and Technology, 8*(2), pp. 151-154.
Šopić, M., & Vukomanović, M. (2019). Video analiza praćenja kretanja i rada kamiona kipera na gradilištu. In: *Zajednički temelji 2019 – Sedmi skup mladih istraživača iz područja građevinarstva i srodnih tehničkih znanosti proceedings of the conference in Rijeka, Hrvatska, 2019,* pp. 107-113.
Šopić, M., Vukomanović, M., & Car-Pušić, D. (2018). Značaj vizualnih tehnologija za praćenje progresa i procjene produktivnosti zemljanih radova. *e-Zbornik, elektronički zbornik radova Građevinskog fakulteta Sveučilišta u Mostaru, 8*(15), pp. 1-9.
Tijanić, K., Šopić, M., Marović, I., & Car-Pušić, D. (2019). Analysis of the construction machinery work efficiency as a factor of the earthworks sustainability. In: *IOP Conference Series: Earth and Environmental Science*, *IOP Publishing, 222*(1), p. 012009.
TLD Vision cooperation. Available at http://www.tldvision.com/ [accessed 3 March, 2020].
Wehle, H. (2017). Machine learning, deep learning, and AI: What's the difference? In: *International Conference on Data scientist innovation day in Bruxelles*, Belgium, 2017.
Xiao, B., & Zhu, Z. (2018). Two-dimensional visual tracking in construction scenarios: A comparative study. *Journal of Computing in Civil Engineering, 32*(3), p. 04018006.
Zou, J., & Kim, H. (2007). Using hue, saturation, and value color space for hydraulic excavator idle time analysis. *Journal of Computing in Civil Engineering, 21*(4), pp. 238-246.