

# Realization of Autonomous Detection, Positioning and Angle Estimation of Harvested Logs

Songyu Li, Håkan Lideskog

## Abstract

To further develop forest production, higher automation of forest operations is required. Such endeavour promotes research on unmanned forest machines. Designing unmanned forest machines that exercise forwarding requires an understanding of positioning and angle estimations of logs after cutting and delimiting have been conducted, as support for subsequent crane loading work. This study aims to improve the automation of the forwarding operation and presents a system to realize real-time automatic detection, positioning, and angle estimation of harvested logs implemented on an existing unmanned forest machine experimental platform from the AORO (Arctic Off-Road Robotics) Lab. This system uses ROS as the underlying software architecture and a Zed2 camera and NVIDIA JETSON AGX XAVIER as the imaging sensor and computing platform, respectively, utilizing the YOLOv3 algorithm for real-time object detection. Moreover, the study combines the processing of depth data and depth to spatial transform to realize the calculation of the relative location of the target log related to the camera. On this basis, the angle estimation of the target log is further realized by image processing and color analysis. Finally, the absolute position and log angles are determined by the spatial coordinate transformation of the relative position data. This system was tested and validated using a pre-trained log detector for birch with a mean average precision (mAP) of 80.51%. Log positioning mean error did not exceed 0.27 m and the angle estimation mean error was less than 3 degrees during the tests. This log pose estimation method could encompass one important part of automated forwarding operations.

*Keywords:* log detection; autonomous forwarding; log grasping

## 1. Introduction

The demand for wood, an important natural material in our society, has for centuries brought people out to the forest to use tools to fell trees and transport out logs for use, which in broad terms is called »logging«: the process of acquiring wood from the forest. It covers all operations from cutting, processing, converting trees into logs, collecting, loading, and eventually moving them out from the forest.

Cut-to-length (CTL) logging system is a highly mechanized system, which consists of a harvester with a forwarder that work together to complete all the tasks of logging. The harvester is used to fell, delimit and cut the trees into logs, and the forwarder is used to accumulate and transport the logs. The working

form of the CTL logging system was gradually formed since the invention of the harvester in Sweden and Finland in 1972–1973 (Drushka and Konttinen 1997), having become mainstream in Swedish forests in the 1980s and 1990s (Fyrk et al. 1991). Since then, this dual-machine system has remained the same. With the high mechanization of forestry operations in the Nordic region, almost 100% of the current harvesting in Sweden and Finland is carried out by CTL systems (Gellerstedt and Dahlin 1999).

Harvester productivity in final felling is about 24 m<sup>3</sup>/PMH (productive machine hours), while the mean productivity of forwarders is 21.4 m<sup>3</sup>/PMH (Eriksson and Lindroos 2014). Moreover, large forwarders currently on the market carry around 20 tonnes of load (Komatsu 2022, PONSSE 2022), indicating that

the productivity of current forest machines is actually quite high, and as such the insufficient production capacity of the machinery no longer hinders further improvements in production efficiency. However, the problem nowadays is that the machine operators are not able to fully realize the potential of the machinery (Hellström et al. 2009). The machines can work at a fast pace, but all the actions of the machine require the operator to constantly make decisions and operations. At the same time, the working environment of forest operators is challenging, and the workload is heavy. Working under such conditions for the long term can have a serious adverse impact on the health and safety of the personnel (Ahola et al. 2013). Moreover, being a highly productive forest machine operator requires high professionalism; to become skilled in the work requires long-term training. Nevertheless, the operators themselves become the bottlenecks to production efficiency, but such problems can be avoided and production efficiency enhanced by improving the automation level of forest machines.

With ongoing research and continuous development of computer assistance, communication, and sensing, some entry-level automation products have recently appeared on the market. For instance, John Deere became the first forestry machine manufacturer to produce a mature commercial boom-tip control system for forwarders in 2013 (John Deere 2013). This system is called »intelligent boom control« (IBC), which allows the operator to focus on controlling the grapple, or end-effector, rather than the boom joint movements (John Deere 2022a). In 2017, this system was also introduced for harvesters (John Deere 2022b). Meanwhile, PONSSE has launched their active suspension system called »ActiveFrame«, which is a cabin suspension system for eight-wheel machines that can level out the roughness of the terrain and keep the cabin horizontal (PONSSE 2017). Around the same time, Cranab released their intelligent forward crane, called »Cranab Intelligent System« (CIS), comprising built-in sensors in the crane (CRANAB 2015). Moreover, research on further automation of forest machines is already underway and has yielded some results, such as the »Besten med virkeskurir« system, which is a prototype system consisting of a driverless harvester controlled remotely from a manned forwarder. This system was tested in 2006 in experiments conducted by Skogforsk (Skogforsk 2006).

At present, a forwarder needs the operator to control the movement of crane and grapple to manipulate logs and complete the loading process. However, with continuous improvement of the forwarder automation level, possibilities have emerged to assist in the crane

work. For example, researchers have evaluated the use of structured light to realize the reconstruction of log morphology to derive log poses (Yeonchool Park et al. 2011). There is also related research on automatic detection of log grab points (Gietler et al. 2022) and log pose estimation based on image segmentation (Fortin et al. 2022), which rely on powerful deep learning algorithms; a research field that currently undergoes rapid development.

AORO: Arctic Off-Road Robotics Lab, which is a cooperation between Luleå University of Technology, Swedish University of Agricultural Sciences and the Cluster of Forest Technology, launched their unmanned forest machine experimental platform in October 2021, and at that time demonstrated their autonomous forest machine capable of autonomous forwarding (AORO 2021) in a simplified scenario, that is, retrieving and transporting roundwood without human intervention. For such unmanned forwarding systems, a necessary capability is to automatically identify, locate, detect the state of the target logs, and grasp and load them. In the landing, logs then need to be unloaded close to the road for further transport. This comprises the needed operations that autonomous forest machines need to consider. An indispensable part of these operations is the realization of automatic forwarding. Based on this consideration, we have realized, deployed, and tested a log position and angle estimation system.

This paper focuses on the implementation method of a log position and pose estimate function which was used on the AORO machine platform. The design, implementation, testing, and final realization of the entire system is described below. The hardware for the system includes the AORO machine platform and some mature products on the market, with a generic underlying software architecture for deep-learning object detection. With this system, an effective, real-time log pose estimation and positioning function is realized and it can be deployed on unmanned machines for the purpose of lifting harvested logs during e.g. forwarding operations. To that end, the feasibility and shortcomings of its practical application have been explored, and the direction of further improvement of this function has been determined.

## 2. Materials and Methods

The target objects considered in the study are logs with a certain length after delimiting and cutting, while the system function is to realize the detection of the logs, positioning under relatively static conditions as well as estimating their pose relative to an absolute

coordinate system. The study includes several composite solutions developed in order to have the entire real-time system functioning:

- ⇒ an unmanned machine platform with dual GNSS antennas
- ⇒ a vision system based on a Zed2 stereo camera (Zed2, Stereolabs Inc, Montrouge, France) with NVIDIA JETSON AGX XAVIER (Jetson AGX Xavier, NVIDIA, Santa Clara, California, USA)
- ⇒ a series of script files that realize detection, localization, and pose estimation for the target logs.

## 2.1 Hardware Preparation

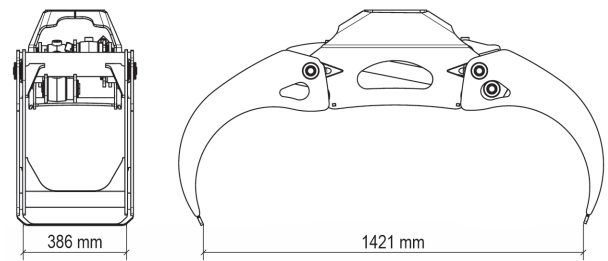
### 2.1.1 Unmanned Forest Machine Experiment Platform

The unmanned forest machine experiment platform used (Fig. 1) is a stage result of a long-term research by AORO aimed at realizing unmanned, autonomous forest operations. For now, the potential of this unmanned forest machine experimental platform in realizing unmanned forwarding has been fully verified.

This machine platform was originally designed and constructed at Luleå University of Technology in collaboration with the Swedish University of Agricultural Sciences. The main purpose of the build is to use it as a platform for in situ forest machine technology research and development, mainly focused on automation. This 10 tonne platform mainly consists of two chassis parts: the front part is the location of the controller, engine and hydraulic pump, and the rear part where auxiliary equipment is attached. Connecting these two parts are an articulated joint manipulated through hydraulic cylinders. The machine is equipped with a 4.5 liter 129 kW diesel engine, which propels



**Fig. 1** Unmanned forest machine experiment platform used in AORO



**Fig. 2** Grapple size parameters

hydraulic pumps that in turn drive four hydraulic hub motors, pendulum arms, articulate joint, and a crane. At the same time, the engine is also connected to the generator, which is responsible for the electric power supply.

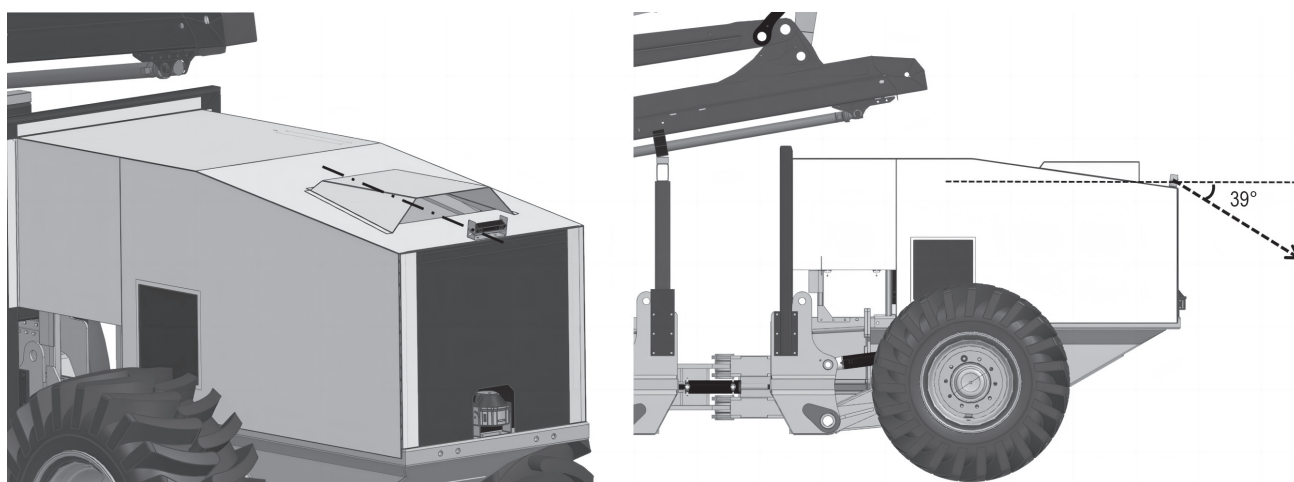
The crane mounted at the back of the machine is supplied by Cranab AB, mounted with a rotator supplied by Indexator and Cranab CR250 grapple. The four first joints slewing, inner boom, outer boom, and telescope are all measured through high precision rotational (slewing, inner and outer boom) and positional (telescope) sensors, allowing for high precision positioning of the crane tip. The size of the grapple being used (Fig. 2.), i.e. the end-effector, will set the constraints on the precision of the crane tip positioning and grapple rotation if the goal of autonomously lifting a harvested log is to be conducted by the crane.

The control units on the machine comprise two main computers that have different jobs to conduct. One is an Ueidaq i/o computer run as a Simulink i/o target on the machine to achieve low-level motion control such as controllers for crane and pendulum arm manipulation. The other is the Nvidia JETSON XAVIER, which handles visual computing and robot localization, onto which a stereo camera and a Leica iCON GPS80 with dual antenna units are connected.

### 2.1.2 Vision System

The vision system is responsible for image perception and image-based analysis and computational processing of the environment. It is composed of two parts: data processing unit and the camera. The Zed2 stereo camera from STEREO LABS was used as the camera to obtain the colour and depth information of the field of view. It is mounted at the top-front of the machine with a slight downward tilt (Fig. 3).

During our tests, there was a downward angle of about 39 degrees in the orientation of the camera, but the orientation of the camera is adjustable. The data from the camera is processed by the NVIDIA JETSON XAVIER, an embedded computing board from



**Fig. 3** Camera mounted on machine front

NVIDIA as the image processing unit on the unmanned forest machine experiment platform. The JETSON is mounted in the computer area on the front part of the machine and connected to the camera by a USB 3.0 cable. Considering the subsequent image processing and computing power, as well as the requirements for image clarity and real-time, the output specification of the camera image was set to 720p, 15FPS. Related parameters for JETSON and the camera are shown in Table 1.

## 2.2 Software Implementation

The JETSON XAVIER minicomputer is running on Ubuntu 18.04 LTS, and the Robot Operating System (ROS) (Quigley et al. 2009) was selected and installed

**Table 1** Technical specifications and running parameters of Zed2 and JETSON

RUNNING PARAMETERS OF Zed2 CAMERA	
Video Mode	720p
FPS	15
Depth Resolution	Native video resolution
Depth Range	0.2 – 20 m <sup>1</sup>
Depth FOV	110° (H) x 70° (V) x 120° (D) max. <sup>1</sup>
TECHNICAL SPECIFICATIONS OF NVIDIA JETSON AGX XAVIER Developer Kit	
GPU	512-core NVIDIA Volta™ GPU with 64 Tensor cores
CPU	8-core ARM® v8.2 64-bit CPU, 8MB L2 + 4MB L3

<sup>1</sup> According to the official product parameters

as a software framework for the system. ROS is a set of open-source software libraries and tools that help users build robot applications. We use ROS as our software architecture because we focus on its highly customizable message definition and transceiver mechanism, with flexible calling and running of user-custom functional scripts. This allows us to complete the reception of the camera signal in ROS and publish output results in a custom data format to be processed in any other part of the ROS infrastructure.

## 2.3 System-Wide Workflow

The vision system uses a stereo camera as imaging sensor and a JETSON as computing unit, which is the core to realize automatic detection, positioning, and angle analysis of target objects. It is part of the entire automated forest machine platform. In Fig. 4 and Fig. 5, the structure of the vision system part and the whole automatic forest machine platform are depicted.

The unmanned forest machine experiment platform is the carrying platform and executing mechanism for verifying the constructed log detection, positioning and angle estimation system. The platform receives the output signal and carries out corresponding actions on the basis of the output signal from the vision system, and realizes the final actions of the end-effector.

## 2.4 Log Detection, Localization and Angle Estimation

The realization of the visual function module is mainly divided into three parts: detection module, positioning module, and angle estimation module. The detection module is used to find the area of the

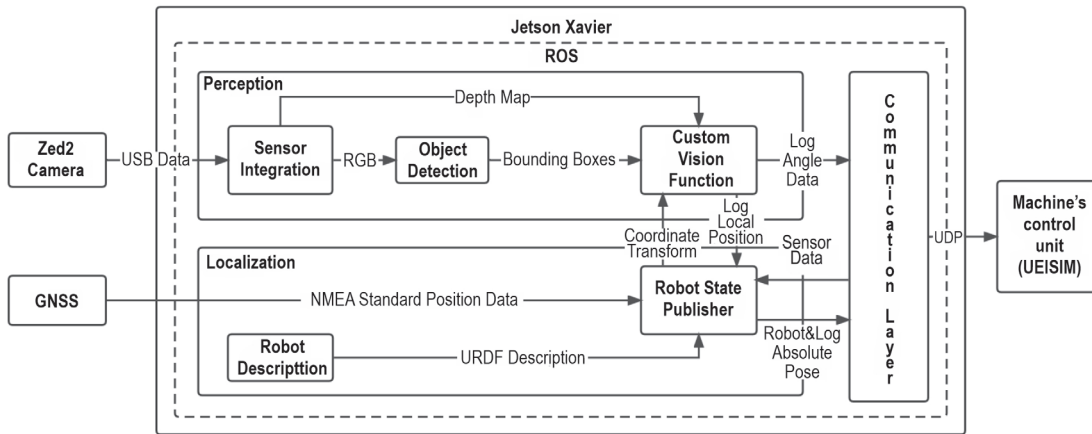


Fig. 4 Structure of the vision system for detecting, locating and estimating the angle and position of target logs, as well as conducting robot localization

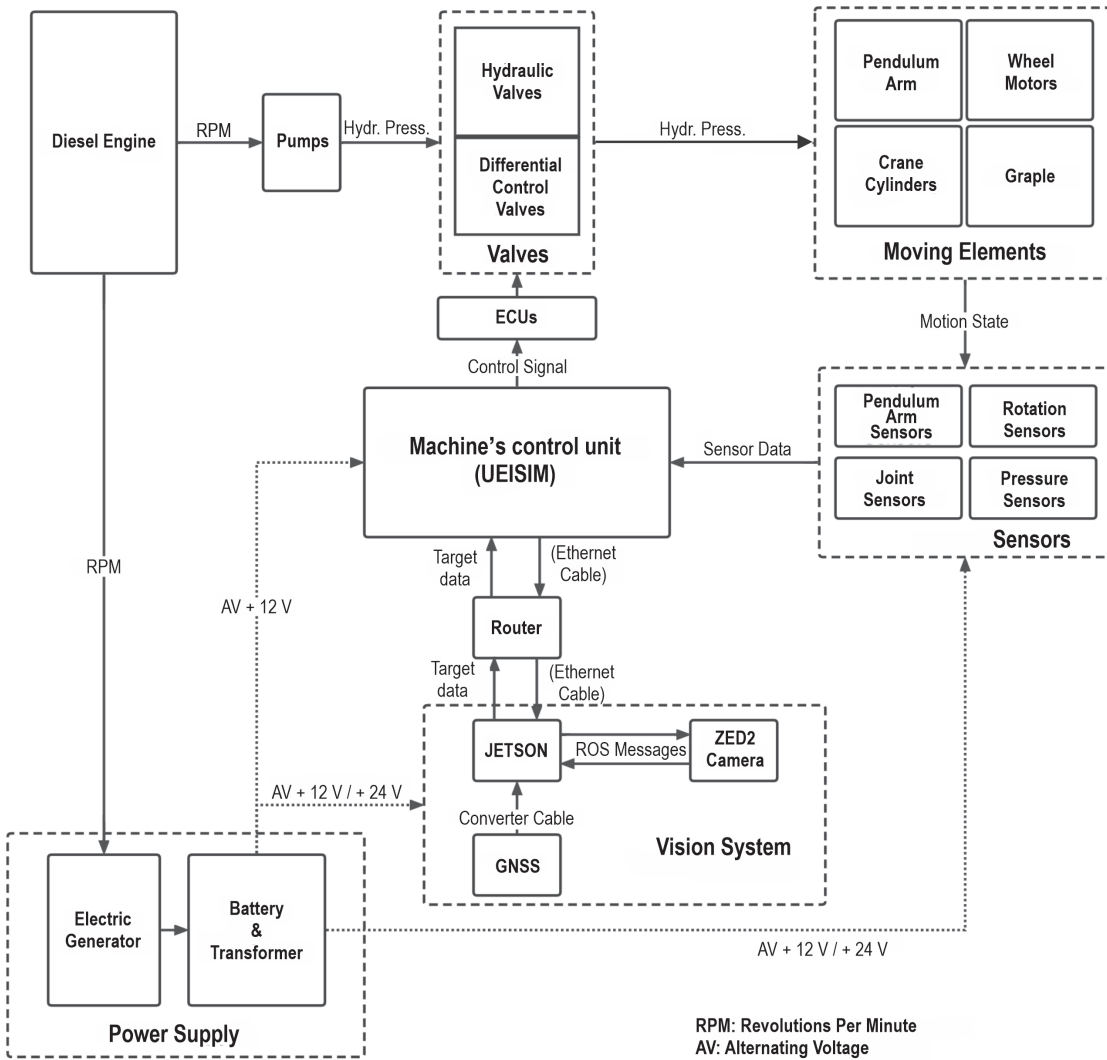


Fig. 5 Architecture of unmanned forest machine experiment platform

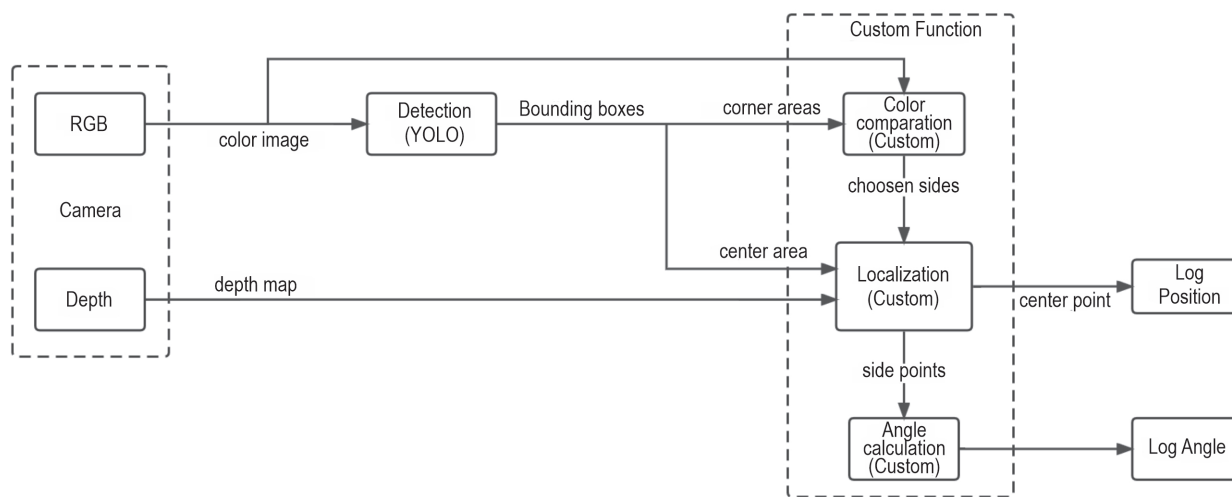


Fig. 6 Realization flow of each function of the vision system

target from the RGB (colour) signal sent from the camera and outputs the position of bounding boxes of target logs. Bounding boxes are an imaginary rectangle that surrounds a detected object in an RGB image and are defined by its corner coordinates. The positioning module calculates the spatial position of the target based on bounding box data from the detection module and the synchronized depth map from the camera. The angle estimation module also calculates the pose of the target log based on the bounding box and the depth map. This process is illustrated in Fig 6.

### 2.4.1 Detection of Logs

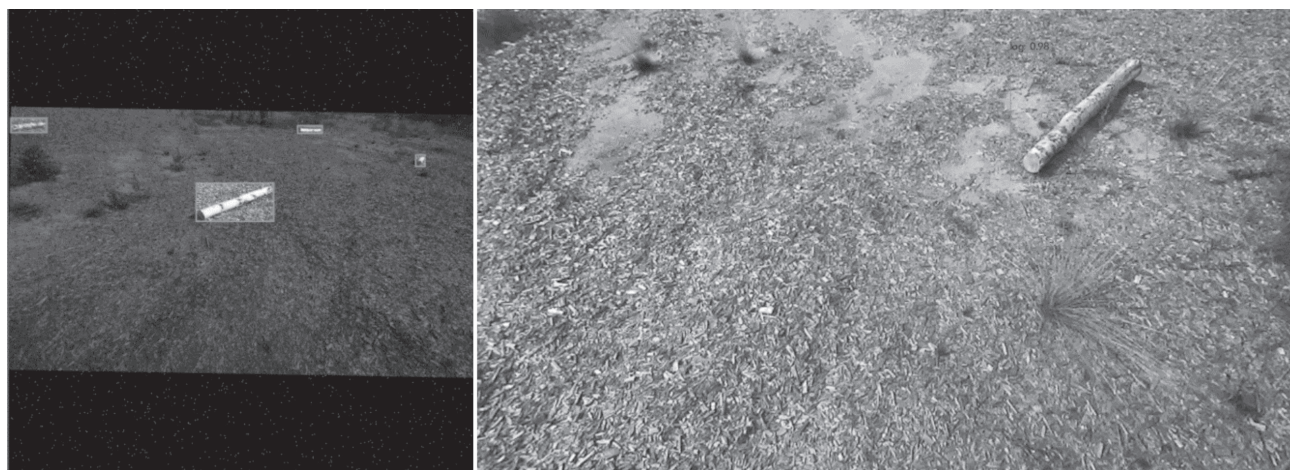
After the camera transmits the video signal to JETSON, the system identifies and determines the location of the target log in the video, and marks the detected target log with a bounding box. The detection algorithm used to obtain the bounding box is not specifically limited to our choice, and can be changed to other relevant algorithms. The YOLO-ROS (Bjelonic 2018) package was introduced into the system to de-

tect target logs in the RGB data stream under the ROS framework. The YOLO-ROS is a transplant of YOLOv3 (Redmon and Farhadi 2018) detection architecture in the ROS environment, which can realize the target detection function of YOLOv3 in ROS. YOLO-ROS, the ROS package of YOLOv3 transplant, is very mature and easy to use. It directly outputs bounding boxes of the detection results in the form of ros topics and, at the same time, meets our test performance needs. Meanwhile, considering the limited computing power of JETSON, computing resources are allocated to deal with spatial location analysis and related calculations. In order to reduce the computational burden of the target detection stage and meet the real-time requirements of detection, we choose to call YOLOv3-Tiny (Adarsh et al. 2020) encapsulated in YOLO-ROS to achieve the target detection function. Further details of the neural network structure can be found in their paper.

The training data set is composed of real photos. In the autumn of 2021, in Västerbotten County, Sweden,

Table 2 Partition of data sets and augmentation

Partition of data sets, number					
Total	Validation set	Initial training set	Initial labeling	Augmented training set	Augmented labeling
1739	171	1568	3504	6272	14,016
Training set Augmentation					
Outputs per picture	Rotation	Saturation	Exposure	Blur	Noise
4	-3° to +3°	-5% to +5%	-5% to +5%	Up to 0.5 px	Up to 2% of pixels



**Fig. 7** A sample of resized and labeled image of augmented dataset (left) and a detection result of log in a picture (right) with detection confidence printed above bounding box

randomly placed birch logs were placed in simplified terrain and were depicted by the Zed2 camera. In total, 1739 photos were obtained, of which 1568 photos were used to make the training set, and 171 photos (about 10%) were retained as the verification set. For the 1568 photos used to make the training set, the logs were manually labeled, processed, resized, and augmented to comprise the full training set. The final training set contains 6272 images with over 14,000 labeled regions. Table 2 shows the data set information.

Fig. 7 shows an example of the augmented data set and an example of a visual detection result.

Considering the number of pictures in the data set prepared, the yolo-tiny network was trained with learning rate 0.001, burn in 1000, Max batch 12,000 and setps 9600, 10,800 as the input parameters. Finally, a detector that achieved a mean average precision (mAP) of 80.51% was used for log detection. YOLO-ROS realizes the function of log detection based on YOLOv3-Tiny by using the weight file obtained by training, and the detection result is released to the ROS framework in real time in the form of ROS messages, outputting coordinates of the corresponding bounding boxes.

Before outputting (in ROS called »publishing«) bounding box data, to improve the validity of the data, the bounding boxes are filtered. Combined with our actual needs in the test, the detection results that are too close, too far away, and too close to the edge of the field of view are removed. Consequently, boxes within the image are required to be more than 30 pixels wide or high, comprise more than 2500 pixels, and be more than 2 pixels from the image border to be considered.

#### 2.4.2 Localization of Logs

After the target log is found, its actual position in space is determined and output. The color and depth signals in the field of view are synchronously obtained from the Zed2 camera, and thus the depth values of desirable pixels in the color image (the bounding box pixel positions) are determined.

Then, to calculate the three-dimensional coordinates corresponding to the two-dimensional pixel points, it is necessary to model the camera according to the pinhole camera model to obtain the corresponding relationship between cartesian and pixel coordinates.

It can be deduced that the relationship between the coordinates of a three-dimensional space point  $[X, Y, Z]$  and the pixel coordinates of its two-dimensional image projection  $[u, v]$  is given by:

$$Z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} [R \quad t] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

Where:

$R$  and  $t$  are the rotation and translation, which relate the world coordinate system to the camera coordinate system

$f_x$  and  $f_y$  denote the focal length of the camera in X-axis and Y-axis

$c_x$  and  $c_y$  denote the center of the camera aperture. It can further be deduced from the above formula that:

$$\begin{bmatrix} X \\ Y \end{bmatrix} = Z \cdot \begin{bmatrix} (u - c_x) / f_x \\ (v - c_y) / f_y \end{bmatrix} \quad (2)$$

In (2), except that the  $X$  and  $Y$  of the target point in the spatial coordinates are unknown, other parameters including  $Z$ , the depth of the pixel, the pixel coordinates  $u$  and  $v$ , and the camera intrinsic parameters  $f_x$ ,  $f_y$ ,  $c_x$  and  $c_y$  are all known (The camera parameters  $f_x$ ,  $f_y$ ,  $c_x$  and  $c_y$  are usually provided by the camera manufacturer). The calculated coordinates are related to the camera position, meaning that forward kinematics can be used to relate the coordinates to any coordinate system on the machine.

Using the pixel coordinate data provided by the bounding box from log detection, the area represented by the bounding box in the depth map is determined. The depth information of this bounded region reflects the spatial characteristics of the target object. Then, according to the above calculation method, combined with the bounds generated by the log detection function, the position corresponding to the central point of the bounding box can be calculated. The location of the center point of the bounding box is used to represent the actual location of the target log. Considering that there is a lot of depth information loss in the original depth image, the original depth map is preprocessed. Linear interpolation is used to fill the missing data in the depth map to ensure the availability of the depth information in the subsequent calculation. At the same time, the average value of the effective depth data within a small range (11×11 pixels) around the center of the box was also used to calculate the position to reduce the influence of fluctuation of the depth data caused by errors in the position calculation process. The visual representation of possible fluctuation errors

and missing information in an unprocessed depth map are shown in Fig 8.

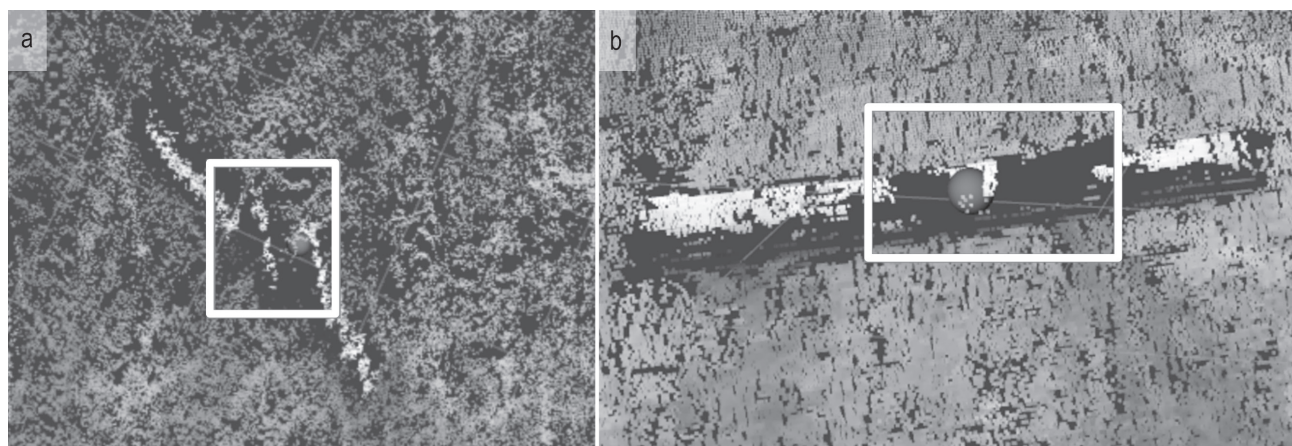
The log position obtained is the position within the coordinate system with the position of the camera as the coordinate origin, based on which this coordinate can be converted into any coordinate information needed by using the coordinate system transformation.

#### 2.4.3 Log Angle Estimation

When the middle position of the target log is represented by a three-dimensional point in the cartesian space, the judgment of the log pose is added to obtain its angle relative to the camera.

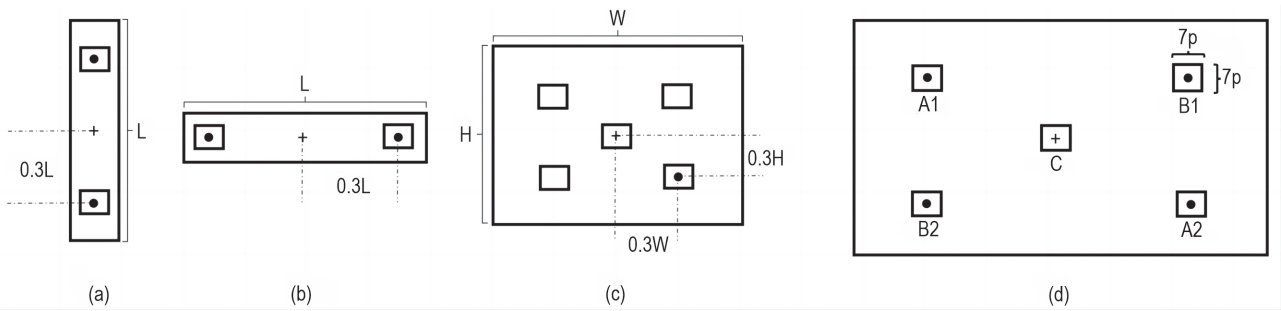
The log angle estimation combines two aspects, the shape of bounding box and the color difference. First, the filtered bounding boxes are divided into three different classes according to the different height-width ratio of the bounding box: vertical type, horizontal type, and the type with no obvious difference between height and width (Fig. 9a, b and c, respectively). In practice, this ratio is set to 5 to 1, that is, if the height is 5 times more or larger than the width, it is classified as vertical type; and contrarily with respect to the horizontal type. Those in between fall into the third. The three different classes of bounding boxes, shown in Fig. 9 (a–c), are marked as type a, type b, or type c.

For the two types a and b of bounding boxes, the center point of the bounding box is taken as the starting point, and points along the  $v$  direction are defined at  $\pm 30\%$  of the length for type a, and similarly along the  $u$  direction for type b, as the center point for constructing a depth calculation area with dimensions of 11×11 pixels. Then, the cartesian position for the two points is calculated in the same way as for the log localization method for both a and b type bounding boxes.



**Fig. 8** Fluctuation (a) and missing depth information (b) in unprocessed depth map





**Fig. 9** Three different types of bounding boxes with chosen points: (a) vertical type where logs are situated entirely vertical in the image, (b) horizontal type where logs are situated entirely horizontal in the image, (c) the type with no obvious difference between height and width, (d) grouping of selected areas in the bounding box: A1, A2 as one group, B1, B2 as another group, C as the center area

For type c bounding boxes, for the determination of the center point of the candidate depth calculation areas, we start from the center point of the bounding box and create areas A1, A2, B1, and B2 (Fig 9. (d)) of size 7x7 pixels, respectively, whose centers are positioned at ±30% of the total height or width from the center point. Then color data are extracted from areas A1, A2, B1, and B2, taking each diagonal area as a pair, where the difference between the color information in the two pairs of diagonal areas and the color information in the central area are evaluated.

In RGB color space, two 7x7 pixel regions of the same group are fused into a new 7x7 pixel image with a weight of 0.5 for each. After that, the two groups of fused images are compared with the central 7x7 pixels image area in CIELAB (also known as the  $L^*a^*b^*$ ) color space (CIE 1976  $L^*A^*B^*$  COLOUR SPACE, ISO/CIE 11664-4:2019(E) ). Now, the difference between the two colors can be judged in the form by comparing the Euclidean distance between the two colors. This difference can be given as:

$$\Delta E^*_{ab} = \sqrt{(L_2^* - L_1^*)^2 + (a_2^* - a_1^*)^2 + (b_2^* - b_1^*)^2} \quad (3)$$

Where:

$(L_1^*, a_1^*, b_1^*)$  and  $(L_2^*, a_2^*, b_2^*)$  are two colors in  $L^*a^*b^*$  color space.

After the comparison of color difference is completed, the group whose color is closer to the central area is selected, whereby the cartesian positions corresponding to the selected areas are calculated in the same way as for the log localization method, and maintaining the method of 11x11 pixels averaging.

After having calculated the cartesian value of the two points, the height between the two points ( $\Delta Y$ ) is disregarded, whereby the angle value for the log is calculated related to the camera X-Z plane. The angle

obtained is in the range of 0 to 180 degrees, and this angle is used as the estimated angle of the target log.

#### 2.4.4 Validity of Log Localization and Angle Estimation

To ensure that the location and angle data can be used for the log grasping operation, it is necessary to judge the effectiveness of the data based on the actual operation after obtaining the log location and angle. In our case, the log to be grasped was approximately regarded as a cylinder about 2 m in length and 25 cm in diameter, and, according to experience, successful grasping requires that the opened grapple is within one meter from the center of gravity, and that the edge of the log should not exceed the grapple gap. During the grasping operation, the position and rotation of the grapple related to the log center and angle are influenced mutually by each other. For example, the closer the grapple center is to the log center, the greater the tolerable rotation error.

When performing log localization and angle estimation, the target log is artificially placed, whereby the midpoint position is measured by CORS network GNSS. The angle is manually measured with respect to the machine. This data is used as the reference data and regarded as the real state of the log, while the positioning and angle estimation results through the system are regarded as test data. The difference between the test data and the reference data is regarded as an error, which is used as a basis for judging the validity of the test data.

#### 2.4.5 Testing

The testing took place in October in Luleå, Norrbotten county, Sweden. At that time, the weather was overcast with light snow, and the selected ground was flat, and the surface was partially covered by ice and snow. The machine and the target logs were placed in a pre-marked fixed position. The logs were cut to

about 2 meters with the stem diameter of about 25 centimeters. The tree species was birch.

During the testing, a fixed position is manually selected as the origin of the system coordinates, with the due east direction defined as the coordinate  $X$  axis, and the due north direction defined as the coordinate  $Y$  axis. The position and orientation of the machine are always fixed, while the target log is placed at fixed positions in the area in front of the machine. The position of the machine, the coordinate origin, and the log position are all measured by the vehicle-mounted GNSS system at once to ensure that the relative position between the three points will not produce errors through multiple measurements. To evaluate the general performance of the developed system, a total of four rounds of tests were conducted, with the center position of the log kept fixed. The position of the log relative to the camera is:

- ⇒ horizontal
- ⇒ diagonally near left and far right
- ⇒ vertical
- ⇒ diagonally near right and far left.

The angle between the log and the machine was approximately 0 (180) degrees, 45 degrees, 90 degrees, and 135 degrees. The localization and angle estimation results of the log under the corresponding tests were recorded separately.

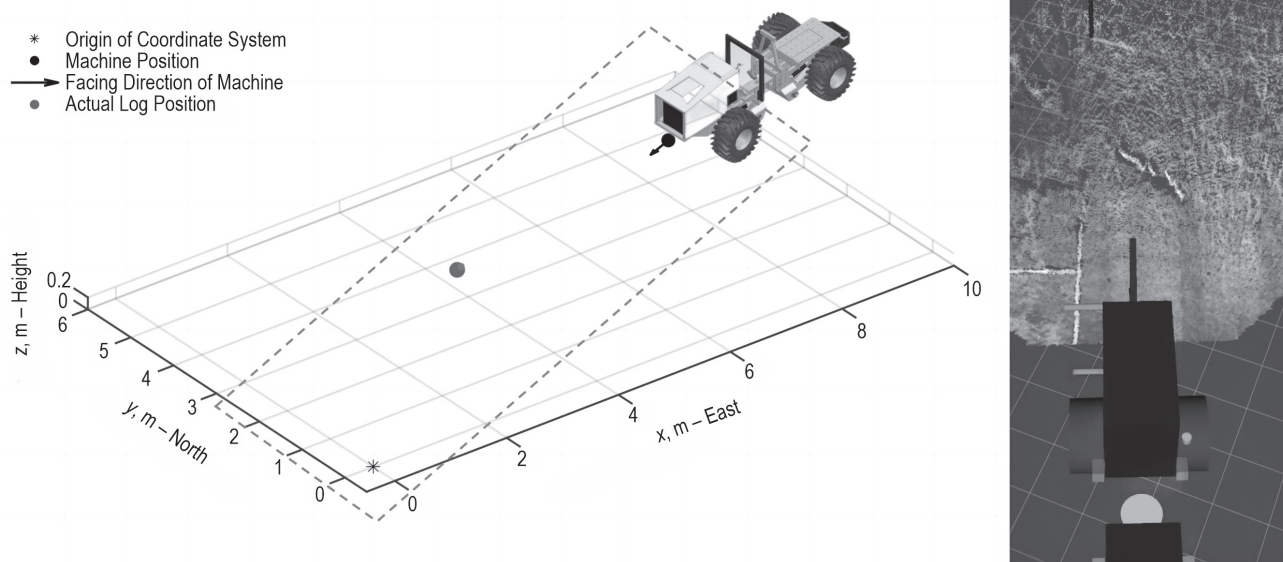
### 3. Results

The results of the test are published in ROS in the form of data stream, and the publishing frequency reaches around 1 Hz. In order to effectively analyze the results, the data generated in the test process were intercepted and saved to be analyzed in the form of images.

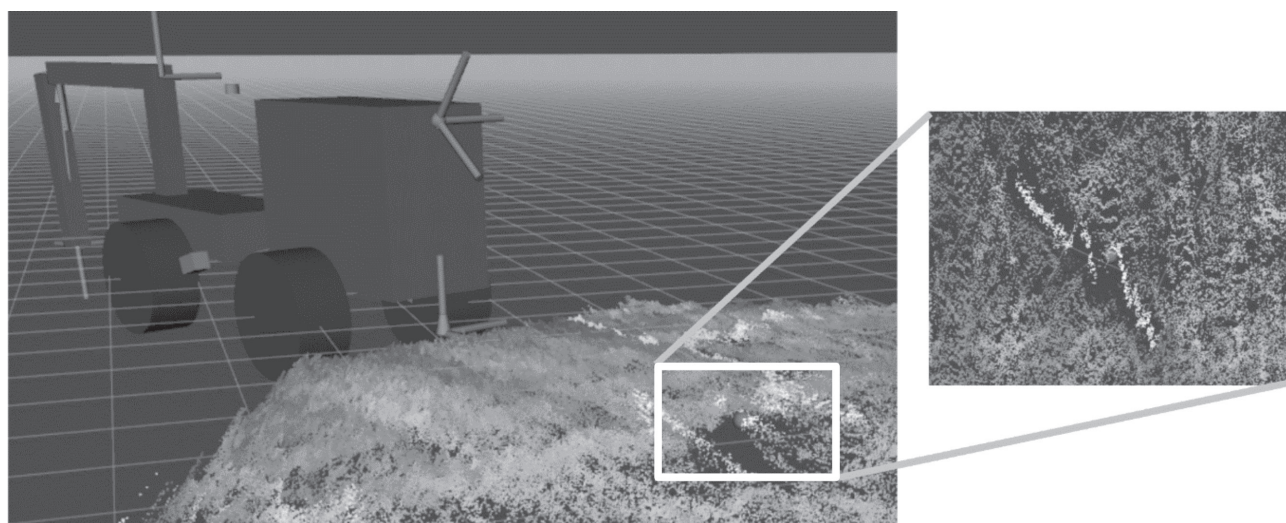
#### 3.1 Localization of Target Log

In four rounds of testing, the machine is allowed to run automatically and record data for a fixed period of time in each round, resulting in four sets of data, each containing 21 measurement positions. All the measured results obtained in each round of measurement and the average value of all the results of each round of measurement are compared with the actual position measured by GNSS.

According to the implementation principle of the positioning function, the error of the position obtained by the final measurement is not only caused by the selection of pixels determined by the algorithm but also by the GNSS system and the depth data of the stereo camera. According to the camera datasheet, the error of depth measured by the camera will increase with the increase of detection distance. It can be considered that before the error caused by GNSS begins to be greater than the depth error caused by the measurement of the stereo camera, within the working range of the camera, the closer the detection object is,



**Fig. 10** Schematic figure and visualization of actual working state to show the position of the origin set during test, actual position of the target log, position of the machine (projection position from the front camera to the ground) and orientation of the machine



**Fig. 11** An example visualizing the positioning result of the target log. The sphere depicts detected log position

the more accurate the result is. However, the actual working state of the system determines that the distance between the target log and the machine should not be too close to each other, so it is necessary to leave enough distance for operation. Finally, combined with the actual needs of the demonstration, the distance between the log and the vehicle was set to about 5 meters during the test.

**Table 3** Comparison of all measured positions of the log in each round with the actual position along  $X$ ,  $Y$ ,  $Z$  axes in pre-set ground coordinate system, and mean and standard error deviation

Axis		$x$	$y$	$z$
Actual log position		4.208	3.528	0.125
Round 1	Mean measured log position, m	4.130	3.617	0.091
	$Means_{error}$ m	-0.078	0.089	-0.034
	$SD_{error}$ m	0.037	0.019	0.015
Round 2	Mean measured log position, m	4.254	3.645	0.147
	$Means_{error}$ m	0.046	0.117	0.022
	$SD_{error}$ m	0.141	0.013	0.058
Round 3	Mean measured log position, m	4.402	3.360	0.114
	$Means_{error}$ m	0.194	-0.168	80.011
	$SD_{error}$ m	0.056	0.035	0.027
Round 4	Mean measured log position, m	4.011	3.443	0.018
	$Means_{error}$ m	-0.197	-0.085	-0.107
	$SD_{error}$ m	0.024	0.008	0.009

Given the reach of a forwarder crane and the operating principle of a forwarder, the approximate distance between the camera and the log was set to 5 meters.

The position of the origin set during the test, the actual position of the target log, the position of the machine (the projection position from the camera to the ground), and the orientation of the machine are shown in Fig. 10.

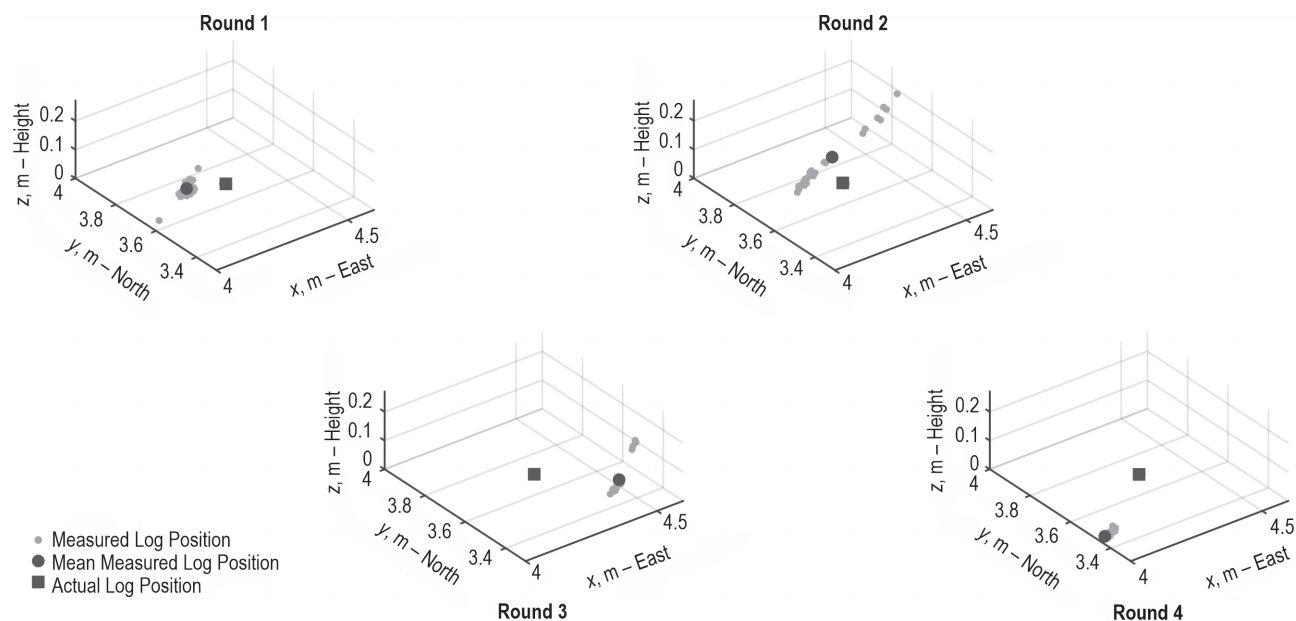
An image example after the test process is visualized as shown in Fig. 11. The system will however not generate and output similar images in the actual working process due to unnecessary overhead.

Table 3 shows the measured and actual positions of the log, and the mean and standard deviation ( $SD$ ) of the errors along the  $X$ ,  $Y$ , and  $Z$  axes, for four round tests with 21 samples for each round with the angle at approximately 0 (180), 45, 90, and 135 degrees between the target logs and the machine.

The distribution of all the measured points obtained in each round of testing, the mean value of the measured positions, and the comparison with the actual positions are shown in Fig. 12.

Table 4 shows the average value and standard deviation of the difference between the 21 measured positions and the actual position recorded in each of the four rounds of measurement.

The test results show that regardless of the relative angle between target log and machine, the distance between the position of the target log measured by the system and the actual position is maintained within 0.26 m. Except for the standard deviation of the first



**Fig. 12** Distribution of all measurement points obtained in four rounds of tests, mean value of measurement positions, and comparison with actual positions

round at 0.078, the standard deviation of the second, third, and fourth rounds of test results all fell within 0.03, with little difference.

### 3.2 Angle Estimation of Target Log

For the angle estimation between machine and target log in four rounds of tests, 26 estimated angles were recorded for each round. By comparing with the relative angle between the target log and the machine obtained by manual measurement, the performance of the angle estimation function of the system can be visually presented. Table 5, together with Fig. 13, shows the actual angle of the target log compared to the estimated angles and the mean over 4 rounds of testing.

Table 6 shows the average value and standard deviation of the difference between the 26 estimated

angles and the actual angle recorded in each of the four rounds of measurement, to show the difference between the estimated angles and the actual angle and the dispersion of this difference.

**Table 5** The mean of estimated angle and actual angle between target log and machine for four rounds (The actual angle is only accurate to an integer because of the manual measurement of ground truth)

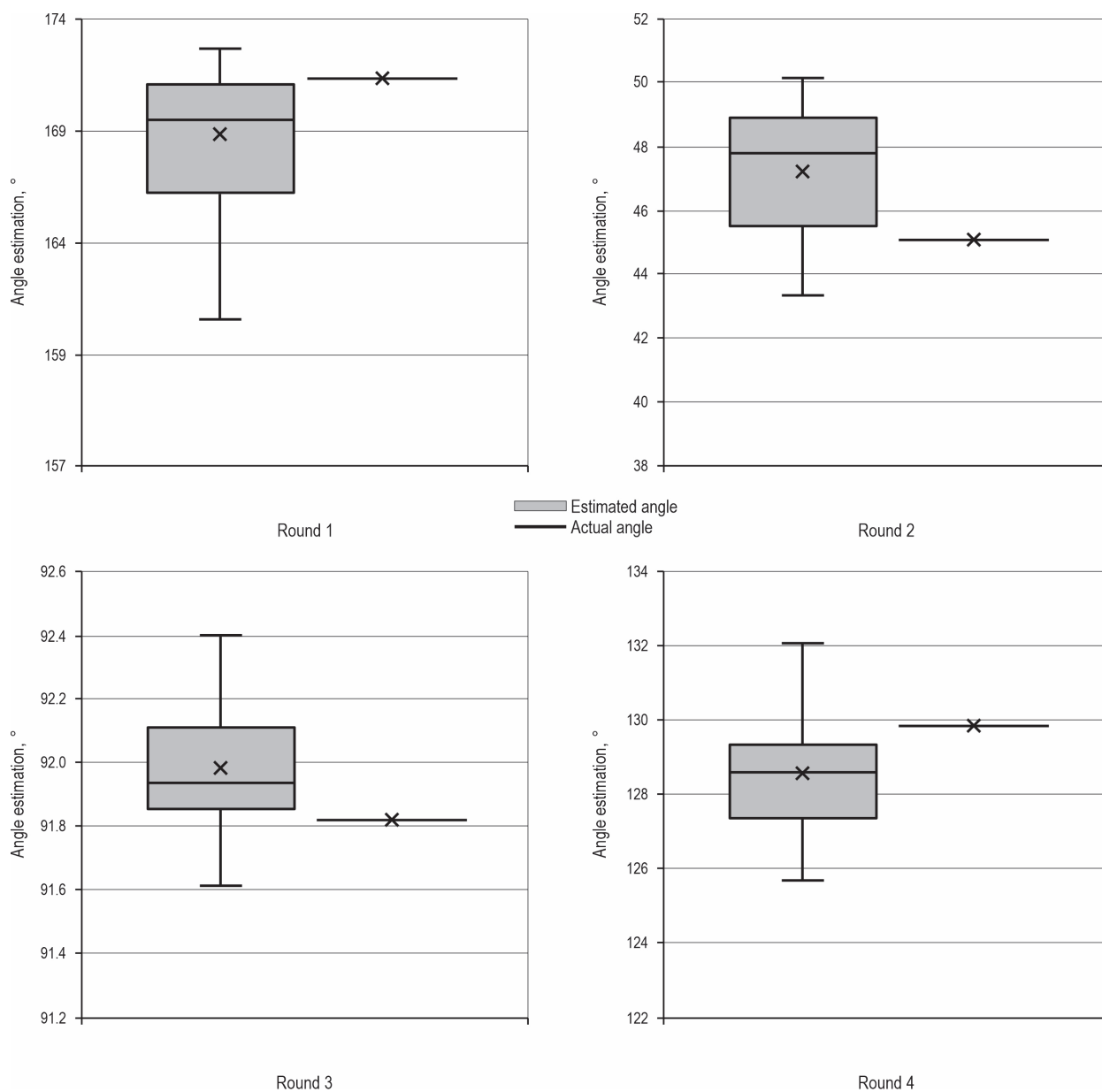
ROUND	Mean estimated angle, °	Actual angle, °
1	168.8	171
2	47.2	45
3	91.9	92
4	128.6	130

**Table 4** Euclidean distances from the actual position of the target log to all measured positions are presented as means and standard deviations (SD) for each round

ROUND	Means <sub>error</sub> , m	SD <sub>error</sub> , m
1	0.180	0.078
2	0.128	0.026
3	0.265	0.029
4	0.241	0.025

**Table 6** Difference between the estimated angle and actual angle presented as means and standard deviations (SD) for all angle estimations in each round

ROUND	Means <sub>error</sub> , °	SD <sub>error</sub> , °
1	2.478	1.46
2	2.83	2.69
3	0.19	0.17
4	1.66	1.03



**Fig. 13** Angle estimation and actual angle between target log and machine for four rounds

The test results show that the mean value difference between the angle estimation and the actual angle is no more than 3 degrees, while the maximum error of a single angle estimation result falls below 5 degrees. Thus, the angle estimation is stable and the data dispersion is low.

#### 4. Discussion and Conclusions

The proposed automatic log detection, localization, and angle estimation system has demonstrated the feasibility of working in conjunction with an auto-

nomous forest machine in our case. The test results show that the difference between the measured position of the log and the actual position of the log center can be kept within the range of log diameter, and is far less than the grasping range. At the same time, the distribution of measurement positions is relatively concentrated, and there is no extreme deviation. The average value of the angle estimation result is used as the output, which can effectively guide the rotation of the grapple in the process of grasping the target log. Combined with the positioning of the target log, it can be considered that the system can effectively guide the

grapple of the crane to successfully grasp the target log.

The successful construction of this system provides a new idea for the realization of unmanned forest machine vision system, and shows that it could be possible to use the vision system to guide an unmanned forest machine to carry out automatic forwarding operations. However, the system is only a preliminary implementation, and there are still some limitations and shortcomings. Although it can work well under preset conditions, the working conditions and the reality still have a gap. For example, the neural network was trained on birch logs of a certain size, so the detector itself is probably limited to those conditions. At the same time, because of the hardware system performance, the results are often difficult to achieve ideally. For example, the depth signal obtained by the stereo camera suffers from missing data and the accuracy is rather low, while being easily affected by changing environmental conditions, which causes fluctuations in depth data. In addition, the computing power of JETSON used for processing and calculation is not ideal, which makes the actual frame rate lower than what may be needed for real-time operation. Nevertheless, the required speed of log pose estimation is contingent to the crane working speed where 1Hz evaluations may not be a bottleneck.

In practice, logs may be covered with snow and ice, which makes detection more difficult and may lead to errors or missed detections, given the color dependence. In addition, different types of trees have different surface colors and textures. In our test, only birch logs are used. It is therefore likely that e.g. pine or fir would result in a more unreliable angle estimation, thus the color comparison method we use for angle estimation need to be tested with a variety of tree species. The current system can only detect, locate and analyze the angle of a single log with a certain precision, but logs lying in piles have not been evaluated with our system. Limited by the implementation method, the shape of the target log needs to be relatively straight and the overall color probably needs to be relatively uniform. It can be assumed that the system will not work effectively for logs with large bends and logs with large changes in surface color; however, this should be further investigated.

To improve our system, detection accuracy speed and robustness could be improved by using a more accurate depth camera, a more powerful computing unit, by preparing adequate data sets, and updating to new detection algorithms. In the future, it will be important to realize the automatic detection and pose analysis of logs in piles, which may also be accompa-

nied by the modification or replacement of the algorithm, but will also make the system closer to the actual state of today's cut-to-length method.

Although there are still many shortcomings, this realization of automatic log detection, positioning, and angle estimation and the successful practice on the unmanned forest machine experiment platform have undoubtedly reached an important milestone in the strive towards unmanned forestry operations.

### Acknowledgments

This research was partly funded by the VINNOVA project *Autoplant*, the Swedish Energy Agency project *Sustainable Automated Offroad Transport of Forest Biomass*, and the EU-Interreg project NUVE.

### 5. References

- Adarsh, P., Rathi, P., Kumar, M., 2020: YOLO v3-Tiny: Object Detection and Recognition using one stage improved model. International Conference on Advanced Computing and Communication Systems, Coimbatore, India, March 06–07, 687–694 p.
- Ahola, K., Salminen, S., Toppinen-Tanner, S., Koskinen, A., Väänänen, A. 2013: Occupational Burnout and Severe Injuries An Eight-year Prospective Cohort Study among Finnish Forest Industry Workers. *J. Occup. Health.* 55(6): 450–457. <https://doi.org/10.1539/joh.13-0021-0a>
- AORO, 2021: Först i världen med autonom skotning. Available online: <https://www.skogstekniskaklustret.se/nyheter/forst-i-varlden-med-autonom-skotning> (accessed 04 March 2022)
- Besten med virkeskurir – ett innovativt och lovande drivningssystem. Available online: <https://www.skogforsk.se/kunskap/kunskapsbanken/2006/Besten-med-virkeskurir-ett-innovativt-och-lovande-drivningssystem/> (accessed 04 March 2022)
- Bjelonic, M., 2018: YOLO ROS: Real-Time Object Detection for ROS. Available online: [https://github.com/leggedrobotics/darknet\\_ros](https://github.com/leggedrobotics/darknet_ros) (accessed 04 March 2022)
- Gietler, H., Böhm, C., Ainetter, S., Schöffmann, C., Fraundorfer, F., Weiss, S., Zangl, H., 2022: Forestry Crane Automation using Learning-based Visual Grasping Point Prediction. *IEEE Sensors Applications Symposium (SAS)*, Sundsvall, Sweden, Aug 01–03;1–6 p. <https://doi.org/10.1109/SAS54819.2022.9881370>
- Drushka, K., Konttinen, H., 1997: *Tracks in the Forest: The Evolution of logging machinery.* Timberjack Group Oy: Helsinki, Finland, 175 p.
- Eriksson, M., Lindroos, O., 2014: Productivity of harvesters and forwarders in CTL operations in northern Sweden based on large follow-up datasets. *J. For. Eng.* 25(3): 56–65. <https://doi.org/10.1080/14942119.2014.974309>

Fortin, J.M., Gamache, O., Grondin, V., Pomerleau, F., Giguère, P., 2022: Instance Segmentation for Autonomous Log Grasping in Forestry Operations. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Kyoto, Japan, 6064–6071 p. <https://doi.org/IROS47612.2022.9982286>

Forwarder cranes with world leading technology. Available online: <https://www.cranab.com/downloads/Forwarder-Cranes/Cranab-FC-brochure-EN.pdf> (accessed 04 March 2022)

Fyrk, J., Larsson, M., Myhrman, D., Nordansjö, I., 1991: Forest Operations in Sweden; Frumerie, G., Forskningsstiftelsen Skogsarbeten: Kista, Sweden; 25 p.

Gellerstedt, S., Dahlin, B., 1999: Cut-To-Length: The Next Decade. 10(2): 17–24. <https://doi.org/10.1080/08435243.1999.10702731>

Hellström, T., Lärkeryd, P., Nordfjell, T.2, Ringdahl, O., 2009: Autonomous Forest Vehicles: Historic, envisioned, and state-of-the-art. J. For. Eng. 20(1): 31–38. <https://doi.org/10.1080/14942119.2009.10702573>

Intelligent Boom Control Quick Facts. Available online: <https://www.deere.com/en/forestry/infographics/intelligent-boom-control/> (accessed 04 March 2022)

Intelligent Boom Control. Available online: <https://www.deere.co.uk/en/forestry/ibc/> (accessed 04 March 2022)

Komatsu 895 En pålitlig och produktiv skotare. Available online: <https://www.komatsuforest.se/skogsmaskiner/v%C3%A5ra-skotare/895-2020> (accessed 03 March 2022)

More wood in a single trip elephant king. Available online: [https://www.ponsse.com/en/web/guest/products/forwarders/product/-/p/elephant\\_king\\_8w#/](https://www.ponsse.com/en/web/guest/products/forwarders/product/-/p/elephant_king_8w#/) (accessed 03 March 2022)

Park, Y., Shiriaev, A., Westerberg, S., Lee, S., 2011: 3D log recognition and pose estimation for robotic forestry machine. 2011 IEEE International Conference on Robotics and Automation, Shanghai, China, May 09–13, 5323–5328. <https://doi.org/10.1109/ICRA.2011.5980451>

Ponsse introduces new technology for productivity and ergonomics. Available online: [https://www.ponsse.com/company/news/-/asset\\_publisher/P4s3zYhpxHUQ/content/ponsse-introduces-new-technology-for-productivity-and-ergonomics?\\_com\\_liferay\\_asset\\_publisher\\_web\\_portlet\\_AssetPublisherPortlet\\_INSTANCE\\_P4s3zYhpxHUQ\\_groupId=20143#/](https://www.ponsse.com/company/news/-/asset_publisher/P4s3zYhpxHUQ/content/ponsse-introduces-new-technology-for-productivity-and-ergonomics?_com_liferay_asset_publisher_web_portlet_AssetPublisherPortlet_INSTANCE_P4s3zYhpxHUQ_groupId=20143#/). (accessed 04 March 2022)

Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., Echele, R., Ng, A.Y., 2009: ROS: an open-source Robot Operating System. ICRA workshop on open source software, Kobe, Japan, May 12–17, 5 p.

Redmon, J., Farhadi, A., 2018: YOLOv3: An Incremental Improvement. abs/1804.02767. <https://doi.org/10.48550/arXiv.1804.02767>

Smooth Boom Control / Intelligent Boom Control. Available online: <https://www.deere.com/assets/pdfs/common/products/forestry-intelligent-boom-control.pdf> (accessed 04 March 2022).



© 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

---

Authors' addresses:

Songyu Li, MSc \*

e-mail: [songyu.li@ltu.se](mailto:songyu.li@ltu.se)

Håkan Lideskog, PhD

e-mail: [hakan.lideskog@ltu.se](mailto:hakan.lideskog@ltu.se)

Luleå University of Technology

Department of Engineering Sciences and Mathematics  
97187 Luleå

SWEDEN

\* Corresponding author

Received: March 25, 2022

Accepted: February 15, 2023