2nd Conference on Production Systems and Logistics

# A Concept for Camera-based Classification of Load Carriers

Dimitrij-Marian Holm[1], Johannes Fottner[1]
[1]Chair of Materials Handling, Material Flow, Logistics , Technical University of Munich, Garching, Germany

**Abstract**

Due to growing environmental awareness, the Circular Economy and in particular the concept of Reverse Logistics (RL) are more and more becoming the focus of industry, yielding ecological as well as economic advantages. However, the successful implementation of the concepts requires that several challenges be met. One of the most common challenges is the lack of information within RL. One proposed solution is to use more Automatic Identification Systems (Auto-ID) to track returning goods and close the information gaps between RL participants. Currently available identification systems are often limited in their field of application, as they can be very expensive, require a huge change in current logistics processes or suffer from physical characteristics, such as electromagnetic absorption or limited visual contact. With this paper, we introduce our novel concept for load carrier classification and quantification using Time-of-Flight (ToF) cameras in combination with color images, beginning with a general overview of the system architecture and process structure. This is followed by an in-depth analysis of the process steps, starting with triggering camera records followed by image pre-processing, classification and finally a quantification of the loaded cargo.

**Keywords**

Circular Economy; Reverse Logistics; Classification; Load Carrier; Cargo Identification

## 1. Introduction

The discussion about the responsible use of natural resources has grown over the last decades. One major concept in addressing this issue is the Circular Economy [1]. It defines a production and consumption system that implements a circular life cycle for goods. This means not only that products, load carriers and packages are delivered from the manufacturer to the customer, but also that used goods are returned from the customer to the manufacturer for reuse or recycling, in a process called Reverse Logistics (RL) [2]. This process allows the sustainable usage of resources and components by returning them back to the commodity cycle [3].

Not only does the implementation of RL reduce the need for raw materials, companies can also gain economic benefits as well as social recognition [4]. Along with these benefits, RL faces several challenges, including social, political and economic barriers [5]. Other frequently mentioned major challenges are the lack of information, such as about the timing, quality or quantity of returned goods, and about the required technological systems needed to acquire this information [5]. If manufacturers implement Reverse Logistics, their production depends more and more on these kinds of information [6]. This paper contributes a technical concept for the classification and quantification of returned packages, especially load carriers, which can be easily adapted to current processes. The aim is to increase the amount of information which will improve knowledge about incoming and stored packages, enabling higher production certainty and better prediction of future material demands.

DOI: https://doi.org/10.15488/11268

## 2. Related Work

### 2.1 Automatic Identification (Auto-ID)

In the Forward Supply Chain identification and tracking of goods are common tasks, mainly solved by using Auto-ID technologies. One of the most established and well used technologies are barcodes [7]. Even if they are easy to use and cheap, they have several disadvantages, such as the need for visual contact, the relatively short reading range [8] and susceptibility to environmental factors such as water and dirt [9]. On the other hand, Radio Frequency Identification (RFID) can solve the problem of visual contact and range, as it uses radio frequency for identification [10]. For the RL concept, RFID demonstrates good applicability for tracking returned products [11],[12]. While RFID seems to be a better choice than barcode-based identification systems, there are several drawbacks. RFID in general is much more expensive than a simple barcode. Besides tags, infrastructure for reading and processing the data is required [11], necessitating huge initial investments [13]. Further disadvantages result from physical characteristics. RFID uses radio frequency waves which are absorbed by liquids and reflected by metals, which reduces the utility of RFID in certain environments and industries [8],[14].

### 2.2 Machine Vision

With the breakthrough of modern Machine Learning (ML) methods such as Convolutional Neural Network (CNN), Machine Vision has experienced a huge acceleration in different areas such as Image Classification and Object Detection and Segmentation [15]. As ML methods are data driven, many datasets are available, either general [16],[17] or task-specific, such as autonomous driving [18] or remote sensing [19]. As logistics involves a large number of different objects and is associated with a high degree of continuous change [20], especially when customized load carriers are used, it is very difficult to find usable datasets. There are only a few datasets available, such as *Logistics Objects in Context (LOCO)* [21], which offers a wide range of general objects used in logistics, such as load carriers, pallets and forklifts, or, on the other hand, datasets that contain only one class (load carrier) [22].

## 3. Problem Description and Motivation

The lack of information within RL cannot be solved only by using currently available Auto-ID technologies, which leads to the need to consider alternative technologies. With the huge breakthroughs that modern machine learning techniques provide to machine vision, new approaches for identification are possible. However, huge datasets are needed for such approaches, and they are often not available within specific areas. For a better understanding, we emphasize an example using the German beverage industry. Here the end customer returns empty bottles as well as the load carriers, mainly crates, to the retailer, which delivers them to sorting companies and then back to the bottler. Future production therefore depends on the number of returned load carriers, which requires a process for classifying and counting. Most of the commonly used Auto-ID technologies such as RFID are not applicable in this industry because of the liquids and the high investment resulting from the huge number of crates in relation to the costs for installing tags. Auto-ID systems using optical identification, such as barcodes, are not applicable, because the composition of the transported goods is often changed by different processes like sorting. Old barcodes therefore need to be removed and new barcodes need to be attached to cargo at every process that changes composition, or else every load carrier would get a barcode, which is not possible due to the lack of visibility. Additionally, most beverage manufacturer design their own load carriers for product branding purposes and change the designs regularly, increasing the difficulty of data mirroring recently used load carriers. Due to the limitations of current Auto-ID systems, stocktaking is often performed by hand, resulting in an inaccurate and inefficient mapping of the inventory, and potentially affecting future production. With this paper, we contribute a concept to address both challenges, combining a method for fast data acquisition with the use of machine vision approaches to increase information density. This eliminates the need for manual stocktaking, reduces inventory differences and improves production forecasting. We aim to increase the scope of application and promote RL, mainly using cameras (ToF and RGB) and CNN for the classification of load carriers.

## 4. Concept Overview and Architecture

### 4.1 Use-Case

To explain our concept, we will focus on the load carrier transportation processes within a German beverage manufacturer. Beverage load carriers are regular and uniform, but vary in geometry depending on the type, so that carriers of one type can be easily stacked and transported on a pallet by forklifts. Depending on the forklift's size, it can simultaneously transport several pallets, which are called a bundle.

### 4.1 Planned Hardware Setup

For our setup we plan to use several sensors and computation units for different purposes. We will use two ultrasonic sensors with a measurement range of 4.5 meters to determine the distance between the forklift and the bundle and the height of the forks above the ground. For depth and color images, we will use a Microsoft Kinect V2, which combines color and a ToF camera in one single device. For process steering and data processing we will use an Nvidia Jetson NX Development Kit because of its small size compared to CPU performance and an integrated GPU to accelerate our CNN execution. All these components will be mounted on forklifts, so no external components are required.

### 4.2 Pipeline Architecture

For our concept, we designed a pipeline which can be subdivided into two parts. A general part provides a fast way to acquire and refine needed data (Triggering, Data Acquisition, Preprocessing) and two different modes – training (Labeling, Training) and operating (Classification, Quantification) – which can either be used for data labelling and training of neural networks or be used for classification and counting of load carriers (Figure 1).

**Triggering and Data Acquisition:** As the forklift approaches the load carriers, we measure the distance between the forklift and the load carriers. If the distance falls below a certain threshold, an event signal is sent to activate the camera and start data acquisition. As the forklift is driving towards the load carriers, the distance continues to decrease. If a second threshold is reached, data acquisition is stopped because the camera modules are too close to the load carriers and the field of view is restricted. The recorded data between start and stop are henceforth referred as time series. In our case our time series contain color and depth images, which are both needed for further processing. The end of data acquisition represents the start of the preprocessing task, which is realized by sending an event signal from data acquisition to preprocessing.

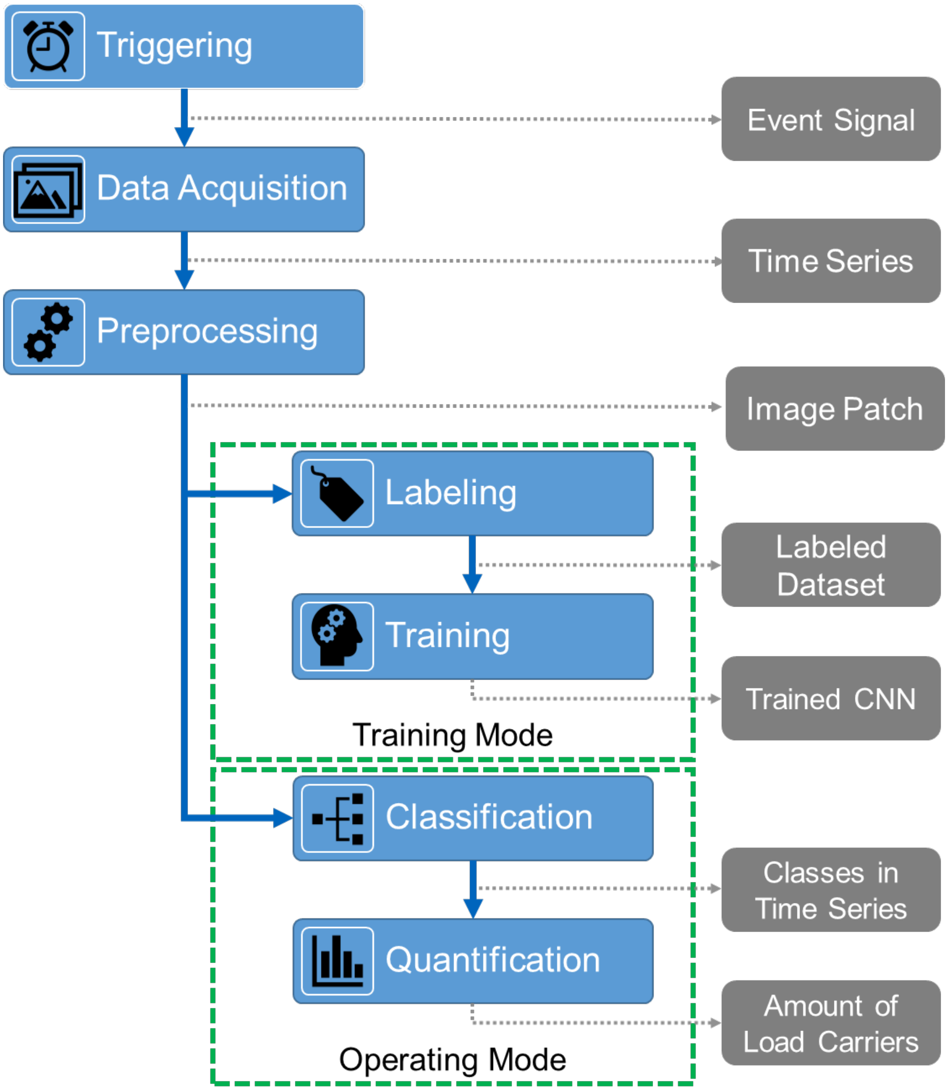


Figure 1: Pipeline overview.
Blue boxes describe the tasks while the blue arrows show transitions between the tasks.
Grey arrows and boxes refer to the results of the tasks.

**Preprocessing:** The preprocessing task will provide necessary data for both modes, training and operating. First, we need to specify the position of our transported load carriers and separate them from other load carriers to prevent misclassification. Therefore, the main task for preprocessing is to calculate the Region of Interest (ROI) in the color image in which the load carriers are contained (Figure 2).

Figure 2: Color image with marked RoI (red rectangle)

First, we identify specific key points within our depth image. As the load carriers are arranged as a cuboid, the most apparent key points are the four edges describing the front of the cuboid (Figure 3). Additional data about the cuboid are not used. Because our camera position is not restricted in height, the front is the only guaranteed visible part independent of the camera height.

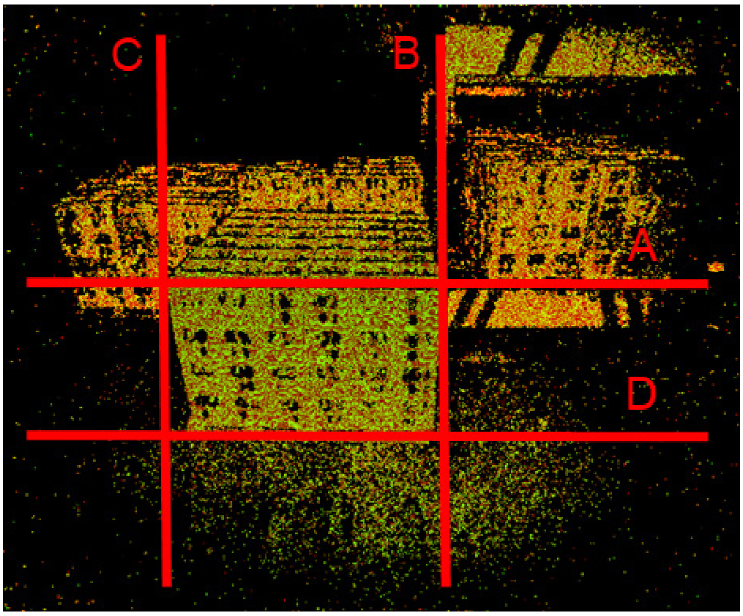


Figure 3: Depth image with marked edges.
A: upper edge, B: right edge, C: left edge, D: lower edge

By calculating the intersection of these edges, we find the corner points for our ROI, which are transformed from a depth image coordinate space to a color image coordinate space and used to crop the image (Figure 4). Since load carriers differ in size and we do not know the size of our selected load carriers, we subdivide our image into predefined areas which can handle different sizes to preserve geometrical and structural information. For example, if we choose one mask too small for our load carrier size, we cut off necessary information. Using a lager mask compensates for these information losses, but contains a small amount of additional information from neighbor load carriers. Therefore, we need to select our mask sizes precisely to find an equilibrium between the loss and additional information. In addition, this is a common practice to reduce problem of scale invariance from which CNN suffer [23]. We therefore cut our cropped image into smaller pieces by using different masks of predefined size (Figure 4). The resulting stack of small images is henceforth called a patch. These patches are used for both modes, training and operating. The patch is useful especially in the classification task to determine whether the bundle contains different types of load carriers. Depending on the current mode, patches are either stored for offline processes like labeling and training or used directly for classification and discarded afterwards.

Figure 4: Cropped color image and the resulting subdivided images using different mask sizes (red rectangle and blue rectangle)

**Training mode:** Training mode will be used for a fast generation of the necessary labeled data needed for the subsequent training of neural networks.

**Labeling:** Thanks to the preprocessing task, we can generate a huge number of cropped images automatically. In addition, since we are working with time series, we need to label only the time series instead of every image, since the images inherit the time series label.

**Training:** For training we will use Pytorch [24] as the framework for implementing and testing our CNNs for classification. Within Pytorch, several predefined CNNs such as ResNet [25] can be selected and trained with the generated patches. As our concept is developed to quickly switch to operating mode, object detection and segmentation are not considered because of the higher manual effort in labeling data.

**Operating mode:** With the operating mode, the picked load carriers can be classified and, by using additional information from the ultrasonic sensors, the amount of loaded cargo can be calculated.

**Classification:** Since we create our patch, we can now exploit the fact that every image contains only a small part of the original cropped image. Although we trained our CNN only with data created using bundles with single types, we can now classify every patch independently. By adding up the classification results of every single image, we can derive information about the whole bundle by evaluating the resulting distribution and making a statement whether there is only one or several types of load carriers within the bundle. In addition, if we use the position information of the patched images within our ROI and combine the information with our classification results, we can define areas within which certain load carriers occur.

**Quantification:** If the classification was successful and we use the lower edge and upper edge height, we can determine the total height of the cargo. Using the classification and area information, we are able to look up historical data related to our loaded cargo, e.g., the single load carrier height and the number of crates per layer, which finally can be used to estimate the total number of crates picked up by the forklift.

### 4.3 Restrictions

Due to the cargo loading process and the sensors mounted on forklifts, we can gain information from only one side of the bundle (Figure 2). Our pipeline can therefore estimate only the class and number of load carriers based on this data. Because this approach reduces the operating area, no further process adjustments or extra components need to be installed. In addition, we have further restrictions regarding the two modes of our concept.

**Training mode:** To accelerate labeling it is necessary that the loaded bundle contain only one class of load carriers. This allows us to label a whole time series with just one single label.

**Operating mode:** In contrast to the training mode restriction, different classes of load carriers can be handled within the bundle. However, with our restricted view of the bundle, we cannot detect mixed types outside the ROI. We therefore assume that, if our classification outcome predicts only one class, the whole bundle will be viewed as containing one type, otherwise the bundle is identified as mixed bundle. To quantify mixed bundles, an additional restriction comes into play. Because we have information about where different load carriers occur in our ROI, we can still estimate the number of load carriers if they are limited to pallets, e.g., every pallet contains only one type of load carrier. This allows us to look up historical data related to single-pallet cargo. If we cannot comply with this restriction, quantification is marked as impossible because of the high spatial mixing of different load carriers.

## 5. Conclusion and Future Work

Reverse Logistics as a part of the Circular Economy is a key concept for enabling a sustainable and environmentally friendly product lifecycle with additional benefits for companies willing to implement it. Beside the benefits, there are several challenges associated with Reverse Logistics, especially the lack of technical systems to acquire the information needed for successful operation. Current technologies such as Barcodes and RFID suffer from different physical, processual or economical limitations, e.g., electromagnetic absorption, limited visibility, major changes in current processes or high financial investments. In this paper, we presented a camera-based data processing pipeline for classification of load carriers to increase information density and overcome technical problems for manufacturers. Furthermore, our approach can be easily integrated into current processes without major changes, and the expense is scalable without huge initial investments, as no expensive external components are needed.

In future work we will implement this concept within a beverage facility to test the system under real conditions. We assume that the quality of images recorded by the camera system will be especially affected by environmental conditions such as rain, snow, the day-and-night cycle or fog, as the system will be used indoors and outdoors. Accordingly, we are defining our requirements and developing the software needed to realize our concept. Afterwards a test phase under real conditions is planned with an evaluation of the concept.

## References

[1] Wautelet, T., 2018. The Concept of Circular Economy: its Origins and its Evolution.

[2] Brito, M.P. de, Dekker, R., 2004. A Framework for Reverse Logistics, in: Dekker, R., Fleischmann, M., Inderfurth, K., van Wassenhove, L.N. (Eds.), Reverse Logistics. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 3–27.

[3] Korhonen, J., Honkasalo, A., Seppälä, J., 2018. Circular Economy: The Concept and its Limitations. Ecological Economics 143, 37–46.

[4] Thun, J.-H., Müller, A., 2009. An empirical analysis of green supply chain management in the German automotive industry. Bus. Strat. Env.

[5] Ravi, V., Shankar, R., 2005. Analysis of interactions among the barriers of reverse logistics. Technological Forecasting and Social Change 72 (8), 1011–1029.

[6] SRIVASTAVA, S., 2008. Network design for reverse logistics. Omega 36 (4), 535–548.

[7] McCathie, L., 2004. The advantages and disadvantages of barcodes and radio frequency identification in supply chain management.

[8] Helmus, M., 2009. RFID in der Baulogistik: Forschungsbericht zum Projekt "Integriertes Wertschöpfungsmodell mit RFID in der Bau- und Immobilienwirtschaft", 1. Aufl. ed. Vieweg+Teubner Verlag / GWV Fachverlage GmbH, Wiesbaden, Wiesbaden, 733 pp.

[9] Blumberg, D.F., 2004. Introduction to management of reverse logistics and closed loop supply chain processes. CRC, Boca Raton, Fla., London, 253 pp.

[10] Payaro, A., 2004. The Role of ICT in Reverse Logistics: A Hypothesis of RFID Implementation to Manage the Recovery Process, 7 pp.

[11] Plaza-Úbeda, J.A., Abad-Segura, E., Burgos-Jiménez, J. de, Boteva-Asenova, A., Belmonte-Ureña, L.J., 2021. Trends and New Challenges in the Green Supply Chain: The Reverse Logistics. Sustainability 13 (1), 331.

[12] Lee, C.K.M., Chan, T.M., 2009. Development of RFID-based Reverse Logistics System. Expert Systems with Applications 36 (5), 9299–9307.

[13] Vijayaraman, B.S., Osyk, B.A., 2006. An empirical study of RFID implementation in the warehousing industry. Int Jrnl Logistics Management 17 (1), 6–20.

[14] Koski, K., Koski, E., Virtanen, J., Björninen, T., Sydänheimo, L., Ukkonen, L., Elsherbeni, A.Z., 2012. Inkjet-printed passive UHF RFID tags: review and performance evaluation. Int J Adv Manuf Technol 62 (1-4), 167–182.

[15] Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., Alsaadi, F.E., 2017. A survey of deep neural network architectures and their applications. Neurocomputing 234, 11–26.

[16] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., 2009 - 2009. ImageNet: A large-scale hierarchical image database, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition. 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops), Miami, FL. 20.06.2009 - 25.06.2009. IEEE, pp. 248–255.

[17] Everingham, M., van Gool, L., Williams, C.K.I., Winn, J., Zisserman, A., 2010. The Pascal Visual Object Classes (VOC) Challenge. Int J Comput Vis 88 (2), 303–338.

[18] Geiger, A., Lenz, P., Stiller, C., Urtasun, R., 2013. Vision meets robotics: The KITTI dataset. The International Journal of Robotics Research 32 (11), 1231–1237.

[19] Li, K., Wan, G., Cheng, G., Meng, L., Han, J., 2020. Object detection in optical remote sensing images: A survey and a new benchmark. ISPRS Journal of Photogrammetry and Remote Sensing 159, 296–307.

[20] Arnold, D., Isermann, H., Kuhn, A., Tempelmeier, H., Furmans, K., 2008. Handbuch Logistik. Springer Berlin Heidelberg, Berlin, Heidelberg.

[21] Mayershofer, C., Holm, D.-M., Molter, B., Fottner, J. LOCO: Logistics Objects in Context. Proc. of 2020 IEEE International Conference on Machine Learning and Applications (ICMLA 2020).

[22] Poss, C., Irrenhauser, T., Prueglmeier, M., Goehring, D., Zoghlami, F., Salehi, V., 2019. Perceptionbased Intelligent Materialhandling in Industrial Logistics Environments, in: Proceedings of the 2019 11th International Conference on Computer and Automation Engineering - ICCAE 2019. the 2019 11th International Conference, Perth, WN, Australia. 23.02.2019 - 25.02.2019. ACM Press, New York, New York, USA, pp. 146–151.

[23] Gong, Y., Wang, L., Guo, R., Lazebnik, S., 2014. Multi-scale Orderless Pooling of Deep Convolutional Activation Features, in: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (Eds.), Computer Vision – ECCV 2014, vol. 8695. Springer International Publishing, Cham, pp. 392–407.

[24] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S., 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library, 12 pp. http://arxiv.org/pdf/1912.01703v1.

[25] K. He, X. Zhang, S. Ren, J. Sun, 2016. Deep Residual Learning for Image Recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778.

**Biography**

**Dimitrij-Marian Holm** received his M.Sc. in robotics at the University of Applied Science Munich in 2018. Since 2018 he has been part of the research staff at the Chair of Materials Handling, Material Flow, Logistics, focusing on image processing and machine learning.

**Johannes Fottner** received a Dr.-Ing. degree in mechanical engineering from the Technical University of Munich (TUM), Munich, Germany in 2002. For the past 15 years, Johannes Fottner acted in different managing functions in the material-handling sector. Since 2016, he has been the head of the Chair of Materials Handling, Material Flow, Logistics with the TUM.