



(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:  
**18.08.2021 Bulletin 2021/33**

(51) Int Cl.:  
**B07C 5/342 (2006.01) B07C 5/34 (2006.01)**

(21) Application number: **20215996.8**

(22) Date of filing: **21.12.2020**

(84) Designated Contracting States:  
**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**  
Designated Extension States:  
**BA ME**  
Designated Validation States:  
**KH MA MD TN**

(71) Applicant: **IHP Systems A/S**  
**2200 Copenhagen N (DK)**

(72) Inventors:  
• **ANDERSEN, Jesper Stemann**  
**2000 Frederiksberg (DK)**  
• **MENSAL, Lars**  
**2840 Holte (DK)**

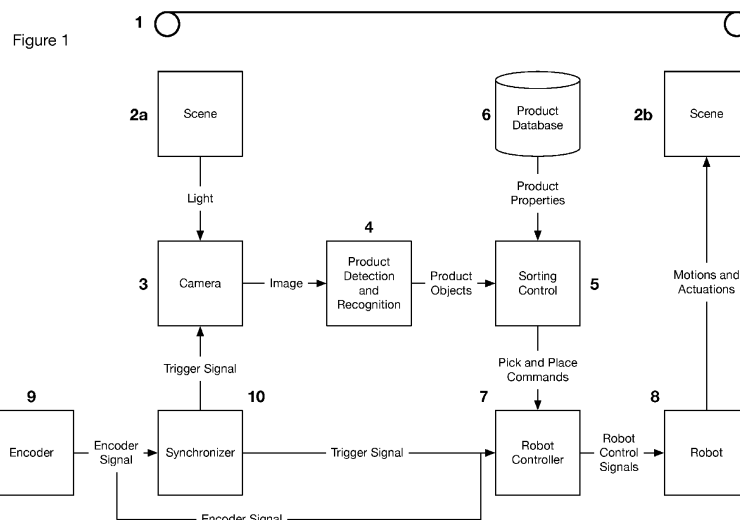
(30) Priority: **20.12.2019 EP 19218995**

(74) Representative: **Plougmann Vingtoft a/s**  
**Strandvejen 70**  
**2900 Hellerup (DK)**

(54) **A METHOD FOR SORTING CONSUMER PACKAGING OBJECTS TRAVELLING ON A CONVEYOR BELT**

(57) The present invention relates to a method for sorting objects, the method includes at least one imaging sensor and a controller comprising a processor and a memory storage, wherein the controller receives image data captured by the at least one imaging sensor; and at least one sorting robot is coupled to the controller, wherein the at least one sorting robot is configured to receive an actuation signal from the controller. The processor executes an object identification module configured to detect objects travelling on a conveyor belt and recognize at least one target item travelling on the conveyor belt as one of at least 10 consumer packaging product objects

and/or recognizing the object as one of at least 40 consumer packaging brand objects. The detection and recognition are based on one or more of the following: the characteristics of the shape of the object, the characteristics of the color/colors of the object, the characteristics of image features on the object in at least three areas on the object, and processing the image data and to determine an expected time when the at least one target item will be located within a diversion path of the sorting robot; and wherein the controller selectively generates the actuation signal based on whether a sensed object detected in the image data comprise the at least one target item.



## Description

**[0001]** The present invention relates to a method for sorting consumer packaging objects travelling on a conveyor belt, where image data is captured by at least one imaging sensor for an image containing at least one object travelling on the conveyor belt and where the at least one imaging sensor provides color image data.

## BACKGROUND ART

**[0002]** In many recycling centers that receive recyclable materials, sortation of materials may be done by hand or by machines. For example, a stream of materials may be carried by a conveyor belt, and the operator of the recycling center may need to direct a certain fraction of the material into a bin or otherwise off the current conveyor. It is also known to use automated solutions using sensors or cameras to identify materials carried on a conveyor belt, which via a controller may activate a sorting mechanism. However, these new solutions do not always function perfectly.

**[0003]** The conventional plastic sorting solutions are based on near-infrared / short-wave-infrared (NIR/SWIR) spectrometry, where e.g., a NIR/SWIR reflection spectrum is collected for each plastic object and the spectrum identifies the material type of the plastic object - which determines the sorting.

**[0004]** The NIR/SWIR-spectrometric sorting systems are unable to handle dark and black plastics as all dark and black plastics return the same flat spectrum in the NIR/SWIR-range regardless of the material type. Moreover, NIR/SWIR-systems also cannot discriminate properly between white and transparent plastics, which is important for proper recycling. Another drawback of the spectrometric systems is that the systems cannot sort waste by application - e.g., they cannot sort food from non-food plastics, which is a major drawback in a sorting process.

**[0005]** Finally, spectrometric systems are also challenged by composite plastic objects, e.g., a bottle with a bottle cap and/or a foil covering the bottle - the spectrometric system might sort the object based on the foil or the cap.

## DISCLOSURE OF THE INVENTION

**[0006]** An object of the present invention is to provide a method for identifying and sorting waste material, in a more precise manner.

**[0007]** A further object is to provide a cost-effective and efficient method of identifying and sorting waste material, in particular consumer packaging objects, and in particular waste material comprising plastic or cardboard.

**[0008]** A further object of the present invention is to provide a method for identifying a waste object as being a specific (packaging) product, such as a product of a specific brand or a product of a specific producer.

**[0009]** Thus, the present invention provides a method which can enable identification and sorting of high-value packaging objects, such as food packaging or, alternatively, the identification and sorting out of contaminating objects, such as hazardous objects or objects containing hazardous materials.

**[0010]** Normally, when waste and garbage is collected, an initial sorting into different material categories is performed. The categories may e.g., be glass, metal, plastic, cardboard, paper and biological waste. Thus, when the waste reaches the recycling center, each material fraction is normally sorted into even finer fractions. The metal fraction may be sorted into aluminum and iron fractions and plastic into fractions based on different plastic types such as PE, PP or fractions with soft and hard plastic.

**[0011]** However, the present invention is capable of detecting and recognizing consumer packaging objects travelling on a conveyor belt among several other different objects such as generic, non-packaging or non-consumer packaging objects or glass objects or metallic objects and sorting out consumer packaging objects from a stream of waste material. Consequently, the present invention is suitable for sorting out consumer packaging objects from a stream of waste material.

**[0012]** The present invention relates to a method for sorting consumer packaging objects travelling on a conveyor belt, the method comprising:

receiving image data captured by at least one imaging sensor for an image containing at least one feature on or of an object travelling on the conveyor belt said imaging sensor providing color image data with a spatial resolution of at least 0.4 px/mm;

executing a product detection and recognition module on a processor, the product detection and recognition module being configured to detect characteristics of the at least one feature on or of the object travelling on the conveyor belt by processing the image data and recognizing the object as one of at least 10 consumer packaging product objects and/or recognizing the object as one of at least 40 consumer packaging brand objects; and

wherein the detection and recognition are based on one or more of the following: the characteristics of the shape of the object, the characteristics of the color/colors of the object, the characteristics of image features on the object in at least three areas on the object; and

when an object has been detected and recognized determining an expected time when the at least one object will be located within a sorting area of at least one sorting device; and

selectively generating a device control signal to operate the at least one device on whether the at least

one object comprises a target object.

**[0013]** In this context the term "sorting device" should include a robot, mechanical actuators, actuators based on a solenoid, air jet nozzles etc.

The terms "object", "item" and "target object" and their plural form are used interchangeably in this text.

**[0014]** Consumer packaging objects is to be understood as an object for packing consumer products such as food products or products for personal care/hygiene, such as soap or toothpaste. The consumer packaging objects may be made from plastic, cardboard, other recyclables or combinations of these.

**[0015]** The term "stream of objects" should herein be taken to mean a stream of objects where the objects for the most part are made up of a primary material type, e.g., plastic. Examples include source-separated post-consumer waste, e.g., a stream of post-consumer plastic waste or a stream of post-consumer packaging waste.

**[0016]** The term "consumer packaging object" should herein be taken to mean an object designed for packaging a consumer product: the object may have one or more properties selected among: shape (e.g., bottle, tray, tub, lid), size, color, opacity/transparency (primary/dominant, e.g., black, blue transparent), material (e.g., PET plastic, cardboard), application (e.g., food, soap/cosmetics/personal care/hygiene) producer (e.g., Acme Ltd.), brand (e.g., Acme Brand), and product/SKU (e.g., Acme Brand Product)

**[0017]** The term "product object" should herein be taken to mean an object with all properties listed above except possibly material and/or size. Examples of product objects include Coca-Cola bottle 2L and Heinz Tomato Ketchup Bottle 580 g.

**[0018]** The term "brand object" should herein be taken to mean an object with properties selected from shape, color, opacity/transparency, application, producer and brand. The object may also have the properties size and material defined.

Examples include a Coca-Cola bottle and a Heinz Tomato Ketchup Bottle.

**[0019]** The names Coca-Cola, Heinz and Heinz Tomato Ketchup are trademarks.

**[0020]** A brand is a name, term, design, symbol or any other feature that identifies one producer's goods as distinct from those of other producers.

**[0021]** A logo (or logotype) is a graphic mark, emblem, or symbol used to aid and promote public/consumer identification and recognition. It may be of an abstract or figurative design or include the text of the name it represents as in a wordmark.

**[0022]** A symbol is a picture which is easily recognizable and have a certain meaning, e.g., the symbol for recycling (see figure 13 for further symbols).

**[0023]** A slogan is a short and easily recognizable text, an example may be the slogan from the company Carlsberg: "Probably the best beer in the world".

**[0024]** The imaging sensor is preferably a camera

which is able to provide color images in an environment with low light intensity, e.g., light intensities around 500 lumen. Preferably, the camera operates at light intensities around 1000 lumen or more, such as 1500 lumen or more.

**[0025]** Spatial resolution is a resolution of an image and is determined by the resolution of the sensor (how many pixels on the sensor), and the size of the area being projected onto the imaging sensor, with the latter being a product of the optical configuration of the imaging system. The camera may have a resolution of 2000 pixel/mm (px/mm, pixel density) at the image plane or image forming surface. However, due to the linear spacing between the image plane and the product surface and the angular spread of the light waves to be reflected on the product surface, the pixel density on the product surface will appear less dense than the resolution on the image forming surface in the camera. Thus, the spatial resolution is the resolution (pixel density) appearing on the product surface. Spatial resolution is a well-known concept, e.g., within the technical field of satellites and satellite photos. In an embodiment the object travelling on the conveyor belt is recognized as one of at least 20 consumer packaging product objects, such as one of at least 50 consumer packaging product objects, such as one of at least 80 consumer packaging product objects, such as one of at least 100 consumer packaging product objects. The method is capable of recognizing and detecting at least 2, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 consumer packaging product objects, such as at least 110, 120, 130, 140, 150, 1000, 10,000, 100,000 consumer packaging product objects.

**[0026]** In an embodiment the object travelling on the conveyor belt is recognized as one of at least 80 consumer packaging brand objects, such as at least 100 consumer packaging brand objects, such as at least 500 consumer packaging brand objects, such as at least 1000 consumer packaging brand objects.

**[0027]** Detection (object detection) is the localization of an object within an image (i.e., "where in the image is the object"). The output of detection is a location where the object is located and possibly also the angular orientation of the principal axis of the object. The location may be a point (where the object is located), a rectangle (e.g., a bounding box including the object) or similar, or a set of points defining the object (a "segmentation mask") or other geometric information identifying the location and possibly the shape of the object. Detection may also include the concept of recognition. Recognition is the attribution of an object within an image to a class, i.e., the attribution of an object to e.g., "white bottle" (i.e., "which object is in the image").

Detection (and detection and recognition) may also refer to object segmentation (where the output is a segmentation mask), known to those skilled in the art of computer vision.

**[0028]** An image feature is a property of an image. A local image feature is a property of an image at a certain

position in the image. An example of a local image feature is a corner or an edge or a line.

**[0029]** In an embodiment the target object is guided to a collection device in the sorting area by means of the sorting device. The sorting device may control e.g., a pusher device, a pick-up device or air jet nozzles which are suitable for guiding the target object to a collection device.

**[0030]** In an embodiment the consumer packaging objects travelling on the conveyor belt are at least partly a post-consumer packaging waste stream comprising packaging materials, such as plastic or cardboard or other recyclables, such as packaging materials made from paper, composites and combinations of the mentioned materials. A consumer packaging waste stream will mainly be composed of post-consumer packaging objects/products, and the stream will mainly contain one or a few packaging materials (e.g., plastic or cardboard).

**[0031]** An object or product in a stream of consumer packaging waste is the packaging product that has been used for packaging fast-moving consumer goods (FM-CG) / consumer packaged goods (CPG), such as foods, beverages, or cosmetics and personal care/hygiene products. A product can be described by a number of properties:

- Shape (e.g., tray, tub, bottle)
- Size
- Colour (primary/dominant colour, e.g. transparent, white)
- Application (e.g., food, cosmetics/personal care/soap, detergent)
- Material (e.g., PET plastic, cardboard)
- Producer (e.g. Unilever, Nestle, Kraft-Heinz)
- Brand (e.g., Heinz Tomato Ketchup)
- Product (e.g., Heinz Organic Tomato Ketchup 580g)

**[0032]** The names Unilever, Nestle, Kraft-Heinz, Heinz Tomato Ketchup and Heinz are company names and/or trademarks.

**[0033]** Although the sorting device may be e.g., a pusher device or an air jet nozzle, in an embodiment the sorting device is a pick-up or lifting device adapted for lifting the consumer packaging object away from the conveyor belt. By picking up or lifting the object off the conveyor belt it is possible to avoid collision or interference with other objects on the conveyor belt when removing the target object from the conveyor belt.

**[0034]** In an embodiment the sorting device is a pick-up or lifting device adapted for lifting the object in such a way that the side facing the conveyor belt can be captured by an image sensor. Thus, it is possible to obtain a more precise detection and recognition of the object on the conveyor belt as the embodiment allows an image sensor (preferably located at the level of or below the conveyor belt) to capture an image of the surface of the object facing the conveyor belt. This surface may comprise a specific characteristics (e.g., logo, trademark or

information about the material of the object) which cannot be captured by an image sensor located above the conveyor belt.

**[0035]** The pick-up or lifting device may e.g., apply suction and vacuum when lifting the object.

**[0036]** In an embodiment at least two image sensors are applied. The image sensors may be arranged in a line above the conveyor belt, substantially parallel with the direction of the conveyor belt, or, alternatively, substantially perpendicular with the direction of the conveyor belt. In an embodiment of the invention the image sensors are arranged such that the image data of the object is captured from different angles.

**[0037]** If a vector V1 defines the traveling direction of the conveyor belt and a vector V2 is perpendicular to vector V1, i.e., perpendicular to the surface of the conveyor, and pointing upwards from the middle of the conveyor belt, and, further, the direction to the image sensor from the middle of the conveyor belt is defined by vector V3, then the angle to the camera is the angle between V2 and V3. The angle may be in the range of 0 to 135 degrees, such as in the range 0 to 90 degrees.

**[0038]** The one or more image sensors may be located at any height or distance from the conveyor belt. However, generally it is preferred that the distance is chosen such that if the width of the conveyor belt is W, the height of the image sensor with respect to the surface of the conveyor belt is between W/2 and 4W. Thus, if the width of the conveyor belt is 1 m, the image sensor should be located between 0.5 m to 4 m above the surface of the conveyor belt, and if the width of the conveyor belt is 2 m the image sensor should be located between 1 m to 8 m above the surface of the conveyor belt. For two image sensors (mounted perpendicular to the conveyor direction - each imaging half of the conveyor width W): the relation is W/4 to 2W distance from the conveyor. For four image sensors (each imaging a quarter of the conveyor width W): the relation is W/8 to W distance from the conveyor. Generally, the relation is W/(2C) to 4W/C (C = number of image sensors along width of conveyor).

**[0039]** Normally, the at least one image sensor is located between about 0.5 m to about 8 m above the surface of the conveyor belt. The distance should be measured in a direction substantially perpendicular to the surface of the conveyor belt. By arranging the at least one imaging device at least 0.5 m above the surface of the conveyor belt interference between the image sensor and objects travelling on the conveyor belt can be avoided.

**[0040]** In an embodiment of the method according to the invention, the characteristics of the at least one object travelling on the conveyor belt is the physical appearance or shape and/or size of the object. Thus, the method is capable of identifying objects based on their design features.

**[0041]** In an embodiment of the method according to the invention, the characteristics of the at least one object travelling on the conveyor belt is the color, patterns/text-

tures of colors and/or transparency/opacity of the object. Thus, the method is also suitable for detecting objects based on their color or transparency. Naturally, the patterns/textures may comprise text, images, pictures, logos, symbols etc.

**[0042]** In an embodiment the characteristics of the at least one object travelling on the conveyor belt is selected from vendor names, brand names, product names, trademarks, logos, symbols, slogans, text or a combination of one or more of the characteristics. The product detection and recognition module may interact with one or more databases containing information about vendor names, brand names, product names, trademarks, and slogans and retract information from these databases to identify objects.

**[0043]** With respect to the three above mentioned embodiments it is clear that the features of these embodiments may be combined in any desirable manner.

**[0044]** For the purpose of obtaining a more precise identification, the product detection and recognition module may apply two or more characteristics in the product detection and recognition process.

**[0045]** Accordingly, the following information may be extracted from the image: object image (the image of an entire object and identifying the entire object as a product of a specific class, e.g., a "product object" or "brand object"). One or more of the following features may also be extracted: logo (identifying logos), symbol (identifying symbols), text (identifying text and matching with a text database related to products). The one or more pieces of information (the object image and optional additional information) are combined to output a single output product (detected at a location), i.e., detection and recognition of a consumer packaging object.

**[0046]** The information provided by the one or more pieces of information (detection and recognition methods) is fused in a statistical framework yielding one single output of the product detected. The statistical framework may exploit prior information (such as a Bayesian statistical prior distribution), e.g., the prior likelihood of a product object being detected - i.e., if a Heinz Tomato Ketchup bottle object is more likely to appear in the stream of objects than other objects or ketchup bottles, this likelihood can affect the single output of the product detection and recognition module.

**[0047]** The characteristics shape, size and color function as input to all detection types (product, logo, symbol and text).

**[0048]** Although the method operates very well with a spatial resolution of 0.4 px/mm (pixel/mm), the invention also provides an embodiment where the imaging system (camera) yields images with a spatial resolution of at least 2 px/mm (pixel/mm). With such a spatial resolution the imaging system is able to provide very detailed images.

**[0049]** In an embodiment the spatial resolution is at least 4 px/mm. When the spatial resolution is about 4 px/mm or more, the imaging sensor is able to detect very small-scale details, such as logos with an extent of about

5 mm or less.

**[0050]** In an embodiment the method is adapted for detecting and recognizing objects used as packaging or containers for food items, such as bottles, trays, tubs and lids. The objects may e.g., be bottles for juice and soft drinks made from plastic, such as transparent plastic. The object may also be a tray used for e.g., meat or fruit or biscuits. The trays may e.g., be made from plastic material in any desired colors. The trays may be marked with a "fork and knife" logo indicating that the tray is for use with foods.

**[0051]** In an embodiment the method is adapted for detecting and recognizing black objects. Black objects are difficult to detect due to the low reflection from the material. However, the method according to the invention has proven to be surprisingly efficient in detecting and recognizing black objects. The black object may e.g., be made from plastic which is desirable to sort properly. Preferably, the black object is a tray for food, such as a plastic tray for meat.

**[0052]** In one aspect of the method the detection and recognition of objects are based on the detection and recognition modules' interaction with one or more databases, such as databases comprising information about e.g., specific products (such as materials used in the product), vendor names, brand names, product names, trademarks, and slogans.

**[0053]** As an example, the imaginary company Acme produces mayonnaise under the name "best mayonnaise" and sells the mayonnaise in containers of white PE with product number 120E. Thus, if the method of the present invention detects the name "Acme" and "mayonnaise", then the system will recognize the container as product 120E and sort accordingly (taking into account that the object is made of white PE plastic).

**[0054]** Thus, in an embodiment of the method, the method further comprise interaction with a product database. The product database may contain information about an identified object, such as which material or materials the object is manufactured from. Such information is very useful in a sorting process.

**[0055]** The interaction with one or more databases may also relate to or include image features. The detection and recognition of objects may include clustering and matching of local image features. First,  $N \geq 3$  image features are computed for a reference image of each object (a reference image may be an image of a clean/new product). If  $M \geq 3$  image features detected on an image for sorting (image of waste object) and matched to  $M \geq 3$  image features on reference image, then the product/object is recognized as represented by the  $M$  image features.

The  $N \geq 3$  image features are stored on a data medium in a database.

**[0056]** The method may also apply a convolutional neural network. The detection and recognition are based on the visual appearance of objects/products. A convolutional neural network is a Machine Learning-method

such as a convolutional neural network for object detection, object segmentation etc. Among others, these methods include:

- One-stage Convolutional Neural Network for Object Detection
- One-stage Convolutional Neural Network for Object Detection with Feature Fusion Network
- Two-stage Convolutional Neural Network for Object Detection
- Anchor-free, bottom-up Neural Network for Object Detection

**[0057]** Anchor-based methods may use a number of anchors per grid cell (image divided into grid cells). The terms "object detection" and "image/object segmentation" may be used interchangeably.

**[0058]** Convolutional neural networks (CNN) for object detection may consist of a backbone CNN that forms a compressed image representation, and an object detection network which predicts bounding boxes and confidence scores for objects.

**[0059]** Convolutional neural networks (CNN) for object detection may also consist of a backbone CNN that forms a compressed image representation, then a feature fusion network passes fused features to an object detection network which finally predicts bounding boxes and confidence scores for objects. Examples of feature fusion networks are the Feature Pyramid Network (FPN), the Path Aggregation Network (PANet), the Neural Architecture Search (NAS-FPN), and the Bi-directional Feature Pyramid Network (BiFPN).

**[0060]** Convolutional neural network bottom-up anchor-free detection methods predict bounding boxes for objects based on key point detections. The key points of interest are first located in the image and then combined to form the boundaries of the detected objects. Points such as the bounding box corner points (as used in the method CornerNet), center points (as used in the method CenterNet) or extreme points (left-most, right-most, top, bottom) may function as key points for predicting bounding boxes. The key points are detected in heatmaps from the output of a Convolutional Neural Network (CNN) and classified using embeddings. The CNN is trained to predict similar embeddings for key points belonging to the same object.

**[0061]** Thus, in an embodiment of the method according to the invention, the product detection and recognition involves a convolutional neural network.

**[0062]** In an embodiment of the method the product detection and recognition involves a convolutional neural network for object detection, such as a one-stage or a two-stage convolutional neural network for object detection.

**[0063]** Moreover, in an embodiment of the method the product detection and recognition involve a convolutional neural network for object detection with a feature fusion network.

**[0064]** In a further embodiment the product detection and recognition involves an anchor-free, bottom-up convolutional neural network for object detection.

**[0065]** For the convolutional neural network to be used for identification of items/objects learned during training operations, the method proceeds with an inference process where during operation the neural network parameters are loaded into a computer processor (such as the processor mentioned above) in a neural network program that implements the convolutional neural network. During operation, the processor may then receive images from the imaging sensor and pass that image through the convolutional neural network program. The convolutional neural network then outputs a decision, indicating, for example, the type of object present in the image with highest likelihood.

**[0066]** In a training operation, the labeled data is used by a training algorithm (which may be performed by a training processor) to optimize the convolutional neural network to identify the object in the captured images with the greatest feasible accuracy. As would be readily appreciated by one of ordinary skill in the art, a number of algorithms may be utilized to perform this optimization, such as Stochastic Gradient Descent, Nesterov's Accelerated Gradient Method, the Adam optimization algorithm, or other well-known methods. In Stochastic Gradient Descent, a random collection of the labeled images is fed through the network. The error of the output neurons is used to construct an error gradient for all the neuron parameters in the network. The parameters are then adjusted using this gradient, by subtracting the gradient multiplied by a small constant called the "learning rate". These new parameters may then be used for the next step of Stochastic Gradient Descent, and the process repeated.

**[0067]** The result of the optimization includes a set of convolutional neural network parameters (which are stored in a memory) that allow the convolutional neural network to determine the presence of an object in an image. During operation, the neural network parameters may be stored on digital media. In an example of implementation, the training process may be performed by creating a collection of images of items, with each image labeled with the category of the items appearing in the image. Each of the categories can be associated with a number, for instance the conveyor belt might be 0, a carton 1, a transparent plastic bottle 2, etc. The convolutional neural network would then comprise a series of output neurons, with each neuron associated with one of the categories. Thus, neuron 0 is the neuron representing the presence of a conveyor belt, neuron 1 represents the presence of a carton, neuron 2 represents the presence of a transparent plastic bottle, and so forth for other categories.

**[0068]** The method may be designed to detect and recognize waste objects using very specific categories, product-specific categories, i.e., to classify each waste object as belonging to a specific vendor, brand, product

and/or application (food, cosmetics, other). This may be enabled by e.g., using a categorization/classification ordering/grouping of objects by application, shape/size, and color, but also by material, vendor/producer, brand and product:

- Food

- Bottle

- Transparent
- White
- Black
- Blue
- Green
- Red
- Other

- Tray

- Transparent
- White
- Black
- Blue
- Green
- Red
- Other

- Other

- Transparent
- White
- Black
- Blue
- Green
- Red
- Other

- Cosmetics

- Bottle

- Transparent
- White
- Black
- Blue
- Green
- Red
- Other

- Other

- Transparent
- White
- Black
- Blue
- Green
- Red

- Other

- Other

- 5
  - Transparent
  - White
  - Black
  - Blue
  - Green
  - 10
    - Red
    - Other

**[0069]** Thus, the method is not only able to detect and recognize a transparent bottle, but also able to identify the transparent bottle as for example a Heinz bottle, such as for example a Heinz Tomato Ketchup bottle, or even as a Heinz Organic Tomato Ketchup bottle, or even as a Heinz Organic Tomato Ketchup bottle 580 g

**[0070]** For the convolutional neural network to be used for identification of items/materials learned during training operations, the method proceeds with an inference process where the neural network parameters are loaded into a computer processor (such as the processor mentioned above) in a neural network program that implements convolutional neural network. During operation, the processor may then receive images from the imaging sensor and pass that image through the convolutional neural network program. The neural network then outputs a decision, indicating, for example, the type of item/material present in the image with highest likelihood.

**[0071]** In an embodiment the object is a plastic object. The object may be made from plastic material such as e.g., PE, PP, PS, PET, PVC, PVA or ABS. The large amounts of plastic used at present day generates large amounts of plastic waste and the present invention provides a method for efficient sorting of plastic material.

**[0072]** The invention also provides a system for sorting objects, the system comprising:

40 at least one imaging sensor;

a controller comprising a processor and a memory storage, wherein the controller receives image data captured by the at least one imaging sensor; and

45 at least one sorting robot coupled to the controller, wherein the at least one sorting robot is configured to receive an actuation signal from the controller;

50 wherein the processor executes an object identification module configured to detect objects travelling on a conveyor belt and recognize at least one target item travelling on a conveyor belt by processing the image data and to determine an expected time when the at least one target item will be located within a diversion path of the sorting robot; and

55 wherein the controller selectively generates the ac-

tuation signal based on whether a sensed object detected in the image data comprise the at least one target item.

#### DETAILED DESCRIPTION OF THE INVENTION

**[0073]** The invention will now be described in further details with reference to drawings in which:

Figure 1: shows an embodiment with a conveyor and a robot;  
 Figure 2: shows an embodiment with just a conveyor;  
 Figure 3: shows an embodiment without a conveyor (nor robot);  
 Figure 4: shows a detailed view of the invention;  
 Figure 5: shows the inference and training process for a neural network method;  
 Figure 6: shows a method using feature matching for logo/symbol detection;  
 Figure 7: illustrates the principles of neural network object detection;  
 Figure 8: illustrates the principles of two-stage neural network object detection;  
 Figure 9: shows a method using feature matching for logo/symbol detection;  
 Figure 10: shows the principles of text detection and recognition;  
 Figure 11: shows the principles of a method for object detection using a bottom-up, anchor-free neural network;  
 Figure 12: shows an embodiment linking high resolution with a neural network; and  
 Figure 13: shows examples of symbols, which can be detected by the method.

**[0074]** The figures are only intended to illustrate the principles of the invention and may not be accurate in every detail. Moreover, parts which do not form part of the invention may be omitted. The same reference numbers are used for the same parts.

**[0075]** Figure 1 is a diagram showing the principles of the invention. Reference number 1 indicates the conveyor belt. Box 2a illustrates the "scene" on the conveyor belt 1, i.e., the conveyor belt with one or a number of items. The scene 2a reflects light, which is registered by the camera 3 and transformed into an image. The image is processed in a product detection and recognition module 4 to identify the item or items present in the scene 2a. The information from the product detection and recognition module 4 is sent to the sorting control 5, which may obtain further information about the identified items from the product database 6.

**[0076]** The sorting control 5 communicates with a robot controller 7 which control a robot 8, which is physically able to intervene in scene 2b in a sorting area on the conveyor belt 1 and sort the item or items into specific categories of waste material.

**[0077]** The speed of the conveyor belt 1 is monitored,

and an encoder 9 sends information about the speed of the conveyor belt 1 to a synchronizer 10. The synchronizer sends signals to the camera 3 and determines when the conveyor is in a position where the camera 3 should capture an image. The synchronizer also sends signals to the robot controller 7 with information about when the scene 2b reaches the sorting area. The encoder 9 may also send signals directly to the robot controller 7.

**[0078]** Scene 2a and scene 2b are in principle identical, and the reference numbers only indicate that the conveyor belt has moved the scene a distance from the point where scene 2a was registered by the camera 3.

**[0079]** Figure 2 illustrates the principles of the conveyor belt information system. The speed of the conveyor belt is monitored, and the information about the speed is transformed by the encoder 9 and send as an encoder signal to the synchronizer 10. The synchroniser 10 sends a signal to the camera 3 when an image of the scene 2a needs to be provided. Depending on the actual speed of the conveyor belt the camera may provide several images of the scene 2a per second. However, if the speed of the conveyor belt is slow the camera 3 only needs to provide a few images per minute.

**[0080]** The images from the camera 3 are sent to the product detection and recognition module 4 to be processed and the items in the image are identified. The information about the identified items is then sent to the visualization and statistics module 5a for further processing to display or otherwise provide the information that can be extracted or accumulated from the detection system.

**[0081]** The visualization and statistics module 5a communicate with the product database 6 to obtain more detailed information about product properties for an identified item. The information about product properties may e.g., be information about material. Figure 3 illustrates the principles of the information system. The information system includes the camera 3, the product detection and recognition module 4, the visualization and statistics module 5a and the product database 6.

**[0082]** The images from the camera 3 are sent to the product detection and recognition module 4 where the items on the images (appearing on the scene 2a) are identified.

**[0083]** The camera 3, the lighting and the conveyor speed must be adjusted to provide images which meet the requirements, e.g., images with sufficient lightning and with little motion blur.

**[0084]** The information about the identified items is then sent to the visualization and statistics module 5a for further processing.

**[0085]** The visualization and statistics module 5a communicate with the product database 6. The visualization and statistics module 5a can search the product database 6 and obtain more detailed information about product properties for an identified item. The information about product properties may e.g., be information about material.



**[0086]** Figure 4 shows the principles of the product detection and recognition module. The image distributor receives and image and distributes the image to one or more of a product detection module(s) (which may comprise a neural network product detection module, and/or a feature-based product detection module), a logo detection module (ditto), a symbol detection module (ditto), and a text detection and text+font recognition module.

**[0087]** The information which is deduced from the product detection module(s) (neural network product detection module, and/or feature-based product detection module), the logo detection module, and the symbol detection module are sent to the recognition module for further processing.

**[0088]** The information from the text detection and text+font recognition module is further processed in the vendor name recognition module, the brand name recognition module, the product name recognition module, the slogan recognition module, and product description recognition module, before the information is sent to the product recognition module for further processing.

**[0089]** The product recognition module may include prior information (such as a Bayesian prior over the likelihood of objects, such as product objects).

**[0090]** The product recognition module is integrated in the product detection and recognition module.

**[0091]** Figure 5 illustrates the principles of inference and training for a neural network object detection module, such as the neural network product detection module.

**[0092]** In a training process, a camera provides images, which are stored in an image database. An annotation process involves human and machine annotation of images in the image database. Each image is annotated with locations/boxes/shape of objects, where each object is annotated with its class. Classes stem from a product classification, which classifies products by properties. The annotated images resulting from the annotation process is stored in database. A neural network object detection model with a minimum number of classes ( $K \geq 10$ ), is optimized in the Neural Network Training Process with respect to a set of annotated images. The trained/optimized neural network is stored in a database.

**[0093]** In an inference process a camera provides images to a neural network inference algorithm which detects objects (e.g., product objects), outputting pairs of locations and classes.

**[0094]** Figure 6 illustrates a method for logo and symbol detection as shown in figure 4.

**[0095]** In the logo detection module and symbol detection module the overall detection principles are generally the same. When the modules receive an image from the image distributor, the image is first processed in a feature extraction module, extracting local features. The information is sent to a feature description module which describes the local features and sent the information to a matching module. The matching module interacts with a feature descriptor database which can provide further information about the features. From the matching mod-

ule, matched local feature descriptors are sent to a clustering module, which determines clusters of features which stem from the same object (using e.g., geometric model verification), before the information is provided to the product recognition module for further processing.

**[0096]** A prerequisite for the logo/symbol detection is a database of reference images of logos/symbols to be detected and recognized. Features from reference images are extracted, and descriptors are computed for each feature, before storing the features and their descriptors in a feature descriptor database.

**[0097]** Figure 7 illustrates the general principles of neural network object detection. The image is sent to the convolutional neural network for processing and the convolutional neural network sends a compressed image representation to a feature fusion network, which in turn sends the fused image features to an object detection module which detects the objects.

**[0098]** During the process the convolutional neural network, the feature fusion network and the object detection module interact with the images and annotations database. Neural network parameters are learned in the training phase from images and annotations. It is the learned model that is extracted from the images and annotations which is interacted with during operation/processing.

**[0099]** Figure 8 illustrates the general principles of two-stage neural network object detection.

**[0100]** An image is distributed from the image distributor module. The image is sent to the convolutional neural network and the object recognition module. The convolutional neural network sends compressed the image representation to the object detection module which detects the objects and sends the information to the object recognition module, which recognizes the objects.

**[0101]** The convolutional neural network, the object detection module, and the object recognition module interact with the images and annotations database during the detection and recognition process. The neural network parameters are learned in the training phase from images and annotations. It is the learned model that is extracted from the images and annotations which is interacted with during operation/processing.

**[0102]** Figure 9 is identical to Figure 6 - Figure 9 illustrates the same feature-based detection and recognition for symbols/logos which Figure 6 shows for product objects.

**[0103]** Figure 10 illustrates in more details the principles of text detection and recognition carried out in the text detection and text+font recognition module.

**[0104]** When the text detection and text+font recognition module receive an image from the image distributor, the image is first processed in a convolutional neural network which send a compressed image representation to a text detection module which again sends text boxes to a text recognition module and font recognition module. The text recognition module 25b and the font recognition module provides information about text and font to the modules in figure 4. After processing in the modules ,

text information is provided to the product recognition module.

**[0105]** During the processing of the image, the convolutional neural network, the text detection module, and the text recognition module interact with a images and annotations database. The images and annotations database is a training database which supports the image the convolutional neural network. Neural network parameters are learned in the training phase from images and annotations. It is the learned model that is extracted from the images and annotations which is interacted with during operation/processing.

**[0106]** Figure 11 shows the general principles of bottom-up, anchor-free neural network for object detection. The image is sent to the convolutional neural network for processing and the convolutional neural network sends compressed image representation to a keypoint pooling network, which in turn sends information about pooled features to a heatmap network which detects the objects.

**[0107]** During the process the convolutional neural network, the keypoint pooling network and the heatmap network interact with the images and annotations database. Neural network parameters are learned in the training phase from images and annotations.

**[0108]** It is the learned model that is extracted from the images and annotations which is interacted with during operation/processing.

**[0109]** Figure 12 illustrates an embodiment where an image with high resolution is linked to a neural network for object detection. The architecture of the network is adapted to the high resolution in the images by neural network layers in the beginning of the network. The embodiment corresponds to the embodiment shown in figure 7 but adapted for images with high resolution.

**[0110]** Figure 13 illustrates examples of symbols which can be detected by the method according to the invention.

## Claims

1. A method for sorting consumer packaging objects travelling on a conveyor belt, the method comprising:

receiving image data captured by at least one imaging sensor for an image comprising at least one feature on or of an object travelling on the conveyor belt said imaging sensor providing color image data with a spatial resolution of at least 0.4 px/mm;

executing a product detection and recognition module on a processor, the product detection and recognition module being configured to detect characteristics of the at least one feature on or of the object travelling on the conveyor belt by processing the image data and recognizing the object as one of at least 10 consumer packaging product objects and/or recognizing the object as one of at least 40 consumer packaging

brand objects; and

wherein the detection and recognition are based on one or more of the following: the characteristics of the shape of the object, the characteristics of the color/colors of the object, the characteristics of image features on the object in at least three areas on the object; and when an object has been detected and recognized determining an expected time when the at least one object will be located within a sorting area of at least one sorting device; and selectively generating a device control signal to operate the at least one device on whether the at least one object comprises a target object.

2. A method according to claim 1, wherein the object travelling on the conveyor belt is recognized as one of at least 20 consumer packaging product objects, such as one of at least 50 consumer packaging product objects, such as one of at least 80 consumer packaging product objects, such as one of at least 100 consumer packaging product objects, such as one of at least 1000 consumer packaging product objects
3. A method according to claim 1 or 2, wherein the object travelling on the conveyor belt is recognized as one of at least 80 consumer packaging brand objects, such as at least 100 consumer packaging brand objects, such as at least 500 consumer packaging brand objects, such as at least 1000 consumer packaging brand objects.
4. A method according to anyone of the preceding claims, wherein the consumer packaging objects travelling on the conveyor belt is at least partly a consumer packaging waste stream comprising packaging materials, such as plastic or cardboard or other recyclables.
5. A method according to anyone of the preceding claims, wherein the method is adapted for detecting and recognizing objects used as packaging or containers for food items, such a bottles and trays, preferably the method is adapted for detecting and recognizing black objects, such as a black tray for food.
6. A method according to anyone of the preceding claims, wherein the sorting device is a lifting device adapted for lifting the consumer packaging object away from the conveyor belt.
7. A method according to anyone of the preceding claims, wherein the sorting device is a lifting device adapted for lifting the object in such a way that the side facing the conveyor belt can be captured by an image sensor.

8. A method according to anyone of the preceding claims, wherein at least two image sensors are applied, said image sensors are arranged such that the image data of the object is captured from different angles. 5
9. A method according to anyone of the preceding claims, wherein the target object is guided to a collection device in the sorting area by means of the sorting device. 10
10. A method according to anyone of the preceding claims, wherein said spatial resolution is at least 2 px/mm, preferably said spatial resolution is at least 4 px/mm. 15
11. A method according to anyone of the preceding claims, wherein the characteristics of image features on the object is detected and recognised, in at least five areas on the object, such as in at least ten areas on the object. 20
12. A method according to anyone of the preceding claims, wherein product detection and recognition involves a convolutional neural network and wherein the convolutional neural network is selected from a convolutional neural network for object detection, a one-stage convolutional neural network for object detection, a one-stage convolutional neural network for object detection with a feature fusion network, a one-stage bottom-up, anchor-free convolutional neural network for object detection and combinations thereof. 25 30
13. A method according to anyone of the preceding claims, wherein the characteristics are selected among one or more of a logo, a symbol, text and font. 35
14. A method according to claim 13, wherein the characteristics of text and font may comprise one or more of a vendor name, a brand name, a product name, a slogan and/or product description. 40
15. A method according to anyone of the preceding claims, wherein the detection and recognition is based on detection and recognition of at least one of the following: product detection and recognition, logo detection and recognition, symbol detection and recognition, text and font detection and recognition and where the information provided by the one or more detection and recognition methods is fused in a statistical framework yielding one single output of the product detected, preferably the statistical framework includes information on the likelihood of products. 45 50 55

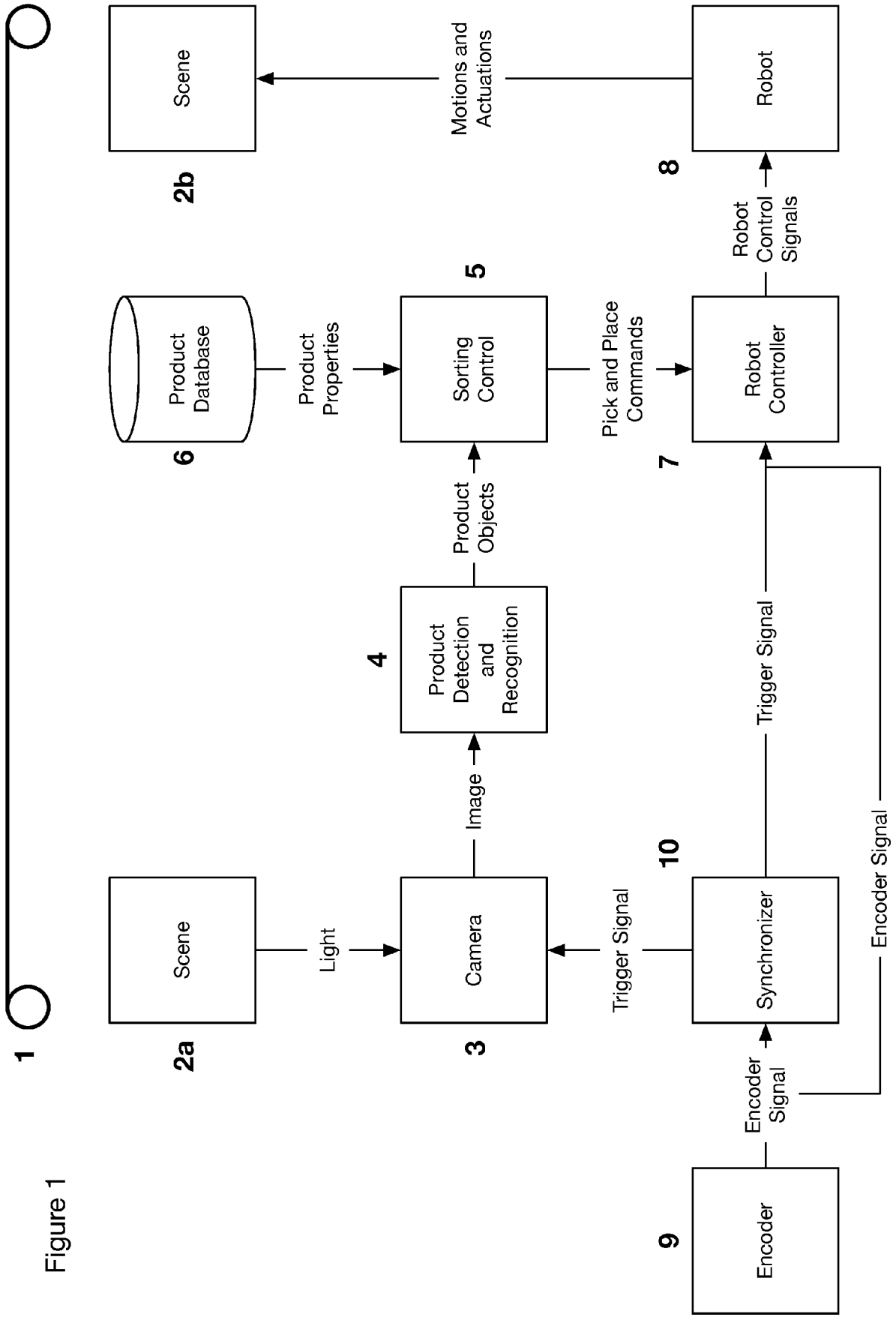


Figure 2

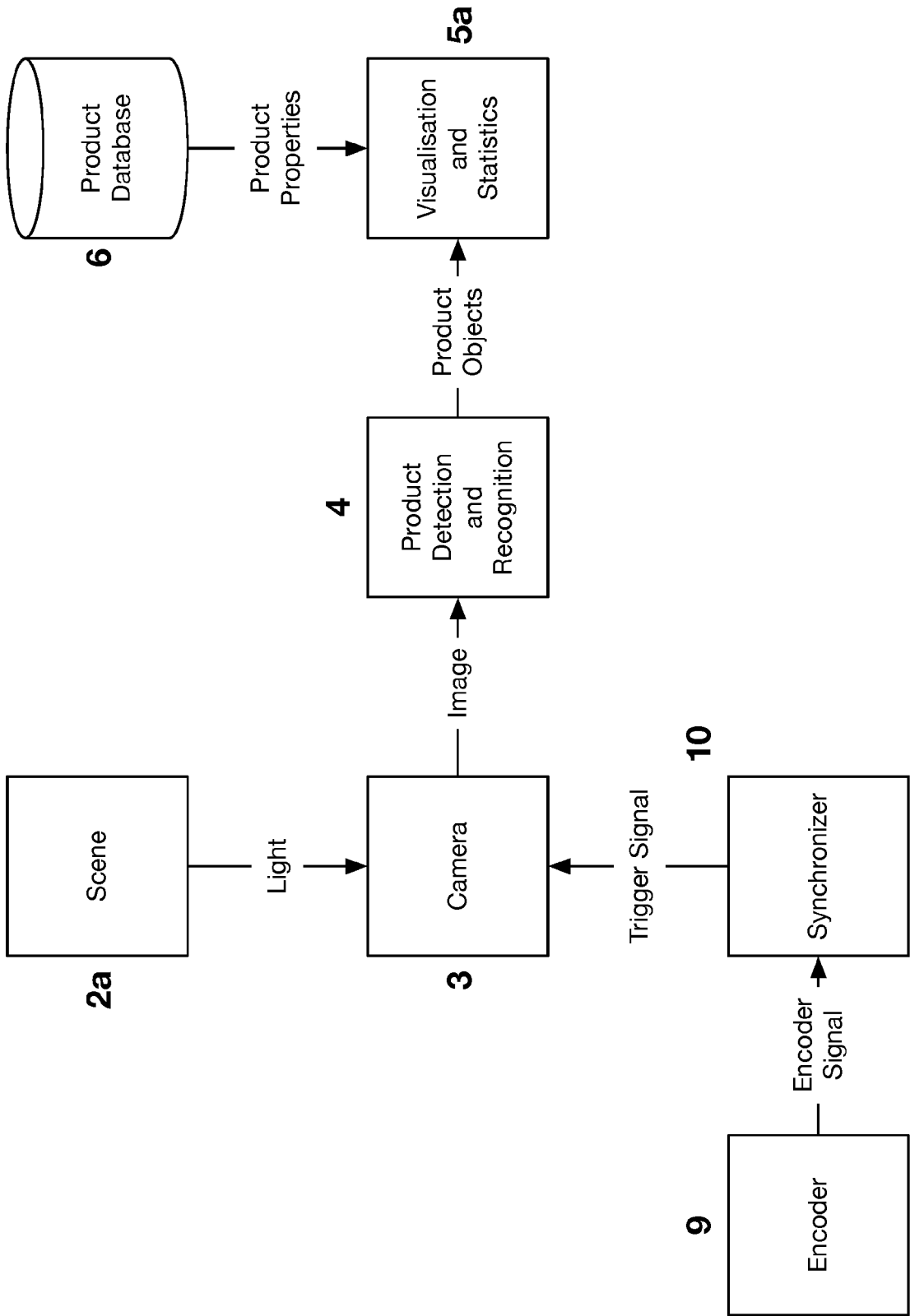
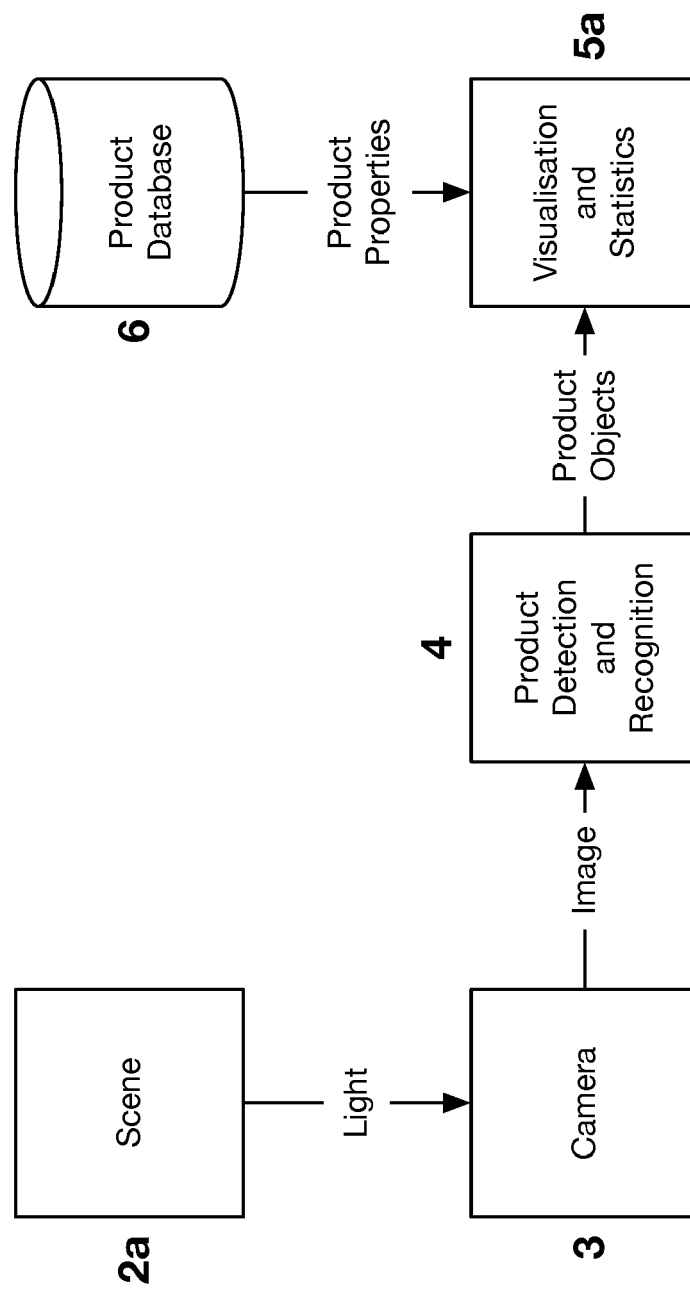


Figure 3



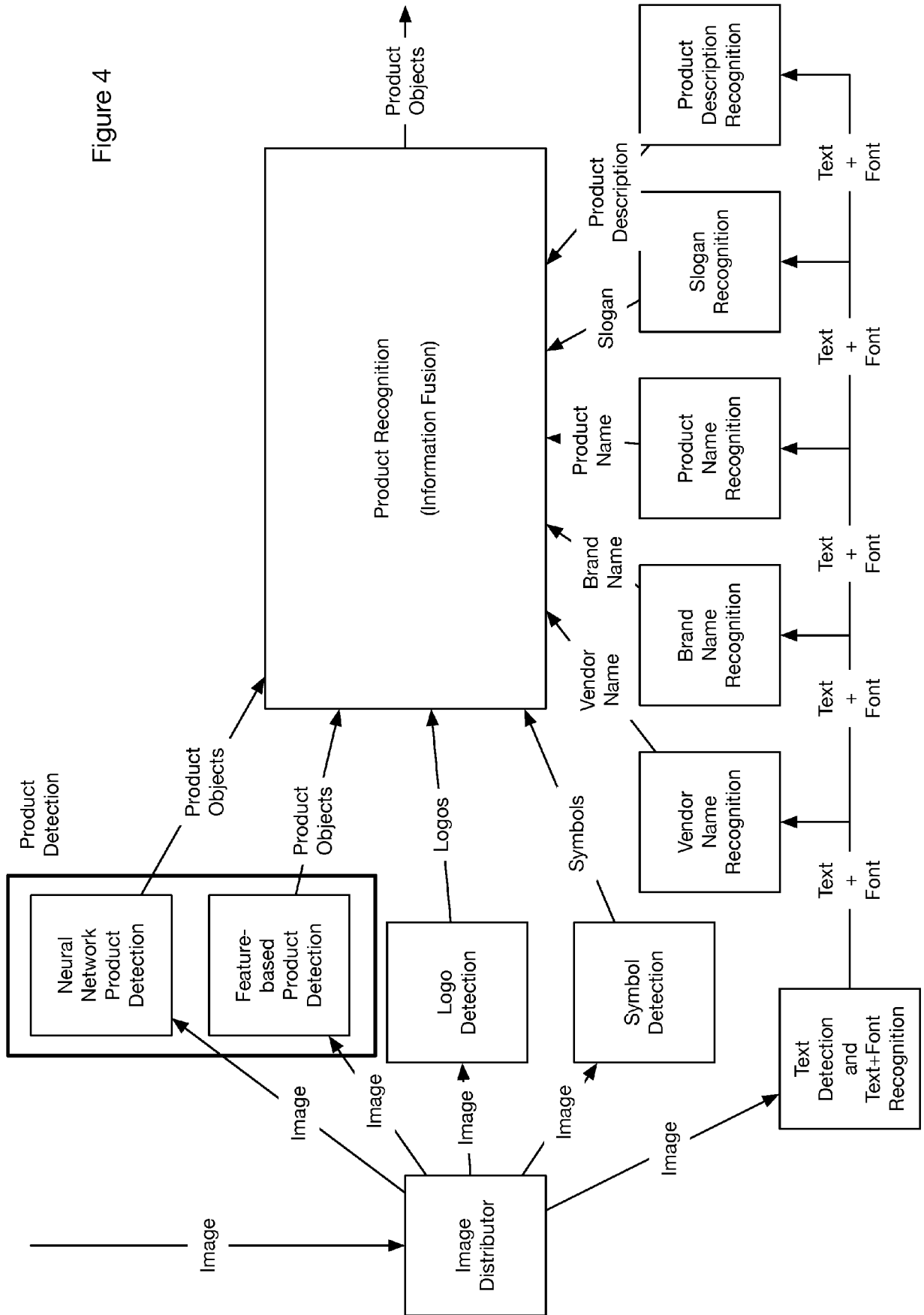
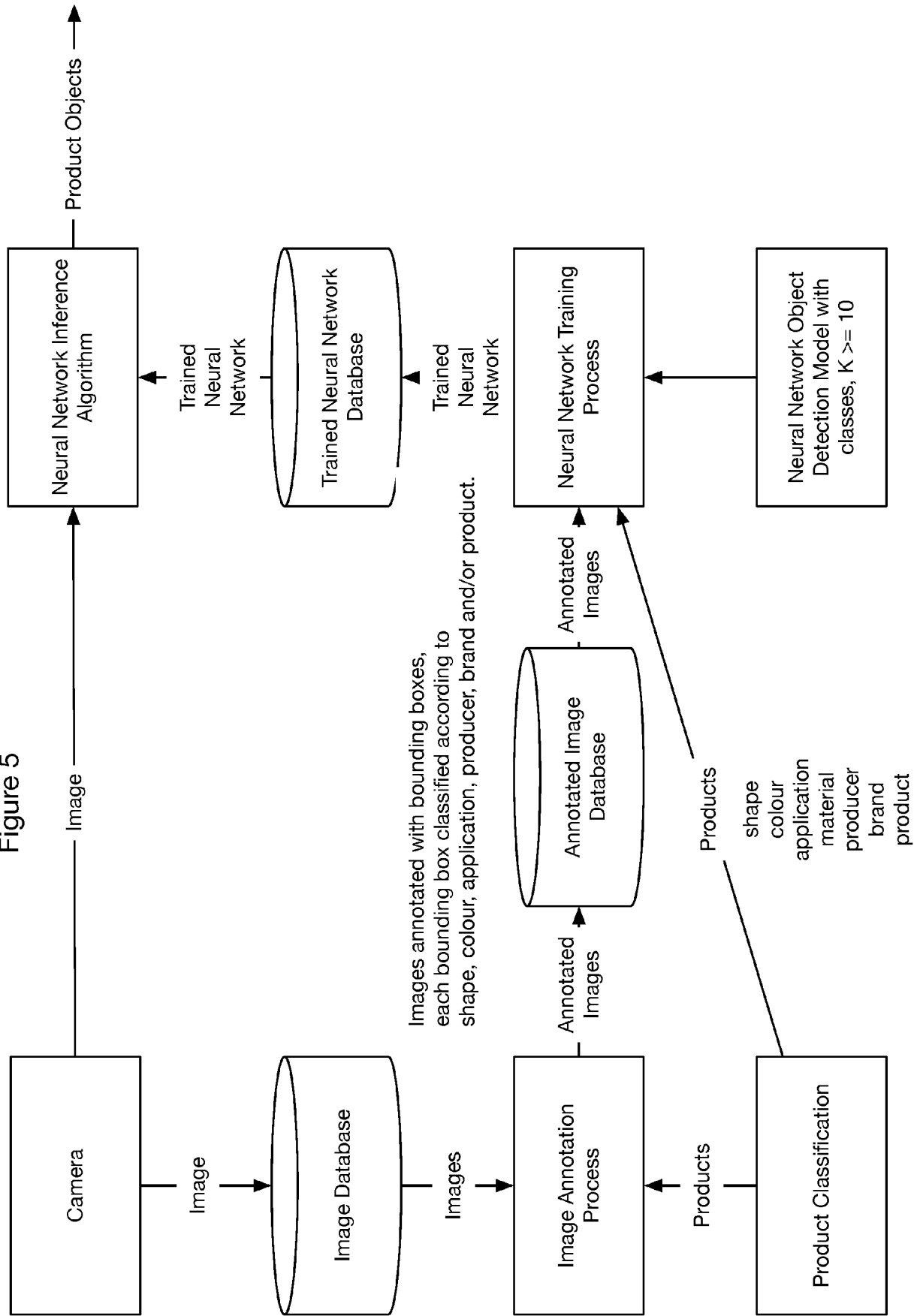


Figure 4

Figure 5





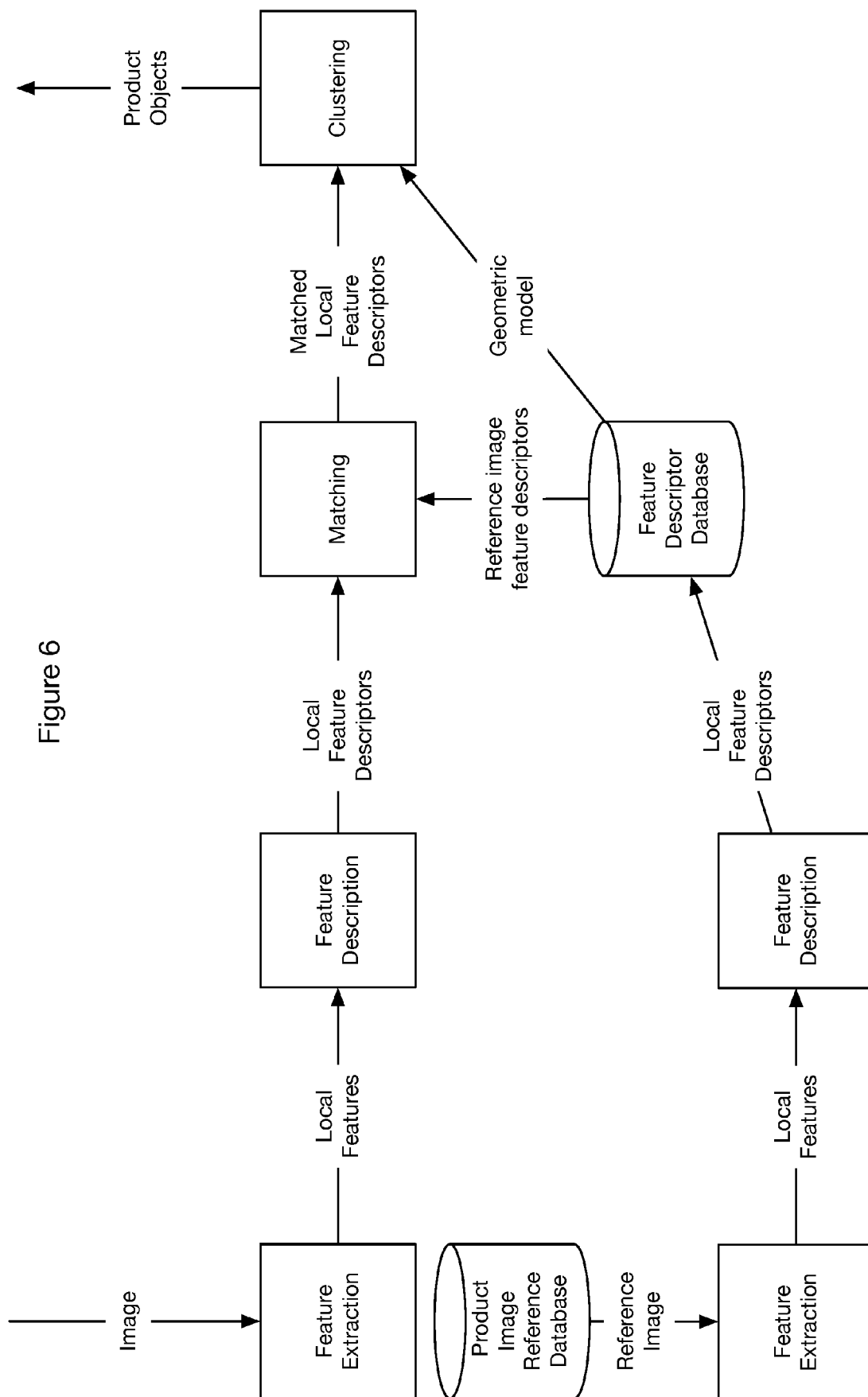
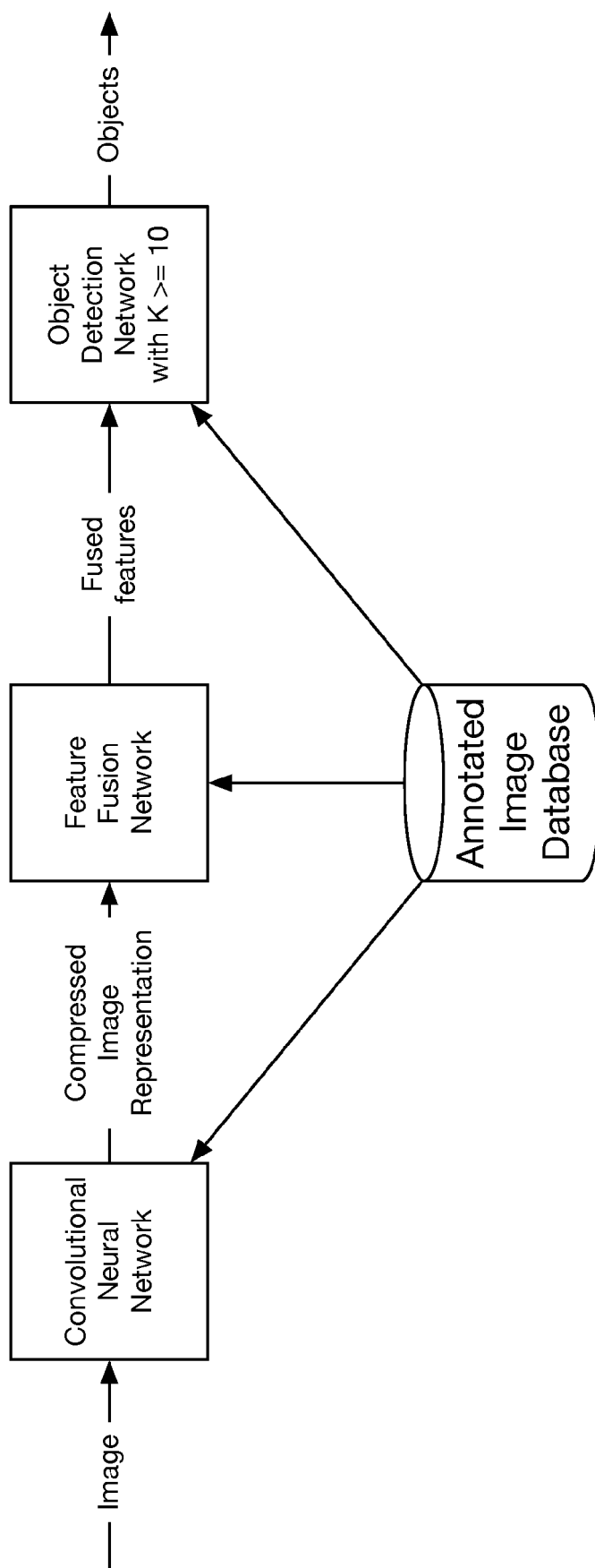


Figure 6

Figure 7



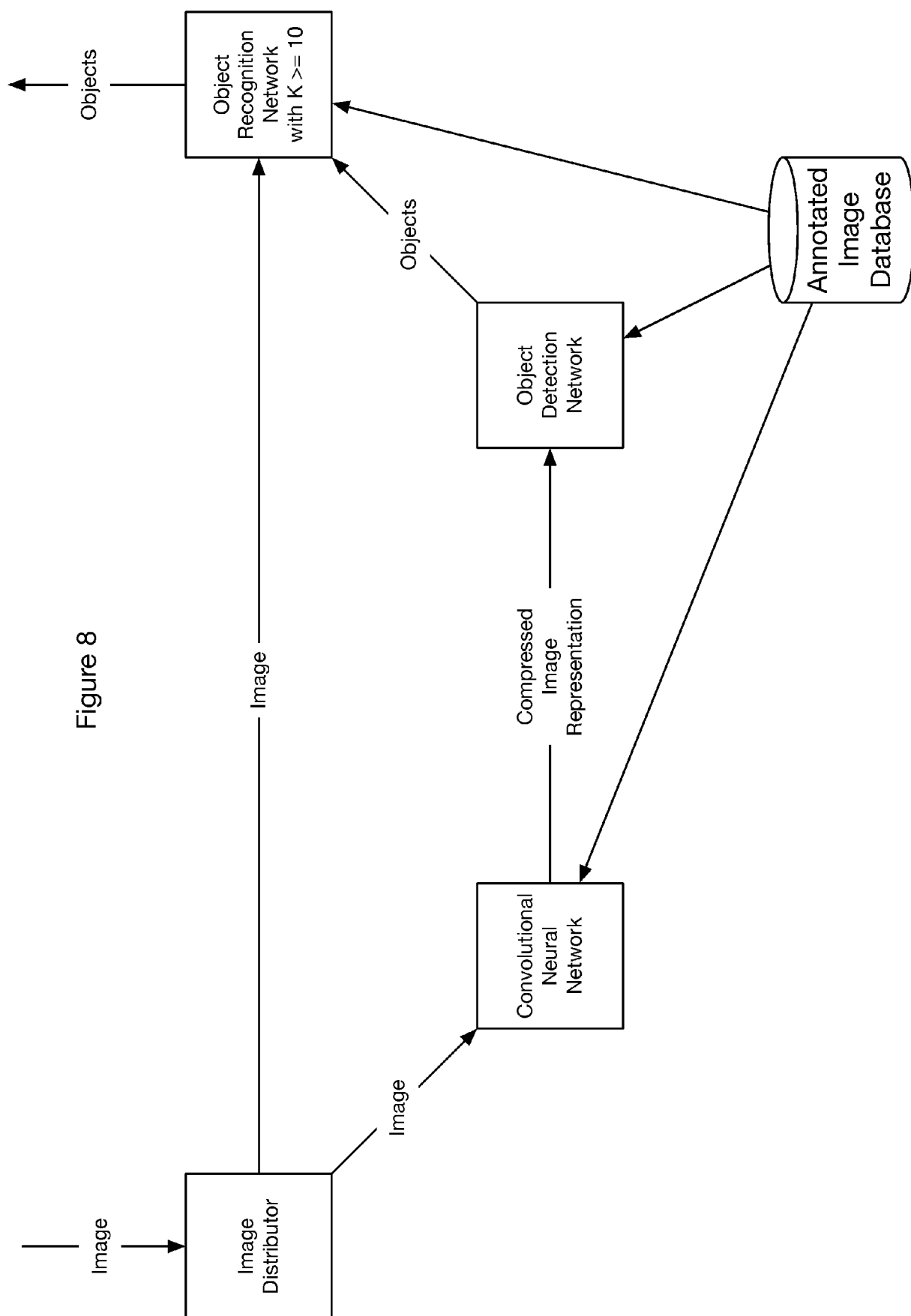


Figure 8

Figure 9

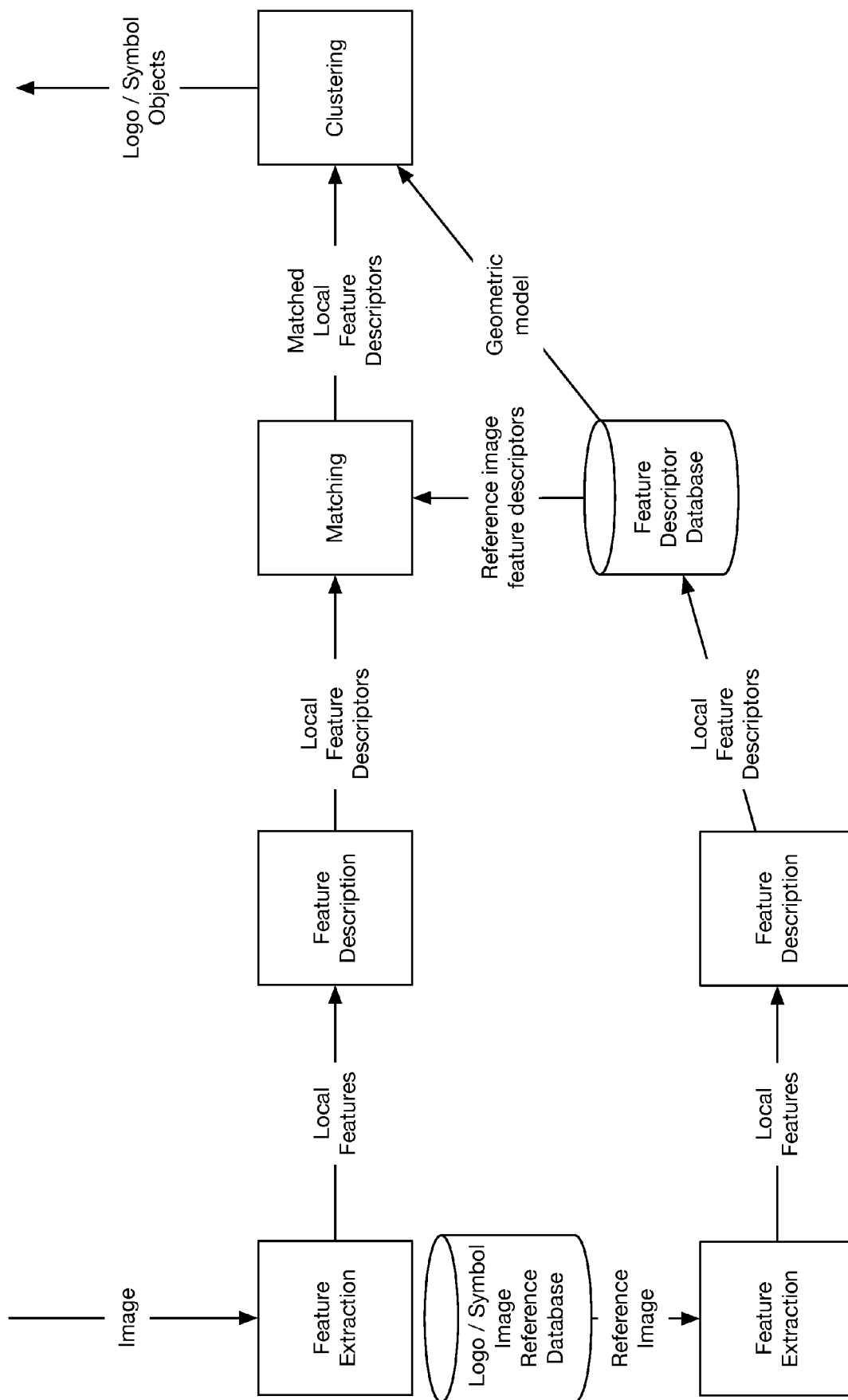


Figure 10

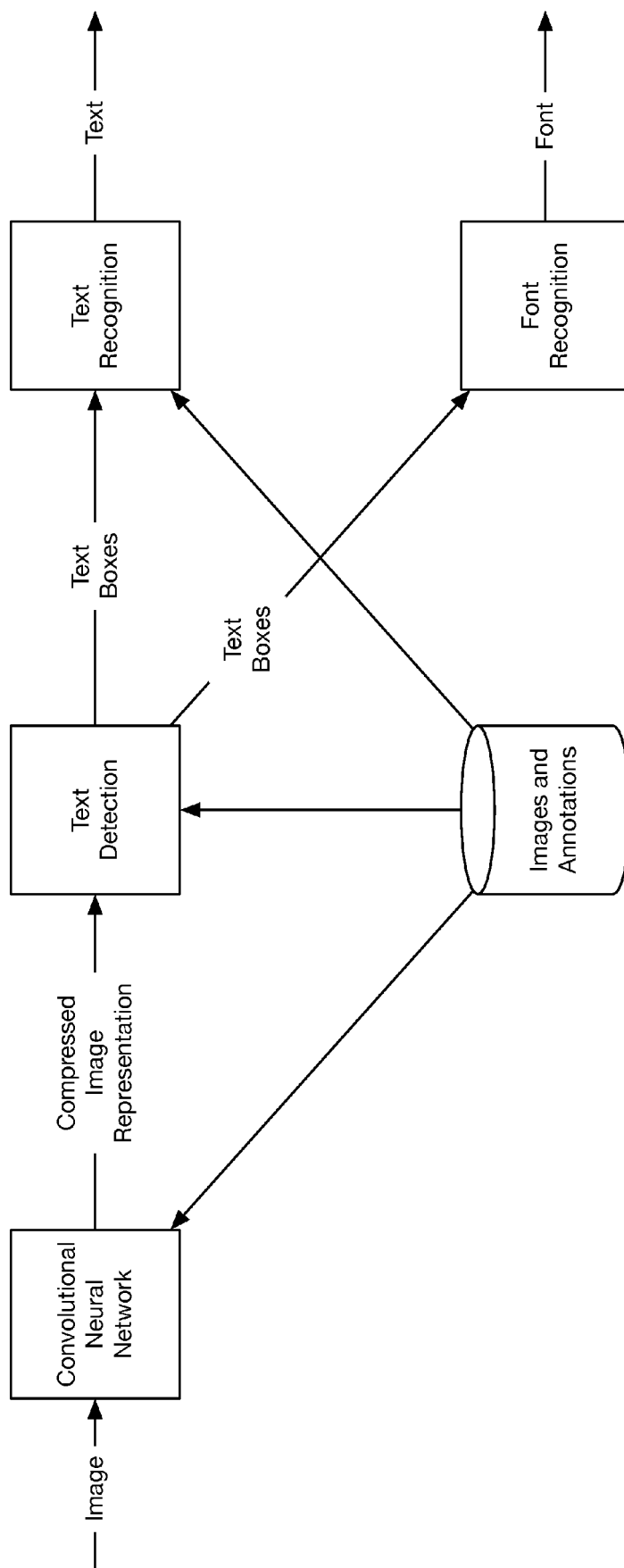


Figure 11

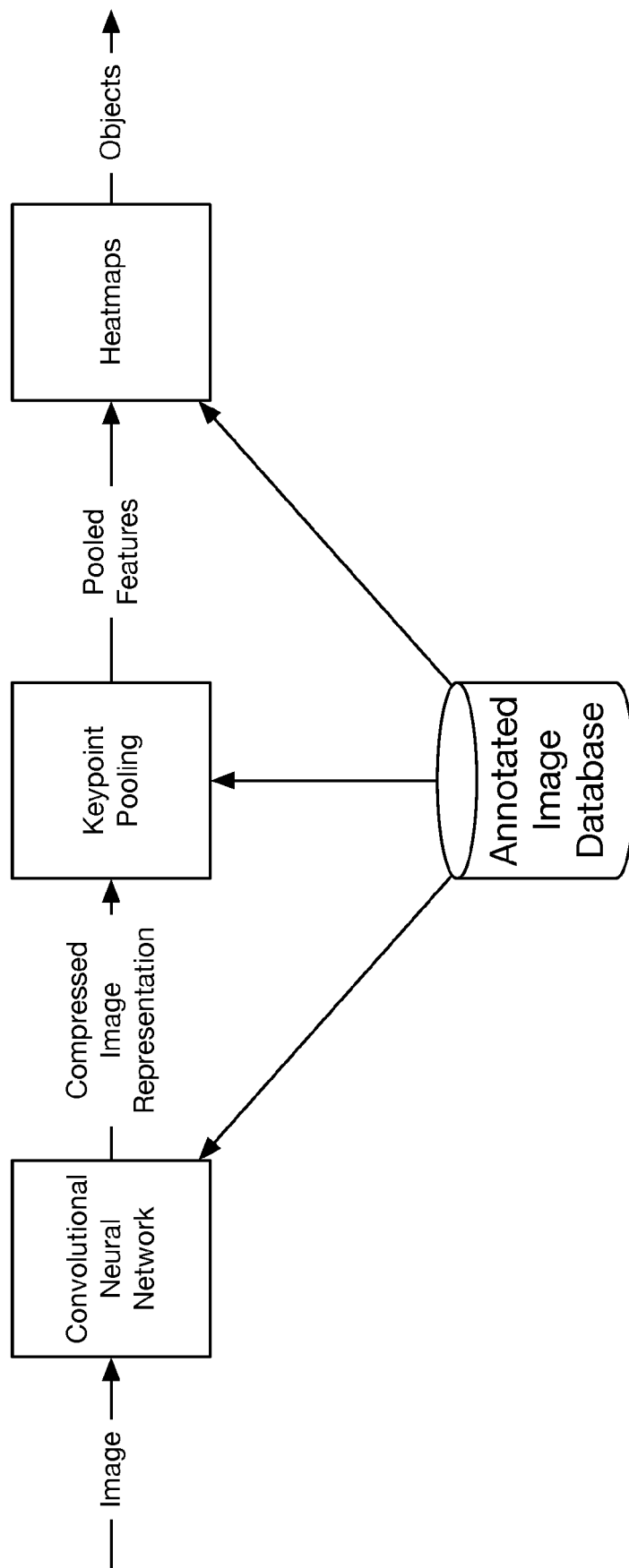
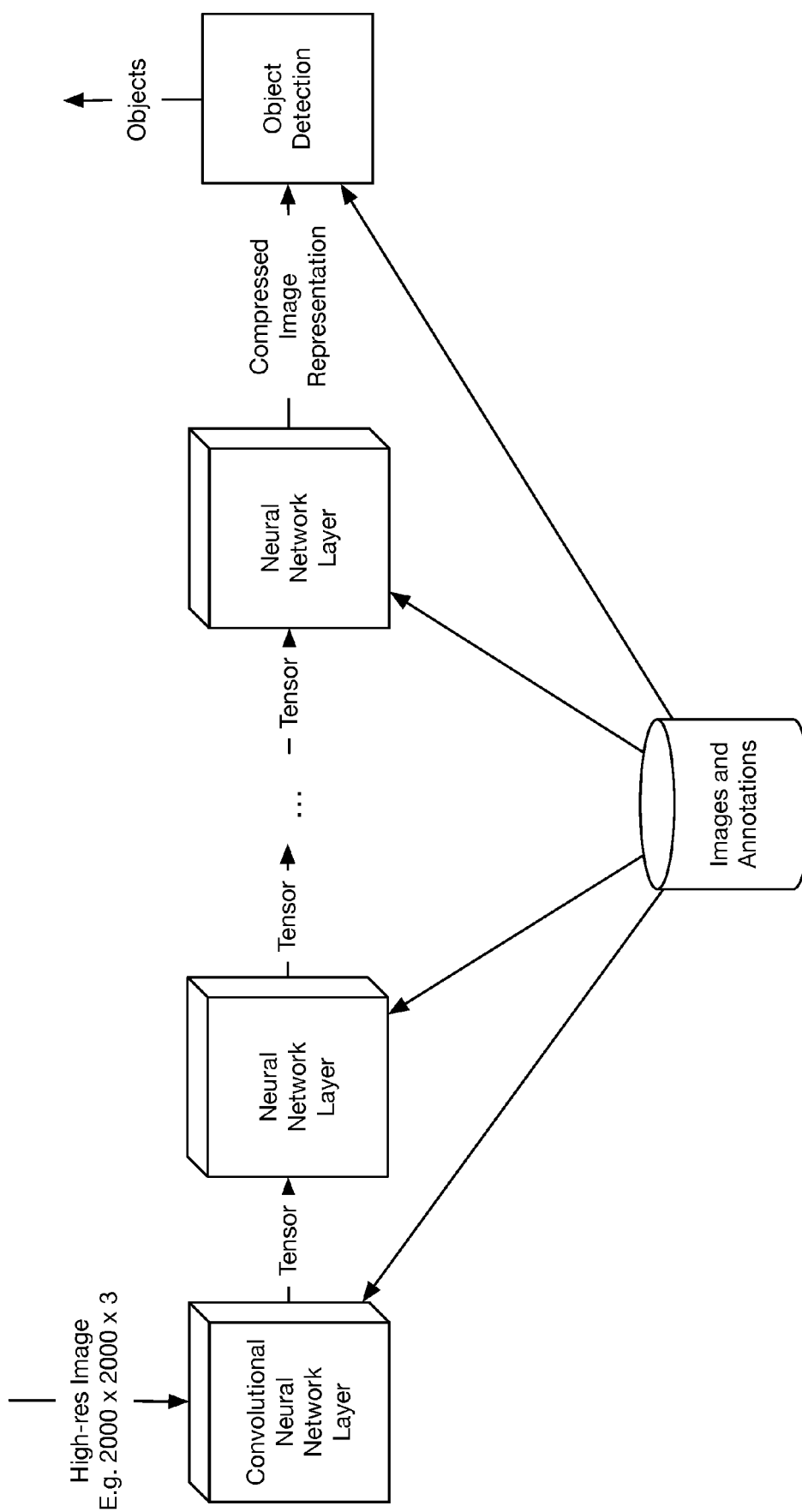
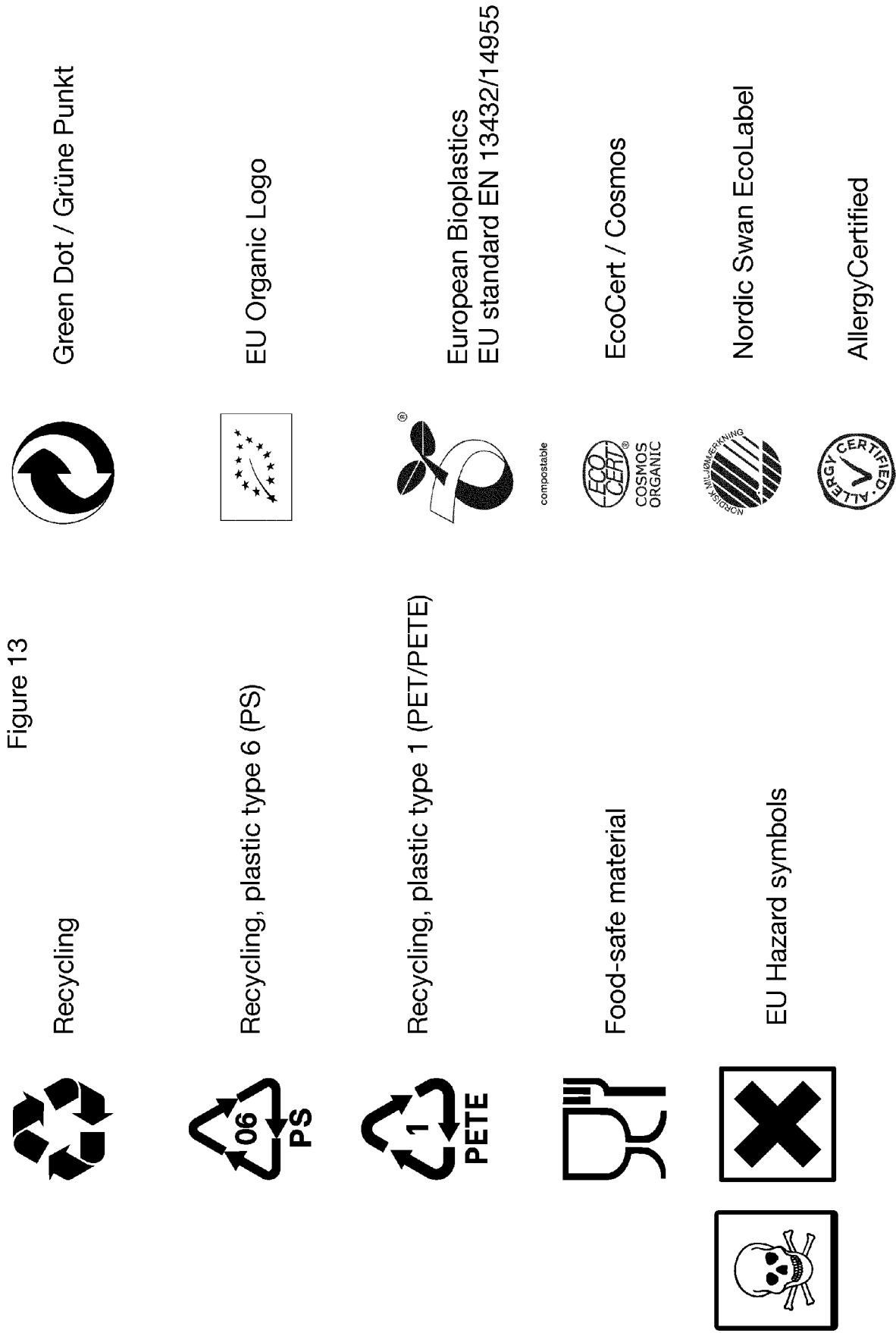


Figure 12









## EUROPEAN SEARCH REPORT

Application Number  
EP 20 21 5996

5

10

15

20

25

30

35

40

45

50

55

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	WO 2017/176855 A1 (WASTE REPURPOSING INT INC [US]) 12 October 2017 (2017-10-12) * figures * -----	1-15	INV. B07C5/342 B07C5/34
			TECHNICAL FIELDS SEARCHED (IPC)
			B07C
The present search report has been drawn up for all claims			
Place of search <b>Munich</b>		Date of completion of the search <b>8 July 2021</b>	Examiner <b>Wich, Roland</b>
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			

 1  
EPO FORM 1503 03.82 (P04C01)

