FusionSystems GmbH | Annaberger Straße 240 | 09125 Chemnitz | www.fusionsystems.de | info@fusionsystems.de

OPTmicro: Labeling Guide

Version: 1.1

Content

1 Introduction

1.1 Scope

For the development of machine vision algorithms high quality annotations of relevant object classes are needed. The project OPTmicro aims to create and publish such high-quality annotations in the context of automatization of micro-mobiles like cargo bikes. The aim of this Labelling Guide is the definition of rules that shall be obeyed at creating such annotations.
The sensor data to be labelled consists of:

- **color images** coming from a color camera
- **point clouds** coming from fused lidar point clouds

The sensor frames are synchronized in time and the sensors are calibrated to each other. An **annotation** consists of one **annotation geometry** (Bounding Box, Polygon, …) and the associated **attribute values** for class-specific attributes.

1.2 Annotation geometries

Annotation geometries are used to mark the location of an object in a sensor frame. The following annotation geometries are used here:
For the color images:

- **2D Bounding Box:** 2D Bounding Boxes are rectangles that enclose an object. Here, the boxes are always axis-aligned to the x- and y-axis of the image.
- **2D Polygon:** 2D Polygons are plane figures consisting of a closed chain of line segments that are used for the annotation of areas. They have at least three and at most 30 anchor points.

For the lidar point clouds:

1. **3D Bounding Box:** 3D Bounding Boxes are cuboids that enclose an object. Here, the boxes can be rotated around the x-, y- and z-axis of the lidar coordinate system.
2. **3D Polygon:** 3D Polygons are spatial figures consisting of a closed chain of line segments that are used for the annotation of areas. They have at least three and at most 30 anchor points.

1.3 Identifiers

A unique identifier (ID) shall be assigned to each **annotation geometry**. This ID is called **geometry ID**. Additionally, a unique identifier shall be assigned to each **real-world object**. This ID is called **object ID**. Each annotation geometry shall be **assigned** to a real-world object by an object ID. All unique identifiers shall follow the **Universally Unique Identifier** (UUID) standard in **version 4**.

1.4 Attributes

Attributes are used to describe some details of an object. The following attributes are used here:

- type: This is a **single-select** attribute, meaning that attribute value is exactly one single value out of a list of pre-defined options of the data type string.
- occlusion: This is a **single-select** attribute. The occlusion shall define the **relative size** (in terms of an area) of the portion of the object that is **hidden** by other objects. Parts

of the object that are truncated by image borders shall not be considered as occluded. For lidars, the **cross-section area** perpendicular to the vector connecting the center of the object with the origin of the 3D reference coordinate system is used for the estimation of the occlusion. After its estimation the occlusion is rounded down to an integer. The occlusion shall be set per annotation geometry.

- connectedObject: This is a **multi-select** attribute, meaning that the attribute value of shall be a subset of a list of pre-defined options of the data type string. The subset can be empty when none of the options fits. connectedObject must be set consistently for all sensor frames of one given point in time.
- connectedPersons: This is a **single-select** attribute, see type.

1.5 Sensor und recording setup

The sensor setup consists of the following sensor:

- 2 Lidars that are fused together and that are recording the surrounding of the ego vehicle in a 360° field of view. The **fused Lidar point** clouds shall be annotated.
- 3 Color cameras that are recording the area in front (left front, central front, right front) of the ego vehicle. All **3 camera frames** shall be annotated.
- 3 Color stereo cameras that are recording the area at the left and right side of the ego vehicle as well as behind the ego vehicle. These 3 camera frames shall **not be annotated** here but serve as an **orientation** for the interpretation of the lidar point clouds.
- Cinematic sensors that shall not be annotated here.
- GNSS sensors that shall not be annotated here.

The sensor data to be annotated is sampled with a frequency of 10 Hz, each sensor data sequence has a length of 10 seconds at most.

1.6 Data format

The annotation data must be stored in the **ASAM OpenLABEL format** (https://www.asam.net/standards/detail/openlabel/).

2 General Annotation Rules

There are rules to be observed for the annotation of the various object classes. In the following, the sensor-independent rules are described first. Then, the camera and lidar specific rules are considered.

2.1 Sensor-Independent Rules

The general annotation rules are applied to all sensors and all object classes. If there is a conflict between the general and sensor-specific rules, the sensor-specific rules have priority. The generally applicable rules are:

a) In principle, **general and sensor-specific** annotation rules shall apply for all annotations. In case of contradiction, class-specific annotation rules shall **overrule** general annotation rules.

b) Objects shall be labelled in each sensor in which they are **visible** regardless of their visibility in other sensors.
c) All points or pixels of a real-world object shall be **enclosed** by the annotation geometry. That means that no point or pixel that belongs to the annotated object shall be outside of the enclosing annotation geometry.
d) The smallest annotation geometry shall be chosen so that the corresponding objects is surrounded **tightly**. That means that the correct annotation geometry has the smallest area/volume of all possible geometries that fulfil rule c).
e) If an attribute has the option "**unknown**", this option shall only be selected when the actual value is hard to determine due to a lack of visual information, e.g. due to high occlusion. If it is uncertain whether to select one value or the other, then the most likely of these two options shall be selected.
f) If an attribute has the option "**other**", this option shall be selected when none of the other options fits.
g) If the borders of an object are **fringed**, then these fringes shall be included in the annotation geometry.
h) The **real size** of an object shall be **estimated** when the object is **occluded** by other objects. The attribute "**occlusion**" must be set appropriately.
i) If the representation of a real-world object (like a vehicle lane) in a camera frame consists of **several disjunct areas** (e.g. due to truncation at image boundaries) then several annotation geometries with the **same object ID** but different geometry IDs shall be annotated.

2.2 Sensor-Specific Rules

Lidar and cameras are the two sensor types used. Due to different properties of the sensor data, specific rules for the two sensor types must be obeyed.
The specific annotation rules for the **lidar point clouds** are:

j) Only objects that consist of **three or more** 3D **points** shall be labelled. If a real-world object consists of less than three points in some M-Frames (and three or more 3D points in other M-Frames), then this object **can** also be annotated in the M-Frames where it consists of less than three points (due to consistency reasons).
k) 3D Bounding Boxes shall be **rotated** around the coordinate axes if necessary to fulfil rule d).
l) Each geometry shall have a **precision of 10 cm**, meaning that the edges of the annotation geometries are not allowed to have a distance of more than 10 cm from the edges of the "true" annotation geometry. In cases where this precision cannot be achieved without violating the maximum number of anchor points of a polygon (see subsection 1.2), the precision can be lower than 10 cm - but should be still as high as possible.
m) When an object shall be annotated with a **3D Polygon**, its **base area** shall be annotated with this Polygon.

The specific annotation rules for the **camera images** are:

n) An object shall be annotated if the enclosing (axis aligned) Bounding Box has an area (width times height) of **at least 500 pixels**. If a real-world object has an area of less than 500 pixels in some M-Frames (and at least 500 pixels in other M-Frames), then

this object **can** also be annotated in the M-Frames where it has an area of less than 500 pixels (due to consistency reasons).

o) Each geometry shall have a **precision of three pixels**, meaning that the edges of the annotation geometries are not allowed to have a distance of more than three pixels from the “true” annotation geometry. In cases where this precision cannot be achieved without violating the maximum number of anchor points of a polygon (see subsection 1.2), the precision can be lower than 3 pixels - but should be still as high as possible.

p) If parts of an object cannot be recognized due to **overexposure**, the true size shall be estimated. The overexposed part does not count as occlusion.

q) If an object protrudes beyond the edge of an image, only the parts that are **inside the image** shall be annotated.

3 Annotation Rules of the Object Classes

In addition to the general annotation rules, further specific rules apply to the individual object classes. If rules contradict each other, the class-specific annotation rules apply, see also rule a) of the sensor-independent rules. The object classes are divided into **traffic areas** and **actual objects**. In the following, the annotation of different traffic areas is discussed first. Then the labeling of different objects is considered.

3.1 Annotation Rules of Traffic Areas

3.1.1 Path

The class-specific annotation rules for the object class “path” are:

a) The object class “path” shall be used for bike paths and lanes, foot paths as well as joint variants (meaning areas that can be used by both pedestrians and cyclists, but not by cars).

b) Lane boundaries (typically the top surface of curbs or boundary strips) on a “path” shall be included into the annotation of a “path” as long as they are walkable/drivable by the users of the “path”. Therefore, the side surfaces of curbs are never part of a “path”. Boundary strips that separate a bike lane from a “vehicle lane” shall be part of both annotations. Additionally, they shall be labeled separately, see section 3.1.4.

Table 1: Geometry requirements of the object class “path”

sensor	geometry
camera	2D Polygon
lidar	3D Polygon

Table 2: Attributes and attribute values of the object class “path”

attribute	attribute value
type	bike path foot path

	joint path unknown
occlusion	0-33 % 34-67 % 68-99 % 100 %

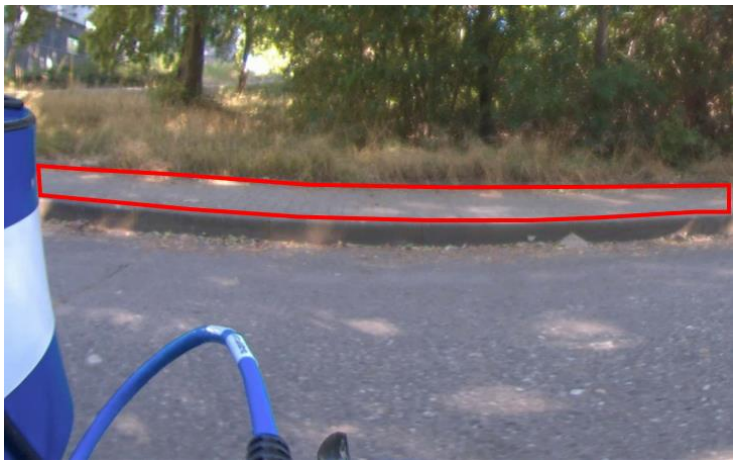

Figure 1: Annotation example of the object class “path” with “type” = “footpath” and occlusion = “0-33 %”. The top surface of the curb is part of the annotation, the side surface is not.

3.1.2 Vehicle Lane

The class-specific annotation rules for the object class “vehicle lane” are:

a) The object class “vehicle lane” shall be used for traffic lanes and parking areas.
b) Special traffic lanes such as a bus lane or a bus stop bay shall be included into the annotation of a “vehicle lane”.
c) Separate parking spaces, e.g. in front of a supermarket, are labelled as “parking areas”.
d) If a “vehicle lane” consists of several not structurally separated lanes, these should be labelled as one single vehicle lane. Consequently, intersections are labelled as one “vehicle lane”, too. However, to avoid annotation geometries with a large amount of anchor points, such lanes can also be split into several annotation objects if necessary.
e) Lane boundaries (typically center lines and lane boundary strip) on a “vehicle lane” shall be included into the annotation of a “vehicle lane” as long as they are drivable by the users of the “vehicle lane”. Boundary strips that separate a “vehicle lane” from a bike lane (see “path”) shall be part of both annotations. Additionally, they shall be labeled separately, see section 3.1.4.
f) Traffic areas like traffic-calmed zones that can be used by pedestrians, cyclists and cars shall also be labeled as “vehicle lanes” of “type” “traffic lane”.

Table 3: Geometry requirements of the object class “vehicle lane”

sensor	**geometry**
camera	2D Polygon

lidar	3D Polygon

Table 4: Attributes and attribute values of the object class “vehicle lane”

attribute	attribute value
type	traffic lane parking area unknown
occlusion	0-33 % 34-67 % 68-99 % 100 %

Figure 2: Annotation example of the object class “vehicle lane” with “type” = “vehicle lane” and occlusion = “0-33 %”

3.1.3 Open Space

The class-specific annotation rule for the object class “open space” is:

a) Areas are labelled as “open space” if they are generally passable for bicycles but cannot be classified as “paths” or “vehicles lanes”.
b) If the distinction between “vehicles lane” with “type” = “parking area” and “open space” is unclear, the region is labelled as “open space”.

Table 5: Geometry requirements of the object class “open space”

sensor	geometry
camera	2D Polygon
lidar	3D Polygon

Table 6: Attributes and attribute values of the object class “open space”

attribute	attribute value
occlusion	0-33 % 34-67 % 68-99 % 100 %

Figure 3: Annotation example of the object class “open space” with an occlusion of “0-33 %”

3.1.4 Lane Boundary

The class-specific annotation rules for the object class “lane boundary” are:

a) The object class “lane boundary” shall be used for curbs, lane boundary strips, center lines and other markings of the boundary of a lane.
b) “Lane boundaries” of type “center line” and “lane boundary strip” are only annotated if they are clearly recognizable as such.
c) A dashed lane boundary of type “center line” or “lane boundary strip” is annotated as if it was a solid line.

Table 7: Geometry requirements of the object class “lane boundary”

sensor	geometry
camera	2D Polygon
lidar	3D Polygon

Table 8: Attributes and attribute values of the object class “lane boundary”

attribute	attribute value
type	curb lane boundary strip center line other

	unknown
occlusion	0-33 % 34-67 % 68-99 % 100 %

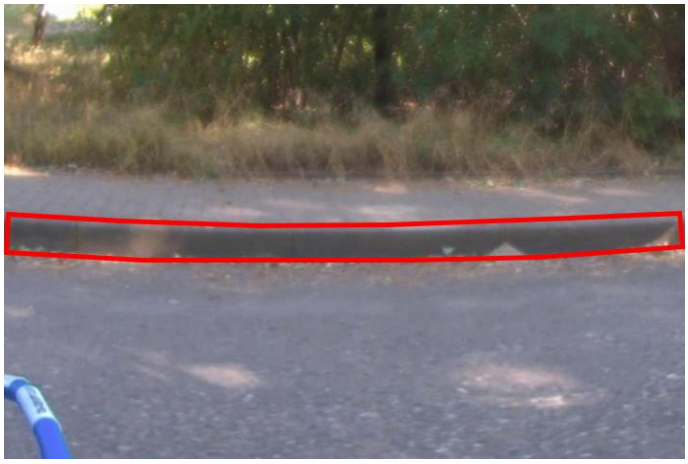

Figure 4: Annotation example of the object class “lane boundary” with “type” = “curb” and occlusion = “0-33 %”

3.1.5 Restricted Area

The class-specific annotation rule for the object class “restricted area” is:

a) The class “restricted area” shall be used to all areas that are not passable by a bike. Additionally, also railway crossings and tracks that are integrated into the street shall be labeled as “restricted area” with “type = track”, although they are passable by bikes.
b) For “restricted areas” with “type” = “track” the whole railroad shall be annotated, not just the rails.

Table 9: Geometry requirements of the object class “restricted area”

sensor	geometry
camera	2D Polygon
lidar	3D Polygon

Table 10: Attributes and attribute values of the object class “restricted area”

attribute	attribute value
type	meadow water tracks other unknown
occlusion	0-33 % 34-67 % 68-99 % 100 %

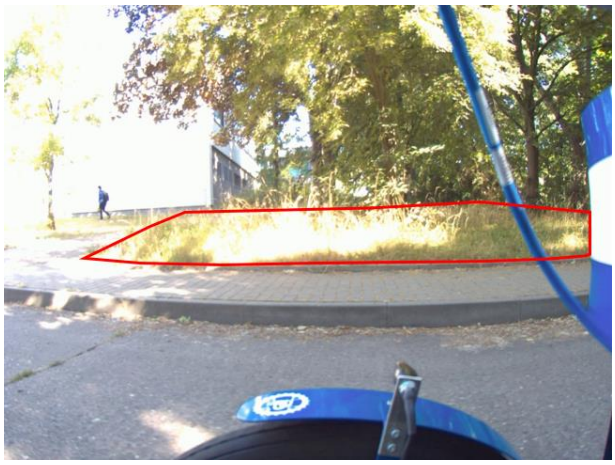

Figure 5: Annotation example of the object class “restricted area” with “type” = “meadow” and occlusion = “0-33 %”

3.2 Annotation of Objects

3.2.1 Person

The class-specific annotation rules of the object class “person” are:

a) The object class “person” includes all human beings.
b) If a “person” is inside an object that is closed from all sides (e.g., a car or a tram), the person shall be ignored in the labeling process.
c) A “person” riding a two-wheeler is not labelled as a “person” but is included into the annotation of object class "two-wheeler".
d) If the person is carrying an item (bag, hat, backpack, suitcase, etc.), this item shall be included into the annotation.
e) Large objects that have an influence on the “person's” movement shall be included into the annotation. These objects shall be listed under the attribute “connectedObject”.
f) A person that is pushing or pulling a two-wheeler shall be annotated as an object of object class “person” and “connectedObject” = “other”.
g) For persons younger than 14 years, the attribute “type” should be set to “child”. For persons older than 14 years, “type” is set to “adult”. If the age of a person is hard to determine, for a person smaller than 1.60 m, the attribute “type” shall be set to “child”,

otherwise, it shall be set to “adult”. If, in such cases, the height is hard to determine, the attribute “type” shall be set to “unknown”.

Table 11: Geometry requirements of the object class “person”

sensor	geometry
camera	2D Bounding Box
lidar	3D Bounding Box

Table 12: Attributes and attribute values of the object class “person”

attribute	attribute value
type	adult child unknown
connectedObject	buggy mobility aid wheelchair other unknown
occlusion	0-33 % 34-67 % 68-99 % 100 %

Figure 6: Annotation example of the object class “person” with “type” = “adult”, “connectedObject” is empty and occlusion = “0-33 %”

3.2.2 Vehicle

The class-specific annotation rules of the object class “vehicle” are:

a) The object class “vehicle” includes all types of wheelers that have at least four wheels or that are designed to be pulled by a vehicle with at least four wheels (e.g. trailers).

Table 13: Geometry requirements of the object class “vehicle”

sensor	geometry
camera	2D Bounding Box
lidar	3D Bounding Box

Table 14: Attributes and attribute values of the object class “vehicle”

attribute	attribute value
type	car truck bus tram ambulance police fire truck trailer other unknown
occlusion	0-33 % 34-67 % 68-99 % 100 %

Figure 7: Annotation example of the object class “vehicle” with “type” = “car” and occlusion = “0-33 %”

3.2.3 Two-Wheeler

The class-specific annotation rules of the object class “two-wheeler” are:

a) The object class “vehicle” includes all types of wheelers that have exactly two wheels.
b) The persons riding a two-wheeler shall be included into the annotation of the two-wheeler.

c) The attribute “connectedPersons” shall describe the number of persons that are riding on a two-wheeler. Here, “other” means more than two persons.

Table 15: Geometry requirements of the object class “two-wheeler”

sensor	geometry
camera	2D Bounding Box
lidar	3D Bounding Box

Table 16: Attributes and attribute values of the object class “two-wheeler”

attribute	attribute value
type	bicycle motor bike scooter other unknown
connectedPerson	none one two other unknown
occlusion	0-33 % 34-67 % 68-99 % 100 %

Figure 8: Annotation example of the object class “two-wheeler” with “type” = “motorbike”, “connectedPerson” = “none” and occlusion = “0-33 %”

3.2.4 Animal

The class-specific annotation rule of the object class “animal” is:

a) The object class “animal” shall include all living beings that can potentially collide with a bicycle with a non-insignificant consequence for the bicycle and its rider.

Table 17: Geometry requirements of the object class “animal”

sensor	geometry
camera	2D Bounding Box
lidar	3D Bounding Box

Table 18: Attributes and attribute values of the object class “animal”

attribute	attribute value
type	dog other unknown
occlusion	0-33 % 34-67 % 68-99 % 100 %

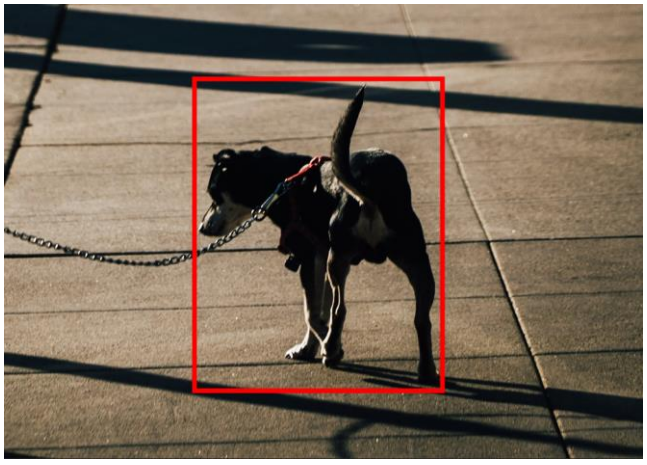

Figure 9: Annotation example of the object class “animal” with “type” = “dog” and occlusion = “0-33 %”

3.2.5 Sign

The class-specific annotation rules of the object class “sign” are:

a) The object class “sign” shall include all shields that regulate traffic or that carry other information (directions of cities, city names, …).
b) “Signs” shall only be annotated if they can be seen from the front.
c) The pole of a sign shall be ignored in the label process (see Figure 10).
d) Different “signs” on the same pole shall be labelled separately.

Table 19: Geometry requirements of the object class “sign”

sensor	geometry
camera	2D Bounding Box
lidar	3D Bounding Box

Table 20: Attributes and attribute values of the object class “sign”

attribute	attribute value
type	traffic regulation other unknown
occlusion	0-33 % 34-67 % 68-99 % 100 %

Figure 10: Annotation example of the object class “sign” with “type” = “traffic regulation sign” and occlusion = “0-33 %”

3.2.6 Traffic Light

The class-specific annotation rules of the object class “traffic light” are:

a) “Traffic lights” shall only be annotated if they can be seen from the front.
b) The pole of a traffic light shall be ignored in the label process (see Figure 11).
c) Different “traffic lights” on the same pole shall be labelled separately.
d) The attribute “type” describes for which kind of traffic participants the “traffic lights” apply.

Table 21: Geometry requirements of the object class “traffic light”

sensor	geometry
camera	2D Bounding Box
lidar	3D Bounding Box

Table 22: Attributes and attribute values of the object class “traffic light”

attribute	attribute value
type	vehicle pedestrian bicycle unknown
occlusion	0-33 % 34-67 % 68-99 % 100 %

Figure 11: Annotation example of the object class “traffic light” with “type” = “vehicle” and “occlusion” = “0-33 %”

3.2.7 Vertical Barrier

The class-specific annotation rules of the object class “vertical barrier” are:

a) The class “vertical barrier” describes all kinds of objects that prevents a bicycle passing through it.
b) If there are holes within the “vertical barrier”, these holes are included into the annotation (e.g., a fence made of several fence slats).
c) A ”vertical barrier” shall be at least 1.0 m long (maximum extend in driving direction of the ego vehicle, in a height interval of 0 to 2.5 m). Consequently, single poles and trees does not count as “vertical barriers” while hedges and dense rows of trees do.
d) There is one exception to the previous requirement: A group of bollards shall always be labeled as “vertical barrier”.

e) Large vertical barriers that significantly deviate from a cuboidal shape should be split into separate objects with approximately cuboidal shape.

Table 23: Geometry requirements of the object class “vertical barrier”

sensor	geometry
camera	2D Polygon
lidar	3D Bounding Box

Table 24: Attributes and attribute values of the object class “vertical barrier”

attribute	attribute value
type	building fence wall vegetation other unknown
occlusion	0-33 % 34-67 % 68-99 % 100 %

Figure 12: Annotation example of the object class “vertical barrier” with “type” = “building” and occlusion = “0-33 %”

Figure 13: Annotation example of the object class “vertical barrier” with “type” = “vegetation” and occlusion = “0-33 %”

3.2.8 Rumble Strip

The class-specific annotation rule of the object class “rumble strip” is:

a) This class included specifically created obstacles to reduce the velocity of vehicles.

Table 25: Geometry requirements of the object class “rumble strip”

sensor	geometry
camera	2D Bounding Box
lidar	3D Bounding Box

Table 26: Attributes and attribute values of the object class “rumble strip”

attribute	attribute value
type	speed bumper unknown
occlusion	0-33 % 34-67 % 68-99 % 100 %

Figure 14: Annotation example of the object class “rumble strip” with “type” = “speed bumper” and occlusion = “0-33 %”<