nature machine intelligence

Article https://doi.org/10.1038/s42256-023-00625-5

Labelling instructions matter in biomedical image analysis

Supplementary Information

1 PRISMA statement

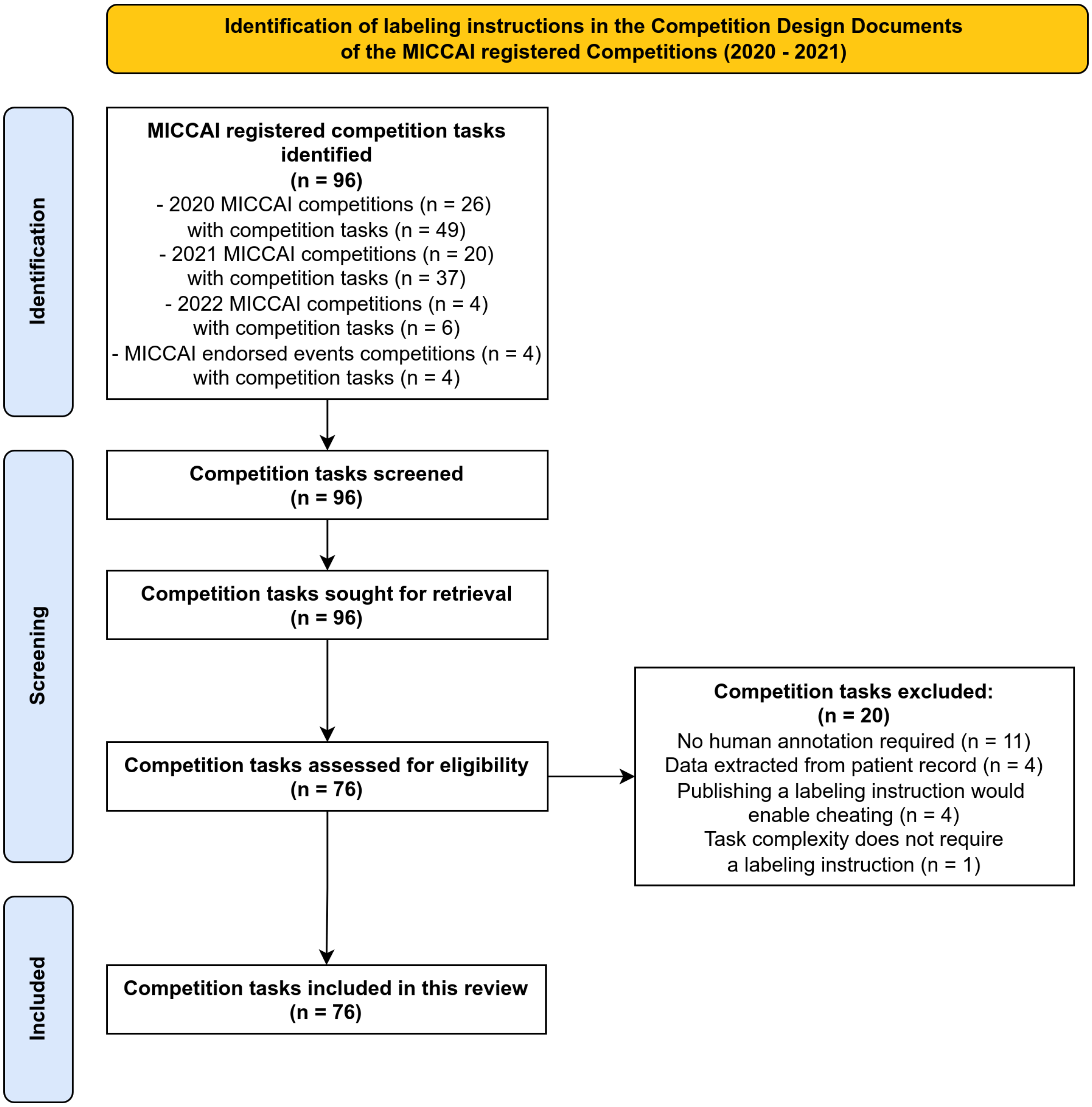

Identification of labeling instructions in the Competition Design Documents
of the MICCAI registered Competitions (2020 - 2021)
Identification
MICCAI registered competition tasks
identified
(n = 96)
- 2020 MICCAI competitions (n = 26)
with competition tasks (n = 49)
- 2021 MICCAI competitions (n = 20)
with competition tasks (n = 37)
- 2022 MICCAI competitions (n = 4)
with competition tasks (n = 6)
- MICCAI endorsed events competitions (n = 4)
with competition tasks (n = 4)
Screening
Competition tasks screened
(n = 96)
Competition tasks sought for retrieval
(n = 96)
Competition tasks assessed for eligibility
(n = 76)
Competition tasks excluded:
(n = 20)
No human annotation required (n = 11)
Data extracted from patient record (n = 4)
Publishing a labeling instruction would
enable cheating (n = 4)
Task complexity does not require
a labeling instruction (n = 1)
Included
Competition tasks included in this review
(n = 76)

Figure SN1: **Structured identification of labeling instructions in the Challenge Design Documents of the Medical Image Computing and Computer Assisted Intervention Society (MICCAI) registered competitions (2020 - 2021).** After the identification of all relevant competition tasks, each individual competition task was screened and excluded, if the organizers provided a valid reasoning why their labeling instructions cannot be provided.

2 DSC scores for the category with the least and the most present image characteristics on the instrument level

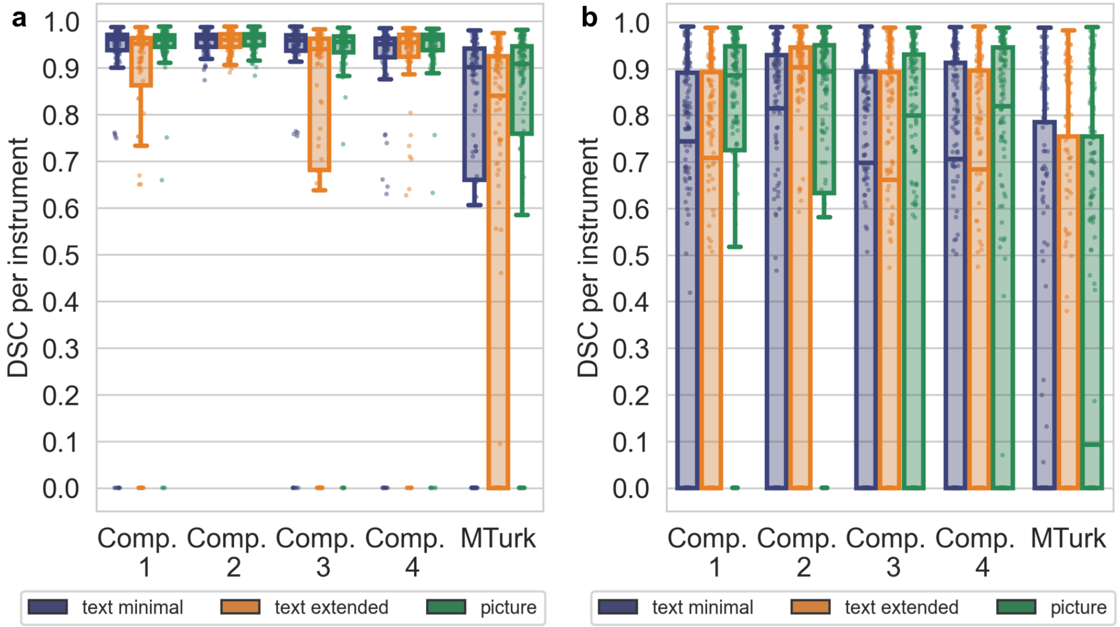


Figure SN2: **DSC scores for the category with the least and the most image characteristics present on the instrument level.** For each annotation provider and type of labeling instruction, the Dice Similarity Coefficient (DSC) scores per instrument are aggregated. (**a**) displays the category with images that do not contain any artifact on the instruments. (**b**) displays the category with images that contain at least three different artifacts on the instruments. The DSC scores for (a) and (b) are displayed as a dots- and boxplot (the band indicates the median, the box indicates the first (25th percentile) and third (75th percentile) quartiles and the whiskers indicate ±1.5× interquartile range, the DSC score maximum is 1 and the minimum is 0 for each image). A combined total of 3,771 instrument annotations were observed for the category with the least and the most present image characteristics.

3 Aggregated MICCAI competition analysis results

Table SN3: **Aggregated results of the Medical Image Computing and Computer Assisted Intervention Society (MICCAI) competition analysis.** For each competition, the individual competition task submission documents were evaluated.

Analysis results of MICCAI registered competitions		
Date of evaluation	Nov 10 2021	
Total number of competitions*	53	
Total number of competition tasks	96	
Competition tasks which..	**absolute**	**relative**
do not need to publish LIs or have valid justification**	20	20.83%
Could/should publish	76	79.17%
In scope competition tasks (76) which ...	**absolute**	**relative**
publish LIs	18	23.68%
do not publish / have LIs	58	76.32%
Inter-rater agreement:	90.63%	

* LI = Labeling Instruction
* All information based on the Challenge design docs, mainly question 23 b/a, of the MICCAI registered competitions (Link — accessed: 2021-11-10) between 2020 and 2021.

4 Labeling for testing/validation

Experimental design

High-quality validation/test data is crucial in safety-critical applications. The purpose of this experiment was to investigate the performance variability resulting from different labeling instructions and annotation providers. To this end, we leveraged the existing output of six algorithms that participated in the instance segmentation task of the Robust Medical Instrument Segmentation (ROBUST-MIS) challenge [45] and that were trained on expert-generated annotations (n = 5,983). For each annotated image in our labeled dataset (n = 14,040) introduced in the Methods section, we proceeded as follows: We regarded each annotated image obtained with any of the three labeling instruction types and any of the five annotation providers as reference annotation for a tuple (annotation provider, labeling instruction). This resulted in 3 (types of instructions) x 5 (annotation providers) x 4 (annotations per image and provider) = 60 different annotation masks in total for each unique image, as detailed in the Methods section. These new reference annotations and the original annotations, generated by experts with what we henceforth call expert labeling instruction type, were all compared to the fixed model outputs of the six algorithms. We then assigned a "valid" label to all images in which no false positive (FP) or false negative (FN) instruments occurred. For all valid annotations, the mean DSC, averaged over all instances, was determined. This yielded a total of 85,644 valid/invalid scores / 54,475 DSC image scores (for the valid images) with corresponding meta information encoding the image id, the algorithm, the labeling instruction type, the annotation provider and the annotation provider type. Based on the generated data, mixed model analysis enabled us to quantify the performance variability resulting from different labeling instruction types and annotation providers. Specifically, we used a two-part beta mixed model which included a) a logistic mixed model analyzing the probability of valid annotations and b) a beta mixed model analyzing the non-zero DSC values of an image when only valid annotations occurred. The image id and the fixed algorithm outputs were defined as random effects. We defined the labeling instruction type and the annotation provider type as fixed effects, in which we are interested, with the minimal text instruction type and crowdworkers as reference categories, respectively. This enabled us to calculate the effects of these two variables on both the probability of valid annotations and on the DSC scores.

The model was implemented with R (version 4.0.3, packages: lme4 version 1.1.27.1, glmmTMB version 1.1.2).

Results

Figure SN4 shows the effects of the labeling instruction type on (a) the probability of valid annotations and (b) the effect on the DSC scores. In comparison to the valid annotation odds of the minimal text labeling instruction type, the valid annotation odds were substantially higher for the expert labeling instruction type and the picture labeling instruction type, and lower for extended text labeling instruction type (cf. Figure SN4a). We observe a comparable change in the ratio between the expected DSC score and the difference to a perfect DSC score for the expert labeling instruction type with an increase of 6% (1.06, CI: 1.01, 1.10) and a reduction of 3% (0.97, CI: 0.96, 0.99) for the extended text labeling instruction type, in comparison to the minimal text labeling instruction type, once an object was identified. Interestingly, we observed a minor reduction, which was not statistically significant, in the ratio between the expected DSC score and the difference to a perfect DSC score with the picture labeling instruction type, as depicted in Figure SN4b. Under the same conditions, the odds of valid annotations for professional annotators were 2.17 times (CI: 2.08, 2.27) that of crowdworkers, and, for experts, 2.22 times (CI: 1.93, 2.56) that of crowdworkers. Similarly, in comparison to crowdworkers, the ratio between the expected DSC score and the difference to a perfect DSC score for the same predicted images by the different algorithms were increased by 39% (1.39, CI: 1.37, 1.41) for professional annotators and by 41% (1.41, CI: 1.35, 1.48) for the experts. **Thus, test set images annotated by professional annotation companies without additional quality assurance are very similar in annotation quality to expert-annotated test set images with quality assurance for the task at hand.**

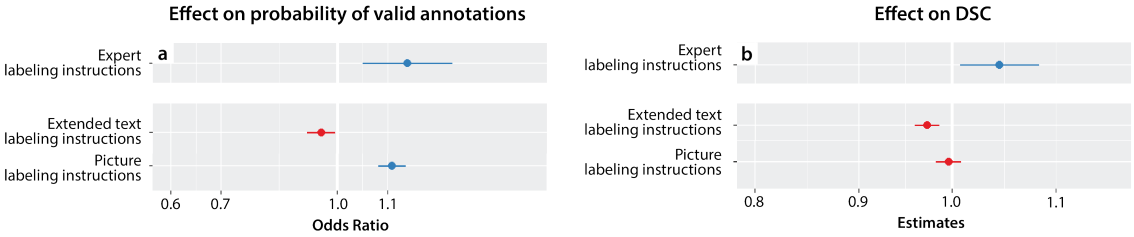
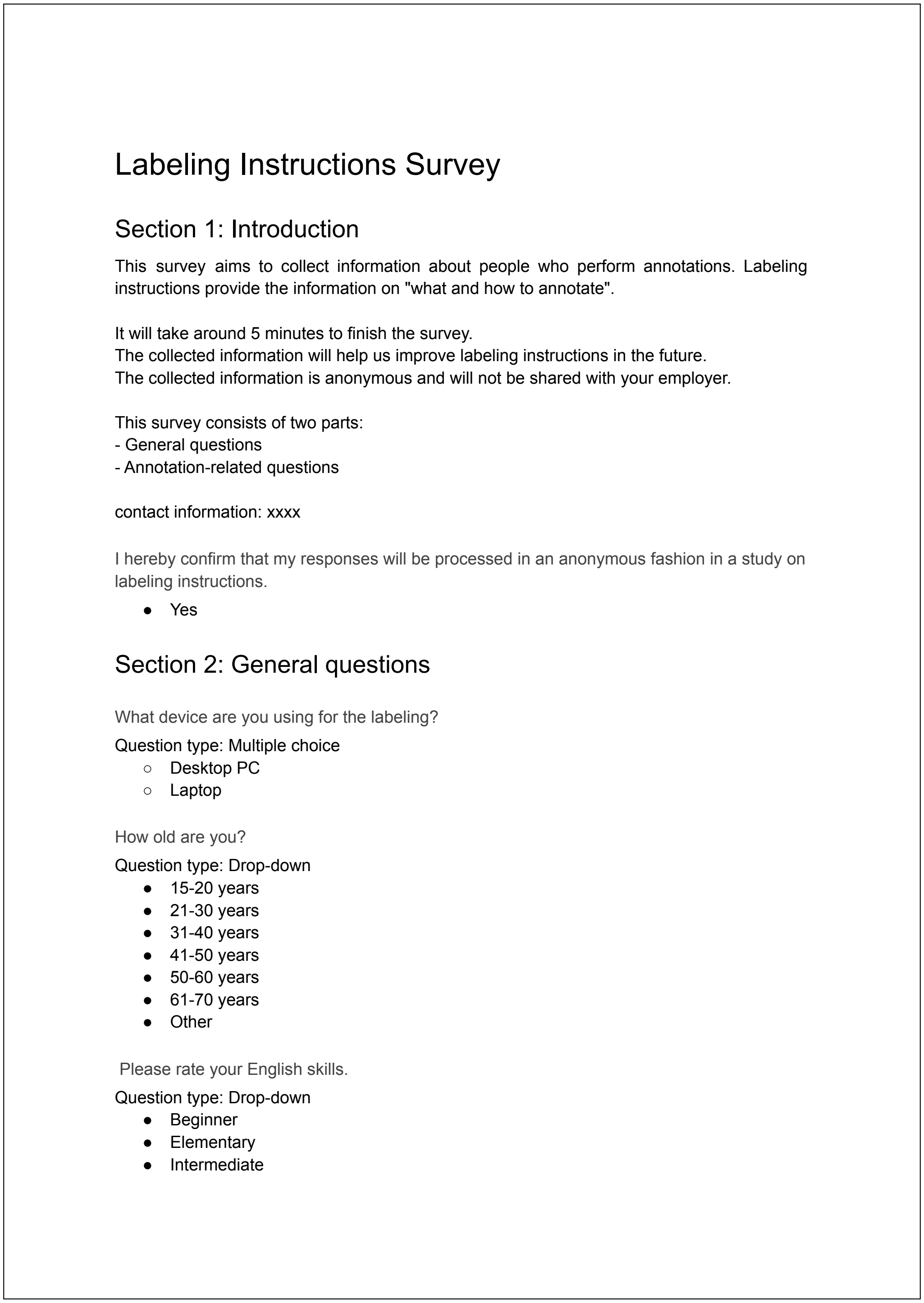


Figure SN4: **Effect of the labeling instruction and the annotator provider type on validation/testing:** The experiments indicate that poor choices of annotators and labeling instructions systematically lead to misreporting the performance of machine learning (ML) models on the test set. (**a**) For the exact same predictions of ML models, extended text descriptions reduce the odds for valid annotations in comparison to minimal text descriptions, while the expert labeling instruction type and including pictures provide a clear benefit. (**b**) Only the expert labeling instructions, which were exclusively used in the 2019 Robust Medical Instrument Segmentation challenge, show an improvement on the Dice Similarity Coefficient (DSC) score. The change (increase: blue, decrease: red) of the valid annotations probability (a) and the perfect DSC score odds (b) is displayed with the minimal text labeling instruction as reference. The two-part beta model was fit on 85,644 DSC image scores from the existing output of six different ML models, our labeled dataset and the expert generated images.

5 Survey: Current status of labeling instructions



Labeling Instructions Survey

Section 1: Introduction

This survey aims to collect information about people who perform annotations. Labeling instructions provide the information on "what and how to annotate".

It will take around 5 minutes to finish the survey.
The collected information will help us improve labeling instructions in the future.
The collected information is anonymous and will not be shared with your employer.

This survey consists of two parts:
- General questions
- Annotation-related questions

contact information: xxxx

I hereby confirm that my responses will be processed in an anonymous fashion in a study on labeling instructions.

● Yes

Section 2: General questions

What device are you using for the labeling?

Question type: Multiple choice

○ Desktop PC
○ Laptop

How old are you?

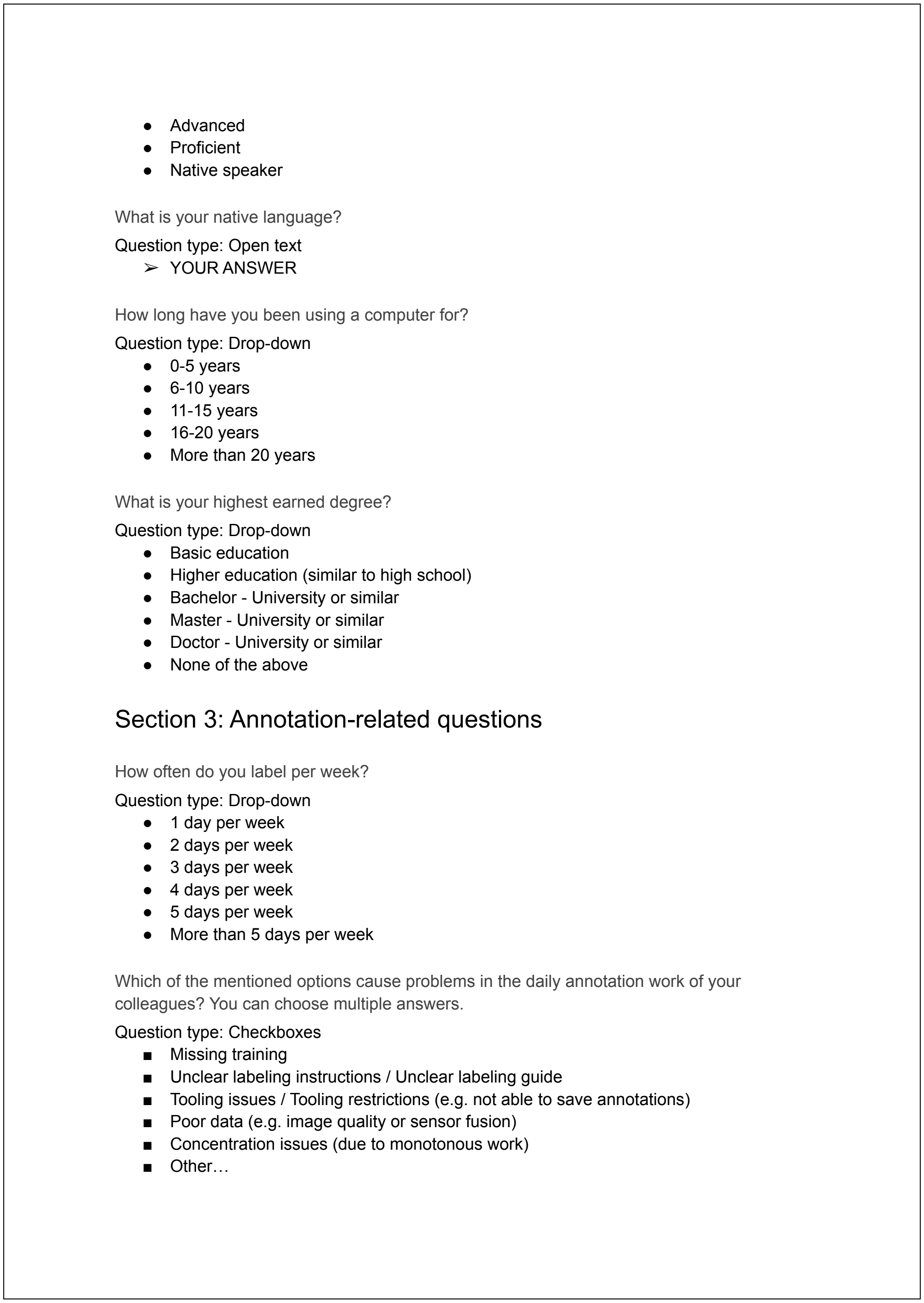
Question type: Drop-down

● 15-20 years
● 21-30 years
● 31-40 years
● 41-50 years
● 50-60 years
● 61-70 years
● Other

Please rate your English skills.

Question type: Drop-down

● Beginner
● Elementary
● Intermediate



- Advanced
- Proficient
- Native speaker

What is your native language?

Question type: Open text

➢ YOUR ANSWER

How long have you been using a computer for?

Question type: Drop-down

- 0-5 years
- 6-10 years
- 11-15 years
- 16-20 years
- More than 20 years

What is your highest earned degree?

Question type: Drop-down

- Basic education
- Higher education (similar to high school)
- Bachelor - University or similar
- Master - University or similar
- Doctor - University or similar
- None of the above

Section 3: Annotation-related questions

How often do you label per week?

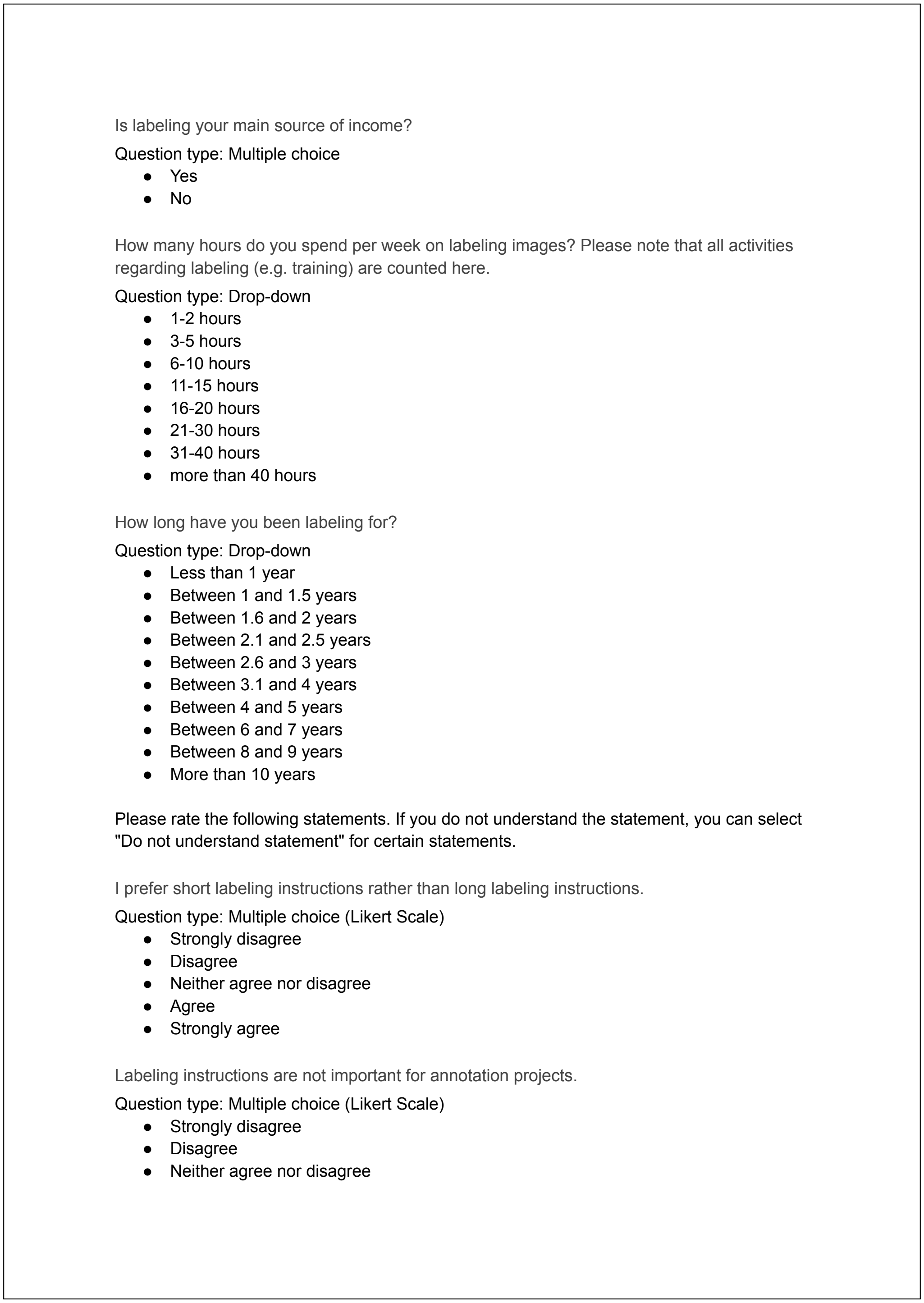
Question type: Drop-down

- 1 day per week
- 2 days per week
- 3 days per week
- 4 days per week
- 5 days per week
- More than 5 days per week

Which of the mentioned options cause problems in the daily annotation work of your colleagues? You can choose multiple answers.

Question type: Checkboxes

■ Missing training
■ Unclear labeling instructions / Unclear labeling guide
■ Tooling issues / Tooling restrictions (e.g. not able to save annotations)
■ Poor data (e.g. image quality or sensor fusion)
■ Concentration issues (due to monotonous work)
■ Other…



Is labeling your main source of income?

Question type: Multiple choice

- Yes
- No

How many hours do you spend per week on labeling images? Please note that all activities regarding labeling (e.g. training) are counted here.

Question type: Drop-down

- 1-2 hours
- 3-5 hours
- 6-10 hours
- 11-15 hours
- 16-20 hours
- 21-30 hours
- 31-40 hours
- more than 40 hours

How long have you been labeling for?

Question type: Drop-down

- Less than 1 year
- Between 1 and 1.5 years
- Between 1.6 and 2 years
- Between 2.1 and 2.5 years
- Between 2.6 and 3 years
- Between 3.1 and 4 years
- Between 4 and 5 years
- Between 6 and 7 years
- Between 8 and 9 years
- More than 10 years

Please rate the following statements. If you do not understand the statement, you can select "Do not understand statement" for certain statements.

I prefer short labeling instructions rather than long labeling instructions.

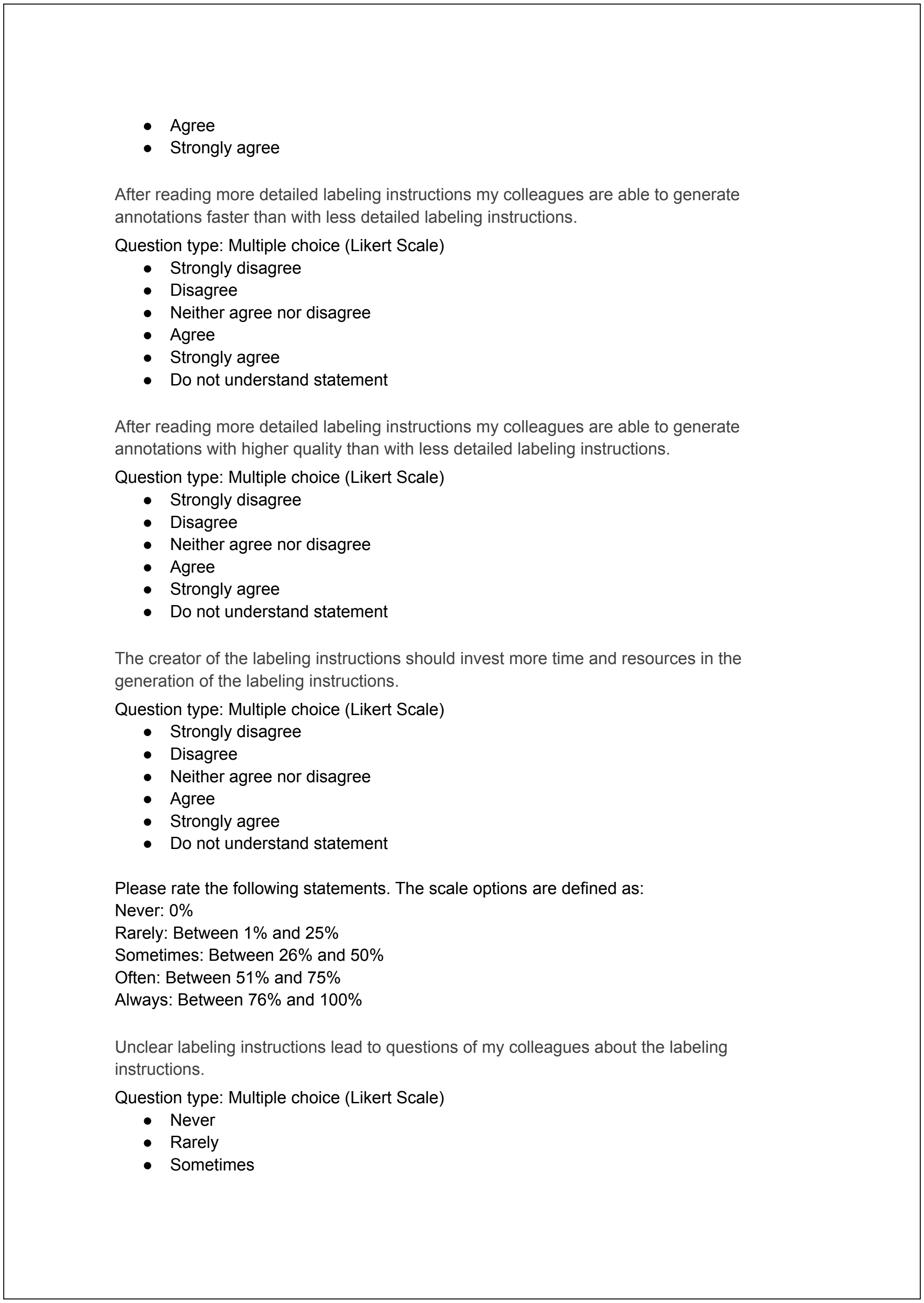
Question type: Multiple choice (Likert Scale)

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

Labeling instructions are not important for annotation projects.

Question type: Multiple choice (Likert Scale)

- Strongly disagree
- Disagree
- Neither agree nor disagree



- Agree
- Strongly agree

After reading more detailed labeling instructions my colleagues are able to generate annotations faster than with less detailed labeling instructions.

Question type: Multiple choice (Likert Scale)

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree
- Do not understand statement

After reading more detailed labeling instructions my colleagues are able to generate annotations with higher quality than with less detailed labeling instructions.

Question type: Multiple choice (Likert Scale)

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree
- Do not understand statement

The creator of the labeling instructions should invest more time and resources in the generation of the labeling instructions.

Question type: Multiple choice (Likert Scale)

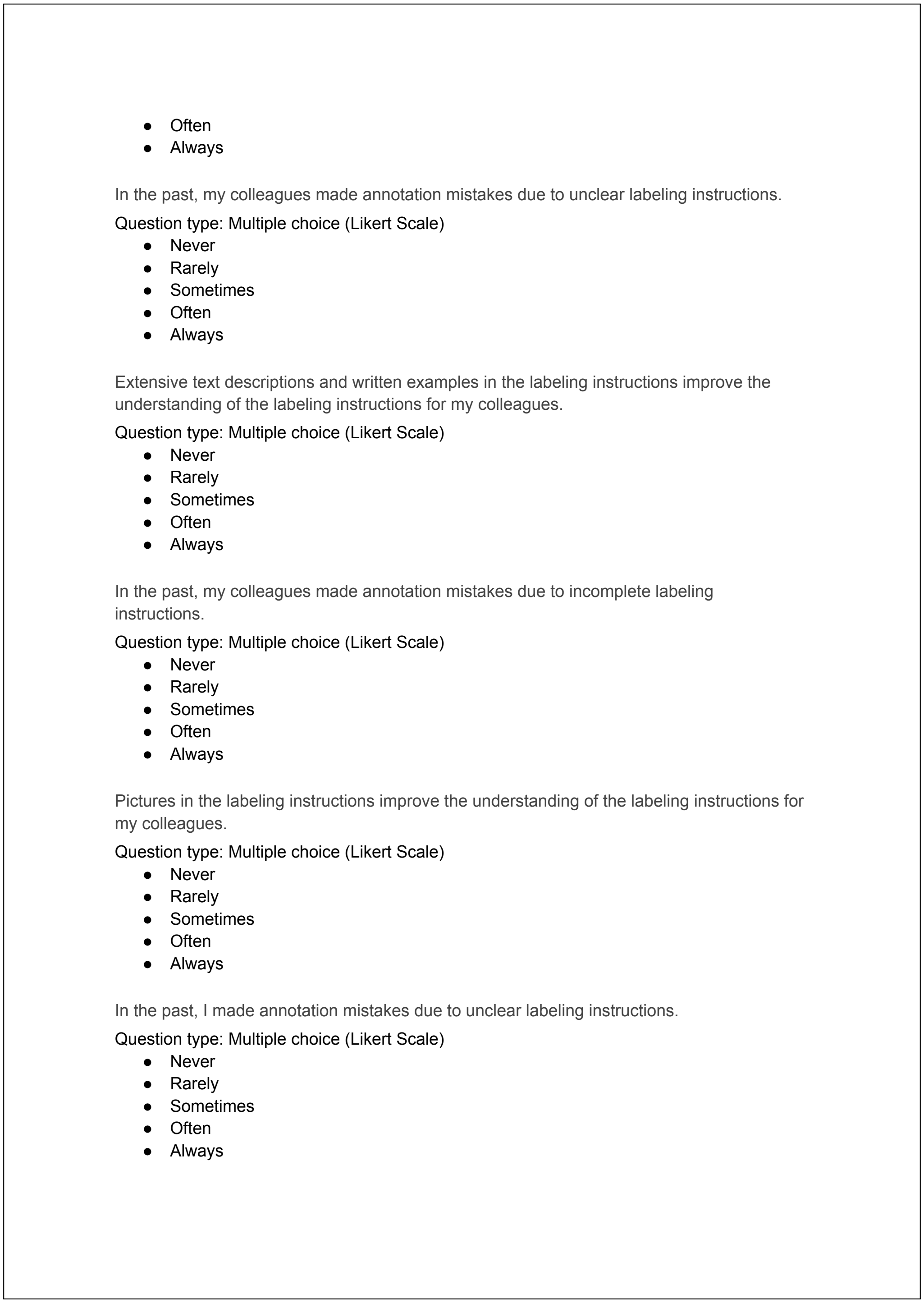
- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree
- Do not understand statement

Please rate the following statements. The scale options are defined as:
Never: 0%
Rarely: Between 1% and 25%
Sometimes: Between 26% and 50%
Often: Between 51% and 75%
Always: Between 76% and 100%

Unclear labeling instructions lead to questions of my colleagues about the labeling instructions.

Question type: Multiple choice (Likert Scale)

- Never
- Rarely
- Sometimes



- Often
- Always

In the past, my colleagues made annotation mistakes due to unclear labeling instructions.

Question type: Multiple choice (Likert Scale)

- Never
- Rarely
- Sometimes
- Often
- Always

Extensive text descriptions and written examples in the labeling instructions improve the understanding of the labeling instructions for my colleagues.

Question type: Multiple choice (Likert Scale)

- Never
- Rarely
- Sometimes
- Often
- Always

In the past, my colleagues made annotation mistakes due to incomplete labeling instructions.

Question type: Multiple choice (Likert Scale)

- Never
- Rarely
- Sometimes
- Often
- Always

Pictures in the labeling instructions improve the understanding of the labeling instructions for my colleagues.

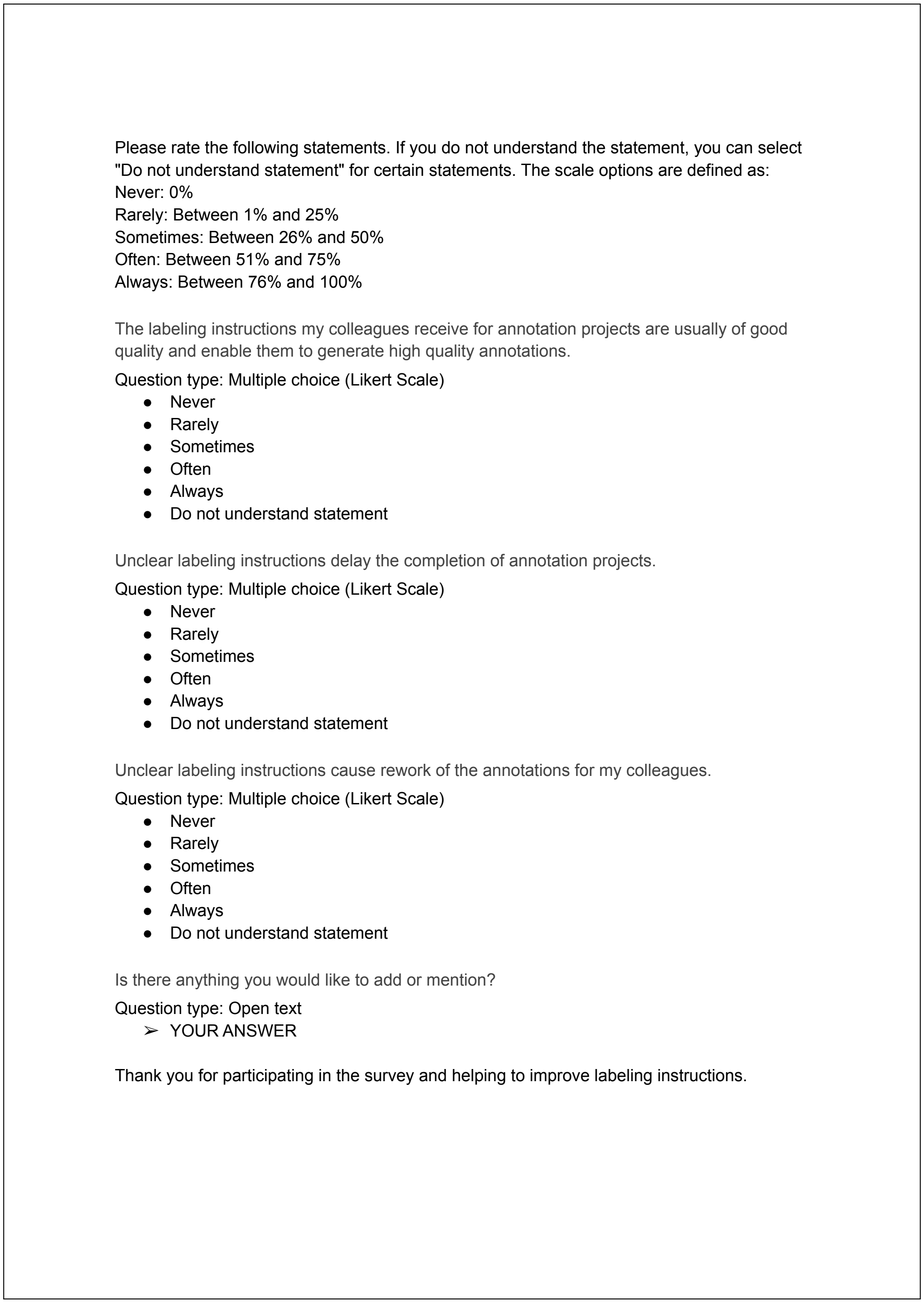
Question type: Multiple choice (Likert Scale)

- Never
- Rarely
- Sometimes
- Often
- Always

In the past, I made annotation mistakes due to unclear labeling instructions.

Question type: Multiple choice (Likert Scale)

- Never
- Rarely
- Sometimes
- Often
- Always



Please rate the following statements. If you do not understand the statement, you can select "Do not understand statement" for certain statements. The scale options are defined as:
Never: 0%
Rarely: Between 1% and 25%
Sometimes: Between 26% and 50%
Often: Between 51% and 75%
Always: Between 76% and 100%

The labeling instructions my colleagues receive for annotation projects are usually of good quality and enable them to generate high quality annotations.

Question type: Multiple choice (Likert Scale)

- Never
- Rarely
- Sometimes
- Often
- Always
- Do not understand statement

Unclear labeling instructions delay the completion of annotation projects.

Question type: Multiple choice (Likert Scale)

- Never
- Rarely
- Sometimes
- Often
- Always
- Do not understand statement

Unclear labeling instructions cause rework of the annotations for my colleagues.

Question type: Multiple choice (Likert Scale)

- Never
- Rarely
- Sometimes
- Often
- Always
- Do not understand statement

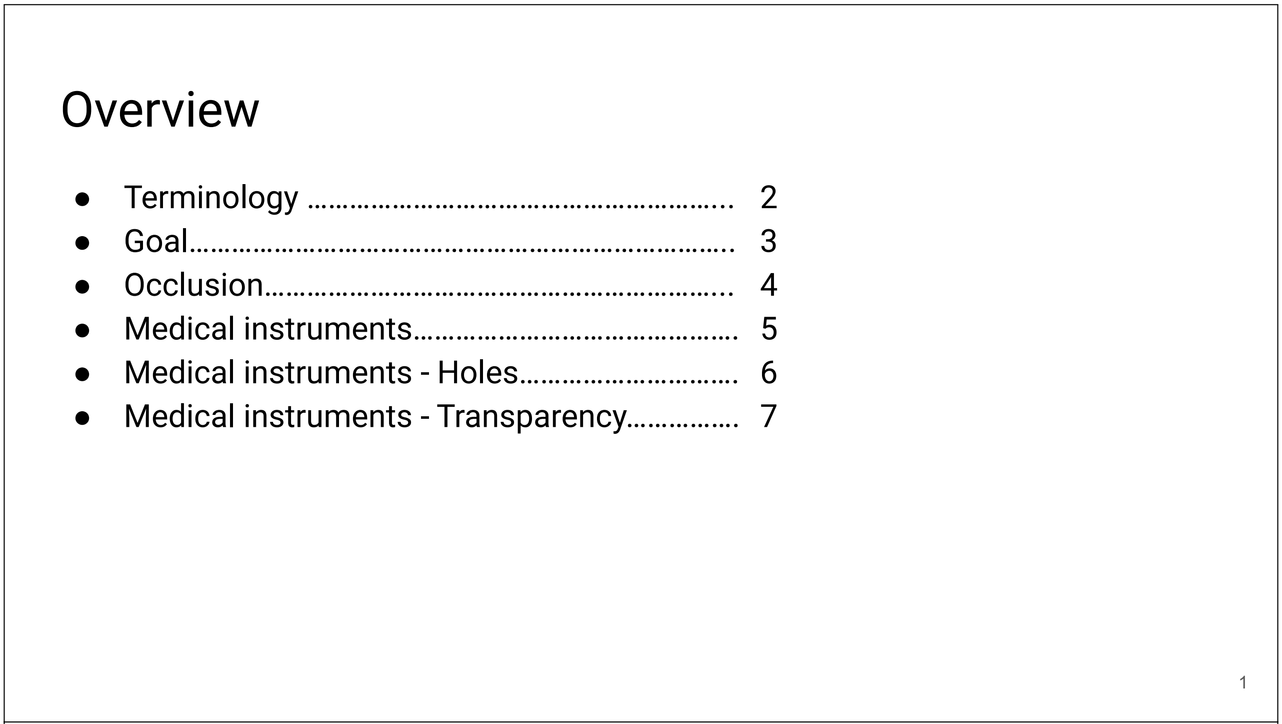
Is there anything you would like to add or mention?

Question type: Open text

➢ YOUR ANSWER

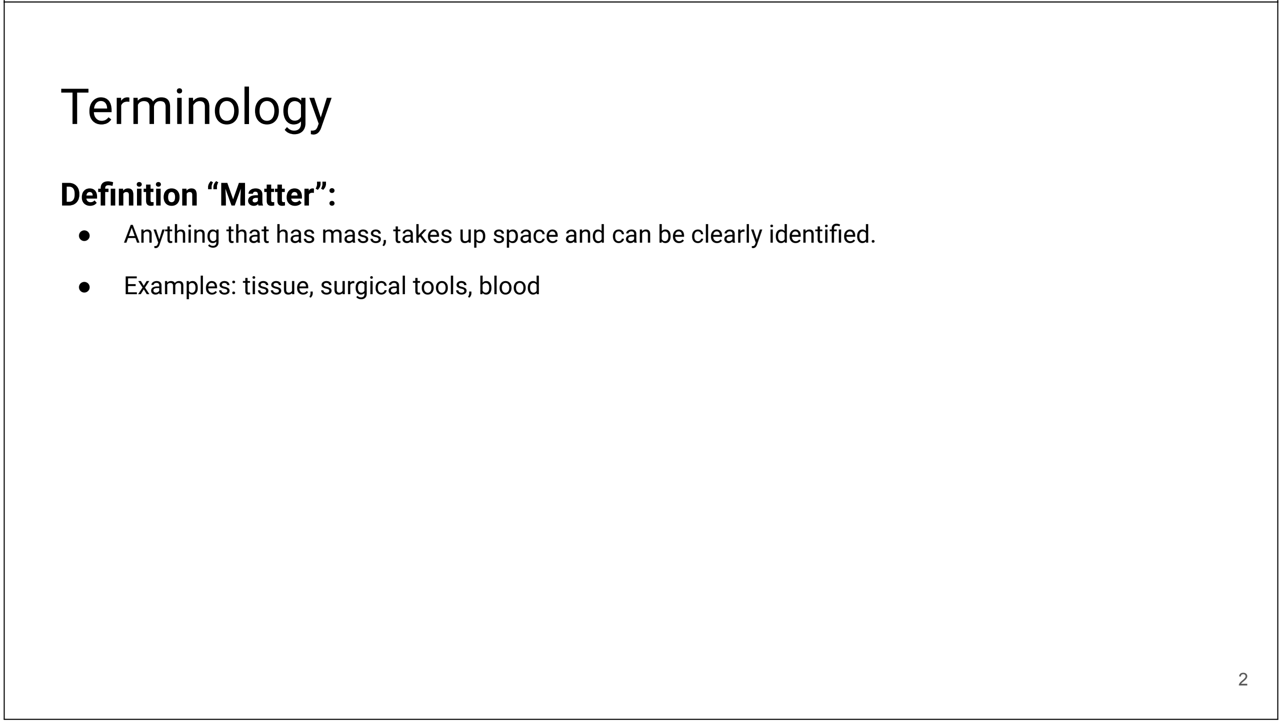
Thank you for participating in the survey and helping to improve labeling instructions.

6 Minimal text labeling instruction



Overview

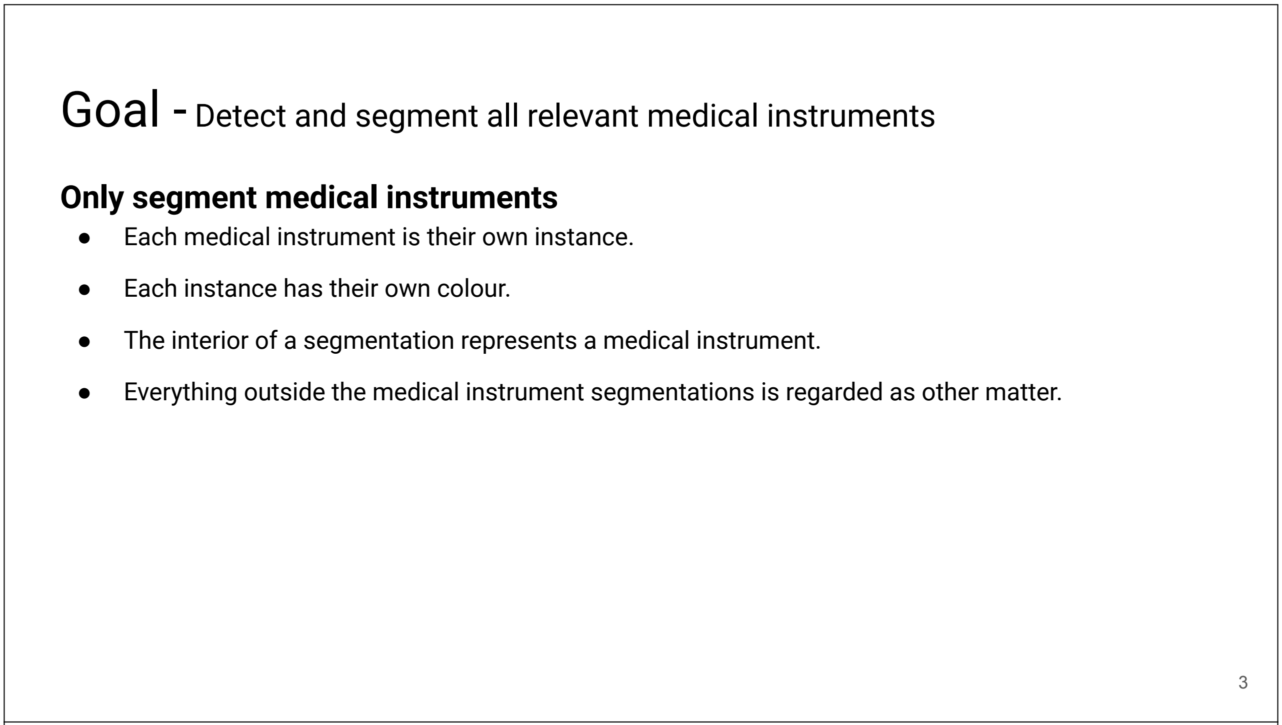
- Terminology ………………………………………………….. 2
- Goal…………………………………………………………….. 3
- Occlusion……………………………………………………... 4
- Medical instruments……………………………………… 5
- Medical instruments - Holes………………………… 6
- Medical instruments - Transparency……………. 7



Terminology

Definition “Matter”:

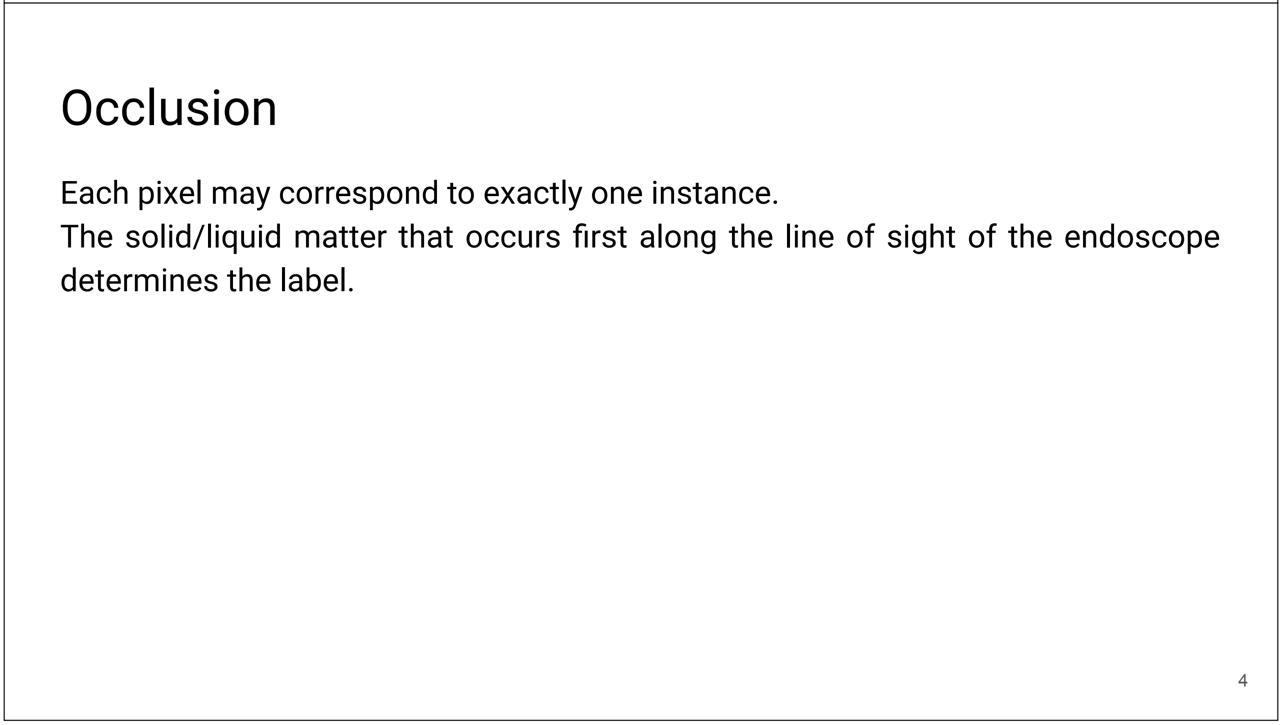
- Anything that has mass, takes up space and can be clearly identified.
- Examples: tissue, surgical tools, blood



Goal - Detect and segment all relevant medical instruments

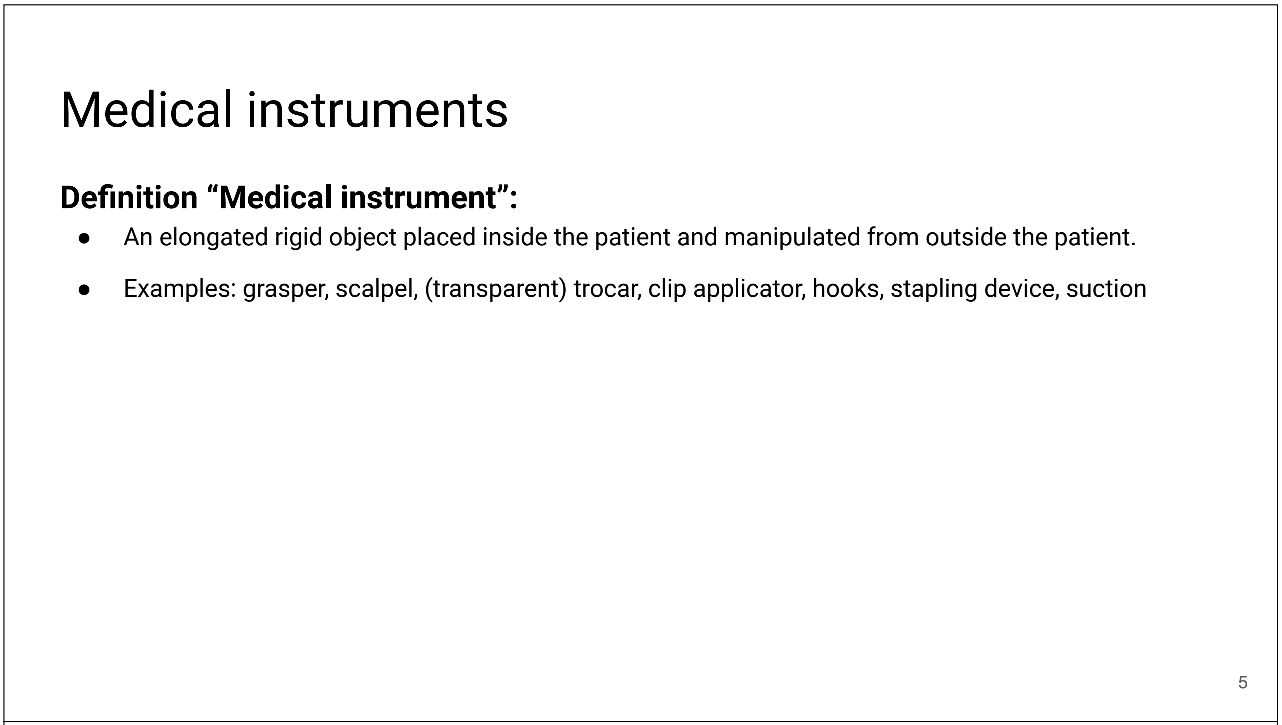
Only segment medical instruments

- Each medical instrument is their own instance.
- Each instance has their own colour.
- The interior of a segmentation represents a medical instrument.
- Everything outside the medical instrument segmentations is regarded as other matter.



Occlusion

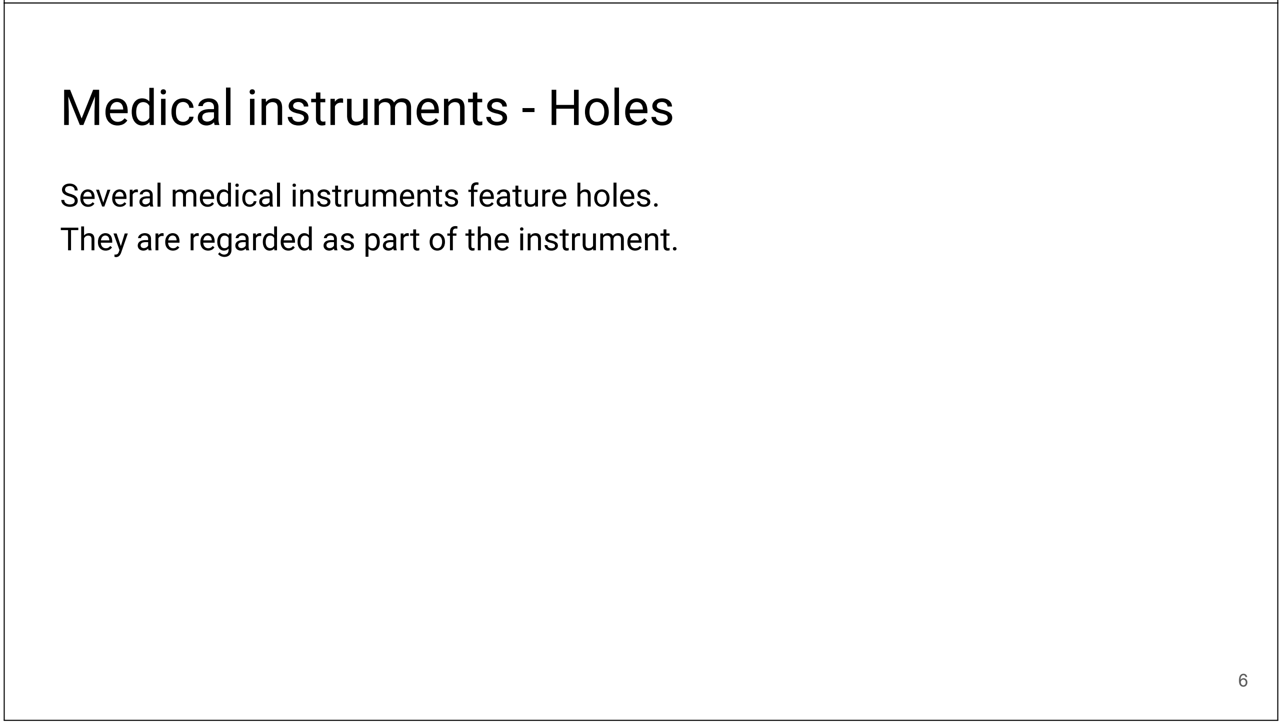
Each pixel may correspond to exactly one instance.
The solid/liquid matter that occurs first along the line of sight of the endoscope determines the label.



Medical instruments

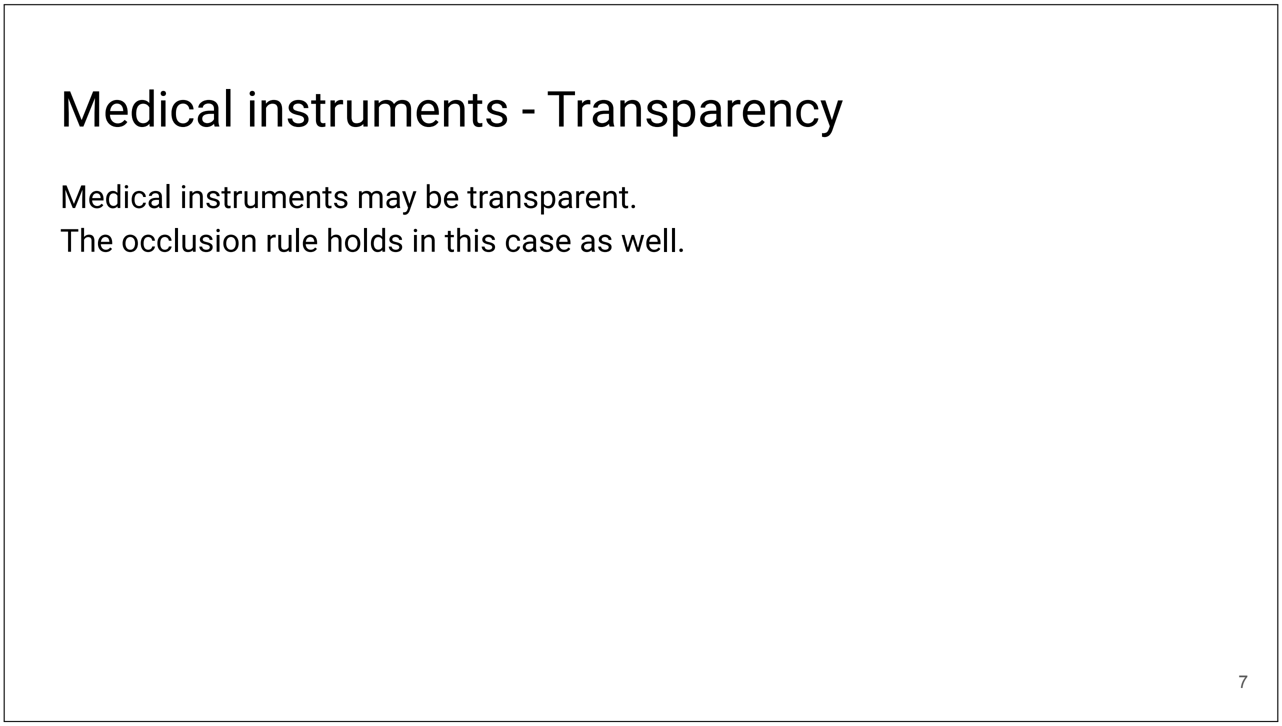
Definition “Medical instrument”:

- An elongated rigid object placed inside the patient and manipulated from outside the patient.
- Examples: grasper, scalpel, (transparent) trocar, clip applicator, hooks, stapling device, suction



Medical instruments - Holes

Several medical instruments feature holes.
They are regarded as part of the instrument.

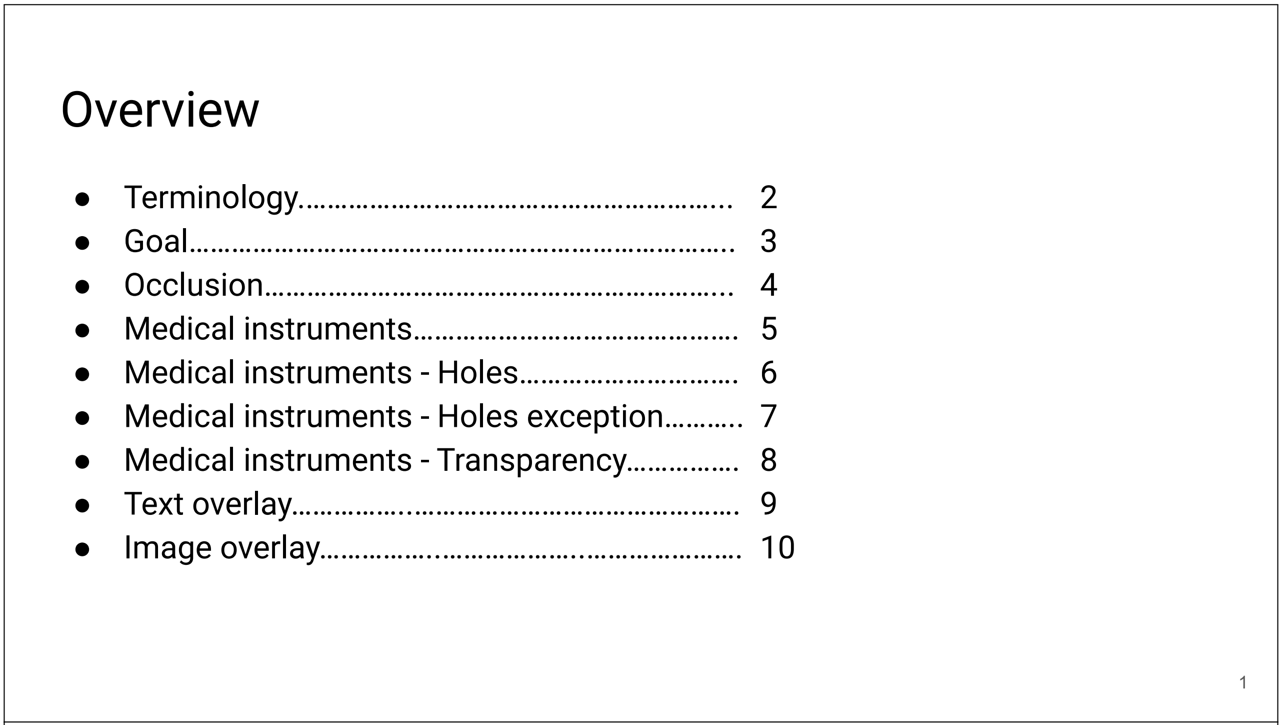


Medical instruments - Transparency

Medical instruments may be transparent.
The occlusion rule holds in this case as well.

7

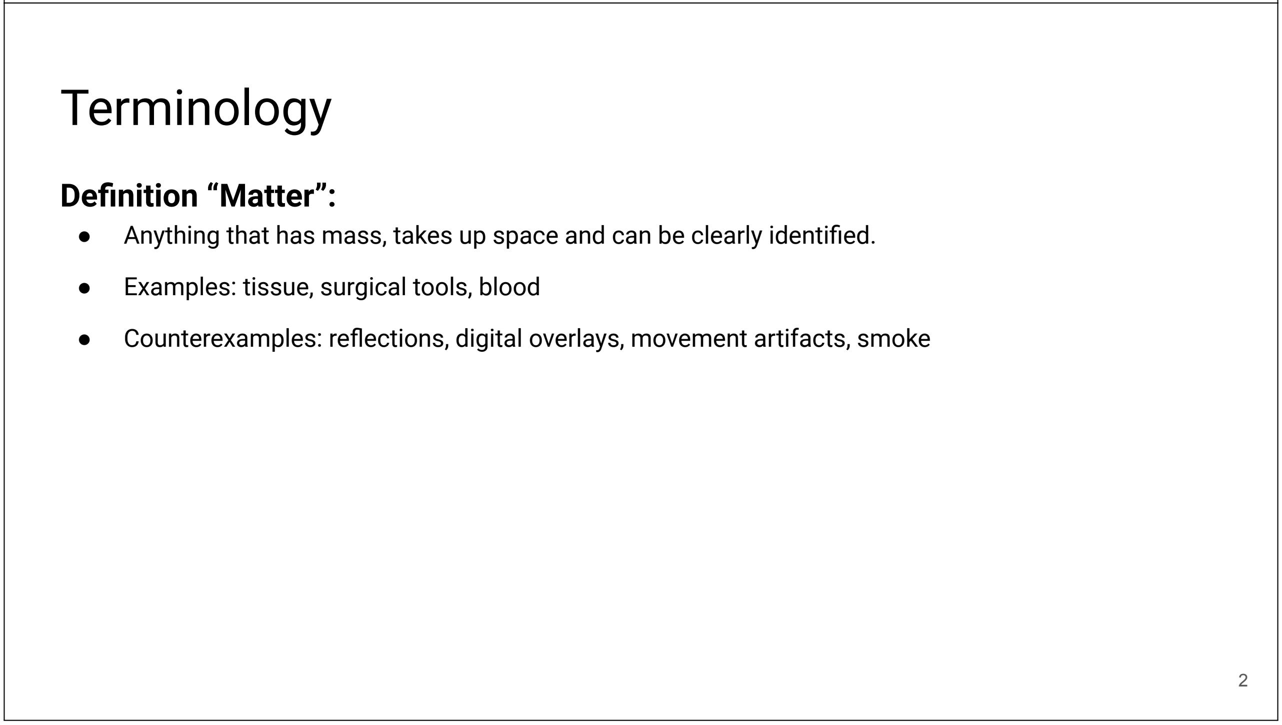
7 Extended text labeling instruction



Overview

- Terminology…………………………………………….. 2
- Goal……………………………………………………. 3
- Occlusion……………………………………………... 4
- Medical instruments…………………………………. 5
- Medical instruments - Holes………………………. 6
- Medical instruments - Holes exception………. 7
- Medical instruments - Transparency……………. 8
- Text overlay……………………………………………. 9
- Image overlay…………………………………………. 10

1

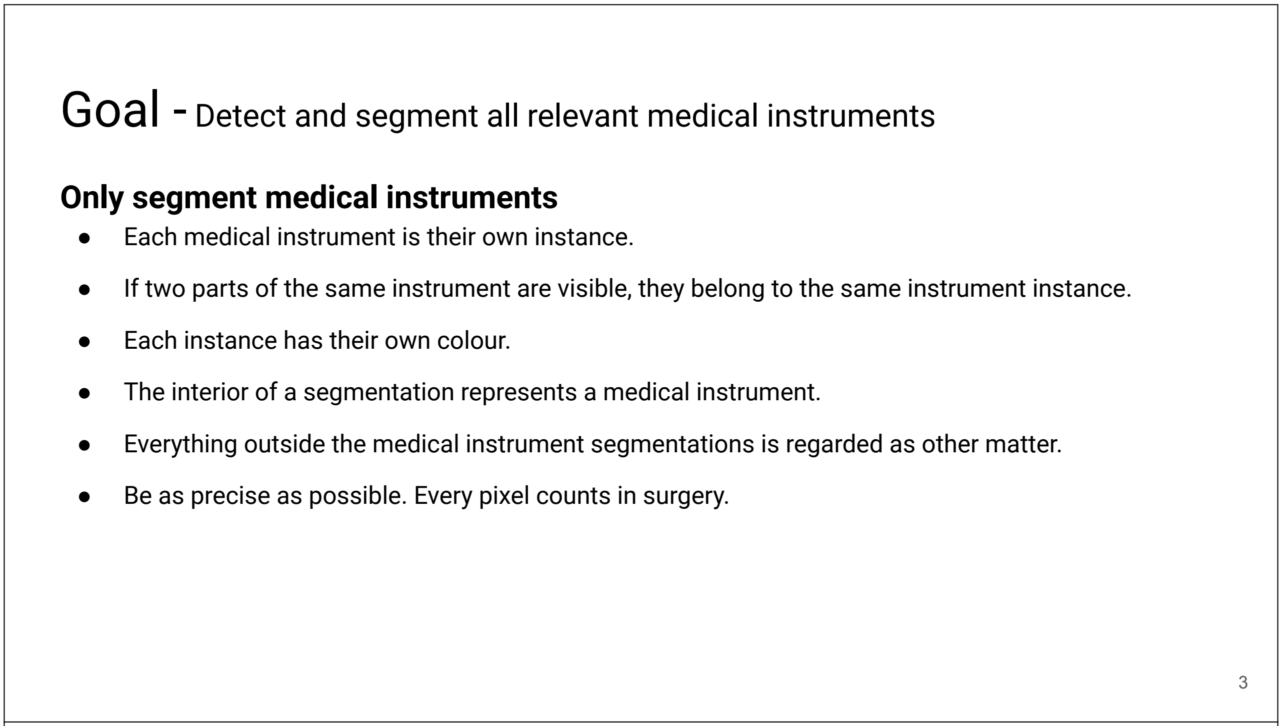


Terminology

Definition “Matter”:

- Anything that has mass, takes up space and can be clearly identified.
- Examples: tissue, surgical tools, blood
- Counterexamples: reflections, digital overlays, movement artifacts, smoke

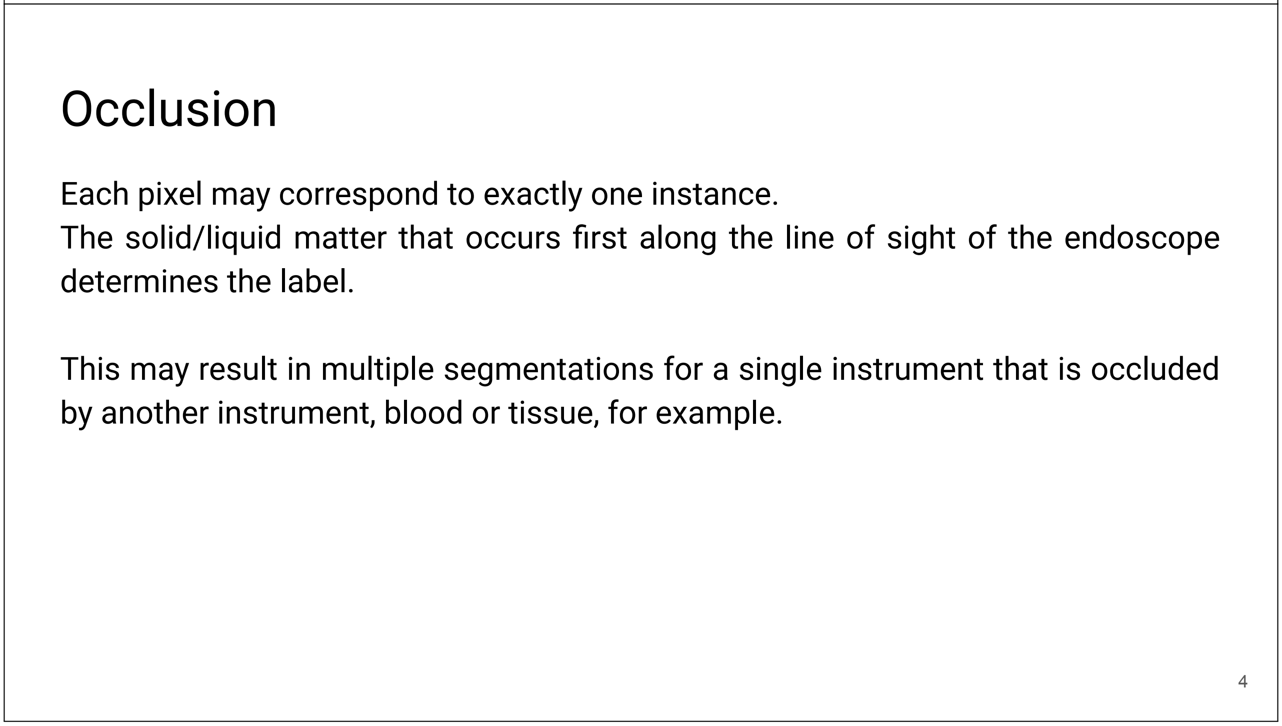
2



Goal - Detect and segment all relevant medical instruments

Only segment medical instruments

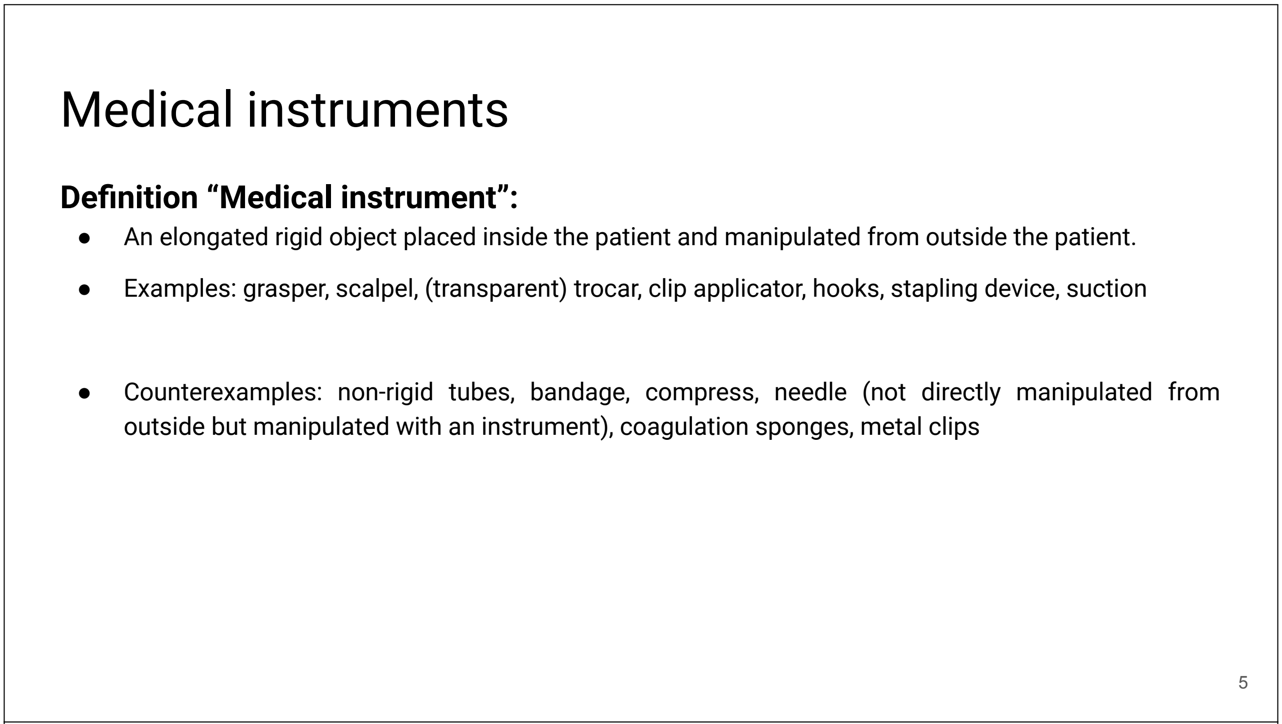
- Each medical instrument is their own instance.
- If two parts of the same instrument are visible, they belong to the same instrument instance.
- Each instance has their own colour.
- The interior of a segmentation represents a medical instrument.
- Everything outside the medical instrument segmentations is regarded as other matter.
- Be as precise as possible. Every pixel counts in surgery.



Occlusion

Each pixel may correspond to exactly one instance.
The solid/liquid matter that occurs first along the line of sight of the endoscope determines the label.

This may result in multiple segmentations for a single instrument that is occluded by another instrument, blood or tissue, for example.

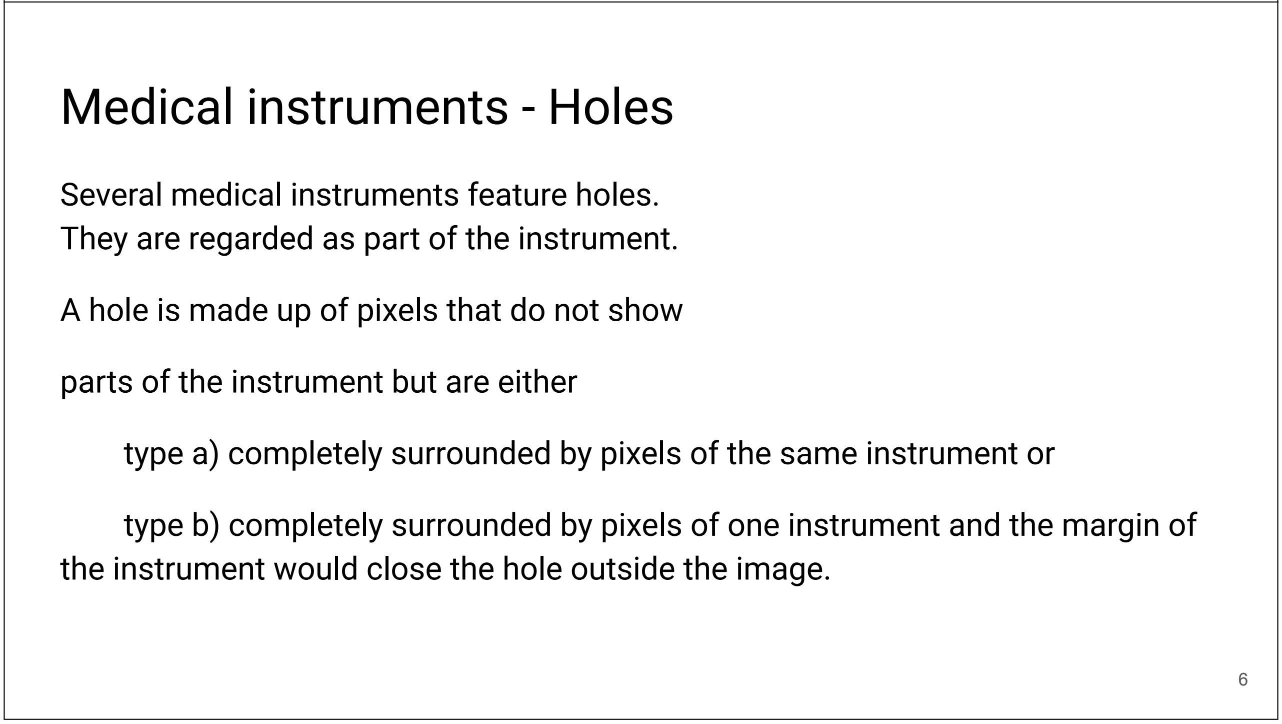


Medical instruments

Definition “Medical instrument”:

- An elongated rigid object placed inside the patient and manipulated from outside the patient.
- Examples: grasper, scalpel, (transparent) trocar, clip applicator, hooks, stapling device, suction

- Counterexamples: non-rigid tubes, bandage, compress, needle (not directly manipulated from outside but manipulated with an instrument), coagulation sponges, metal clips



Medical instruments - Holes

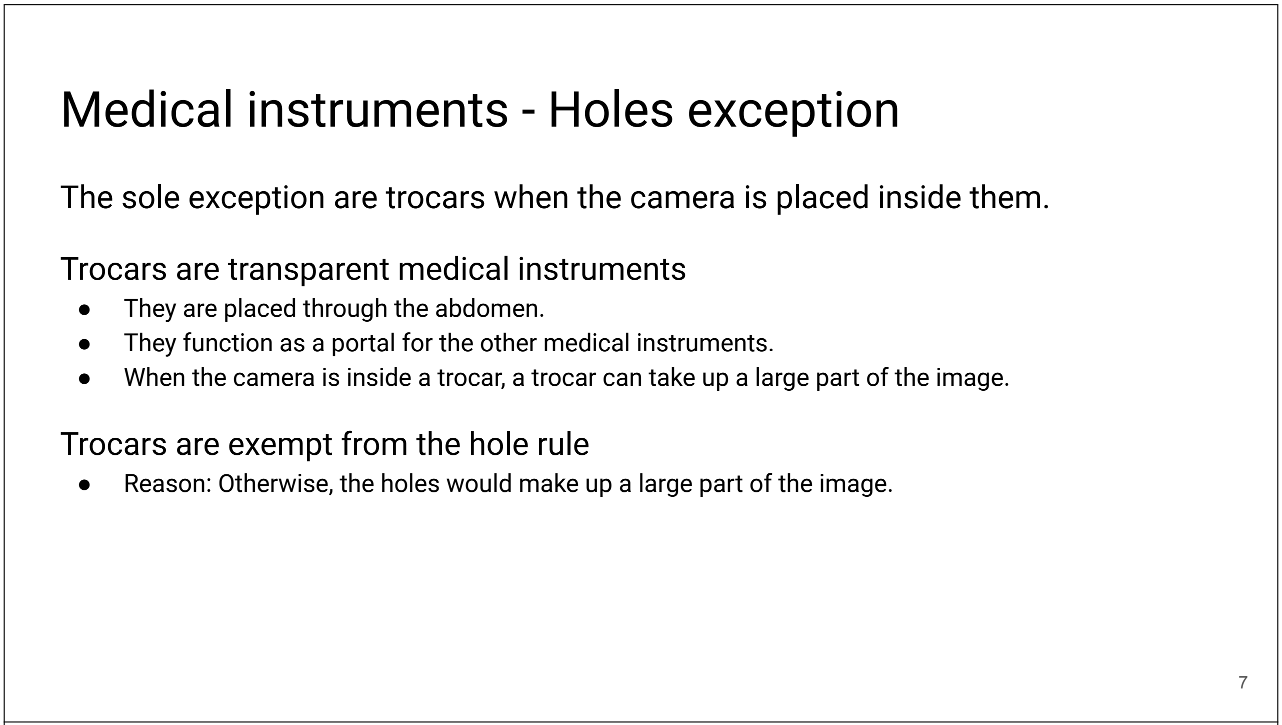
Several medical instruments feature holes.
They are regarded as part of the instrument.

A hole is made up of pixels that do not show

parts of the instrument but are either

type a) completely surrounded by pixels of the same instrument or

type b) completely surrounded by pixels of one instrument and the margin of the instrument would close the hole outside the image.



Medical instruments - Holes exception

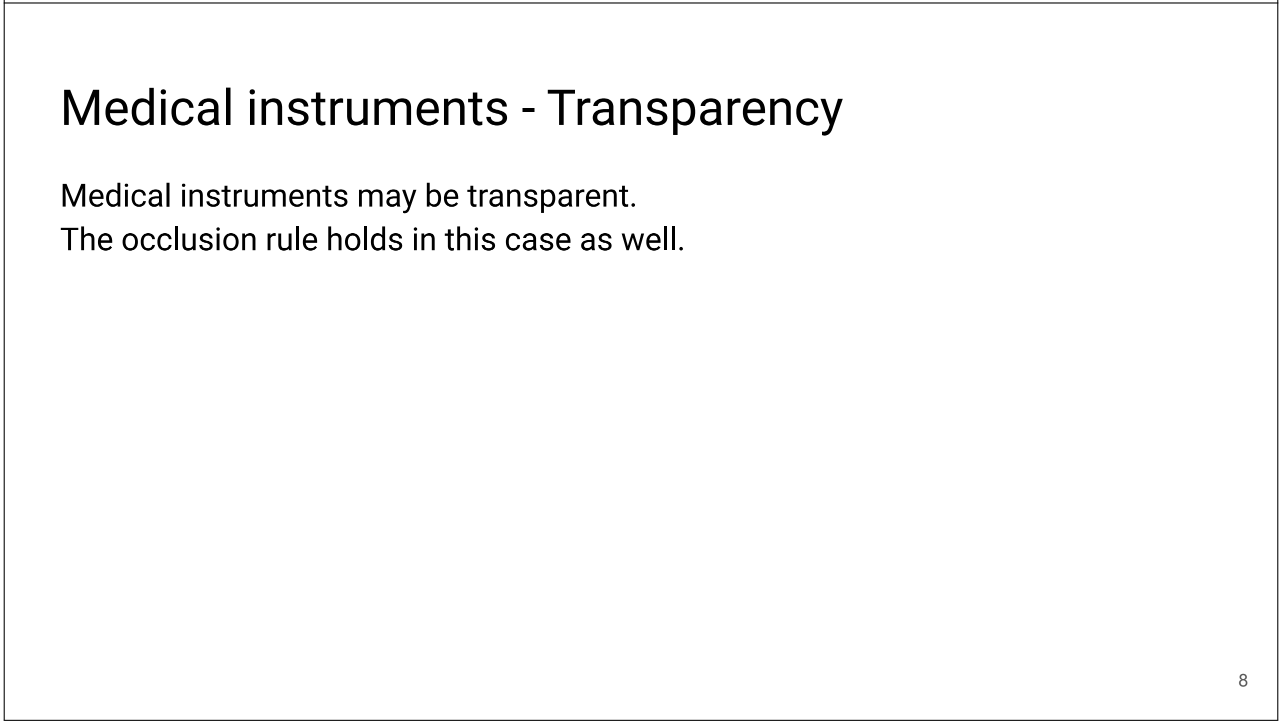
The sole exception are trocars when the camera is placed inside them.

Trocars are transparent medical instruments

- They are placed through the abdomen.
- They function as a portal for the other medical instruments.
- When the camera is inside a trocar, a trocar can take up a large part of the image.

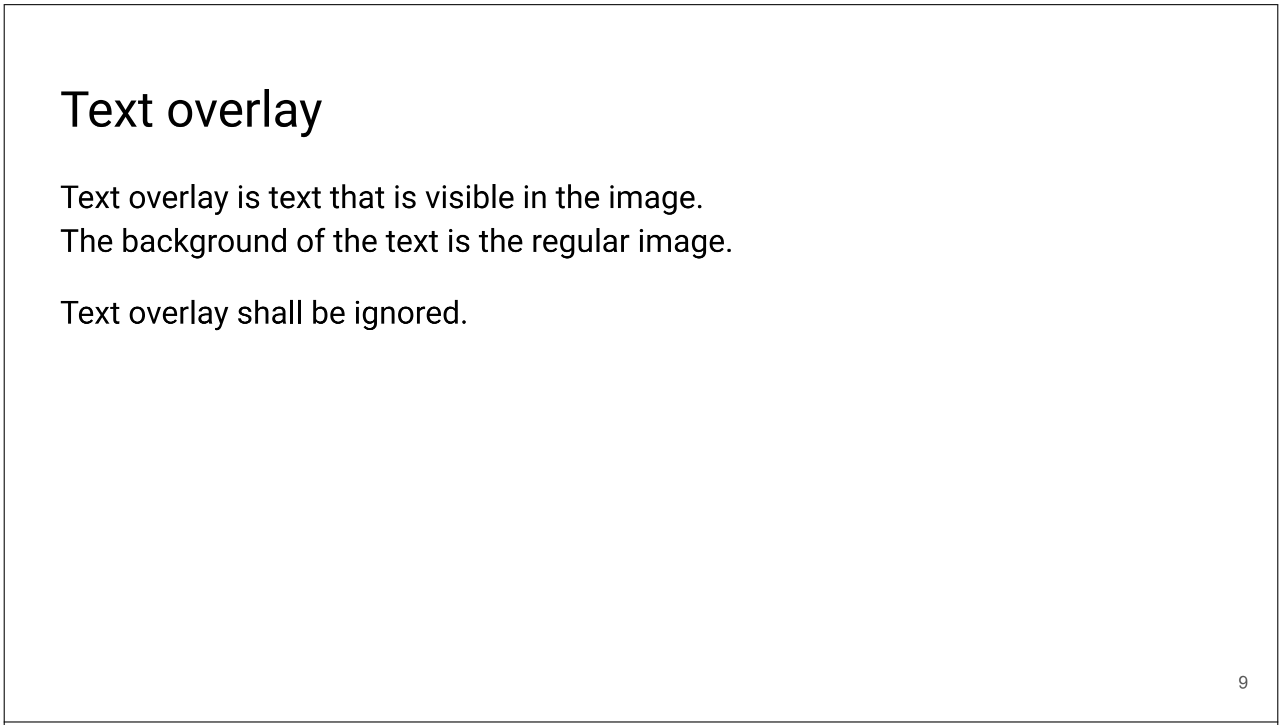
Trocars are exempt from the hole rule

- Reason: Otherwise, the holes would make up a large part of the image.



Medical instruments - Transparency

Medical instruments may be transparent.
The occlusion rule holds in this case as well.



Text overlay

Text overlay is text that is visible in the image.
The background of the text is the regular image.

Text overlay shall be ignored.

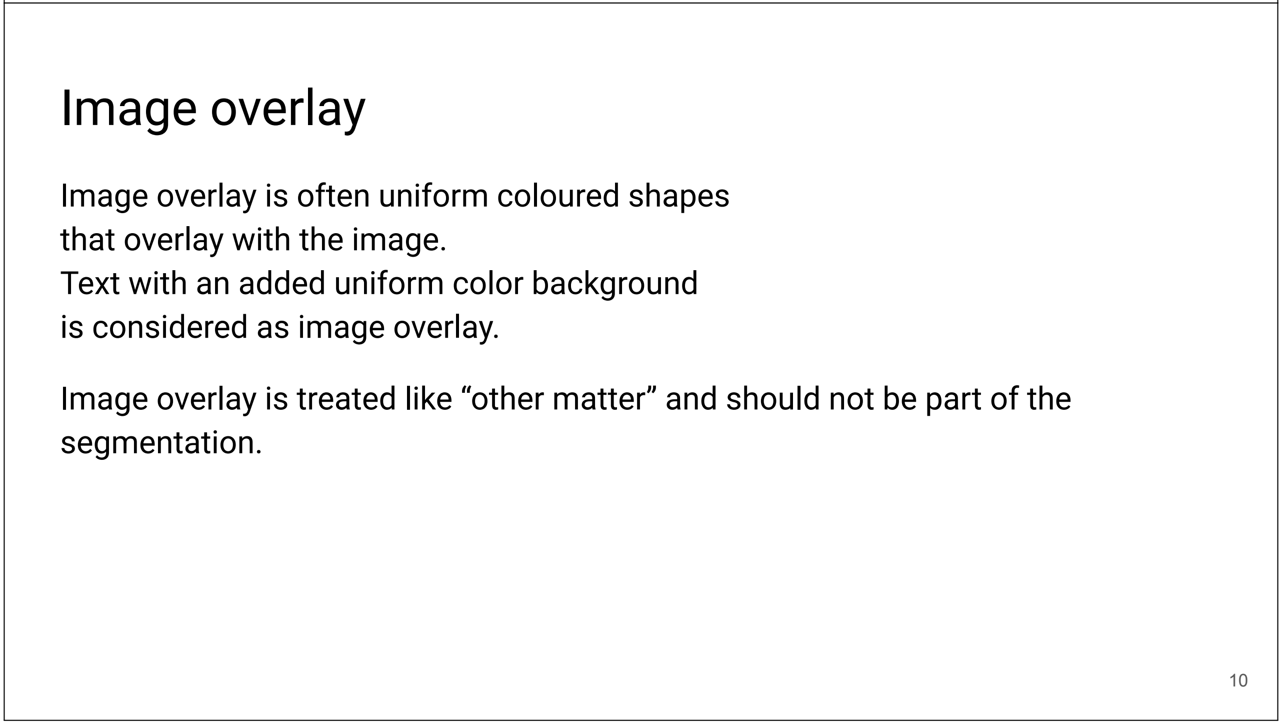
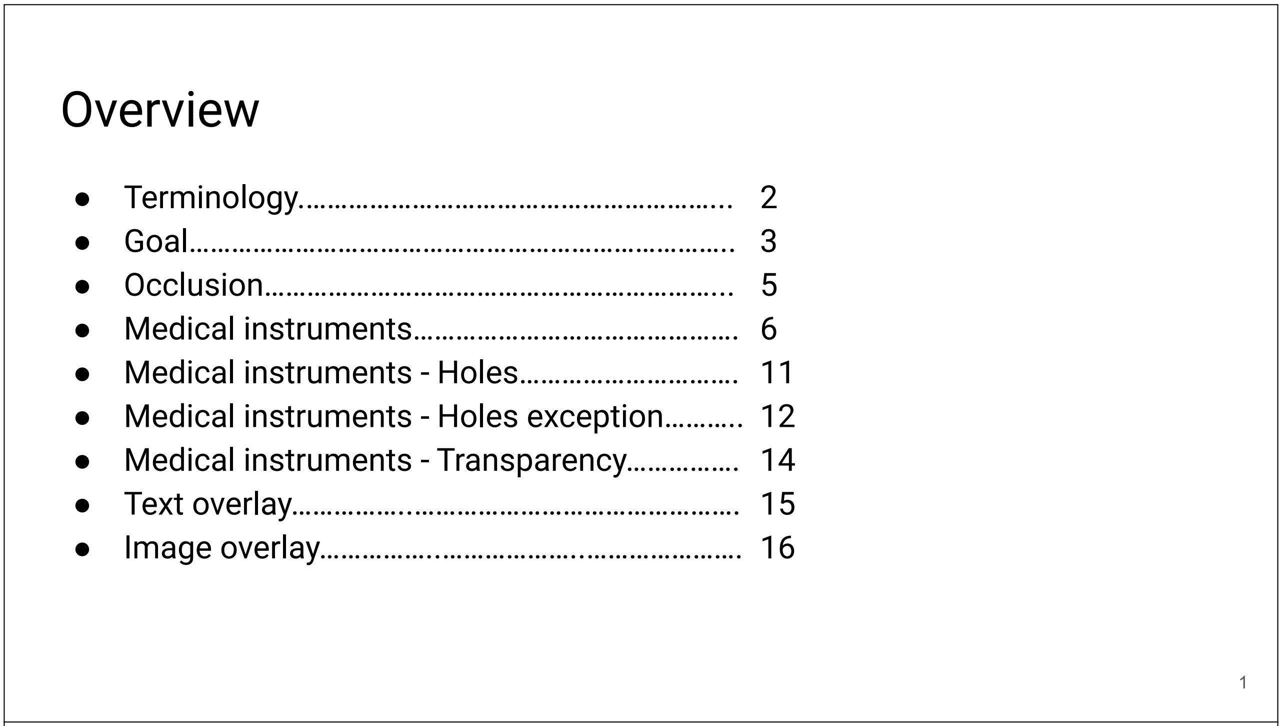


Image overlay

Image overlay is often uniform coloured shapes
that overlay with the image.
Text with an added uniform color background
is considered as image overlay.

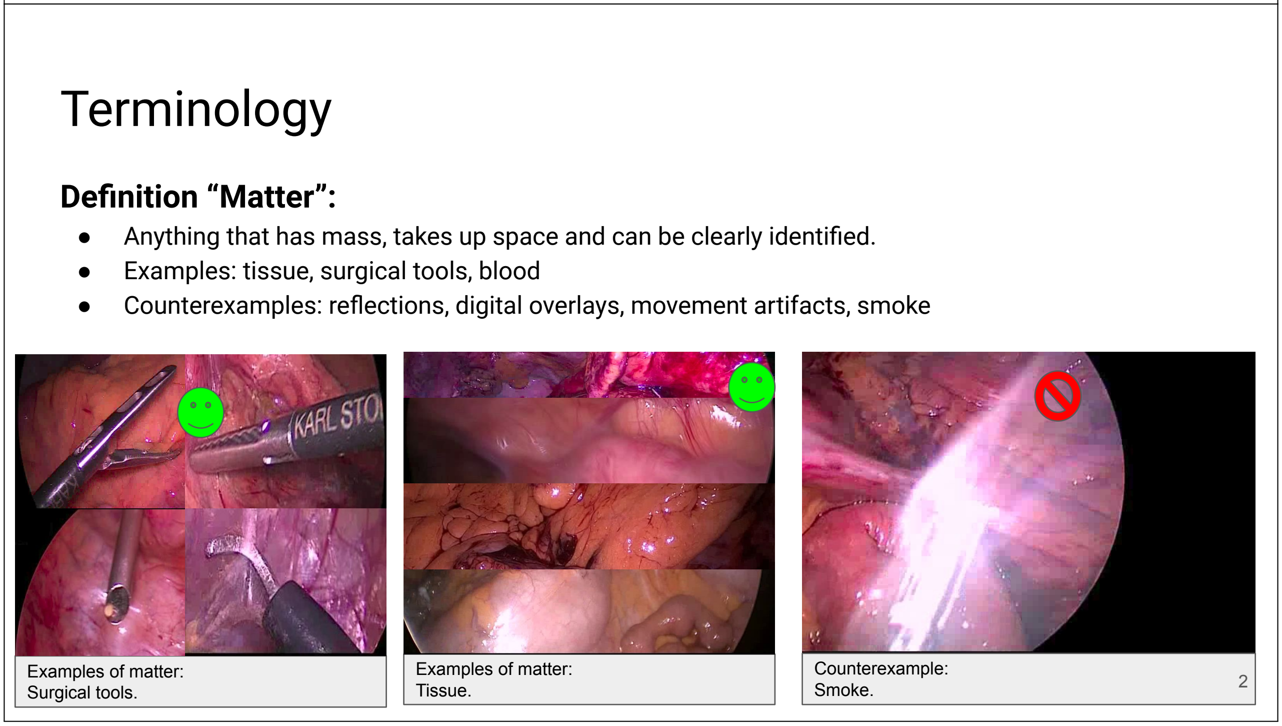
Image overlay is treated like “other matter” and should not be part of the segmentation.

8 Extended text including pictures labeling instruction



Overview

- Terminology.. 2
- Goal.. 3
- Occlusion.. 5
- Medical instruments..................................... 6
- Medical instruments - Holes......................... 11
- Medical instruments - Holes exception.......... 12
- Medical instruments - Transparency............. 14
- Text overlay.. 15
- Image overlay.. 16



Terminology

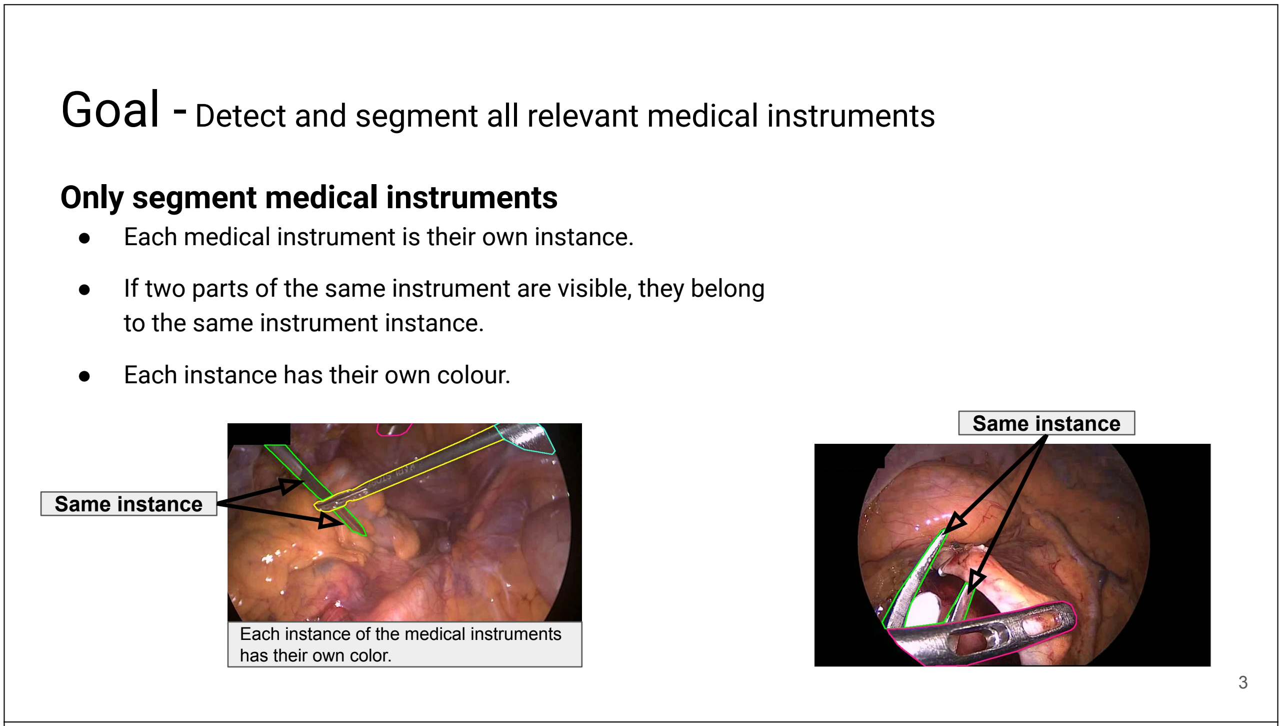
Definition “Matter”:

- Anything that has mass, takes up space and can be clearly identified.
- Examples: tissue, surgical tools, blood
- Counterexamples: reflections, digital overlays, movement artifacts, smoke

Examples of matter:
Surgical tools.

Examples of matter:
Tissue.

Counterexample:
Smoke.



Goal - Detect and segment all relevant medical instruments

Only segment medical instruments

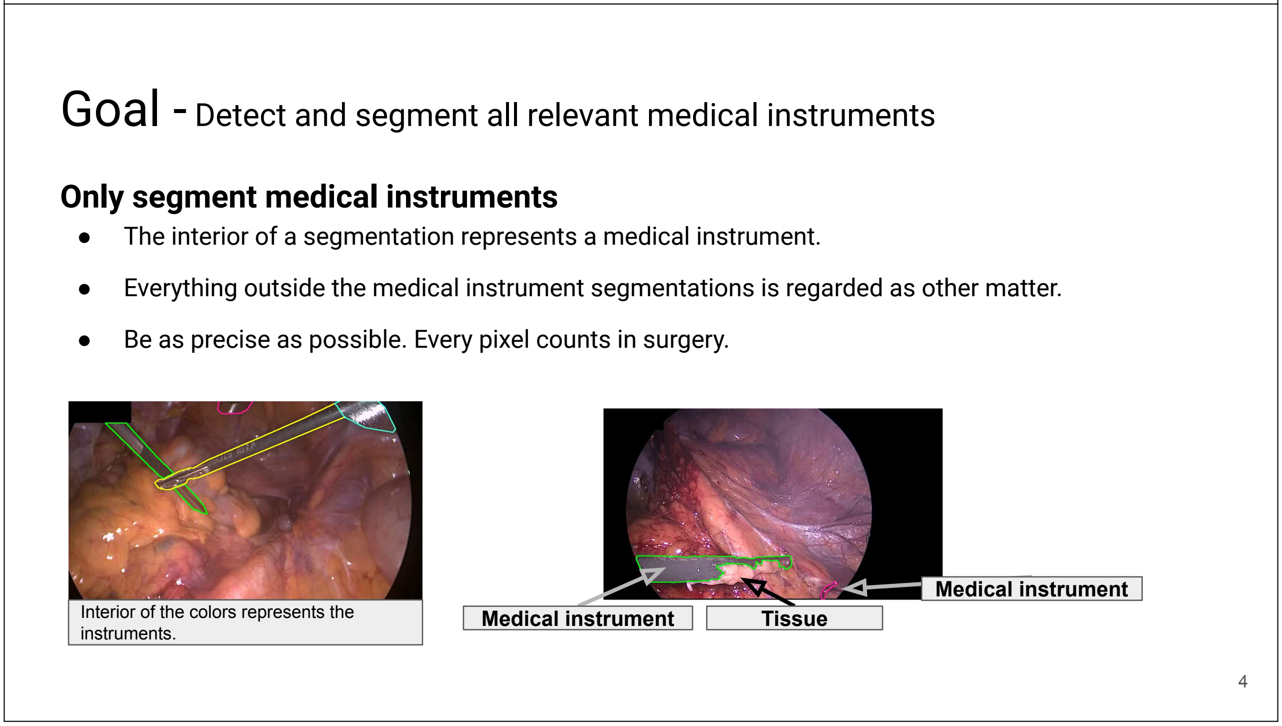
- Each medical instrument is their own instance.
- If two parts of the same instrument are visible, they belong to the same instrument instance.
- Each instance has their own colour.

Same instance

Each instance of the medical instruments has their own color.

Same instance

3



Goal - Detect and segment all relevant medical instruments

Only segment medical instruments

- The interior of a segmentation represents a medical instrument.
- Everything outside the medical instrument segmentations is regarded as other matter.
- Be as precise as possible. Every pixel counts in surgery.

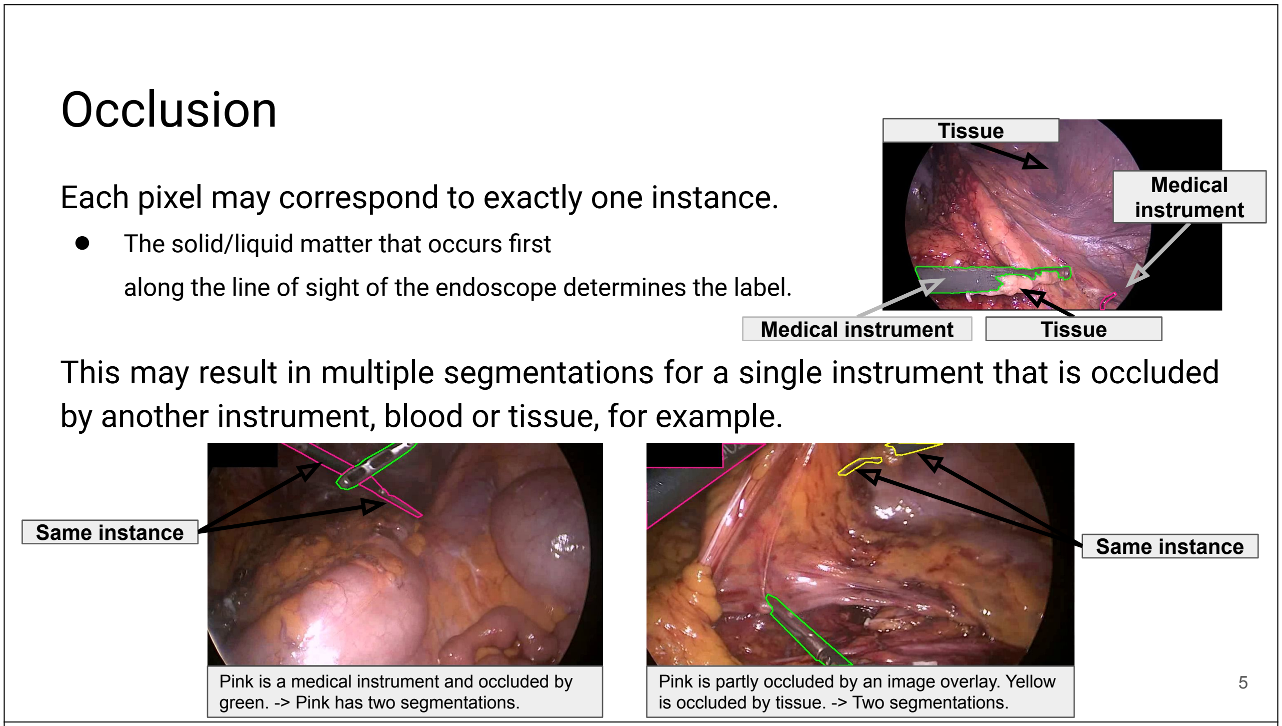
Interior of the colors represents the instruments.

Medical instrument

Tissue

Medical instrument

4



Occlusion

Each pixel may correspond to exactly one instance.

- The solid/liquid matter that occurs first along the line of sight of the endoscope determines the label.

Tissue

Medical instrument

Medical instrument

Tissue

This may result in multiple segmentations for a single instrument that is occluded by another instrument, blood or tissue, for example.

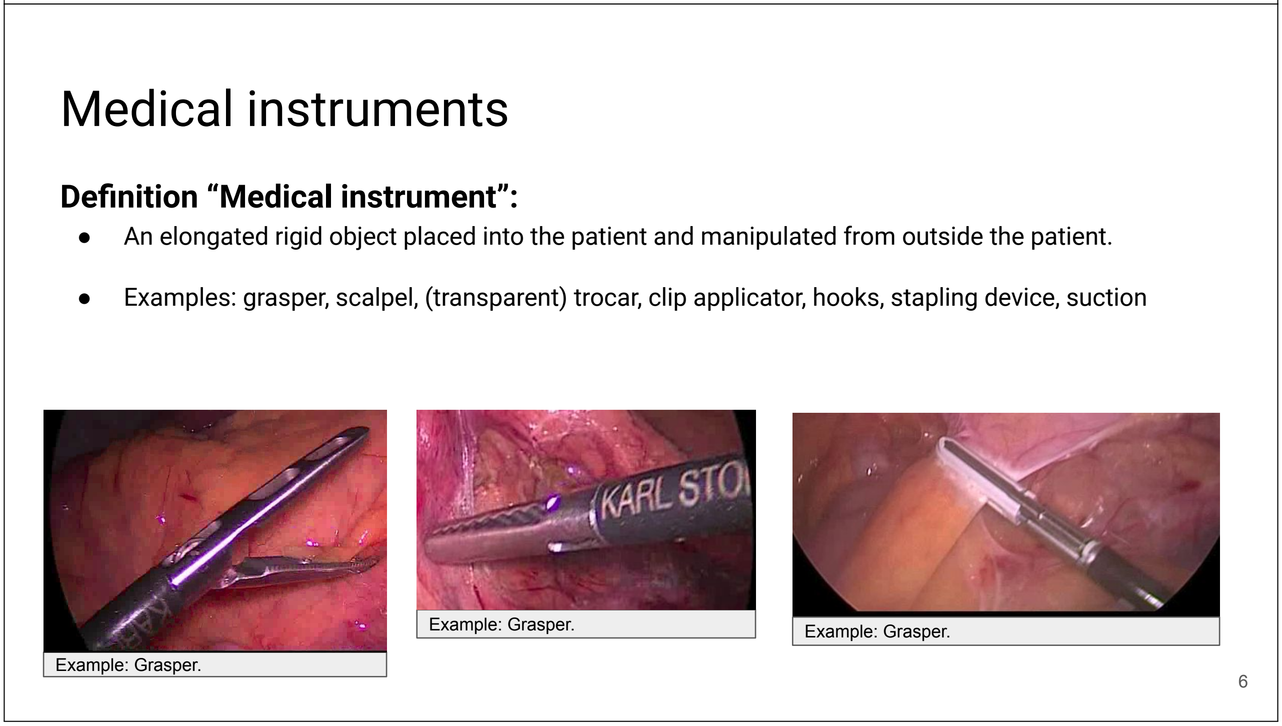
Same instance

Pink is a medical instrument and occluded by green. -> Pink has two segmentations.

Pink is partly occluded by an image overlay. Yellow is occluded by tissue. -> Two segmentations.

Same instance

5



Medical instruments

Definition “Medical instrument”:

- An elongated rigid object placed into the patient and manipulated from outside the patient.
- Examples: grasper, scalpel, (transparent) trocar, clip applicator, hooks, stapling device, suction

Example: Grasper.

Example: Grasper.

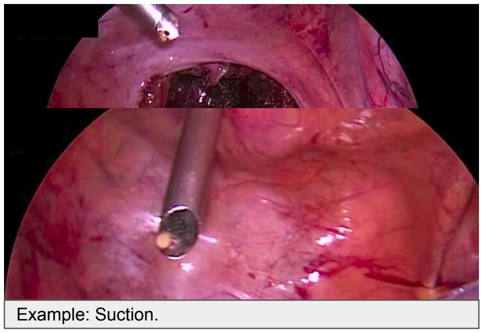
Example: Grasper.

6

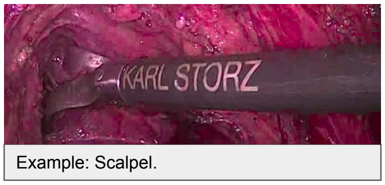
Medical instruments

Definition “Medical instrument”:

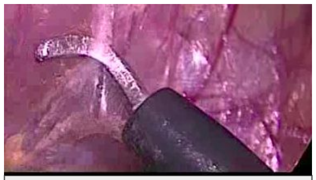
- An elongated rigid object placed into the patient and manipulated from outside the patient.
- Examples: grasper, scalpel, (transparent) trocar, clip applicator, hooks, stapling device, suction



Example: Suction.



Example: Scalpel.

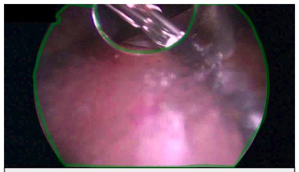


Example: Spreader.

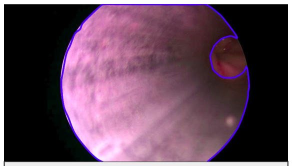
Medical instruments

Definition “Medical instrument”:

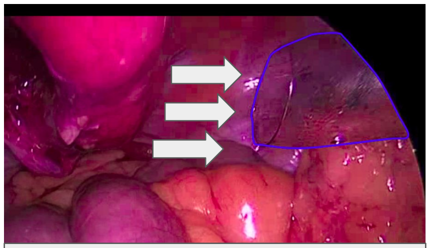
- An elongated rigid object placed into the patient and manipulated from outside the patient.
- Examples: grasper, scalpel, (transparent) trocar, clip applicator, hooks, stapling device, suction



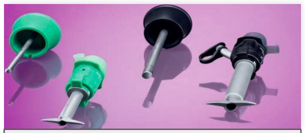
Example: Inside trocar (green). Another instrument is visible through the hole.



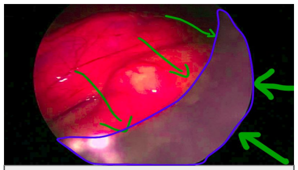
Example: Inside a Trocar.



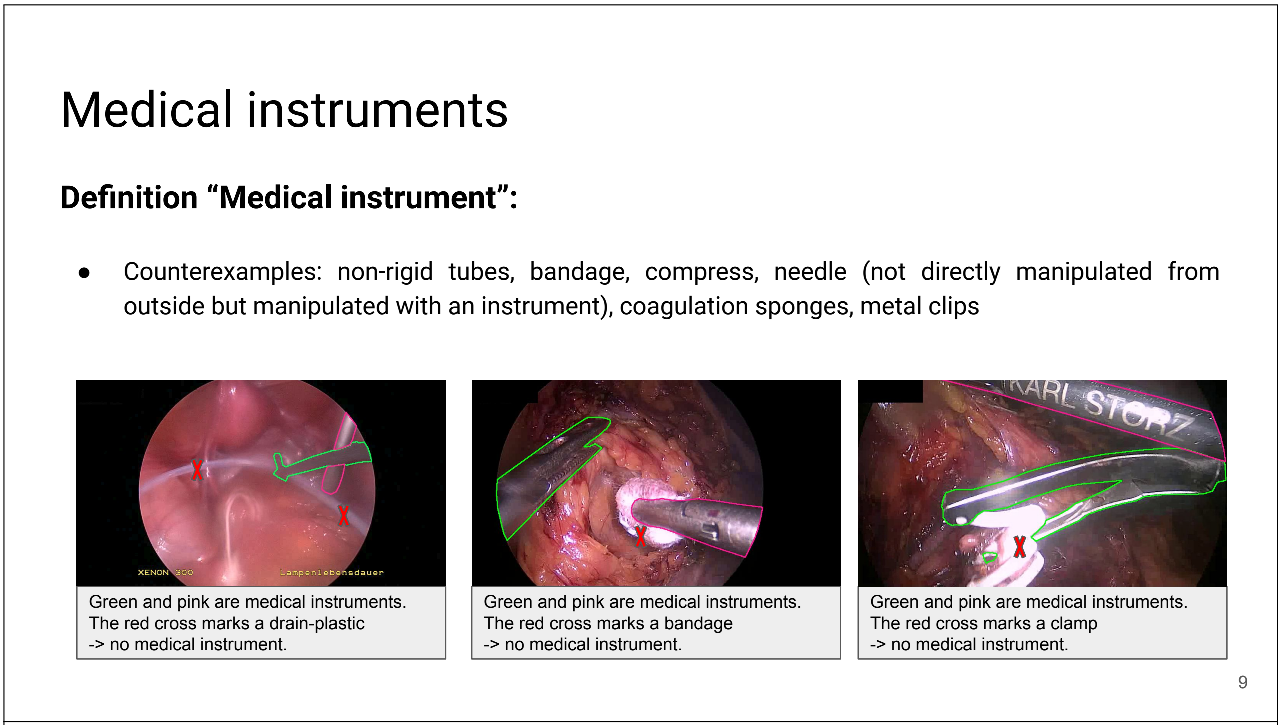
Example: transparent trocar (hard to detect).



Example: Different trocars.



Example: Trocar.



Medical instruments

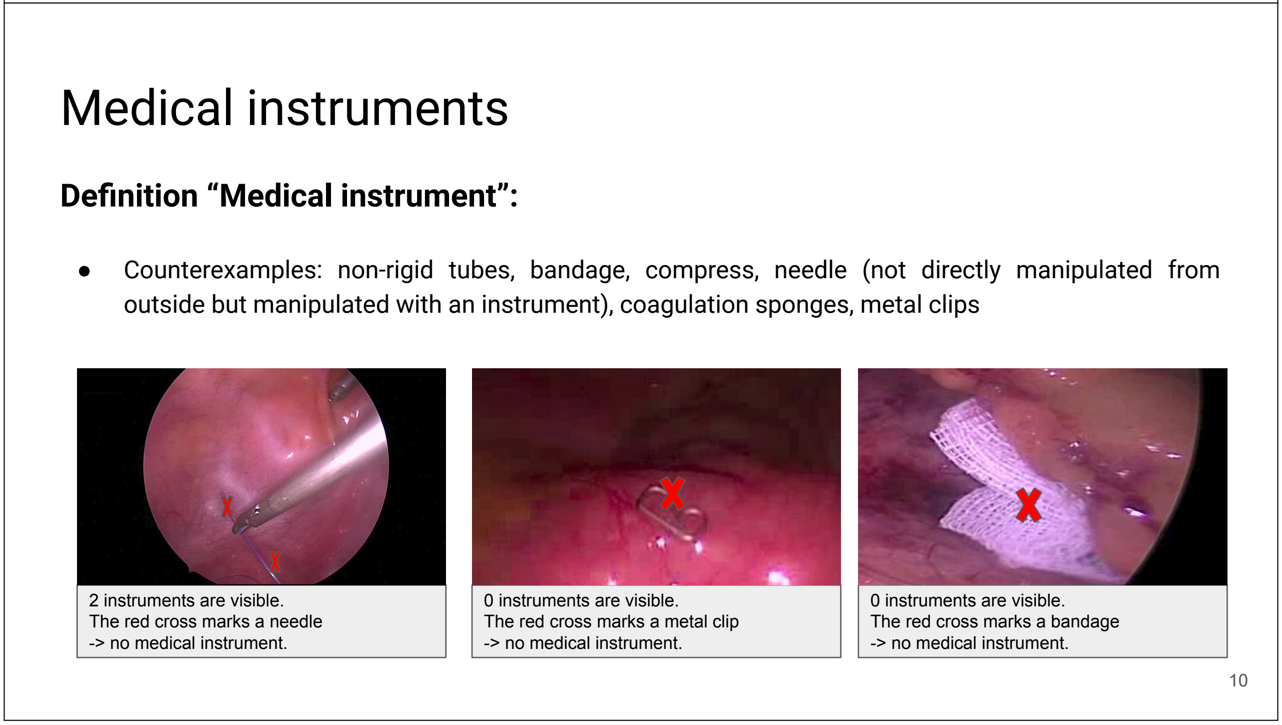
Definition "Medical instrument":

- Counterexamples: non-rigid tubes, bandage, compress, needle (not directly manipulated from outside but manipulated with an instrument), coagulation sponges, metal clips

Green and pink are medical instruments.
The red cross marks a drain-plastic
-> no medical instrument.

Green and pink are medical instruments.
The red cross marks a bandage
-> no medical instrument.

Green and pink are medical instruments.
The red cross marks a clamp
-> no medical instrument.



Medical instruments

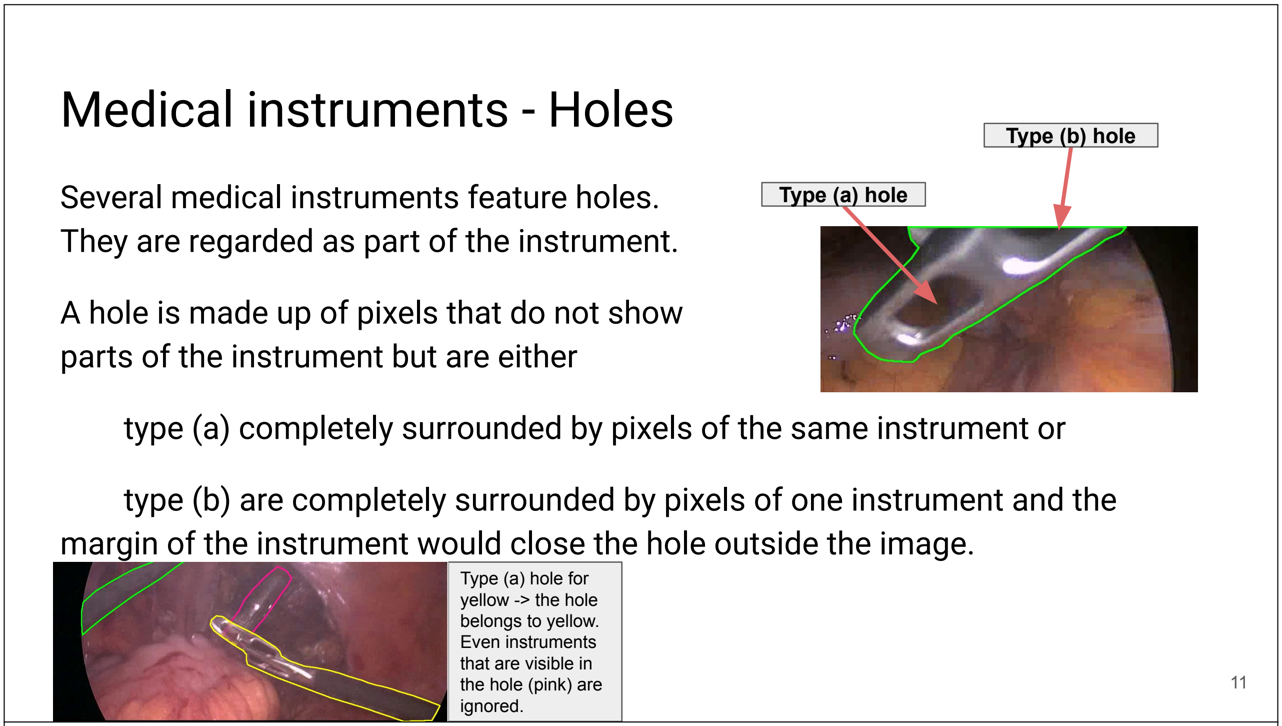
Definition "Medical instrument":

- Counterexamples: non-rigid tubes, bandage, compress, needle (not directly manipulated from outside but manipulated with an instrument), coagulation sponges, metal clips

2 instruments are visible.
The red cross marks a needle
-> no medical instrument.

0 instruments are visible.
The red cross marks a metal clip
-> no medical instrument.

0 instruments are visible.
The red cross marks a bandage
-> no medical instrument.



Medical instruments - Holes

Several medical instruments feature holes. They are regarded as part of the instrument.

A hole is made up of pixels that do not show parts of the instrument but are either

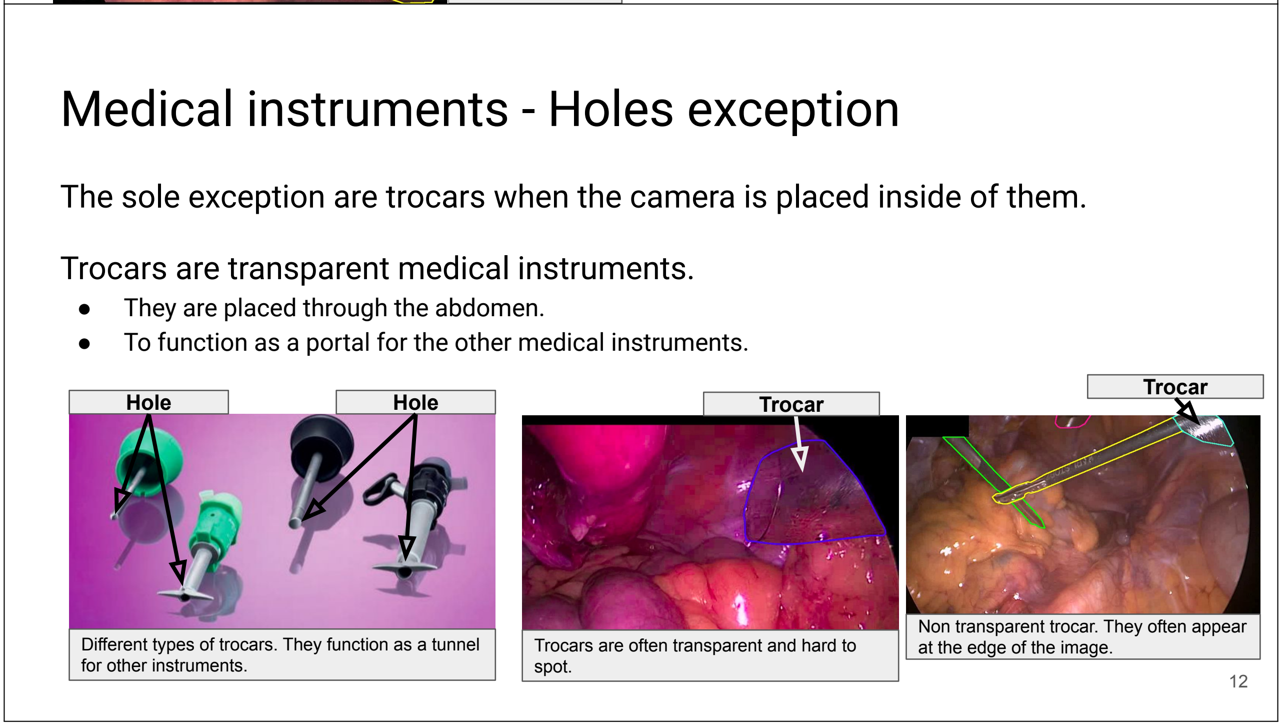
type (a) completely surrounded by pixels of the same instrument or

type (b) are completely surrounded by pixels of one instrument and the margin of the instrument would close the hole outside the image.

Type (b) hole

Type (a) hole

Type (a) hole for yellow -> the hole belongs to yellow. Even instruments that are visible in the hole (pink) are ignored.



Medical instruments - Holes exception

The sole exception are trocars when the camera is placed inside of them.

Trocars are transparent medical instruments.

- They are placed through the abdomen.
- To function as a portal for the other medical instruments.

Hole

Hole

Different types of trocars. They function as a tunnel for other instruments.

Trocar

Trocars are often transparent and hard to spot.

Trocar

Non transparent trocar. They often appear at the edge of the image.

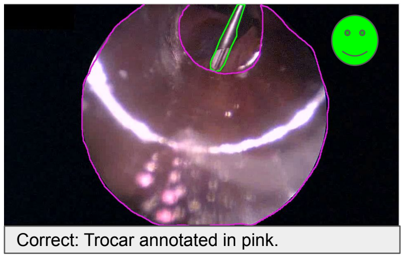
Medical instruments - Holes exception

The sole exception are trocars when the camera is placed inside of them.

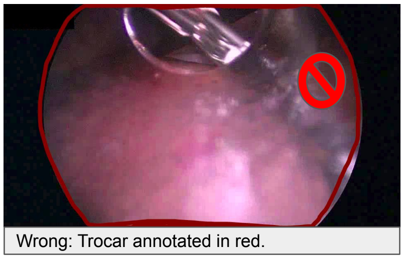
When the camera is inside a trocar, a trocar can take up a large part of the image

Trocars are exempt from the hole rule.

- Otherwise, the holes would contain a large part of the image.



Correct: Trocar annotated in pink.

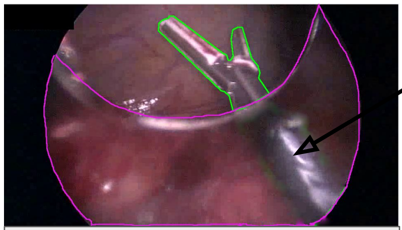


Wrong: Trocar annotated in red.

Medical instruments - Transparency

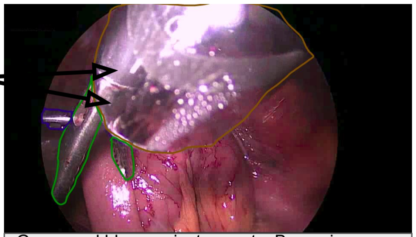
Medical instruments may be transparent.

The occlusion rule holds in this case as well.

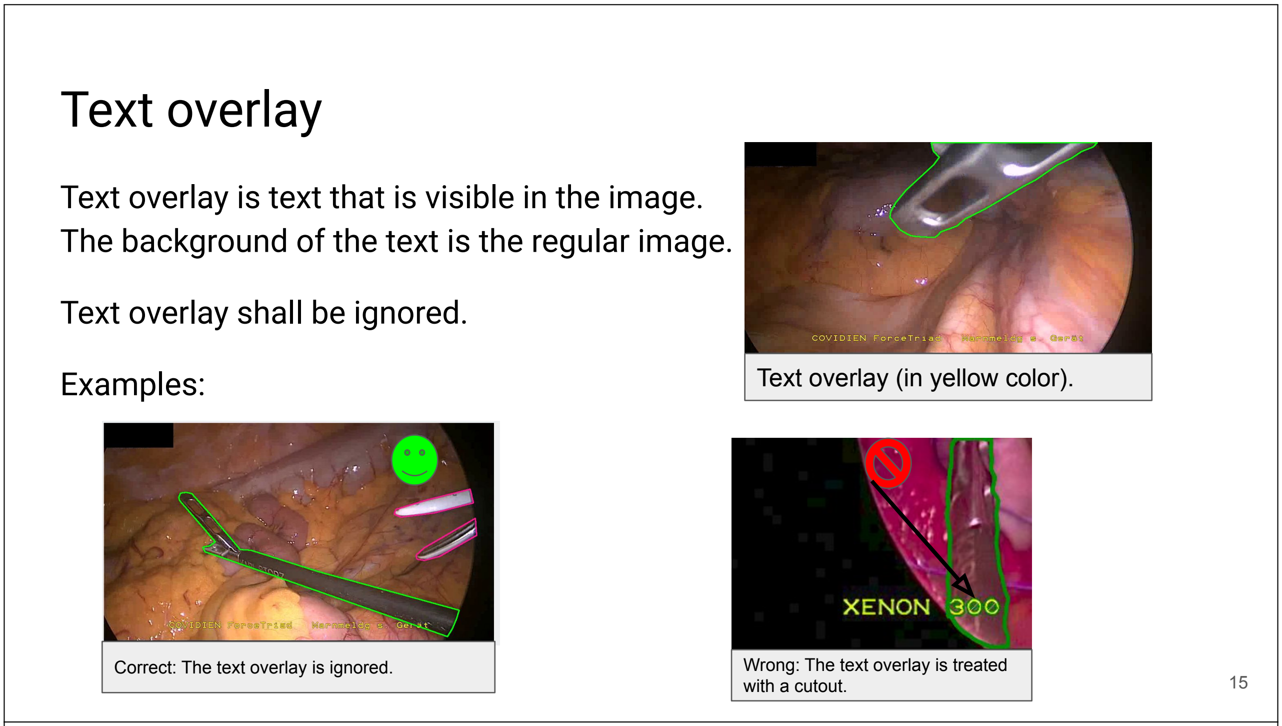


Hole exception for pink (troncar). Pink is transparent. -> The hidden part of green belongs to the pink annotation (occlusion rule).

Pixels only belong to one instrument. Here they belong to the transparent instrument, because the transparent instrument is the first object in line of sight (occlusion rule).



Green and blue are instruments. Brown is a transparent trocar. -> The trocar is the first part in the line of sight. Hidden part of the green instrument belongs to brown (occlusion rule).



Text overlay

Text overlay is text that is visible in the image.
The background of the text is the regular image.

Text overlay shall be ignored.

Examples:

Text overlay (in yellow color).

Correct: The text overlay is ignored.

Wrong: The text overlay is treated with a cutout.

15

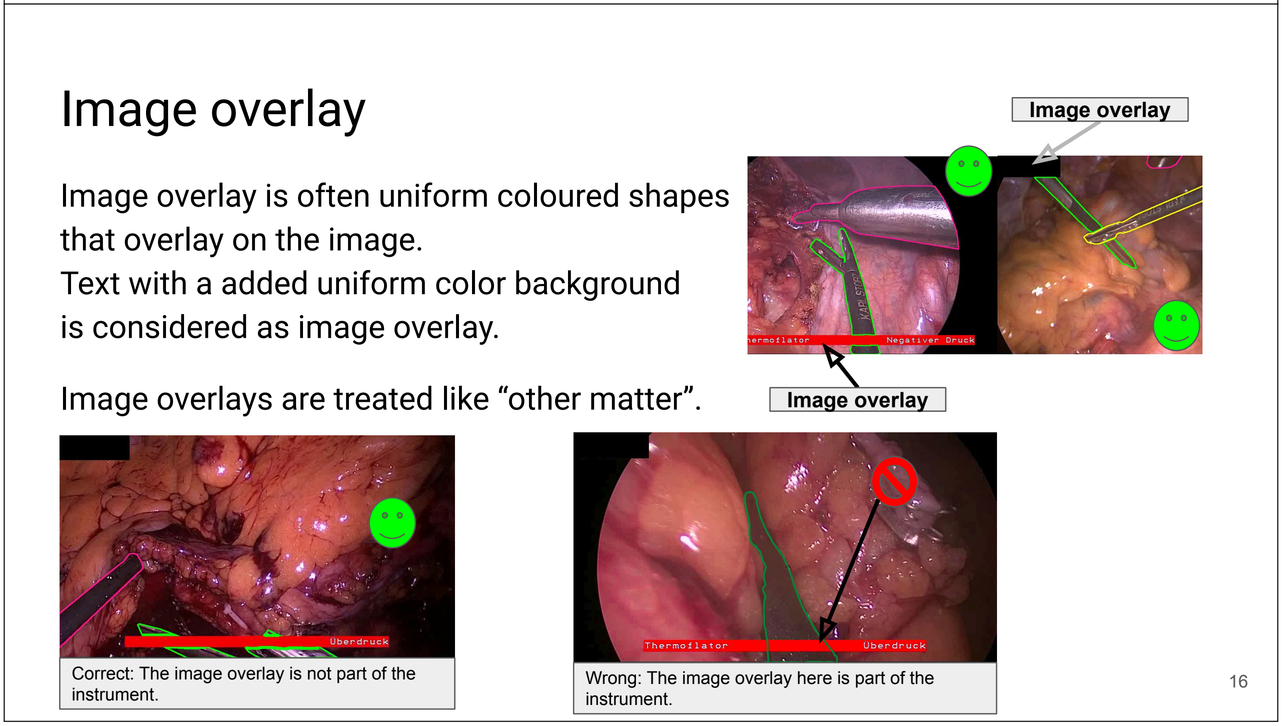


Image overlay

Image overlay is often uniform coloured shapes that overlay on the image.
Text with a added uniform color background is considered as image overlay.

Image overlays are treated like "other matter".

Image overlay

Image overlay

Correct: The image overlay is not part of the instrument.

Wrong: The image overlay here is part of the instrument.

16