Computer Vision –
Identify Surgical Instrument

Challenge Rules & Guidelines

Contents

1. Overview

Computer vision has advanced considerably but is still challenged in matching the precision of human intelligence. Could AI assist human to perform Surgical instrument counting go in and out of Hospital Operating Theatre – a task that take place over hundred thousand times per year!

In this competition, you are challenged to build a machine learning model that identifies the categories of Ophthalmology surgical instrument in an images dataset on HKSTP's latest service addition – Validation Platform where Optimal Performance Metrics can be formed. By automating this task, your participation could help free enormous amount of health professionals time for patient connection.

Phaco and Intraocular, two common instrument sets used in ocular surgery around 10,000 times annually, will be used for the competition. By automating this task, your participation could help free the time of health professionals.

2. Data

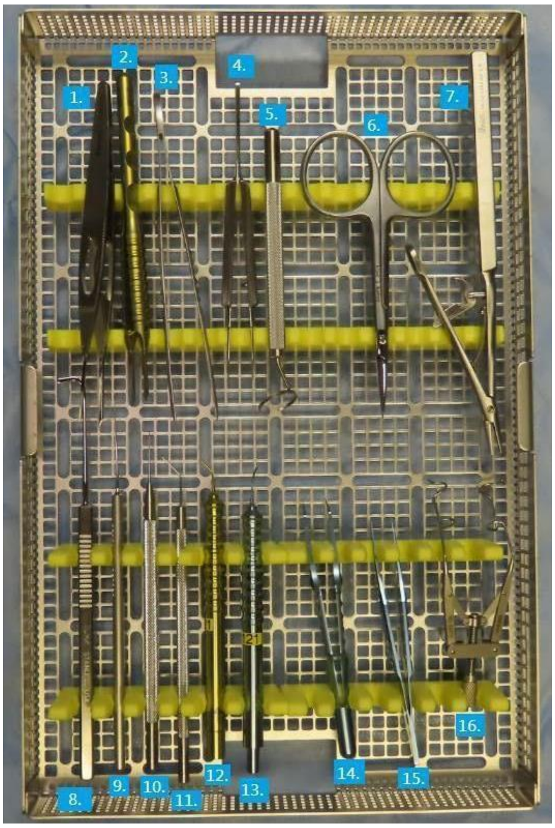
Contestants are expected to train their model to recognise surgical instruments on an instrument tray. Contestants are expected to find the data to train their model.

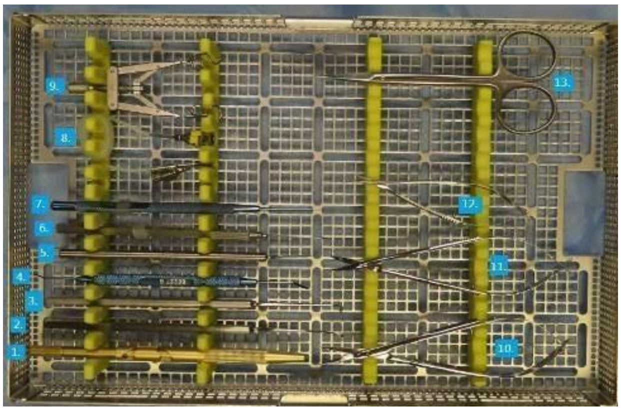
<u>Instrument Trays</u>

The surgical instruments will be placed on an *ophthalmic instrument sterilization tray*. There will be around 100 questions in the Game, as shown in the following diagram.

There will be only 2 different sterilization trays used in the Game, as illustrated in the following diagram 1) left or right bottom photo (same tray) and 2) right top photo. Photo might be in horizontal or vertical.

Instruments will be arranged tidily on the yellow holders. However, the instrument type, the position and the orientation will be changed in different testing photos, e.g. item 8 in the left photo might be replaced with item 1, or item 8 will be taken out from the testing image, or item 8 will be displaced with 180-degree rotation and the needle is pointing downside.



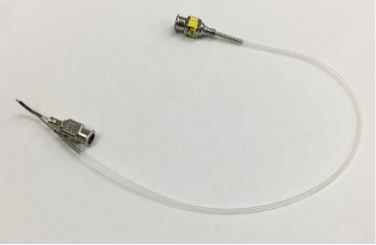


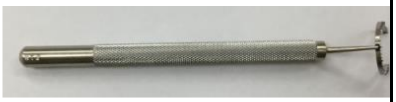
Surgical Instruments

There are **15 categories of surgical instruments** that contestants are expected to annotate the position, identify the **category** name, and count.

There will be additional surgical instruments, **not included within the 15 categories**, that will be included in the trays and are intended to create false positives. These tools should be ignored and should not be identified.

Model names are provided below for each category to aid the data collection process:

Category Name	**Model Name**	**Image**
Cannula	• Simcoe IA, cannula simcoe I/A: katena K7-4300	 
Chopper	• Nagahara Chop phaco, Chop, phaco: Asico AE-2515	 
Dressing Forceps	• Alcon loading fx, Forceps IOL loading: alcon 560.01 • Fx. Conjunctival moorfields, Surtex FR-838-10	 

Fixation Ring	• Ring Nichamin ring, Ring globe fixation: Rheine 8-15217	 
Handpieces	• Aspiration handpiece, Aspiration handpiece, 23G 0.65mm: D&K 8-657 • Irrigation handpiece, irrigation handpiece 23G 0.65mm: D&K 8-652-1	
Hook Surgical	• Hook muscle, Hook muscle Graefe: storz E0592 • Sinskey II lens hook round, hook manipulating lens, Sinskey: katena K3-5232	
Iris Scissors	• Sc Iris straight, Scissors iris: Storz E3404 • Sc Iris straight, Scissors iris: katena K4-7400	 
Knife Scalpel Handles	• Knife diamond 45o, Knife diamond: Meyco ME-600 • Knife tooke's, Knife cornea: katena K2-3650	 
Micro Scissors	• Scissors, Williamson Noble, Scissors conjunctiva: dixey D680,680a • Castroviejo corneal, Scissors cornea: storz E3220	

	• Straight vannas, Scissors capsulotomy: storz E3386	
Needle Holders	• Needle holder Castroviejo, Needle holder: Dixey D605M4 • Needle holder Silcock, Needle holder: dixey D600	 
Pusher	• Pusher, lens (straight), Pusher, IOL , Aker: katena K3-2720	
Spatula Surgical	• Repositor iris, Repositor iris: shandeng 88227B • Repositor iris, Repositor iris: Storz E706	 
Speculum	• Speculum Lieberman, Speculum, k wire, 15mm: katena K1-5671/katena K1-5675	 
Spoon Surgical	• Vectus, lens loop: dixey D261	 
Tissue Forceps	• Mcpherson angled, Forceps tying: storz E1815A • Mcpherson straight, Forceps tying straight: Storz E1815S • Fx, notched st cornea, Forceps notched pierse 0.25mm: D&K 2-100	

	• Fx loading D&K (AMO's), Forceps IOL loading: D&K DK7726 • Fx, capsulorhexis utrata, Forceps utrata capsulorhexis: katena K5-5082 • Mcpherson angled, Forceps tying: D&K 2-522, 2-522E • Fx, notched st cornea, Forceps notched pierse 0.25mm: D&K 2-100	

Additional reference websites are provided:

- http://novosurgical.com/
- https://www.eyecareandcure.com/ECC-Products/Surgical-Instruments
- https://www.storzeye.eu/our-instruments/
- https://www.titanmedicalshop.com/products/
- https://www.accutome.com/

3. Evaluation Criteria

Each contestant's model will be scored based on its localisation, classification, and counting accuracy.

<u>a) Localisation</u>

The accuracy for identifying the position of each surgical instrument.

Contestants are expected to annotate the position of each surgical instrument on the tray using bounding boxes. Intersection over Union (IoU) will be used to evaluate the localisation and contestants will be penalised for false-positive bounding boxes.

<u>b) Classification</u>

The accuracy for identifying which surgical instruments are on the tray.

Contestants are also expected to name all the surgical instruments on the tray. Area Under Curve of Precision-Recall (AUCPR) **per class** will be used to evaluate classification performance.[1] The classification result is considered correct when the IoU is greater than 0.5. The prediction score should be the probability of the class present.

[1] https://www.kdnuggets.com/2020/08/metrics-evaluate-deep-learning-object-detectors.html

$$AUCPR = \sum_{k=0}^{1} P(k)\Delta r(k)$$

k = Threshold
P(k) = The precision at a cutoff
$\Delta r(k)$ = *The change in recall that happened between cutoff* k-1 *and cutoff* k

$$\text{Precision} = \frac{\text{TP}}{\text{TP + FP}}$$

$$Recall = \frac{TP}{TP + FN}$$

c) Counting Accuracy

The accuracy for counting how many of each surgical instrument is present.

Contestants are expected to count how many of each surgical instrument is on the tray.

The **object precision rate per class** will be used to measure the precision of detected objects on the prediction. OPR = 1 means the object counting prediction perfectly matches the ground truth.

$$OPR = 1 - \frac{|L - A|}{\mathrm{A}}$$

OPR - Objects precision rate
A - The number of all sample
L - The total number of prediction objects

Winner Determination

There will be 1 winner from each evaluation criteria and 1 overall winner. To determine the overall winner, the score from each evaluation criteria will be normalised, weighted, and combined, and the team with the highest overall score will be chosen.

$$\bar{x} = \sum R_t \times W_t$$

$\bar{x}$ = *Final score*

t = *Tasks, t {classification, localization, counting}*

R_t = *Group ranking in the tasks*

W_t = *Weight of the tasks, W {classification: 0.4, localization: 0.35, counting: 0.25}*

4. Submission

Each contestant's model will be tested by the Hospital Authority. Also, each contestant should place their model and program under ~/submission.

<u>Expected Output</u>

Contestants will need to read the file paths specified within *input_data.csv* to receive the questions (images).

For each question (image), contestants should output their answer to one row respectively in 2 files: *prediction_result.csv* and *counting_result.csv*. Notice that contestants should place the two files in a folder, namely result.

The *prediction_result.csv* file should contain:

> **First Column:** Image File Path
> **Second Column:** List of semi-colon separated values of <Double quoted Label Class> <prediction score> <xmin> <ymin> <xmax> <ymax>

The *counting_result.csv* file should contain:

> **First Column:** Image File Path
> **Second Column:** List of semi-colon separated values of <Double quoted Label Class> <prediction count>

The prediction program should have the following arguments (Python example):

> python <path-to-prediction-program> <input-data-csv-filename> <path-for-output-prediction-csv> <path-for-output-counting-csv>
>
> Your program should be able to executed by the following command:

```
python ~/submission/main.py /enigma/datasets/HA-Sample/input_data.csv  \
/enigma/local_storage/result/prediction_result.csv \
/enigma/local_storage/result/counting_result.csv
```

The .csv files should look like this:

input_data.csv (provided question):

ImagePath
test/intraocular1.jpg
test/phaco1.jpg

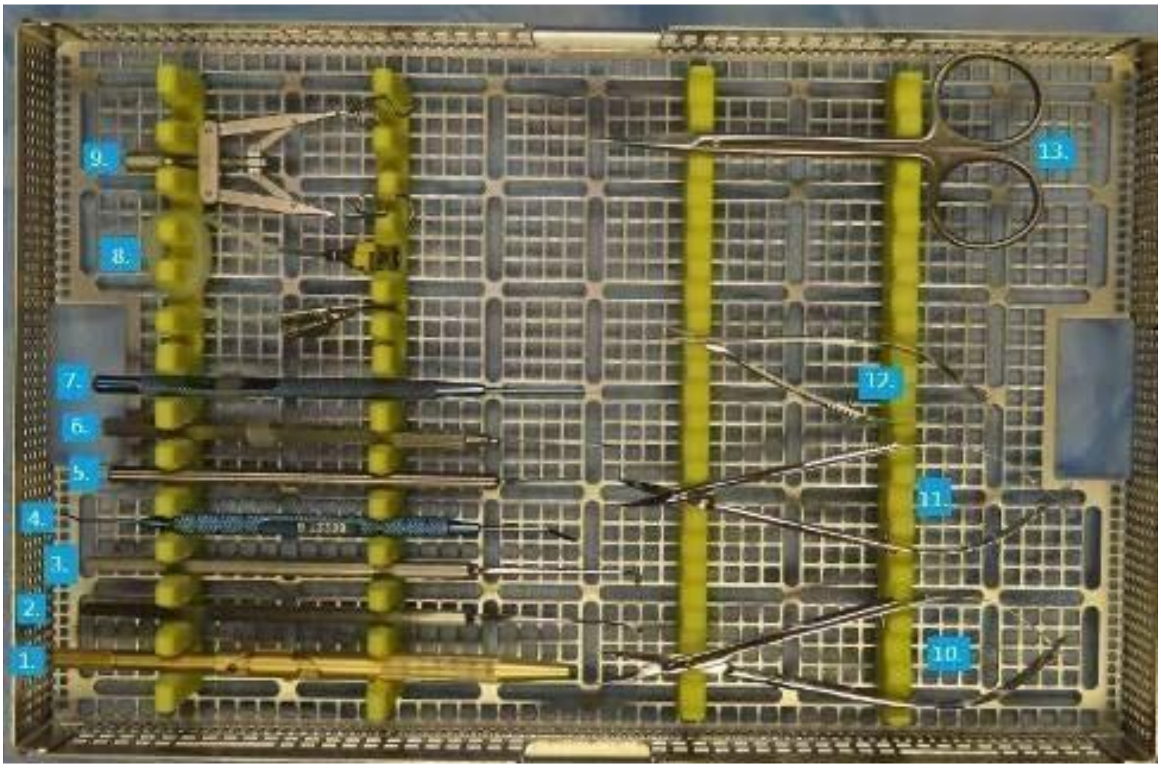
prediction_result.csv (contestant answer):

ImagePath	PredictionString (class prediction score xmin ymin xmax ymax;)
test/intraocular1.jpg	"Knife Scalpel Handles" 0.57 127 1668 1473 1755;"Hook Surgical" 0.45 202 1548 1662 1628;"Spoon S
test/phaco1.jpg	"Dressing Forceps" 0.68 1344 316 1453 832;"Tissue Forceps" 0.7 1478 301 1520 885;"Dressing Force

counting_result.csv (contestant answer):

ImagePath	PredictionString (class count;)
test/intraocular1.jpg	“Knife Scalpel Handles” 2;“Hook Surgical” 2;“Spoon Surgical” 1;“Spatula Surgical” 1;“Pusher” 1;“C
test/phaco1.jpg	"Dressing Forceps" 2;"Tissue Forceps" 4;"Fixation Ring" 1;"Iris Scissors" 1;"Needle Holders" 1;"Ho

The following is an example of a question and answer:



<u>Prediction Result</u>
"Knife Scalpel Handles" 0.57 127 1668 1473 1755;
"Hook Surgical" 0.45 202 1548 1662 1628;
"Spoon Surgical" 0.69 152 1316 1493 1395;
"Spatula Surgical" 0.90 270 1195 1507 1256;
"Knife Scalpel Handles" 0.87 256 1070 1379 1154
"Pusher" 0.71 232 963 1519 1022;
"Hook Surgical" 0.83 721 771 1027 876;
"Cannula" 0.92 711 620 1032 698;
"Speculum" 0.56 309 239 1090 559;
"Micro Scissors" 0.59 1553 1510 2770 1832;
"Micro Scissors" 0.64 1605 1145 2741 1426;
"Micro Scissors" 0.66 1720 835 2608 1092;
"Iris Scissors" 0.73 1551 129 2708 639;

<u>Counting Result</u>
“Knife Scalpel Handles” 2;
“Hook Surgical” 2;
“Spoon Surgical” 1;
“Spatula Surgical” 1;
“Pusher” 1;
“Cannula” 1;
“Speculum” 2;

“Micro Scissors” 3;
“Iris Scissors” 1;