Development of an AI solution for surgical gauze management

Ang, Soon Kim

2021

Ang, S. K. (2021). Development of an AI solution for surgical gauze management. Final Year Project (FYP), Nanyang Technological University, Singapore.
https://hdl.handle.net/10356/150810

https://hdl.handle.net/10356/150810

Development of an AI solution for surgical gauze management

Ang Soon Kim

SCHOOL OF MECHANICAL AND AEROSPACE ENGINEERING
NANYANG TECHNOLOGICAL UNIVERSITY

2020/2021

Project No: C057

Development of an AI solution for surgical gauze management

Submitted by: Ang Soon Kim

Supervisor: Assoc Prof Cai Yiyu

School of Mechanical and Aerospace Engineering

A final year project report presented to the Nanyang Technological University
in partial fulfilment of the requirements of the degree of
Bachelor of Engineering

2020/2021

Abstract

Gossypiboma/textiloma is an avoidable medical complication caused by human error. It has been inferred to be under reported due to the medicolegal nature of the situation with cases reported only for insurance claim. The surfacing of such cases causes unnecessary damage to the patient well-being as well as the surgeon and hospital reputation. Several solutions ranging from manually counting of the gauze to the physical tracking of the gauze using Radio Frequency Identification (RFID) tags has been implemented with varying success with each method having their own benefit and costs. This project aims to tackle the problem using the ever-evolving field of Artificial Intelligent (AI) which can be cheaper to implement compared to the current RFID solution as well as being less labor intensive compared to manual counting. A prototype system will be conceptualized and constructed by implementing the solution into common devices such as smart phone which could be easily deployed to various location. By assisting in the accounting of the gauze, human error can be minimized, and the prototype can be easily scaled up and distributed for practical usage to other organization that requires such a solution.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my project supervisor Prof. Cai Yiyu and Singapore General Hospital (SGH) surgeon, Dr. Luke Tay for providing me with an opportunity to work on a topic regarding AI. I would also like to thank Nicole Goh and Aaron Tham who have provided the explanation regarding how the current operating theatre works as well as its problem and solutions.

Secondly, I would like to thank the Ooi Jing Xuan for the raw dataset that she has accumulated that act as the basics where the I could build the new model on. I would also like to thank my friends and family for the mental support throughout the project.

Lastly, I would like thank Nanyang Technology University (NTU) School of Mechanical and Aerospace Engineering for this learning opportunity as well as the funds needed for the project prototype.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

RFID	Radio Frequency Identification
AI	Artificial Intelligent
SGH	Singapore General Hospital
NTU	Nanyang Technology University
WHO	World Health Organization
FPN	Feature Pyramid Network
SSD	Single Shot Detection
GPU	Graphic Processing Unit
Mask-RCNN	Masked-Region Based Convolutional Neural Networks
COCO	Common Objects in Context
API	Application Programming Interface
SOP	Standard of Procedure
IoU	Intersection over Union
mAP	mean Average Precision
mAR	mean Average Recall
MB	Mega-Bytes
SDK	Software Development Kit

Chapter 1
Introduction

1.1 Background

Gossypiboma or textiloma is a human caused medical complication where the surgical gauze is left within the patient body after a surgical procedure [1]. Patients with gossypiboma can remain asymptomatic [2-4] for different duration of time depending on their body conditions and the sterility of the gauze during insertion [1]. As patients starts to develop symptoms such as abdominal pain, nausea and vomiting due to bacterial overgrowth on the gauze [5] and inflammatory reaction [6] within the body, doctors would often misdiagnose the illness due to other potential causes such as tumor recurrence, postoperative intraabdominal adhesions, or small bowel invagination [4]. Common diagnostic tools such as X-ray and ultrasound examination are not effective in identifying the gauze as it only shows an abdominal mass [3] and cases are usually discovered during live operation [4]. Due to the medicolegal nature of the problem, widespread publication would be limited [5] with cases estimated from documented cases based on malpractice insurance claim [7]. Re-operation would be needed for patients diagnosed with gossypiboma to prevent further complication [1] and it puts unnecessary burden on all party involved. Hence, proposed solutions would be needed to help prevent the situation from occurring before the patients is discharged from their initial surgical operation.

1.2 Current Methodology

There are currently 3 unique solutions being practiced which prevents gossypiboma that can be used simultaneously to improve their performance. These methods are the manual counting of gauze, use of RFID gauze or detectable gauze. These methods use both low and high technology solutions to prevent gossypiboma with varying successes and shortcoming which would be further elaborated under literature review. However, there are still cases of gossypiboma occurring despite the introduction of these solutions which indicates the need for another layer of prevention.

1.3 Objectives

The objective of the project is the implementation of AI within the current system that is being implemented to prevent gossypiboma. This is a continuation of the previous project which uses AI as the means to track the gauze within a video feed. The aim of the current project is as follows:

1. Improvement to the previous project software to a more practical level.
2. Deployment of software to a functional platform.
3. The building of system to house the software for practical application.

1.4 Scope of Work

The project will work on the implementation of an improved AI within a portable device such as smart phone which can be used for deployment to test the usability of the AI model. Data collected from the previous project will be used as the base for additional model training with more images added if necessary.

Suitable accessory will be chosen with the smartphone to build a prototype that uses the AI model effectively and shows the potential implementation of such solution within the current system. The implementation of such a solution would also need to be more practical compared to the current solution that is implemented.

Chapter 2

Literature Review

2.1 Current Gossypiboma Measures

The current gossypiboma prevention methods consist of manual counting of gauze, use of RFID gauze or detectable gauze which can be expanded as follows.

2.1.1 Manual Counting

The principle behind manual counting of gauze is such that the if the amount of gauze used and collected are equal, there will not be any gauze left within the patient. This solution can be easily implemented in any location with no additional cost needed for specific tools. It is recommended to count the gauze a total of 4 times during the set-up phase, before the operation, after closure and before the patient skins are stitch together [7]. To further reduce accounting error, gauzes are to be manufactured in packs of 10 and any open pack are to be completely used during operation. Any deviation in the number would be considered deflect and the packet would be reported without any usage. The World Health Organization (WHO) has also recommended for the count to be verified by two persons with their name recorded in count sheet/nursing records [5]. However, this solution still introduces human error and requires high manpower and strict discipline especially if amount of gauze used increases during operation.

2.1.2 RFID Tags

Figure 1 From RF Surgical System Inc, RF detection console and Blair-Port Wand (Left), Surgical Sponge (Top-Right) and RFID Tag (Bottom-Right) [8]

RFID gauze uses cheap RFID tags embedded within the gauze to uniquely identify the gauze using various devices such as X-ray films, magnets, and specific colours [5]. Figure 1 shows a marketed product by RF Surgical System Inc which uses such technology. Ideally, it would be able to both pinpoint the gauze location within the patient body [9] as well as accounting for the number of gauzes [5]. However, medical practitioner still needs to remain highly vigilant as signal received from the gauze depends on the size of the RFID tags which could also malfunction depending on the distance from the detector [9]. This hardware approach also requires a higher start-up and maintenance cost such as the scanner that could cost up to $400 with the single use gauze price increasing to $0.50 [10] which might not be suitable for all patients with different financial background.

2.1.3 Detectable Gauze

A cheaper alternative to RFID is a radio-opaque sponge that is made visible using an X-ray scan. Instead of tracking individual gauze, the imbedded black string would be visible to the surgeon using X-ray and the procedure would be carried out if gauze count does not match up [1]. Since the use of X-ray expose the patient to radio-active rays, it should be used sparingly with the accounting of gauze being the primary safety net to prevent gossypiboma.

2.2 Deep learning

Deep learning is the generalization of the function from its linear counterpart using hidden layers between the input and output layer of a model [11]. The idea of a deep learning network can be date back to 1957 by Frank Rosenblatt at the Cornell Aeronautical Laboratory and only start to blossom recently due to the advancement in computer hardware [12].

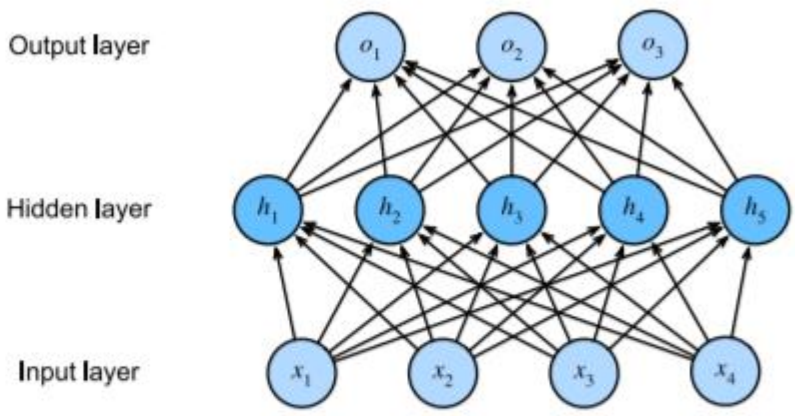

Figure 2 Simple neural network with hidden layer of five units [11]

Each layer within the model would consist of analogue dynamic processor array that interact within a finite local neighborhood as shown in the figure above [13]. As more layers are added to the system due to increased hardware capacity, the system would not only able to approximate any function but create a system of formulars which are used to solve complex problem like image recognition [12]. Combined with the release of publicly available annotated images online such as ImageNet [14], many different models have been constructed with varying accuracy and speed for different applications [14] and MobileNet being one of the models that is optimized for the mobile/embedded system.

2.3 MobileNet

Many AI model has been built to increase the accuracy of detection without evaluating the size and speed needed to process the data. For practical application that requires close to real-time processing and using limited hardware capacity, MobileNet model has been developed [15]. Although other model can reduce their size through compression or directly training using a smaller network [15], MobileNet would be able to operate in an embedded system by reducing the quota of calculation and configuration variables that is inside the model [16,17] hence matching the resource that it is given. This is achieved through depth wise separable convolutions that separate a standard convolution into a depth wise convolution and a 1×1 convolution called a pointwise convolution that is used to generate new features [15].

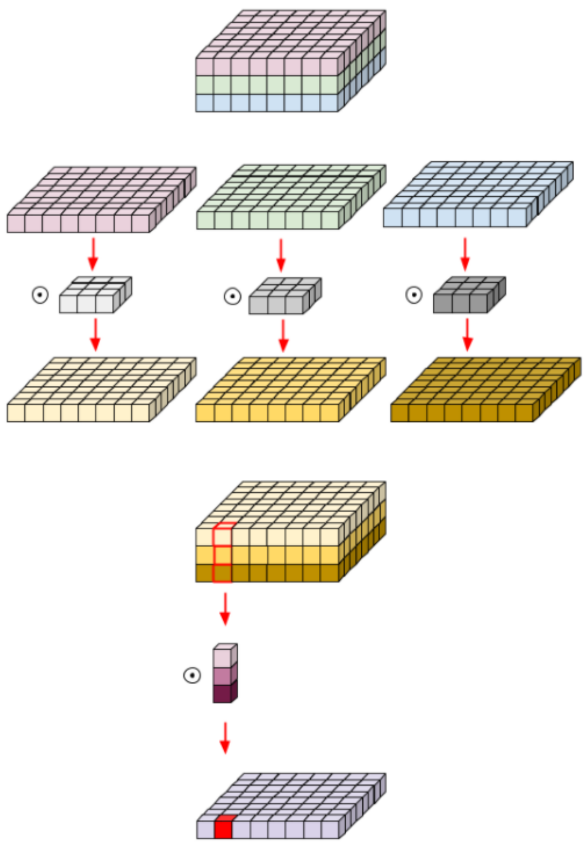

Figure 3 Depth-wise convolution with Depth-wise separable convolution [18]

The figure above shows how the MobileNet model process data using a simple 3-layer input with 8 unit of height and width. The MobileNet model has also been improved in version 2 where feature maps are extracted at a scale of 1/16, 1/32, 1/64, 1/128, 1/256, and 1/512 which are enhanced by the Feature Pyramid Network (FPN) module compared to the traditional 1/16-scale convolution layer. This is designed to overcome the poor and unstable detection of small object within the picture and improving real time performance [17].

2.4 Single Shot Detection

For the detection of object within the picture, Single Shot Detection (SSD) approach is one of faster method due to the fast calculation speed that it provided [17, 19]. SSD is a feed-forward based convolutional network that produces a fixed-size collection of bounding boxes and scoring for the presence of object class instances in said boxes rather than having the algorithm extract it from the input. Techniques such as classify regression and boundary box regression allows for it to achieve good detection precision and speed [20]. It is also more suitable in detecting large number of objects with different scales and types since it considers multiple feature maps of different scales [20].

2.5 Tflite and Quantization

Several techniques can be used to optimize the use of AI model on mobile devices such as converting it to another file format or applying quantization. Tflite leverage on the homogeneous coordinate system used by modern Graphic Processing Unit (GPU) by splitting the data type into a 4-element vector which can be process by

GPU as normal. If number of channels does not match the multiple of 4, it will be padded with zeros so that the order of memory read remain unchanged and smoother process of the data [21]. Furthermore, quantization of the model would also allow for a low latency inference on mobile since they convert the floating-point models to fixes-point which is more efficiently process by the mobile phone [22]. By having the model run on a low-power device such as mobile phones, data would not need to be sent out to be processed, making it more private [23].

2.6 Data Augmentation

Data augmentation also known as data warping [24] is an old approach in manipulating data to increase the amount available within training dataset as well as to overcoming the problem of overfitting due to small data sizes [25]. It could help the network to distinguish between common physical phenomena such as lighting and scale within a single picture by feeding the augmented data into the training network [24]. This approach also reduces the chance of network overfitting provided that more information can be extracted from the augmented dataset [24]. These data augmentation can include the changing of data geometry by rotating, cropping, shifting, or flipping the images as well as the changing of image color [24,26]. These approaches make the network more robust when encountering inputs which are rotated at different angles [26].

Chapter 3 Methodology

3.1 Problem Analysis and Proposed Solution

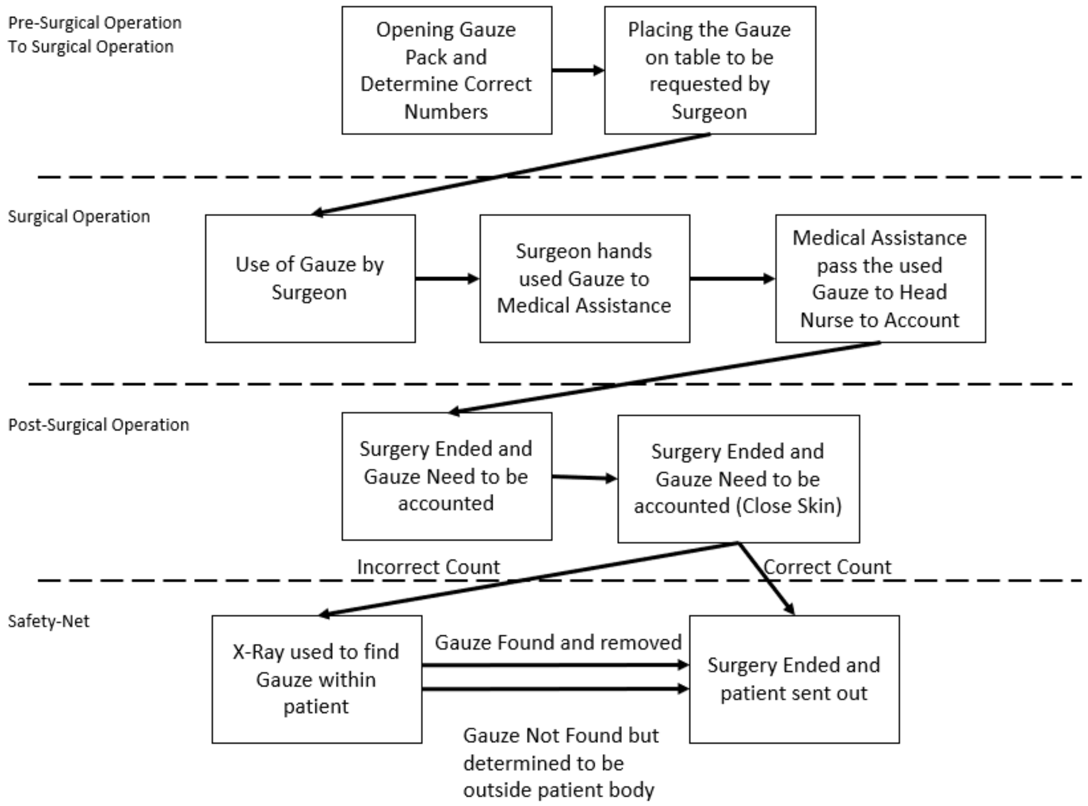

Figure 4 Current Gauze Accounting Flowchart

The current system used to prevent gossypiboma relies on the accounting of gauze before and after the operation as shown in figure 4. The RFID or black string imbedded in the gauze would provide the detection within the body if there are unaccounted gauze. To prevent the laborious task of finding missing gauze at the end of the operation, surgical personal could try to keep track of the small amount of surgical gauze within the operation table by communicating with the head nurse regarding the number of gauzes collected. However, as the gauze moves in and out of the operation table, there is room for human error to occur since the vision of each personal in the theatre is limited and the accounting procedure of gauzes are not standardized during the operation.

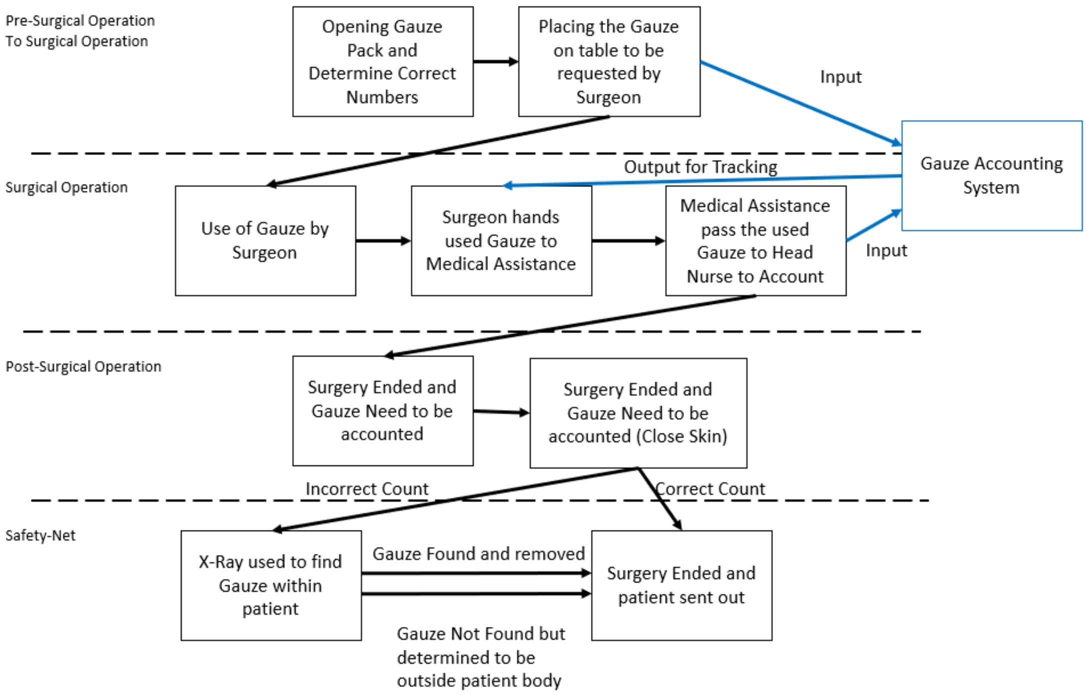

Figure 5 Suggested improvement to the system.

To improve on the current system, a gauze counting system is proposed and implemented as shown in figure 5. The volunteering process of counting the gauze within the operation table would be digitalized and standardized with a standard display being shown to the medical personal. To maintain the cost of operation for the hospital and reduce the labor involved, the proposed system would need to be semi-autonomous and cost efficient to install and maintain.

3.2 Software Design

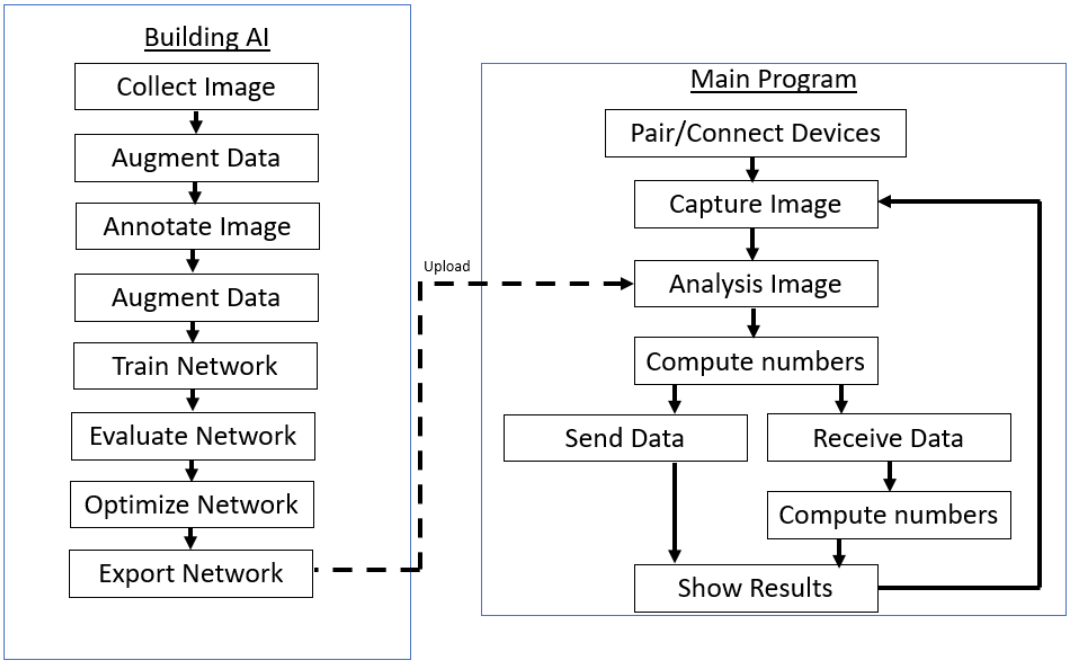

Figure 6 Software Design Overview

To implement the proposed solution, customized software would need to be implemented as shown in figure 6. A specialized AI model would need to be build following the steps from collecting images to exporting the model to be used in another platform. The other software would use the network to perform the computation needed to account for the gauze.

3.3 Software Implementation

The Masked-Region Based Convolutional Neural Networks (Mask-RCNN) was studied in the previous project which place emphasis on the accuracy of the detection. Although it is precise, the device that stores the models needs to have a high processing power to run the program. Hence, a new model would be needed to be introduced which is the MobileNet-v2 SSD model and build using the latest tensorflow-v2 program.

3.3.1 Data Collection

To start training the model, raw dataset would need to be annotated and used as input. The Common Objects in Context (COCO) format has been chosen due to the high usage as a training and evaluation tools by different model. The original COCO dataset consists of 80 types of labeled object that can be used for object detection or segmentation purposes [27]. By converting the raw gauze dataset into COCO format, it allows for subsequence development of different AI model if needed. To convert the dataset, docker would need to be installed along with COCO annotator by Justin Brooks [28]. By setting up the program, user would be able to annotate the picture using various tools.

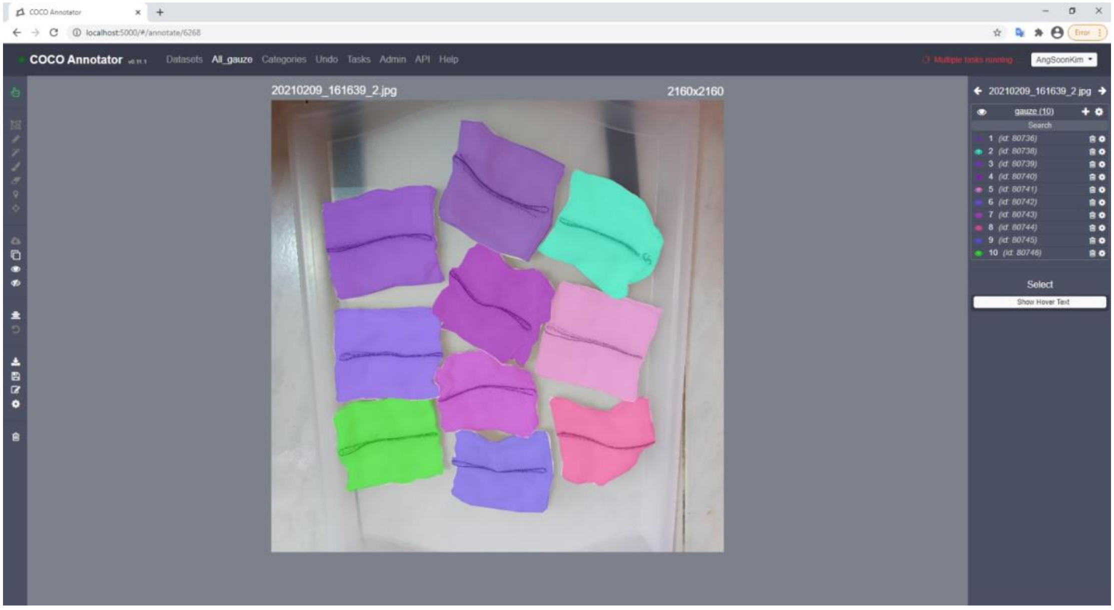

Figure 7 Annotated Gauze Image

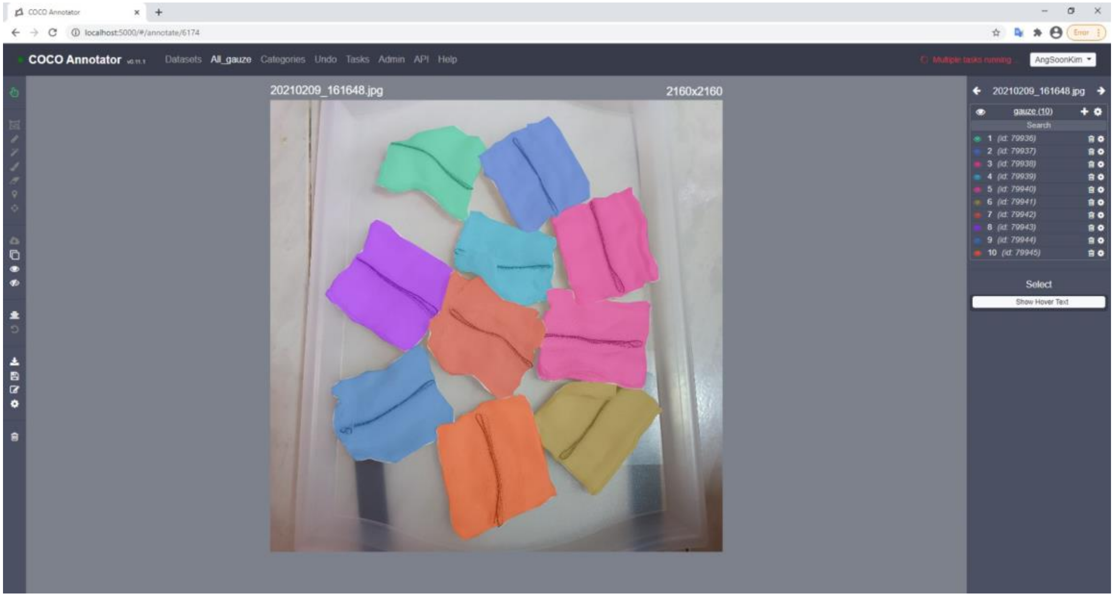

Figure 8 Annotated Gauze Image

Figure 7 and 8 shows the screenshot of a finished annotated image completed using the COCO annotator. Each gauze is labelled using a unique ID under the category of gauze as shown on the left column of the image. A total of 1574 image from the previous project has been relabeled using COCO annotator with an additional 184 images being captured for the model training. The newly added images consist of clean gauze which the previous dataset lacked as well as gauze of different shape. Each image is labeled manually annotated using the polygon tools by connecting the edges of each gauze individually.

After the annotation, the file could be exported into a json file that can be processed further after splitting into the training and evaluation datasets. The exported file is uploaded onto the Google drive where it can be further processed into the TFRecord format which uses a specific way to feed the data into the model for higher efficiency [29]. The transformed dataset would also be more space efficient for the limited

Google drive space [23]. The files are sharded into smaller file sizes of 100MB to optimize the fetching of data by tensorflow-v2 during training.

3.3.2 Model Training

As the dataset of 1000 plus images can still be considered small, a pre-trained model will be used for its feature extractor parameters by reloading pretrained checkpoint of the model. It also helps to speed up the time needed to train the model [30]. The number of workers is also adjusted to make full use of the hardware available.

Figure 9 Datasets from previous report. Original (Left), Grey Scale (Middle), Rotated (Right)

To artificially increase the number of datasets for the training phase, traditional method of rotating the images has been performed as well as having various data augmentation methods selected and implemented in the configuration files with different results as shown in figure 9.

3.3.3 Implementation to Mobile Phone

After the AI model has finish building, it would be converted to a model that can be process by the mobile phone processing unit. To reduce the model size and keep the

GPU acceleration function in place, quantization has been applied to the model which will be implemented in the mobile app iGauze.

IGauze uses the Camera2 Application Programming Interface (API) to continuously run a preview image onto the screen. This function would run separately from the main code while displaying its result. The function also passes an image to the main activity periodically where the AI model would process the image and produces results that can be displayed in a meaningful interface. In addition, the main activity contains a simple function which calculate and constantly display the amount of gauze that is not accounted for. This would be achieved through Bluetooth communication between two mobile devices which keep track of the amount of clean and used gauze during operation.

To ensure that the phone display can be easily viewed by people within the operation table, scrpy program has been used to display the phone screen onto a bigger monitor such as a laptop while a landscape mode has been adopted to make full use of the screen [31]. This would be applied to the output tray that collects the used gauze as it would be able to determine the gauze remaining in operation table using data from the input tray.

3.4 Hardware Implementation

To build the prototype, the structure would need to have a means to hold the mobile phone, gauzes, and an overall structure to house all the item. The items selected would need to perform their relative functions efficiently of monitoring a viewing area of A3 size.

3.4.1 Handphone Stand

The handphone stand should be able to fulfill the following requirement.

1. Ease of use to the user
2. Stability of stand
3. Consistence of stand structure
4. Ability to view an area of A3 size
5. Cost and ease of purchase

An online search was conducted with 2 candidates selected and ranked accordingly.

Table 1 Handphone Stand Ranking

Criteria/ weightage	 Figure 10 Desktop mobile phone stand photo [32]	 Figure 11 Kphoto Ajustable Tripod [33]
Ease of Use (0.2)	4 (0.8)	4 (0.8)
Stability (0.2)	4 (0.8)	4 (0.8)
Consistency (0.25)	4 (1)	3 (0.75)
Functionally (0.2)	5 (1)	5 (1)
Cost (0.15)	4 (0.6)	3 (0.45)
Total	21 (4.2)	19 (3.8)

3.4.2 Gauze Holder

To hold the used and unused gauze, a A3 size tray has been proposed with only 1 candidate found and hence it will be used for the prototype.

Figure 12 SMULA [34]

3.4.3 Overall Structure

The overall structure for the prototype would need to fulfill the following criteria.

1. Stability of the structure
2. Mobility of the structure
3. Size of the structure
4. Compatibility of the structure with the other component
5. Cost of the structure

With the large selection of trolley available, 3 candidates have been chosen and ranked accordingly.

Table 2 Overall Structure Ranking

Criteria/Weight	 Figure 13 SUNNERSTA [35]	 Figure 14 NISSAFORS [36]	 Figure 15 KUNGSFORS [37]
Stability (0.2)	4 (0.8)	3 (0.6)	3.5 (0.7)
Mobility (0.2)	3 (0.6)	4 (0.8)	3.5 (0.7)
Size (0.25)	4 (1)	3 (0.75)	5 (1.05)
Compatibility (0.25)	4 (1)	3 (0.75)	4 (1)
Cost (0.1)	5 (0.5)	4 (0.4)	2 (0.2)
Total	20 (3.9)	17 (3.3)	18 (3.65)

3.5 System Design and Details

The hardware and software component of the project would be assembled into the gauze accounting system which will be illustrated in the following section using different method of analysis.

3.5.1 Physical View

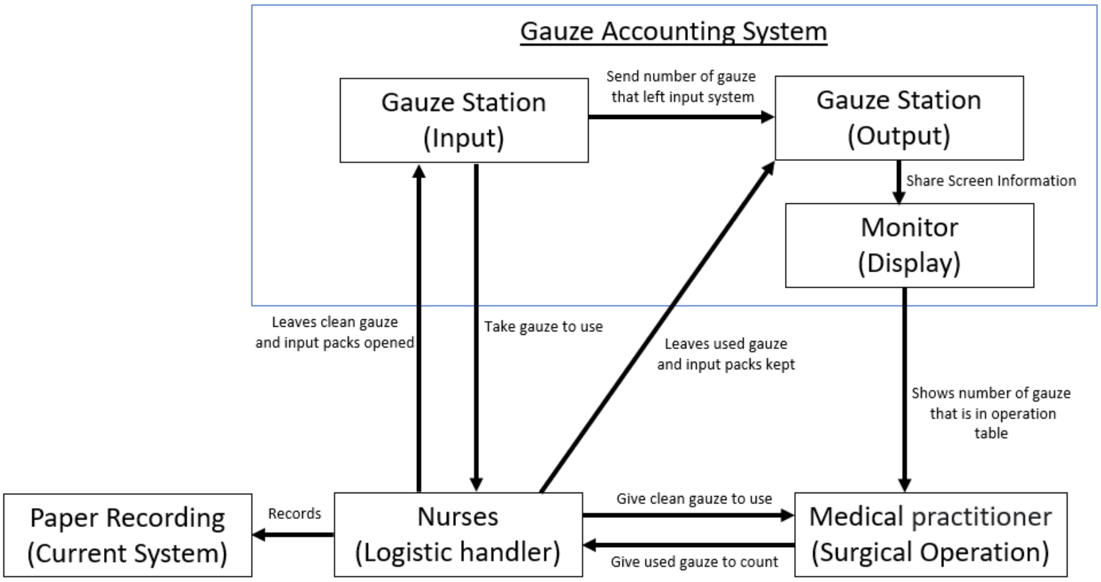

Figure 16 Physical View on how System works in operation.

The figure above shows the overview of how the system will be operated. The nurse in the operation theater would play the central role of moving the gauze around which is similar to the current system. However, the implemented system would remove the need for the medical personal to communicate with the nurses regarding the gauze number that is accounted for as this number will presented to them on the monitor. Using the information, the medical practitioner would be able to keep tackle of a smaller amount of gauze within their field of views and identify missing gauze as soon as it occurs rather than accounting for it at the end of the operation.

3.5.2 Standard of Procedure

To implement the prototype into the current system and improved on its performance, some Standard of Procedure (SOP) would need to be implemented in addition to the physical prototype.

1. The mobile devices used during the operation would need to be paired before using the app and linked to the monitor to ensure smooth operation.
2. The total amount of gauze package open would need to be updated on to the system to be recorded digitally in addition to the current manual recording.
3. The amount of gauze that is not accounted by the system should be track manually by comparing the number displaced on the screen as well as the gauze that could be counted on patients.
4. The gauze that is placed into the collection bin should be spread open, evaluated, and folded back to a 10x10cm configuration before placing onto the collection area. This is to prevent the collection of any other medical devices that is stuck onto the gauze.
5. Once the collection bin reached over 10 used gauze, the gauze would need to be pack away in packet of 10 and accounted for at the end of the operation.
6. In the event in which the liquid starts to flood the tray, it should be replaced with another tray during operation. The remaining uncounted gauze should be place into the new tray while the flooded tray would need to be dried once before reusing.
7. After the whole operation, the trays would need to be wipe clean once to remain sterile for the next operation.

Chapter 4 Result and Discussion

4.1 Evaluation of AI model

AI models are evaluated through another set of well-defined datasets. The common terms used during the evaluation includes Intersection over Union (IoU). mean Average Precision (mAP) and mean Average Recall (mAR).

4.1.1 Terms Definition

IoU is the standard evaluation metric used for object detection. By determining the true positives and false positives in a set of predictions, a quantitative approach can be determined the accuracy of the model in identifying objects [38]. The formular for this approach is as follows.

$$IoU = \frac{|A \cap B|}{|A \cup B|}$$

mAP and mAR are calculated with the following definition for precision and recall for object detection.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positve}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

By taking the mean average out these values, the effectiveness of the model could be evaluated and compare with one another.

Lastly, the definition of small, medium, and large detection is defined as having area less than 32x32 pixels, between 32x32 and 96x96 pixels and more than 96x96 pixels [39].

4.1.2 Evaluation Data between 320x320 and 640x640 model

Both models have been built and evaluated using the same dataset on google collab.

Table 3 320x320 vs 640x640 model evaluation

Method of evaluation	320x320 model	640x640 model
Average Precision (AP) @ [IoU = 0.50:0.95|area = all | maxDets = 100]	0.855	0.87
Average Precision (AP) @ [IoU = 0.50|area = all | maxDets = 100]	0.989	0.996
Average Precision (AP) @ [IoU = 0.75|area = all | maxDets = 100]	0.986	0.990
Average Precision (AP) @ [IoU = 0.50:0.95|area = small | maxDets = 100]	0.000	0.423
Average Precision (AP) @ [IoU = 0.50:0.95|area = medium | maxDets = 100]	0.580	0.816
Average Precision (AP) @ [IoU = 0.50:0.95|area = large| maxDets = 100]	0.865	0.875
Average Recall (AR) @ [IoU = 0.50:0.95|area = all | maxDets = 1]	0.063	0.064
Average Recall (AR) @ [IoU = 0.50:0.95|area = all | maxDets = 10]	0.624	0.633
Average Recall (AR) @ [IoU = 0.50:0.95|area = all | maxDets = 100]	0.890	0.908
Average Recall (AR) @ [IoU = 0.50:0.95|area = small | maxDets = 100]	0.000	0.422
Average Recall (AR) @ [IoU = 0.50:0.95|area = medium | maxDets = 100]	0.667	0.844
Average Recall (AR) @ [IoU = 0.50:0.95|area = large | maxDets = 100]	0.900	0.911

As shown in table 3, the 640x640 resolution model perform better than the lower resolution input model. The result also indicate that small object could not be detected using 320x320 input model and hence 640x640 model is used as the basics for further improvement.

4.1.3 Evaluation of model with different learning rate

During the training phase of the 2 different model, it has been observed that after 2000 steps, the losses detected by the model becomes relatively stagnant.

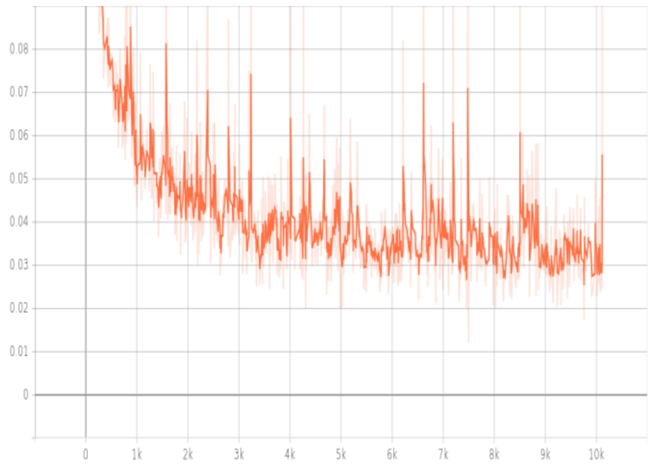

Figure 17 Classification Loss by 640x640 Mobilenet-v2 SSD Model during training

The training was also cut short from its pre-planned 50000 steps training phases to prevent overfitting of the model. Since the value of pre-planned steps affects the learning rate during training, an identical training phase is conducted with a change in the total steps.

Table 4 Evaluation of 640x640 Model with different total steps

Method of evaluation	640x640 model (original)	640x640 model (reduced steps)
Average Precision (AP) @ [IoU = 0.50:0.95|area = all | maxDets = 100]	0.870	0.902
Average Precision (AP) @ [IoU = 0.50|area = all | maxDets = 100]	0.996	0.990
Average Precision (AP) @ [IoU = 0.75|area = all | maxDets = 100]	0.990	0.990
Average Precision (AP) @ [IoU = 0.50:0.95|area = small | maxDets = 100]	0.423	0.402
Average Precision (AP) @ [IoU = 0.50:0.95|area = medium | maxDets = 100]	0.816	0.856
Average Precision (AP) @ [IoU = 0.50:0.95|area = large| maxDets = 100]	0.875	0.910
Average Recall (AR) @ [IoU = 0.50:0.95|area = all | maxDets = 1]	0.064	0.065
Average Recall (AR) @ [IoU = 0.50:0.95|area = all | maxDets = 10]	0.633	0.653
Average Recall (AR) @ [IoU = 0.50:0.95|area = all | maxDets = 100]	0.908	0.936
Average Recall (AR) @ [IoU = 0.50:0.95|area = small | maxDets = 100]	0.422	0.433
Average Recall (AR) @ [IoU = 0.50:0.95|area = medium | maxDets = 100]	0.844	0.880
Average Recall (AR) @ [IoU = 0.50:0.95|area = large | maxDets = 100]	0.911	0.939

As shown in table 4, there is a significant improvement regarding detection at all levels except in the detection of smaller gauze.

4.1.4 Evaluation of Model using different Built in Data Augmentation

To further improve the accuracy of the model, the data sets has been artificially increased using the available data augmentation pipeline in tensorflow-v2. Using the data pipeline to augment the data allows for smaller dataset size to be used within the limited drive space which google Collab provides. The chosen augmentation types to vary include changing to grey scale, varying contrast within pictures, and introducing patch gaussian augmentation [40]. These parameters have been adjusted or combined to produce different results.

Constant data augmentation such as horizonal flips and cropping of images has been included to vary the datasets as the training phase requires more images that was prepared.

Grey scales parameter has been chosen to reduce the dependency on colour identification and performed twice with different conversion rates. The changing contrast is selected to highlights the differences between the images and patch gaussian augmentation was selected to produce noise within the datasets. All models are trained using the reduces steps 640x640 resolution model as explored in section 4.1.3. The final model is trained with all the different data augmentation with increased steps to account for the different combination of data augmentation.

Table 5 Evaluation of different data augmentation types

Method of evaluation	Baseline model	Default grey scale conversion	Increased grey scale conversion	Default Varied Contrast	Default Gaussian Patch	Mixed of listed default type augmentation with increase steps
Average Precision (AP) @ [IoU = 0.50:0.95|area = all | maxDets = 100]	0.902	0.901	0.900	0.898	0.897	0.909
Average Precision (AP) @ [IoU = 0.50|area = all | maxDets = 100]	0.990	0.999	0.990	0.997	0.990	0.990
Average Precision (AP) @ [IoU = 0.75|area = all | maxDets = 100]	0.990	0.990	0.990	0.990	0.990	0.990
Average Precision (AP) @ [IoU = 0.50:0.95|area = small | maxDets = 100]	0.402	0.400	0.306	0.343	0.355	0.335
Average Precision (AP) @ [IoU = 0.50:0.95|area = medium | maxDets = 100]	0.856	0.854	0.856	0.850	0.847	0.857
Average Precision (AP) @ [IoU = 0.50:0.95|area = large| maxDets = 100]	0.910	0.908	0.909	0.904	0.906	0.917
Average Recall (AR) @ [IoU = 0.50:0.95|area = all | maxDets = 1]	0.065	0.065	0.065	0.064	0.065	0.065
Average Recall (AR) @ [IoU = 0.50:0.95|area = all | maxDets = 10]	0.653	0.651	0.651	0.654	0.651	0.658
Average Recall (AR) @ [IoU = 0.50:0.95|area = all | maxDets = 100]	0.936	0.933	0.934	0.931	0.933	0.942
Average Recall (AR) @ [IoU = 0.50:0.95|area = small | maxDets = 100]	0.433	0.400	0.333	0.411	0.378	0.333
Average Recall (AR) @ [IoU = 0.50:0.95|area = medium | maxDets = 100]	0.880	0.865	0.880	0.873	0.869	0.877
Average Recall (AR) @ [IoU = 0.50:0.95|area = large | maxDets = 100]	0.939	0.936	0.937	0.934	0.936	0.946

The table above shows the evaluation results using the different data augmentation method. Although the models with the all the different data augmentation scores the highest among the rest of the model, errors do shows up when passing data which are not included during training.

4.1.5 Classification errors on new images

After training each model, a new dataset which has been set aside for testing was used to determine the accuracy of the model. Although the precision and recall score are high with the model accurately detecting pictures with distinct gauzes, there are still noticeable and fatal error found during testing phase.

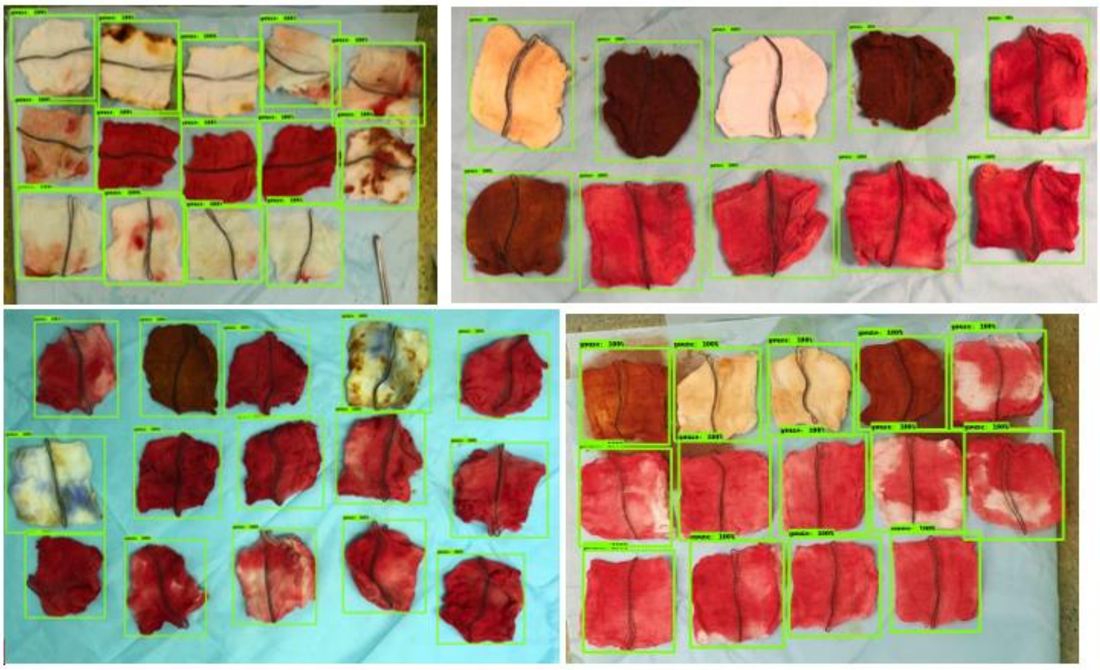

Figure 18 Correct Identification of Gauze from Test Datasets

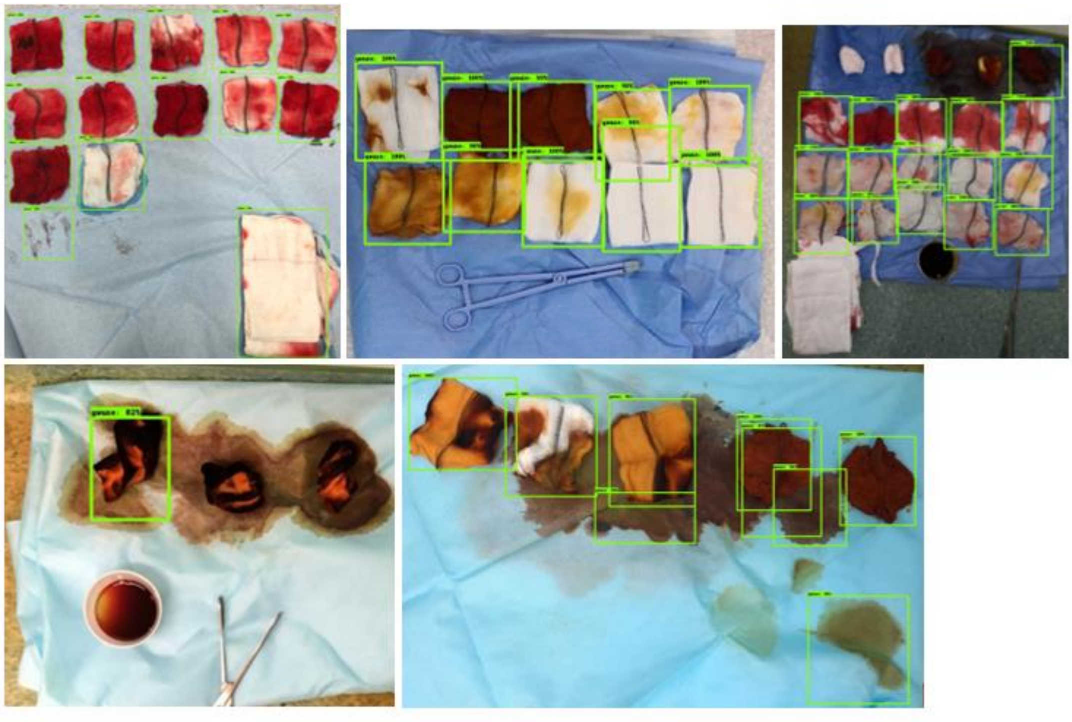

Figure 19 Misclassification of gauze (Top-left, Bottom-right), Overlap bounding box (Top-middle, Bottom-right), Failed Classification (Top-right, Bottom-left)

The results shows that a clean background would be necessary for the correct identification of the gauze. This would need to be considered for when implementing the AI on the physical prototype.

4.2 Integration of different components

After selecting the optimal AI model, it would need to be saved in format which is optimized to be used on mobile application. The model size is further reduce using quantization method which reduces the app size from 24.69 Mega-Bytes (MB) to 12.58 MB. The time needed to process each frame ranges from 400-800 milli-seconds which rarely exceed 1 second.

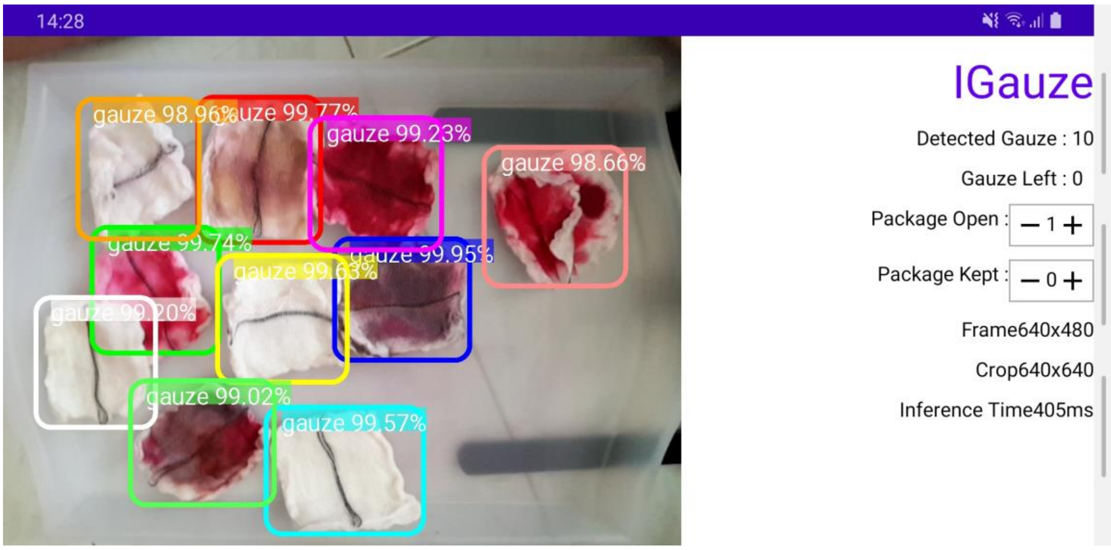

Figure 20 Detection of Gauze using iGauze

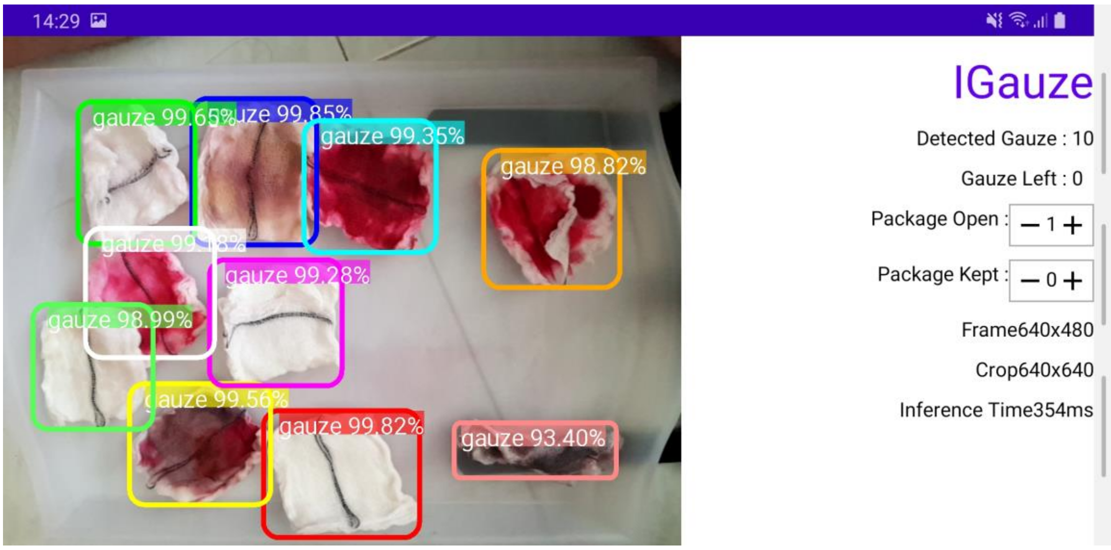

Figure 21 Detection of Gauze with different shapes

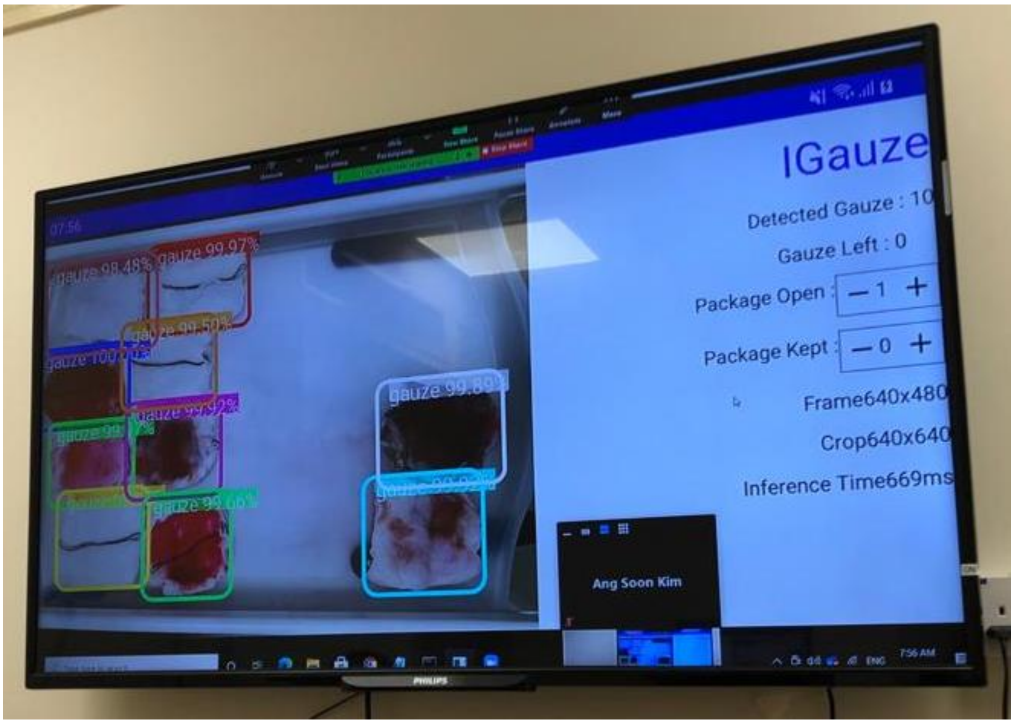

Figure 22 Example of projection of phone display to monitor.

The figures above show the detection of the gauze as well as the projection of the display to a wall monitor. The purchased items which house all the components are compatible with each other with the trolley securing the tray effectively due to the trolley edges.

Figure 23 Fitting of Hardware components

The figure above shows the assembly of the hardware. The Bluetooth connection between two mobile devices is also stable when interacting with the image processing program.

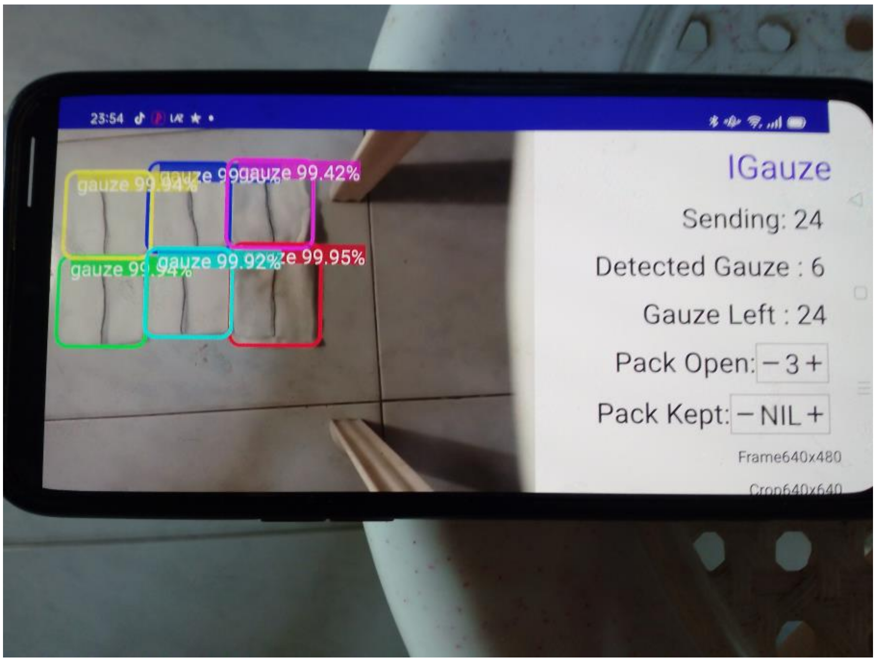

Figure 24 IGauze Input Screen Display

Figure 25 IGauze Output Screen Display

The figures above show the sending and receiving of data between iGauze application as it performs the image processing. The input side of the system would calculate the amount of gauze package opened and detected and send the information to the output device which is shown in the top right-hand corner of both devices. The output device would consider the number of used gauzes packed as well as the detected number to show the theorical number of gauzes that is used on the operation table. This value would be displayed to the surgeon via a monitor screen who can keep track of the gauze that is on the operation table.

IGauze has also been adjusted such that only the input side of the device could interact with the number of package open and output side would interact with package kept button to prevent mis-click from the user during operation.

4.2.1 Limitation of Prototype and Proposed solutions

The completed prototype was able to perform the primary function of detecting and counting the number of gauzes within the capture image. However, there are some limitations found during the testing of the prototype.

1. Smartphone battery was found to be draining at an unsustainable rate during operation with efficiency of detection dropping proportionally to the battery life.
2. The smartphone used will also requires a minimum Software Development Kit (SDK) version of 21 to run the application.

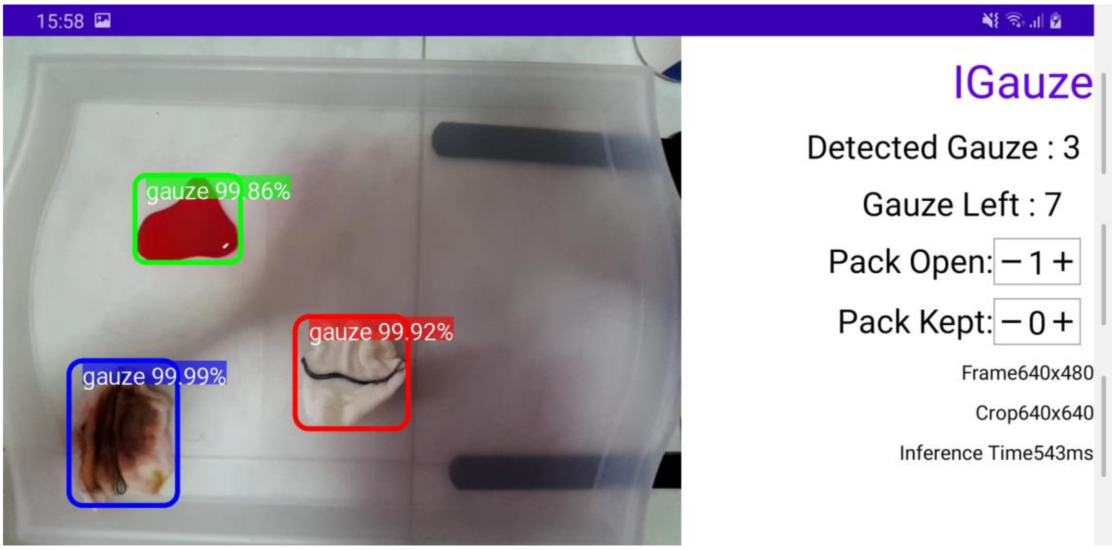

Figure 26 Misclassification of the AI model

3. The AI model would miss classify pool of water as gauze as shown in figure 26 which would introduce calculation error within the system.

To counter the limitation found during the testing of prototype, adjustments can be made to reduce negative impact of the system.

1. Smartphone should be connected to the wall socket during the operation period to retain its battery life.
2. A permanent smartphone would need to be used for the system or having the AI model be transferred to a Raspberry Pi which can be permanently housed within the system.

Figure 27 Use of paper as liquid absorber.

3. The tray would need to house a suitable liquid absorber such as paper which the AI model would have a lower probability to miss classify. This would reduce the error regarding the stain found when using the system.

Chapter 5 Conclusion and Future Works

The objective of the project is to make a functional prototype which improves on the previous software. This has been achieved by having the AI model run on a smartphone which can be operated with its own camera and processing unit. The speed for detecting the gauze has also been improved for real time application. Although the tracking of the gauze has been removed from the previous project to reduce the processing power needed on the smart phone. The newly developed prototype would be able to tackle the problem in a different angle where it reduces the need for constant manual tracking of every gauze by focusing on those that are left out of the accounting system. It also reduces the need for constant communication regarding the amount of gauze by having the number displayed on screen to the personal within the operation table.

5.1 Contribution

Throughout the project, progress has been made to increase the functionality of the AI model to suit the real-world environment. These contributions are as follows:

1. Reformatted dataset to COCO format which could be reused by anyone who continues with the research.
2. Making most of the program executable on google collab which could be

access by any individual with a Gmail account.

3. Building a mobile application which houses the AI model to perform the accounting of gauze and communicating the data through Bluetooth and wifi.
4. Prototyping the system using the software as the base which can be connected to other system.

5.2 Recommendation in Future Work

Although the completed prototype is able to assist in the accounting of the gauze, further improvement and work could be made to improve efficiency and function as shown below:

1. Real-world test to collect the actual data that is captured by the prototype to increase the accuracy of the AI model. Different factors which is not accounted for when building the prototype would need to be considered.
2. A more complex solution can be implemented to completely remove the human involvement by replacing every step in the system with machinal muscle and mind.
3. The hardware of the system can also be improved to contain more gauze of different sizes to make the system more versatile in the number of objects it can detect. Lighting can also be added to make the system more robust when operating in a different environment.
4. The implementation of the software on different operating system can ensure that users are not limited by the hardware they personally possess. Other neuro network such as Nanodet [41] can also be explored for faster image processing

on mobile devices.

5. Other usage for image processing regarding the gauze can also be explored such as looking at the amount of blood lost based on the input.
6. Improvement to the software such as the saving of past images and the management of such dataset would need to be considered. These datasets would need to be easily access by the nurse who will be operating the system.

References

[1] Bani-Hani, K. E., Gharaibeh, K. A., & Yagha, R. J. (2005). Retained Surgical Sponges (Gossypiboma). Asian Journal of Surgery, 28(2), 109–115. https://doi.org/10.1016/s1015-9584(09)60273-6

[2] Cheng, T. C., Chou, A. S. B., Jeng, C. M., Chang, P. Y., & Lee, C. C. (2007). Computed Tomography Findings of Gossypiboma. Journal of the Chinese Medical Association, 70(12), 565–569. https://doi.org/10.1016/s1726-4901(08)70063-7

[3] Stoll, A. (1988). Retained surgical sponge 40 years after laminectomy. Surgical Neurology, 30(3), 235–236. https://doi.org/10.1016/0090-3019(88)90278-9

[4] Serra, J., Matias-Gulu, X., Calabuig, R., Garcia, P., Sancho, F. J., & Calle, J. P. L. (1988). Surgical gauze pseudotumor. The American Journal of Surgery, 155(2), 235–237. https://doi.org/10.1016/s0002-9610(88)80702-5 [5]

[5] Biswas, R. S., Ganguly, S., Saha, M. L., Saha, S., Mukherjee, S., & Ayaz, A. (2012). Gossypiboma and Surgeon- Current Medicolegal Aspect – A Review. Indian Journal of Surgery, 74(4), 318–322. https://doi.org/10.1007/s12262-012-0446-3

[6] Şentürk, Mustafa. (2019). Endoscopic Removal Of Gastric Penetration Of Gossypibom. Laparoscopic Endoscopic Surgical Science. 10.14744/less.2019.47560.

[7] Sarda AK, Pandey D, Neogi S, Dhir U. Postoperative complications due to a retained surgical sponge. Singapore Medical Journal. 2007 Jun;48(6):e160-4.

[8] Thomas Jefferson University Hospital, Philadelphia, Pennsylvania, United States, Setren, A. J., Khan, S., & Kennedy, T. L. (2017). Magnetic Interference With Radiofrequency ID-Tagged Surgical Sponges. American Society of Anesthesiologists. Published. http://www.asaabstracts.com/strands/asaabstracts/abstract.htm?year=2017&index=8&absnum=4124

[9] Macario, A. (2006). Initial Clinical Evaluation of a Handheld Device for Detecting Retained Surgical Gauze Sponges Using Radiofrequency Identification Technology. Archives of Surgery, 141(7), 659. https://doi.org/10.1001/archsurg.141.7.659

[10] Wiederkehr, J. C., Gama, R. R., Wiederkehr, H. A., Stelmasuk, K., Carvalho, C. A. & Wiederkehr, B. A. (2014). Radio-frequency identification of surgical sponges in the abdominal cavity of pigs. Annals of Medicine and Surgery, 3(2), 31–33.

[11] Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (n.d.). Dive into Deep Learning (0.14.3 ed.). https://d2l.ai/d2l-en.pdf

[12] Mueller, J., & Massaron, L. (2019). Deep learning / by John Paul Mueller and Luca Massaron. (1st edition). For Dummies.

[13] Chua, L., & Roska, T. (1993). The CNN paradigm. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 40(3), 147–156. https://doi.org/10.1109/81.222795

[14] Shin, H. C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., & Summers, R. M. (2016). Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. IEEE Transactions on Medical Imaging, 35(5), 1285–1298. https://doi.org/10.1109/tmi.2016.2528162

[15] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.

[16] Younis, Ayesha & Shixin, Li & Jn, Shelembi & Hai, Zhang. (2020). Real-Time Object Detection Using Pre-Trained Deep Learning Models MobileNet-SSD

CCS Concepts •Computing methodologies→Artificial intelligence→

Computer vision→Computer vision problems→Object detection Keywords.

978-1-4503-7673-0. 44-48.

[17] Y. -C. Chiu, C. -Y. Tsai, M. -D. Ruan, G. -Y. Shen and T. -T. Lee, "Mobilenet-SSDv2: An Improved Object Detection Model for Embedded Systems," 2020 International Conference on System Science and Engineering (ICSSE), Kagawa, Japan, 2020, pp. 1-5, doi: 10.1109/ICSSE50014.2020.9219319.

[18] Pandey, A. (2018, September 9). Depth-wise Convolution and Depth-wise Separable Convolution. Medium. https://medium.com/@zurister/depth-wise-convolution-and-depth-wise-separable-convolution-37346565d4ec

[19] Biswas, D., Su, H., Wang, C., Stevanovic, A., & Wang, W. (2019). An automatic traffic density estimation using Single Shot Detection (SSD) and MobileNet-SSD. Physics and Chemistry of the Earth, Parts A/B/C, 110, 176–184. https://doi.org/10.1016/j.pce.2018.12.001

[20] Li, Y., Huang, H., Xie, Q., Yao, L., & Chen, Q. (2018). Research on a Surface Defect Detection Algorithm Based on MobileNet-SSD. Applied Sciences, 8(9), 1678. https://doi.org/10.3390/app8091678

[21] Lee, J., Chirkov, N., Ignasheva, E., Pisarchyk, Y., Shieh, M., Riccardi, F., ... & Grundmann, M. On-Device Augmented Reality with Mobile GPUs.

[22] Feng, C., Sheng, T., Liang, Z., Zhuo, S., Zhang, X., Shen, L., ... & Lu, Y. H. (2019). Low Power Inference for On-Device Visual Recognition with a Quantization-Friendly Solution. arXiv preprint arXiv:1903.06791.

[23] Baranwal, A., Khatri, A., & Baranwal, T. (2019). What's New in TensorFlow 2. 0: Use the New and Improved Features of TensorFlow to Enhance Machine Learning and Deep Learning. Packt Publishing, Limited.

[24] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6(1), 1-48.

[25] Perez, L., & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. arXiv preprint arXiv:1712.04621.

[26] Mikołajczyk, A., & Grochowski, M. (2018, May). Data augmentation for improving deep learning in image classification problem. In 2018 international interdisciplinary PhD workshop (IIPhDW) (pp. 117-122). IEEE.

[27] Kim, D. H. (2019). Evaluation of COCO validation 2017 dataset with YOLOv3. Evaluation, 6(7), 10356-10360.

[28] J. (2020, Sept 17). jsbroks/coco-annotator. GitHub. https://github.com/jsbroks/coco-annotator

[29] Wan, F. (2018). Deep Learning Method used in Skin Lesions Segmentation and Classification.

[30] T. (2020b, July 11). tensorflow/models. GitHub. https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/tf2_training_and_evaluation.md

[31] G. (2021, February 1). Genymobile/scrcpy. GitHub. https://github.com/Genymobile/scrcpy

[32] 【Happliances】网课桌面手机支架照片静物拍摄翻拍架台摄影高俯拍拍摄录视频直播 TB. (n.d.). Ezbuy.Sg. https://ezbuy.sg/

[33] captcha | lazadacom. (n.d.-b). Lazada. https://www.lazada.sg/products/kphoto-ajustable-tripod-with-cellphone-holder-overhead-phone-mount-table-top-teaching-online-stand-for-live-streaming-and-online-video-and-food-crafting-demo-drawing-sketching-recording-i1111860450-s4295220203.html?spm=a2o42.searchlist.list.1.15247dadAuDaf6&search=1

[34] SMULA Tray, transparent, 52x35 cm. (n.d.). IKEA. https://www.ikea.com/sg/en/p/smula-tray-transparent-90133944/

[35] SUNNERSTA Trolley, 56x33 cm. (n.d.). IKEA. https://www.ikea.com/sg/en/p/sunnersta-trolley-50303721/

[36] NISSAFORS Trolley, white, 50.5x30x83 cm. (n.d.). IKEA. https://www.ikea.com/sg/en/p/nissafors-trolley-white-90465735/

[37] KUNGSFORS Kitchen trolley, stainless steel, 60x40 cm. (n.d.). IKEA. https://www.ikea.com/sg/en/p/kungsfors-kitchen-trolley-stainless-steel-90334933/

[38] Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., & Savarese, S. (2019). Generalized intersection over union: A metric and a loss for bounding box regression. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 658-666).

[39] COCO - Common Objects in Context. (n.d.). Cocodataset. https://cocodataset.org/#detection-eval

[40] Lopes, R. G., Yin, D., Poole, B., Gilmer, J., & Cubuk, E. D. (2019). Improving robustness without sacrificing accuracy with patch gaussian augmentation. arXiv preprint arXiv:1906.02611.

[41] R. (n.d.). RangiLyu/nanodet. GitHub. https://github.com/RangiLyu/nanodet