Emergency Exit detection for the visually impaired people

- 자율지능 IOT 팀 프로젝트 최종 발표 -

20520052 이도현   20520053 이새롬

1. Introduction (1/2)

■ Motivation

- □ Blind People can’t identify the sign
- □ Additionally, in workplace there are various obstacles that looks similar with the sign.
  - ▪ doorplate, nameplate, and signboard.
- □ It has been an important issues for safety!
  - ▪ Escape method is unsuitable for blind people.

1. Introduction (2/2)

■Purpose

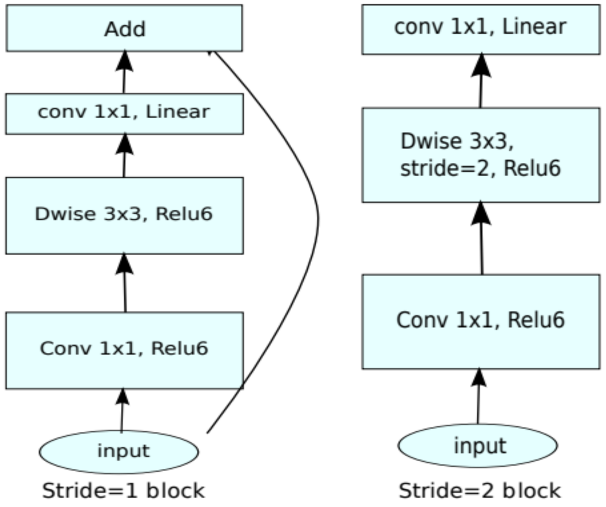
- □Propose effective emergency exit detection system for blind people.
  - ▪ Aims at helping visually impaired people with navigating unfamiliar environments in real-time.

2. Method (1/2)

■ Process

□ Emergency Exit detection

- Transfer Learning based on the pretrained model
  - □ **Mobilenet_V2** [1]
    - ▶ Pre-trained on ImageNet
  - □ Classes referring the emergency exit should be added.
    - ▶ **Emergency exits are classified by street sign according to the case of Real example**
  - □ Re-trained with custom dataset
    - ▶ Indoor Sign Dataset(ISD) [2]
    - ▶ Indoor object detection dataset [3]

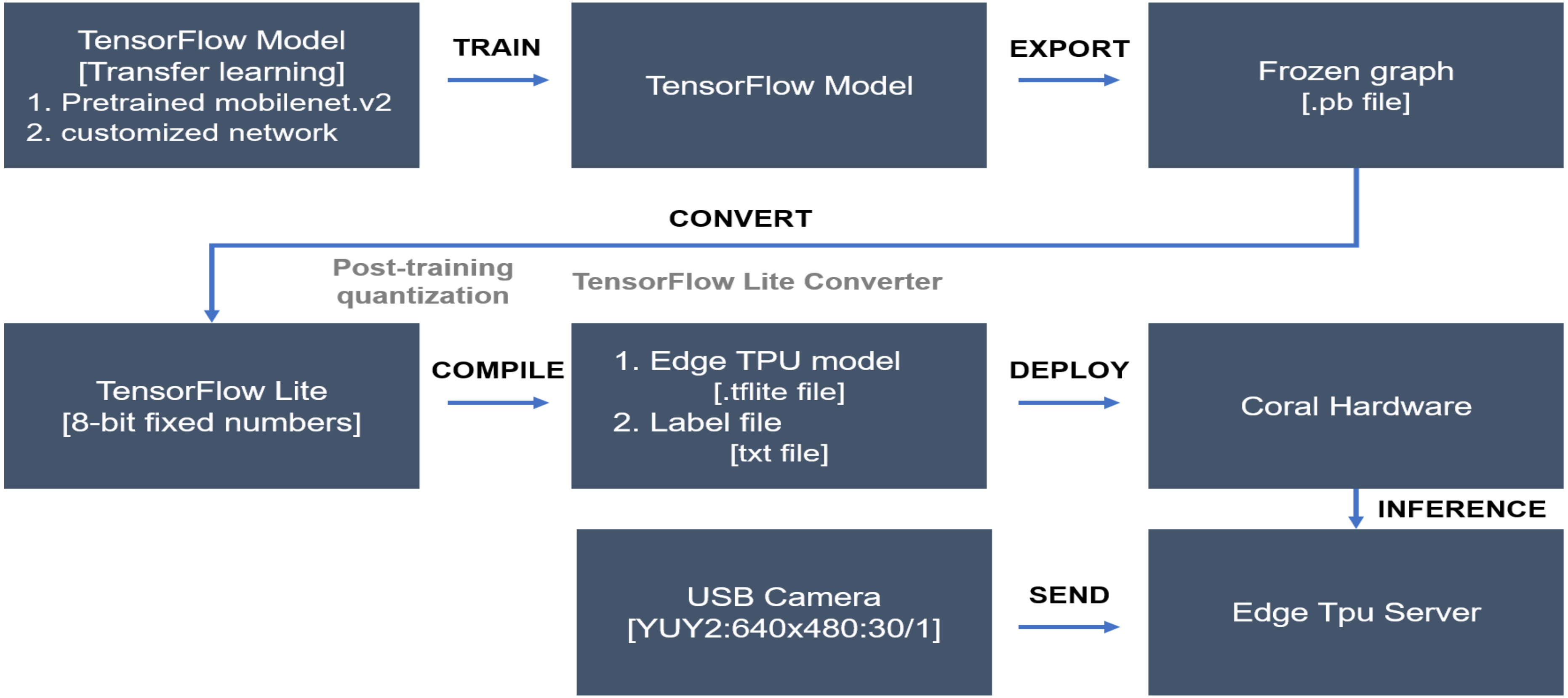


[1] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. -C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510-4520.
[2] Almeida, J. L., Flores, F. C., Roecker, M. N., Braga, M. A., & Costa, Y. M. (2019). An Indoor Sign Dataset (ISD): An Overview and Baseline Evaluation. In *VISIGRAPP (4: VISAPP)* (pp. 505-512).
[3] Adhikari, B., Peltomaki, J., Puura, J., & Huttunen, H. (2018, November). Faster bounding box annotation for object detection in indoor scenes. In *2018 7th European Workshop on Visual Information Processing (EUVIP)* (pp. 1-6). IEEE.

2. Method (2/2)

■ Process (continue)

□ Sequence diagram



TensorFlow Model [Transfer learning]
1. Pretrained mobilenet.v2
2. customized network

TRAIN

TensorFlow Model

EXPORT

Frozen graph [.pb file]

CONVERT

Post-training quantization

TensorFlow Lite Converter

TensorFlow Lite [8-bit fixed numbers]

COMPILE

1. Edge TPU model [.tflite file]
2. Label file [txt file]

DEPLOY

Coral Hardware

INFERENCE

Edge Tpu Server

SEND

USB Camera [YUY2:640x480:30/1]

Figure 1. Sequence diagram of project

3. Experiment Design (1/5)

■Experimental Object and Progress: 66%

● **Step 1**

Binary Classification
1. Emergency sign
2. Others

● **Step 2**

Multiple Classification
1. Emergency sign
- Exit
- Direction
2. Others

Step 3

Multiple Classification
1. Emergency sign
- Exit
- Direction
  - Left
  - Right
  - Bidirectional
2. Others

3. Experiment Design (2/5)

■ Dataset

□ Main issue: Imbalanced dataset

- The number of samples for the class including others is much more than those of emergency sign.
  - □ The class consists of various signs such as accessibility, wi-fi, non-smoking, man-bathroom, woman-bathroom, Warnings
  - □ However, the classes about emergency signs would be smaller following the objective steps.

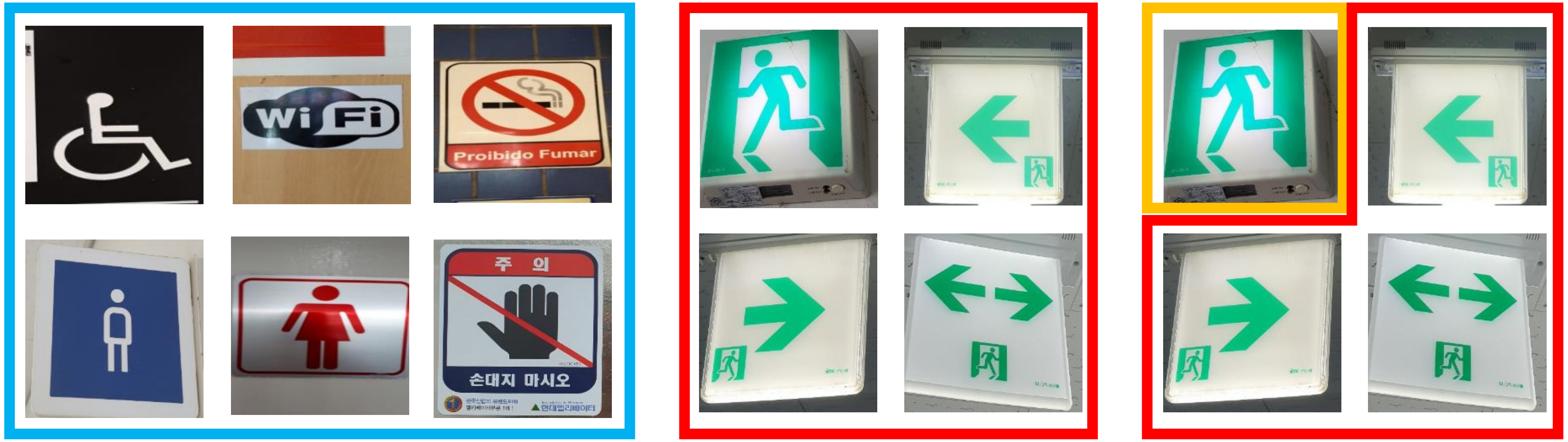


Figure 2. The composition of the classes

3. Experiment Design (3/5)

■Dataset [Step 1]

□Size [Class 1: Emergency, 2: Non-Emergency]

- Train: 100 [>>600] , 590
- Vali: 47, 36
- Test: 10, 10

□Data Preprocessing

- Transform: Non-Cropping
  - □Scaling: 1/255
  - □Resizing: (224, 224)
  - □Horizontal flip, vertical flip

3. Experiment Design (4/5)

■Dataset [Step 2]

□Size [Class 1: Exit_Arrow , 2: Exit_Here, 3:Non-Exit]

- ▪ Train: 65[〉〉1002], 35[〉〉1002], 590
- ▪ Vali: 25, 12, 36
- ▪ Test: 37, 13, 324

□Data Preprocessing

- ▪ Transform: Non-Cropping
  - □Scaling: 1/255
  - □Resizing: (224, 224)
  - □Horizontal flip, vertical flip
  - □Brightness range: [0.7, 1.3]
- ▪ The parts of Emergency signs look similar each other

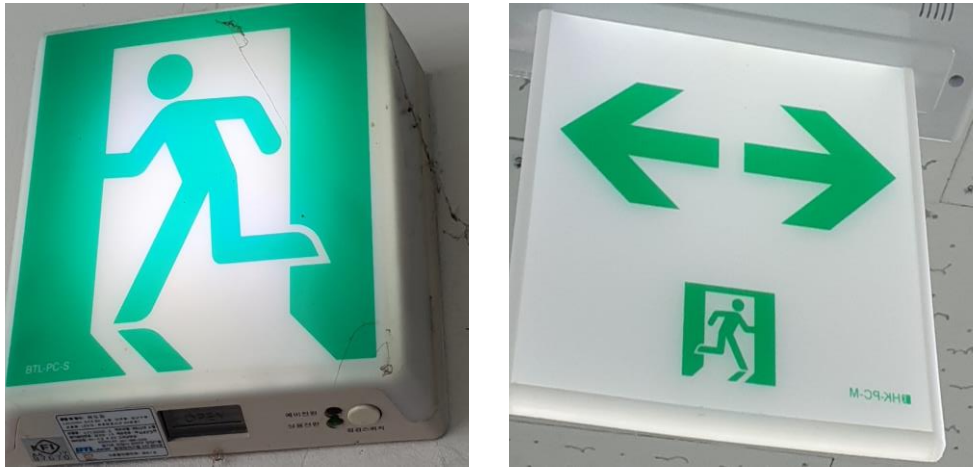


Figure 3. The comparison between the Emergency signs

3. Experiment Design (5/5)

■Setting

- □Batch size: 64
- □Epoch: 20
- □Metric: AUROC
- □Early Stopping: If val_loss kept increasing more than 3 epochs, Stop
  - ▪ Monitor: Val_Loss
  - ▪ Patience: 3
  - ▪ Mode: min

4. Result (1/4)

■ Scores

- □ Train/Validation
  - ▪ Step 1
    - □ Loss kept decreasing
    - □ AUROC kept increasing
  - ▪ Step 2
    - □ Loss kept increasing
    - □ AUROC kept decreasing

Step1

Step2

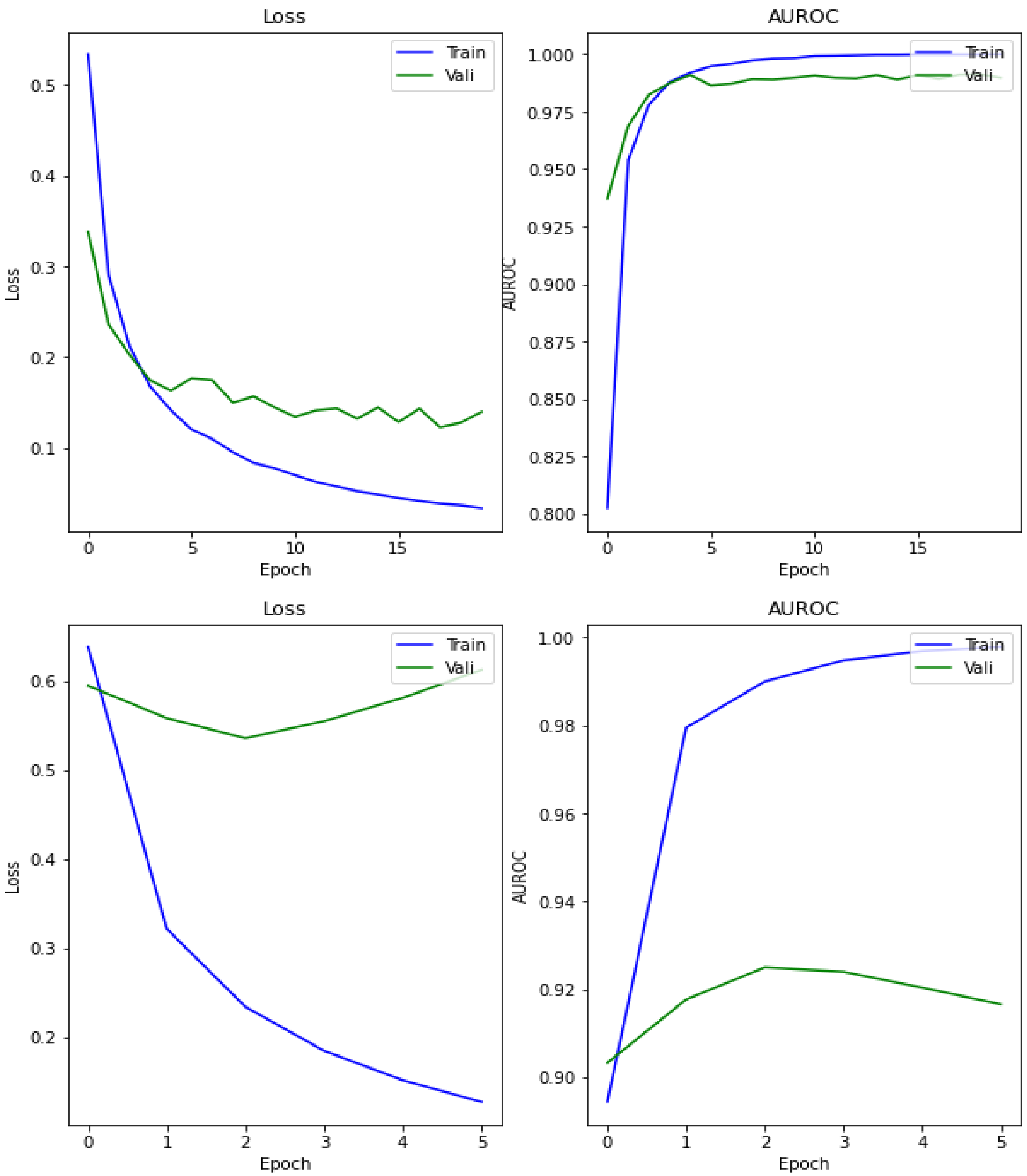


Figure 4. Train Loss and AUROC scores

4. Result (2/4)

■ Scores

□ Test

- ▪ Step 1
  - □ AUROC: 0.95
- ▪ Step 2
  - □ AUROC: 0.9538
    - ▶ Parameters: Sklearn package, average='macro', multi_class='ovo'
    - ▶ Macro: calculate metric for each label 〉〉 find unweighted mean
    - ▶ Ovo: One-vs-One, Computes the average AUC of all possible pairwise combinations of classes

True Predict	Emergency	Non-Emergency
Emergency	10	0
Non-Emergency	1	9

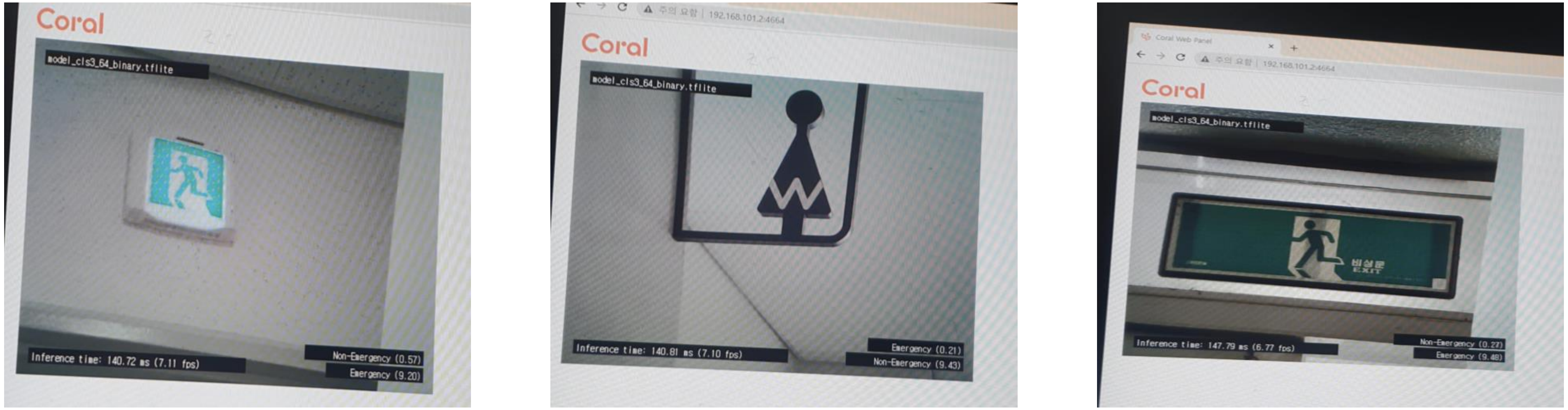
True Predict	Exit_Arrow	Exit_Here	Non-Exit
Exit_Arrow	43	3	1
Exit Here	4	14	0
Non-Exit	10	3	310

Table 1. Confusion matrix based on the predictions

4. Result (3/4)

■Inference [Step 1]

- □Images: 11 images collected manually
- □Inference time: generally, 140-150ms
- □The Sings are recognized well enough to be satisfied.


Coral
model_cls3_64_binary.tflite
Inference time: 140.72 ms (7.11 fps)
Non-Emergency (0.57)
Emergency (9.20)
Coral
model_cls3_64_binary.tflite
Inference time: 140.81 ms (7.10 fps)
Emergency (0.21)
Non-Emergency (9.43)
Coral
model_cls3_64_binary.tflite
Inference time: 147.79 ms (6.77 fps)
Non-Emergency (0.27)
Emergency (9.48)
비상구
EXIT

(a) Emergency
[00:48~00:54]

(b) Non-Emergency
[00:41~00:42]

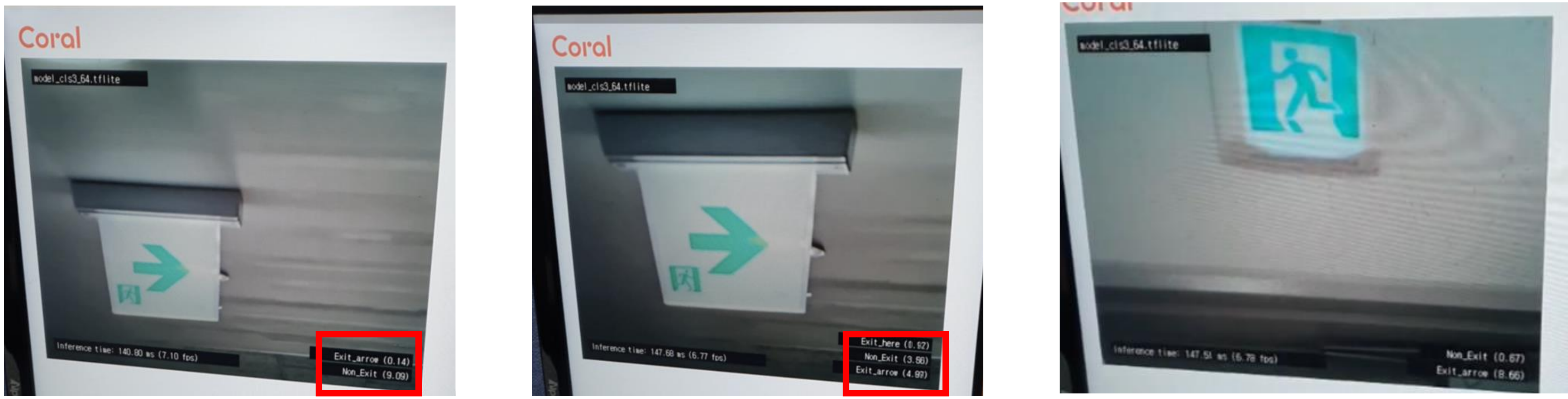
(c) Emergency
[01:16~01:20]

Figure 5. Inference examples of Step 1.

4. Result (4/4)

■Inference [Step 2]

- □Inference time: generally, 140~150ms
- □It takes more time to recognize the signs correctly in the case of images with backgrounds like ceiling, wall, windows.



```
model_cls3_64.tflite
Inference time: 140.80 ms (7.10 fps)
Exit_arrow (0.14)
Non_Exit (9.09)
```

```
model_cls3_64.tflite
Inference time: 147.68 ms (6.77 fps)
Exit_here (0.92)
Non_Exit (3.58)
Exit_arrow (4.97)
```

```
model_cls3_64.tflite
Inference time: 147.51 ms (6.78 fps)
Non_Exit (0.67)
Exit_arrow (8.66)
```

(a) Exit_Arrow
[00:18~00:39]

(b) Exit_Here
[01:01~01:02]

Figure 6. Inference examples of Step 2.

5. Discussion

- ■ Application
  - □ High learnability: easy to use
  - □ Low Usability: Difficult to detect specific sign among signs
- ■ Future work
  - □ Considering directionality of emergency signals for route notification.
    - ▪ More datasets are needed.
  - □ Choosing a target device.

QnA