University of Technology
(Yatanarpon Cyber City)

Eliminating False Detection of Transparent Object by Object Feature Region

Supervisor
Masayuki MUKUNOKI
Co-supervisors
Prof. Kunihito YAMAMORI
Prof. Thi Thi Zin

T1703059
May Phyo Khaing

Outline

- Abstract
- Introduction
- Object Feature Region
- Experimental Results
- Conclusion

1. Abstract

- We propose combination of **transparent object feature region** to SSD model for eliminating false detections from transparent object detection.
- We manually define object feature regions on each transparent objects and **train** along with the glass training data.
- During testing, we **eliminate false detections** by removing the glass regions which do not contain any glass-feature region.
- We do **performance evaluation** of the proposed method and make **comparison** with other four alternative training processes.

2. Introduction

Object detection is the detection of any kind of objects existing in our environment.

2. Introduction

Object detection is the detection of any kind of objects existing in our environment.

Lack of obvious physical feature (e.g. no color)

- Difficult to perform detection with classical computer vision algorithms

Transparent Objects

- Transparent objects are the objects with special features.

glass and background has similar color

Color similarity

Highlights

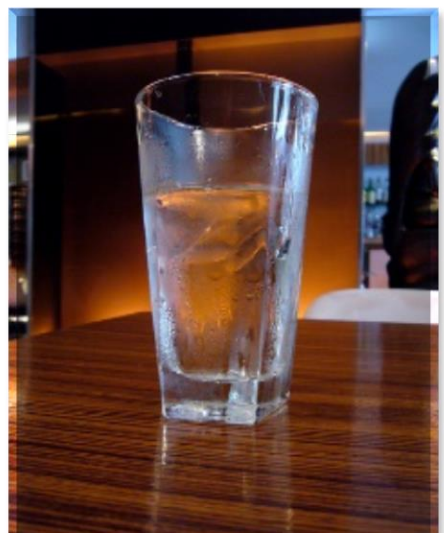
specularities appear on the glass region

Texture of the glass region is blurred

Blurring

intensity distribution on a glass region is constrained by intensity distribution on its background

Overlay consistency

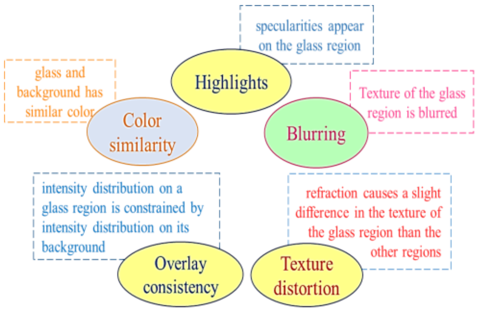


refraction causes a slight difference in the texture of the glass region than the other regions

Texture distortion

Fig. 1. Transparent object and its characteristics

Previous Object Detection Method



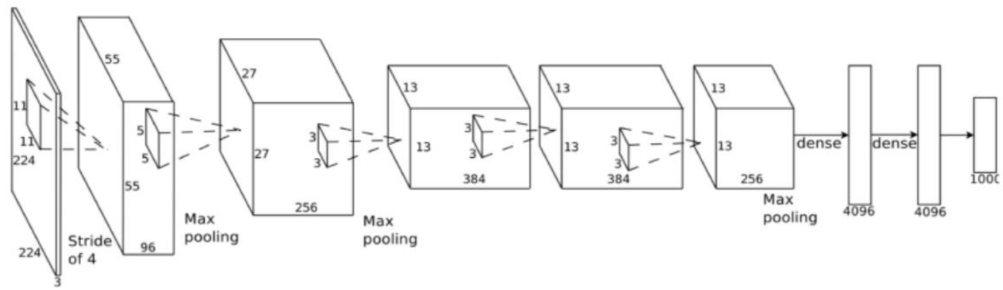
Detect presence of each feature

Input Image → Feature Extractor → Feature → Traditional ML Algorithm → Output

Traditional Machine Learning Flow

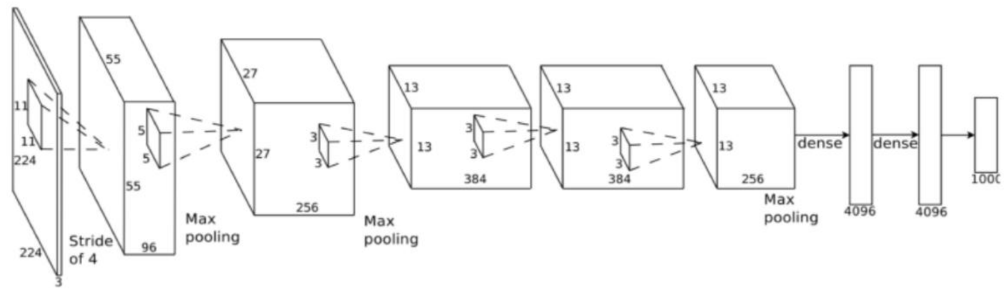
Some limitations are required to detect feature (e.g. glass must be on simple background)

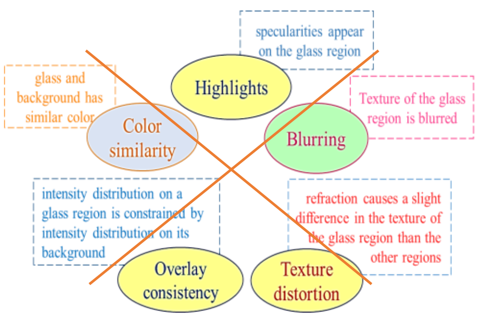
Transparent Object Detection with CNN



CNN : Convolutional Neural Network

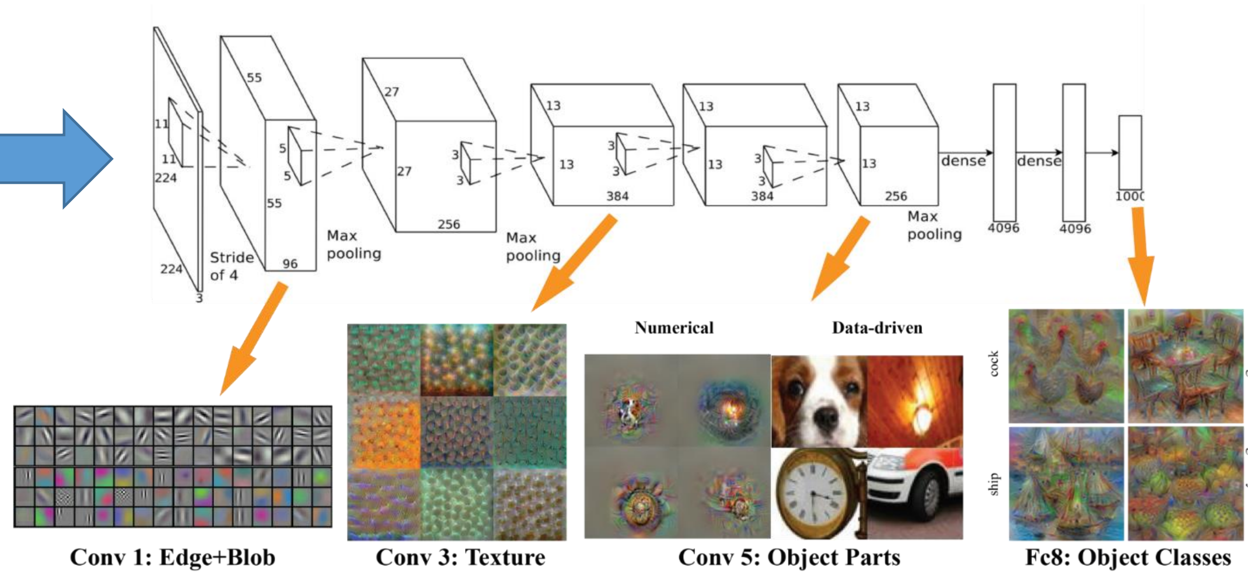
Transparent Object Detection with CNN

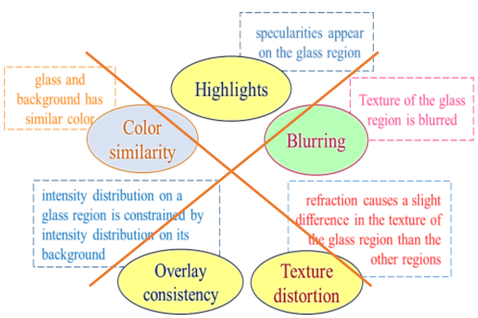




CNN : Convolutional Neural Network

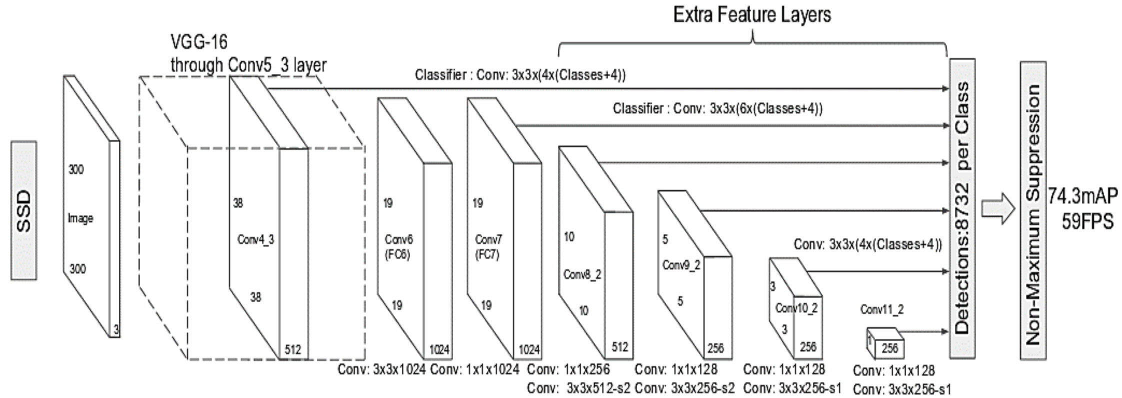
Transparent Object Detection with CNN

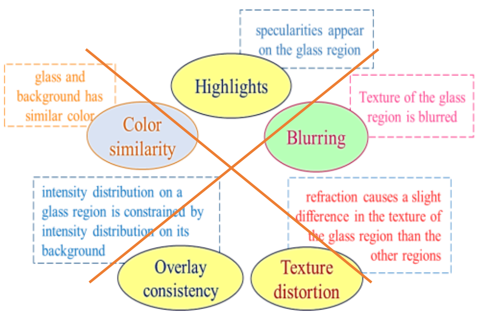




CNN : Convolutional Neural Network

Transparent Object Detection with SSD





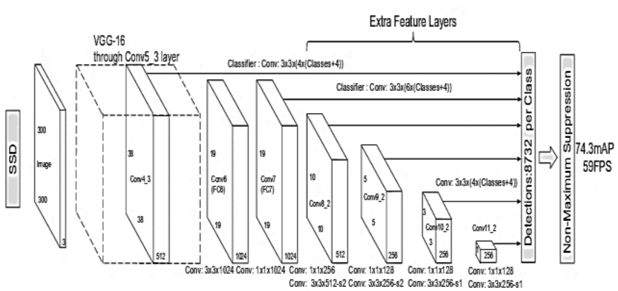
- Therefore, one of the Convolutional Neural Networks called **SSD** [1] is used for transparent object detection.

SSD : Single Shot Multibox Detector [1]

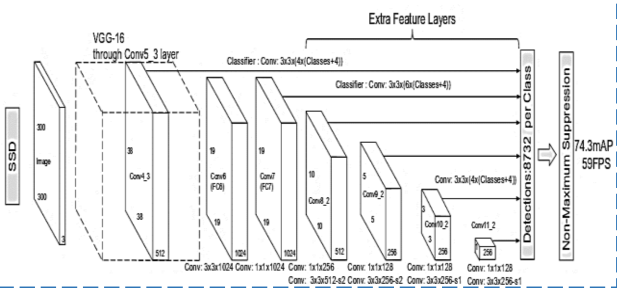
Transparent Object Detection with SSD

Transparent object images taken from **ImageNet ILSVRC** dataset

Train

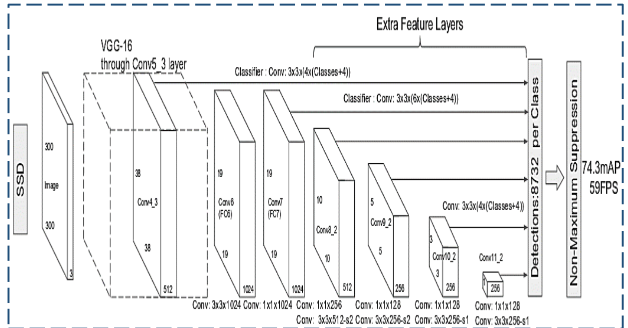


SSD trained with glass objects



Problem

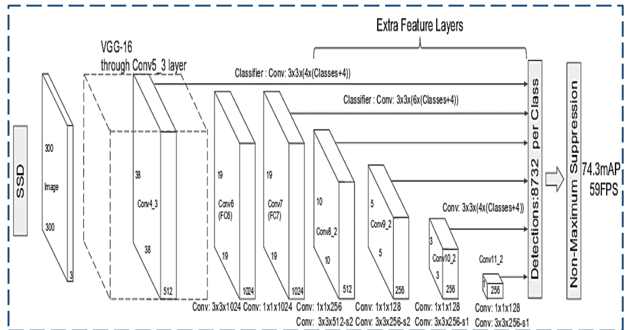
Be
Er
Beryllium
Erbium

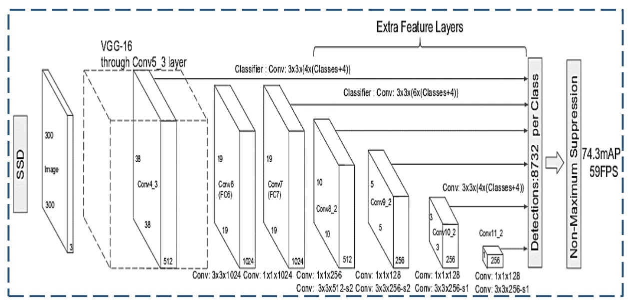


Extra Feature Layers
VGG-16
through Conv5_3 layer
SSD
Classifier : Conv: 3x3x(4x(Classes+4))
Classifier : Conv: 3x3x(6x(Classes+4))
Conv: 3x3x(4x(Classes+4))
300
Image
300
3
38
Conv4_3
38
512
19
Conv6
(FC6)
19
1024
Conv7
(FC7)
10
Conv8_2
10
512
5
Conv9_2
5
256
3
Conv10_2
3
256
Conv11_2
256
Conv: 3x3x1024
Conv: 1x1x1024
Conv: 1x1x256
Conv: 3x3x512-s2
Conv: 1x1x128
Conv: 3x3x256-s2
Conv: 1x1x128
Conv: 3x3x256-s1
Conv: 1x1x128
Conv: 3x3x256-s1
Detections:8732 per Class
Non-Maximum Suppression
74.3mAP
59FPS

class 1: 1.00
class 1: 1.00
Be
Er
Beryllium
Erbium

Problem





Wrongly detects non-transparent objects as transparent objects

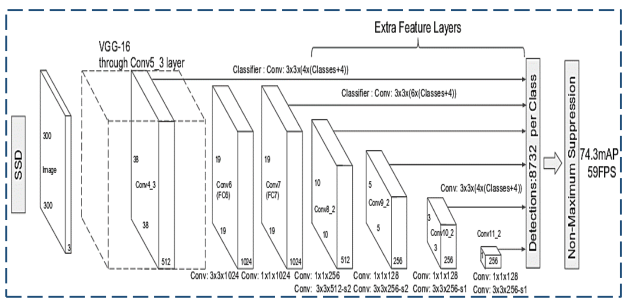
Solution

Training data with glass and glass-feature bounding boxes

Object Feature Region

. . .

Train



Solution

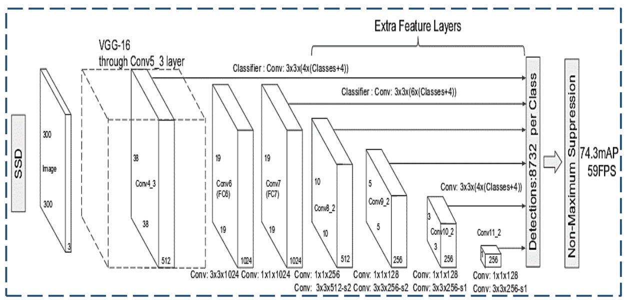
Training data with glass and glass-feature bounding boxes

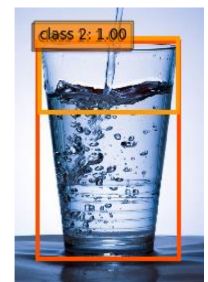
. . .

Object Feature Region

Train

Test





Glass

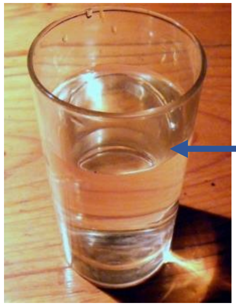
Non-glass

- Eliminate false detections by excluding the detected regions without any glass-feature region.

3. Object Feature Region

Transparency

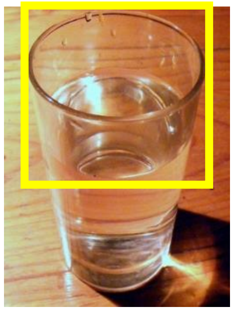
➢ The rays of light can pass through the glass medium.



Water and beer level that can be seen through the glass body

No transparency through paper cup and coffee mug body

Transparent Object Feature Region



Transparent object feature regions are **unique** to transparent object.

- We propose these glass-feature regions to **eliminate false detections**.

Transparent Object Feature Region

- The transparent object features are defined at different regions of the glass.

Empty glass

glass-feature region

Glass with liquid

glass-feature region

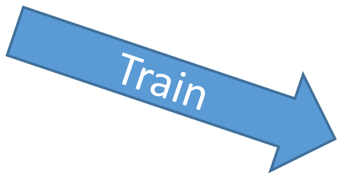
4. Experimental Results

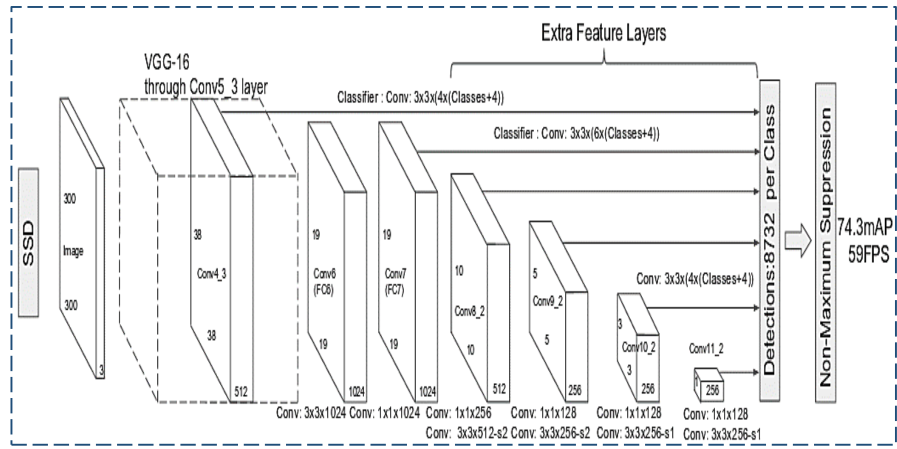
Training Dataset (For proposed method)

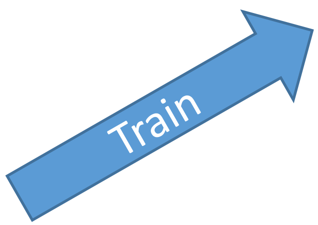
Object Classes		Num. of Images	Num. of object bboxes	Num. of glass-feature bboxes
Transparent Objects	Beaker	411	541	541
	Beer glass	345	379	407
	Water glass	282	302	319
	Wine glass	530	666	666
Total		**1,568**	**1,888**	**1,933**

Network Trained with Proposed Method

class 1 : glass







class 2 : glass-feature

- Train SSD with **glass** and **glass-feature**

Final Detection Results

Glass-feature region

Glass regions which includes glass-feature regions inside

classify as

glass

Glass region

Glass regions which do not includes glass-feature regions inside

classify as

non-glass

Glass region

Comparison Methods

1. Network Trained with only Glass
2. Network Trained with Glass and Augmented Glass Data
3. Network Trained with Glass and Negative Training Data
4. Network Trained with Glass and Non-glass

Training Dataset (For comparison methods)

Object Classes		Num. of Images	Num. of object bounding boxes
Non-transparent Objects	Paper cup	500	648
	Coffee mug	200	230
	Coffee cup	200	223
Total		**900**	**1,101**
Negative training objects	Bicycle	138	158
	Car	150	168
	Airplane	150	168
	Child	85	94
	Cat	163	168
	Dog	150	162
	Table	150	156
	Chair	150	183
	Clock	150	153
Total		**1,286**	**1,410**

Comparison Methods (1/2)

1. Network Trained with only Glass

- Train with images which contain only glass bounding boxes
- Base line network for detecting transparent object

2. Network Trained with Glass and Augmented Glass Data

- Data augmentation increases the number of training data.
- The network can learn more training samples from increased data.
- Train with glass and horizontally flipped glass images

Comparison Methods (2/2)

3. Network Trained with Glass and Negative Training Data

- To decrease FP detections, we use images with false detections.
- Train with glass images and negative training images from 9 classes of common objects.

4. Network Trained with Glass and Non-glass

- To reduce false detections on non-transparent objects of the same shape, these objects themselves are used in training.
- Train with glass images and non-glass images of the same shape as transparent objects

Testing Dataset

Table 1. The number of ground-truth bounding boxes in each class of testing images.

	Image classes	Num. of ground-truth bounding boxes
Transparent objects	Beaker	109
	Beer glass	107
	Water glass	97
	Wine glass	94
	Total	**407**
Non-transparent objects	Paper cup	162
	Coffee cup	129
	Coffee mug	116
	Total	**407**

Performance Evaluation Matrices

- For performance evaluation of the detection results, we use the following matrices to compare different training processes.

1. True Positive (TP) and False Positive (FP)
2. Precision, Recall and F-measure
3. mean Average Precision (mAP)

TP and FP Comparison

Network		TP	FP
Network trained with only glass		394	550
Network trained with glass and augmented training data		**397**	661
Network trained with glass and negative training samples		395	505
Network trained with glass and non-glass		**397**	175
Network trained with glass and glass-feature	Glass-feature th 0.0	**397**	374
	Glass-feature th 0.1	392	198
	Glass-feature th 0.2	392	164
	Glass-feature th 0.3	390	143

TP and FP Comparison

Network		TP	FP
Network trained with only glass		394	550
Network trained with glass and augmented training data		397	661
Network trained with glass and negative training samples		395	505
Network trained with glass and non-glass		397	175
Network trained with glass and glass-feature	Glass-feature th 0.0	397	374
	Glass-feature th 0.1	392	198
	Glass-feature th 0.2	392	164
	Glass-feature th 0.3	390	143

TP and FP Comparison

Network		TP	FP
Network trained with only glass		394	550
Network trained with glass and augmented training data		397	661
Network trained with glass and negative training samples		395	505
Network trained with glass and non-glass		**397**	175
Network trained with glass and glass-feature	Glass-feature th 0.0	397	374
	Glass-feature th 0.1	392	198
	Glass-feature th 0.2	**392**	**164**
	Glass-feature th 0.3	**390**	**143**

Our proposed method has **little lower TP number** but **more FP number is reduced**.

Precision, Recall and F-measure Comparison

Network		Precision $TP/(TP+FP)$	Recall $TP/(TP+FN)$	F - measure 2/ ((1/Precision) + (1/Recall))
Network trained with only glass		41.74 %	96.81 %	58.33 %
Network trained with glass and augmented training data		37.52 %	**97.54 %**	54.19 %
Network trained with glass and negative training samples		43.90 %	97.05 %	60.45 %
Network trained with glass and non-glass		69.05 %	95.09 %	80.00 %
Network trained with glass and glass-feature	Glass-feature th 0.3	**73.17 %**	**95.82 %**	**82.98 %**

Precision, Recall and F-measure Comparison

Network		Precision $TP/(TP+FP)$	Recall $TP/(TP+FN)$	F - measure 2/ ((1/Precision) + (1/Recall))
Network trained with only glass		41.74 %	96.81 %	58.33 %
Network trained with glass and augmented training data		37.52 %	**97.54 %**	54.19 %
Network trained with glass and negative training samples		43.90 %	97.05 %	60.45 %
Network trained with glass and non-glass		69.05 %	95.09 %	80.00 %
Network trained with glass and glass-feature	Glass-feature th 0.3	**73.17 %**	**95.82 %**	**82.98 %**

Our proposed method achieves the **highest precision and F-measure** among all comparison methods.

mAP Comparison

Network		mAP
Network trained with only glass		75.27 %
Network trained with glass and augmented training data		73.71 %
Network trained with glass and negative training samples		77.03 %
Network trained with glass and non-glass		**94.87 %**
Network trained with glass and glass-feature	Glass-feature th 0.3	**87.60 %**

Our proposed method achieves a **near mAP performance** to the highest mAP result.

Discussion

- Our proposed method **reduces more FP detections** than other comparison methods.
- Our proposed method achieves the **highest precision and F-measure** along with a good recall result.
- Although the network trained with glass and non-glass shows the highest mAP result, it is very **difficult** to find its non-transparent training data.
- Our proposed method gives **almost the same mAP result to the highest mAP** result just with **a lower cost** in selecting training data.

6. Conclusion

- We propose transparent object feature region for eliminating false detections.

Our proposed method achieves the **highest precision and F-measure**, and a **near performance to the highest mAP** result just **with a lower cost** in selecting training data.

Future work

- Currently, the proposed transparent object feature region varies according to each kind of transparent objects.
- The future work will be to find another visual property that is common to all kinds of transparent objects and includes in the training processes.

Acknowledgement

- Firstly, we would like to give a special thanks to University of Miyazaki (UoM) for introducing Double Degree Program (DDP) and cooperating with our University of Technology (Yatanarpon Cyber City).
- I would like to appreciate to my supervisor Professor Masayuki Mukunoki for his kindness and guidance.
- Finally, our sincere thanks also goes to JASSO for the financial support for our education.

Thank You for Your Kind Attention !!!