Personal Mobility Safe Driving System with Knowledge Distillation

EWHA WOMANS UNIVERSITY

Damin Yeom*, Heejun Yoon*, Soyeon Lee*, and Kahyun Lee

Department of Electronic and Electrical Engineering, Ewha Womans University

* : contributed equally

Proposed System

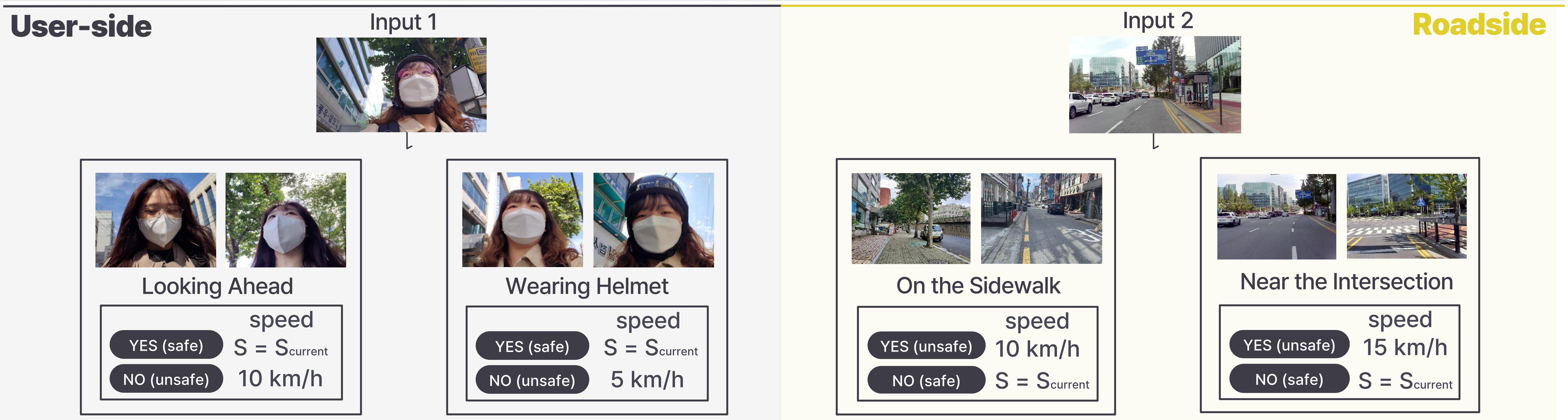

Fig 1. Summary of our proposed system.

Abstract

- **Issue:**
  Increasing number of incidents involving personal mobility devices
- **Key Idea:**
  - ✓ Deep learning (CNN)-based ***safe driving monitoring system for e-scooter*** that detects ***1. Wearing a helmet, 2. Looking ahead & 3. On the sidewalk, 4. Near the intersection.***
  - ✓ Apply additional ***lightweighting*** techniques for real-time inference.

Model Implementation

- Utilize standard ***Knowledge Distillation*** framework for fast and accurate inference for edge devices.
- **Teacher(T):** ***ResNet101 & VGG16*** (highest accuracy and parameters)
- **Student(S):** ***MobileNetV3-Small & EfficientNet-B0*** (fewer parameters for embedding)

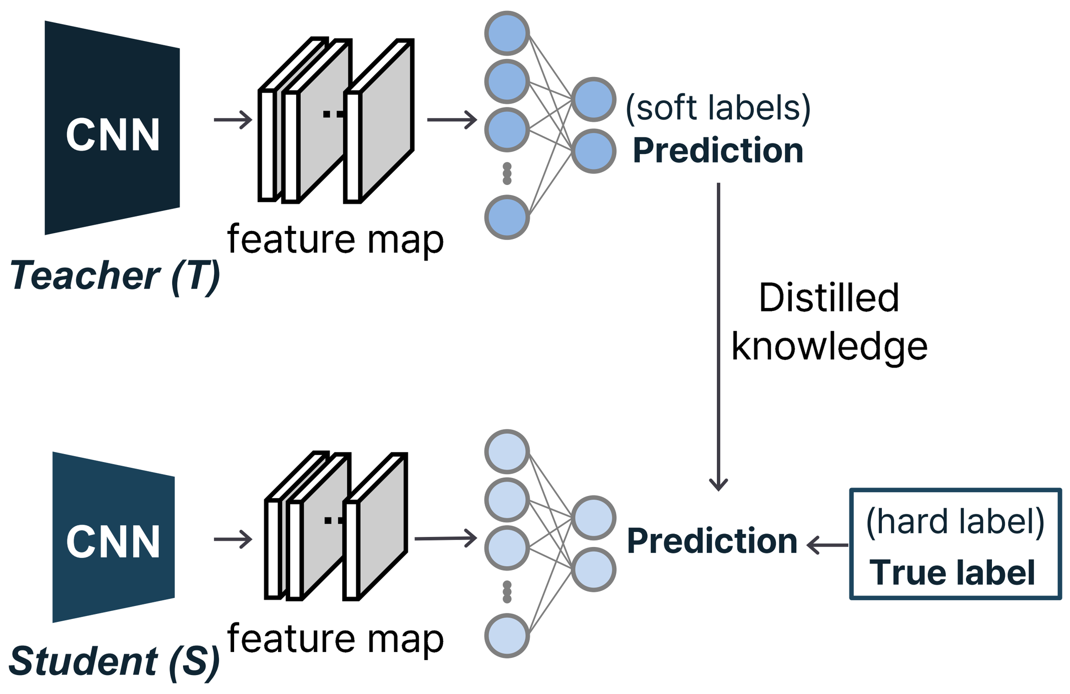

Fig 2. The architecture for the knowledge distillation

Teacher	# Params	Latency	Student	# Params	Ratio	Latency	Ratio
ResNet101	42.7M	1180ms	**EfficientNet-B0**	4.1M	9.6%	364ms	30.8%
			MobileNetV3-S	0.9M	2.1%	144ms	12.2%
VGG16	14.7M	965ms	**EfficientNet-B0**	4.1M	27.9%	364ms	37.7%
			MobileNetV3-S	0.9M	6.1%	144ms	14.9%

Table 1. Comparison number of parameters and latency between teacher and student networks. Latency is measured for one 224×224 image in yahboom jetson nano.

Qualitative Results

(a) (b) (c) (d)

Fig 3. Grad-CAM results of the well-classified & misclassified images : (a) Looking ahead, (b) wearing helmet, (c) on the sidewalk, and (d) near the intersection.

Introduction

- ✓ The autonomous driving market has led to increased usage of personal mobility devices (PMDs) like electric scooters.
- ✓ PMD usage has resulted in an increase in accidents and head injuries due to lack of protective equipment.
- ✓ There is a lack of research on PMD safety compared to automobiles, and current safety guidelines have weaknesses.

Quantitative Results

- **Datasets:**
  User-side - About 15,000 images of the three authors were taken near Gangnam & Seocho Station
  Roadside – Manually annotated Sidewalk pedestrian image public dataset & the SDLane dataset for our tasks
- **Evaluation results:**
  Student networks show ***1-5% higher accuracy*** compared to scratch networks in all four scenarios.

Task			Acc (%)	Task			Acc (%)
Looking Ahead	**Teacher**	ResNet101	94.41%	**On the Sidewalk**	**Teacher**	ResNet101	94.50%
	Student	EfficientNet-B0	90.09%		**Student**	EfficientNet-B0	**91.74%**
		MobileNetV3-S	90.27%			MobileNetV3-S	87.16%
	Teacher	VGG16	91.71%		**Teacher**	VGG16	93.58%
	Student	EfficientNet-B0	89.73%		**Student**	EfficientNet-B0	90.83%
		MobileNetV3-S	**91.89%**			MobileNetV3-S	87.16%
	Scratch	EfficientNet-B0	90.27%		**Scratch**	EfficientNet-B0	88.07%
		MobileNetV3-S	86.85%			MobileNetV3-S	87.16%
Wearing Helmet	**Teacher**	ResNet101	93.58%	**Near the Intersection**	**Teacher**	ResNet101	90.65%
	Student	EfficientNet-B0	91.89%		**Student**	EfficientNet-B0	88.79%
		MobileNetV3-S	89.74%			MobileNetV3-S	86.92%
	Teacher	VGG16	90.27%		**Teacher**	VGG16	86.95%
	Student	EfficientNet-B0	**93.33%**		**Student**	EfficientNet-B0	83.18%
		MobileNetV3-S	88.47%			MobileNetV3-S	**89.72%**
	Scratch	EfficientNet-B0	89.73%		**Scratch**	EfficientNet-B0	87.85%
		MobileNetV3-S	86.31%			MobileNetV3-S	84.11%

Table 2. The overall accuracy of the teacher and student networks. The value marked in bold indicates the accuracy of the highest-performance student model.

Conclusions

- **Contribution:**
  - ✓ Propose a deep learning-based safety system for personal mobility.
  - ✓ Use user-side and roadside images to check helmet usage, forward-looking behavior, and driving on sidewalks or near intersections.
  - ✓ Achieve fast and accurate inference with an accuracy range of 89.72% to 93.33% within 0.36 to 0.14 milliseconds per image.
  - ✓ Applicable to various road applications, including autonomous driving and delivery robots.
- **Discussion:**
  - ✓ Datasets captured in different lighting conditions are required.
  - ✓ Practical performance in actual driving scenarios is needed.