R1-2

# FPGA Implementation of a DPU-Based Facial Expression Recognition System

**Takuto Ando†, Yusuke Inoue ‡**

National Institute of Technology, Oita College, Advanced Course, Electrical, Electronics Information Engineering Major †
National Institute of Technology, Oita College, Department of Information Engineering ‡
aes2301@oita.kosen-ac.jp† y-inoue@oita-ct.ac.jp‡

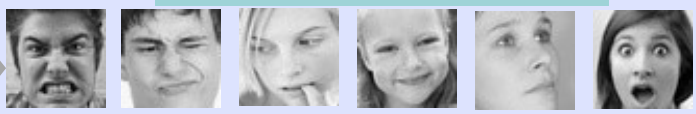
独立行政法人 国立高等専門学校機構 大分工業高等専門学校 National Institute of Technology, Oita College

## INTRODUCTION

### Facial Expression Recognition

- **Facial expressions** are effective in communication
- DNN-based facial expression identification technology is applied to robots

### Paul Ekman 6 Basic Emotions



Anger Disgust Fear Happy Sadness Surprise

Classification is possible on the basis of expression classes

### Problems with Previous Work

- Cascade detectors cannot adapt to changes in head angle and lighting conditions
- DNN based face detection consumes additional FPGA resources

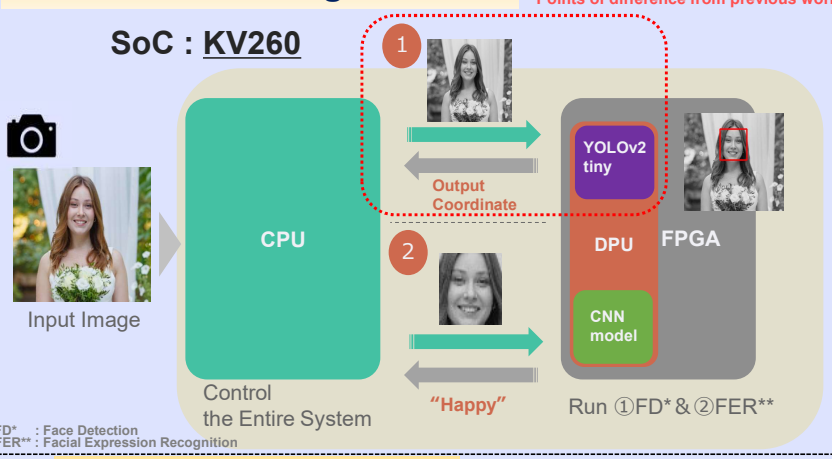
### Objectives

- Implement the system with real-world applications using DNN-based inference
- Use hardware resources efficiently by **DPU**

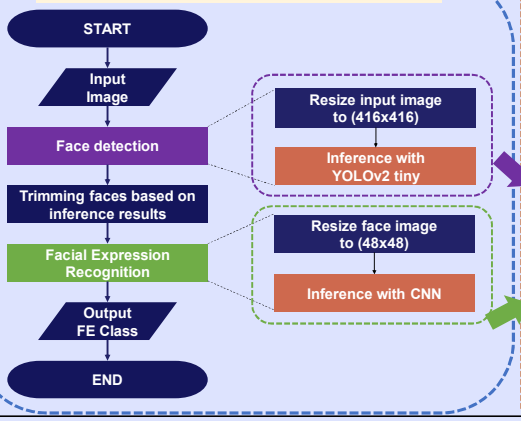
## METHODS & RESULTS

### Hardware Configuration

Points of difference from previous work




### Processing Flow




### DPU

- Dedicated hardware to accelerate inference of DNN
- Select an architecture from B512 to B4096 for each application
  → Our system adopted the B512

### Model

**Face detection : YOLOv2 tiny**
- Lightweight model
- Train dataset : The FDDB Dataset

**Facial Expression Recognition : CNN Model**
- Lightweight model
- Train dataset : FER-2013 Dataset

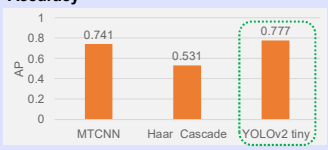
Adaptive quantization to run on DPU

### Experiments and Results

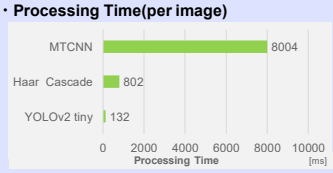
- Verify the effectiveness of offloading to DPU in our system
- FD and FER are evaluated in terms of latency and recognition accuracy

**Face detection** — Test Dataset : AFW Dataset

· Accuracy



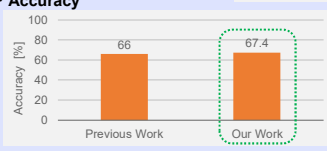

· Processing Time(per image)



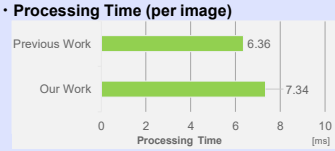

- YOLOv2 tiny Resuls Accuracy : **0.777** (slightly lower but acceptable), Latency : **132** [ms] (shortest)
- YOLOv2 tiny on DPU improved accuracy by 1.46× and reduced latency by 6.08× compared to the Haar Cascade detector
- Model is considered the best in terms of detection accuracy and latency for face detection

**Facial Expression Recognition** — Test Dataset : FER-2013 Dataset

· Accuracy




· Processing Time (per image)




- The quantized facial expression recognition model achieved an accuracy of 67.4%
- Accuracy is slightly higher than previous work
- Processing time is shorter in previous work

Little change in accuracy and latency

## DISSCUTIONS

FPGA resource consumption is higher than in previous work. BUT two DNN inferences can be run.

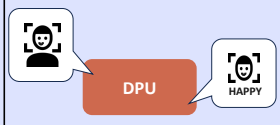



Table 1 Comparison of FPGA resource usage

| | ALM or LUT | DSP |
|---|---|---|
| Previous Work (Intel: ALM) | 22,465 | 112 |
| Our Work (Xilinx: LUT) | 27,023 | 118 |

Previous work only ran FPGA for facial expression recognition

## CONCLUSIONS

- We implemented a stand-alone DPU based facial expression recognition system on SoC FPGA
- Face detection accuracy was improved 1.46×, and the latency was reduced 6.08×
- Configuration using the same DPU achieves better results than the previous work while restraining the increase in the circuit area

**References**

1. Y.-l. Tian, T. Kanade, and J. Cohn. Recognizing action units for facial expression analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 2, pp. 97–115, 2001.
2. P. T. Vinh and T. Q. Vinh. Facial expression recognition system on SoC FPGA. In 2019 International Symposium on Electrical and Electronics Engineering (ISEE), pp. 1–4. IEEE, 2019.

2024.3.11-12 SASIMI 2024