Exploring Cross Modality Feature Fusion for Activity Recognition at OpenPack Challenge 2022

2nd Place

vbu211

Tetsuo Inoshita[1], Yuto Namba[1,2], Yuichi Nakatani[1], Kenta Ishihara[1], Sachio Iwasaki[1], Kosuke Moriwaki[1], Xian-Hua Han[2]

1) NEC Corporation, 2) Yamaguchi University

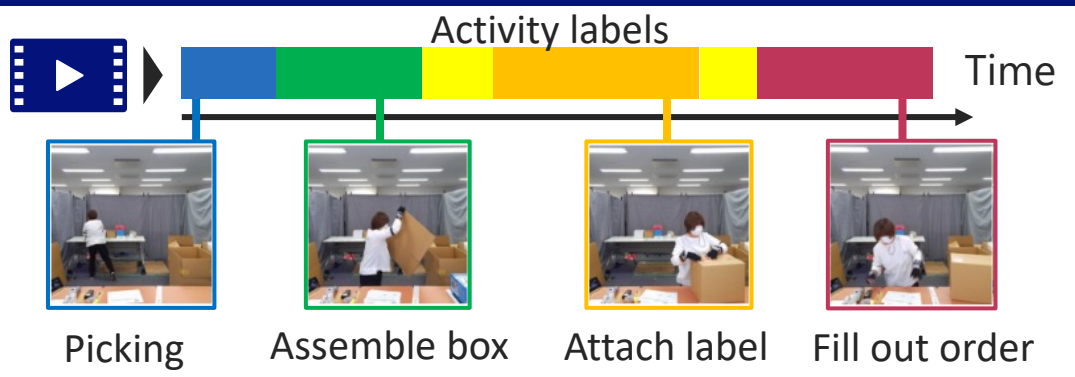
YAMAGUCHI UNIVERSITY

\Orchestrating a brighter world NEC

Introduction

OpenPack Challenge 2022

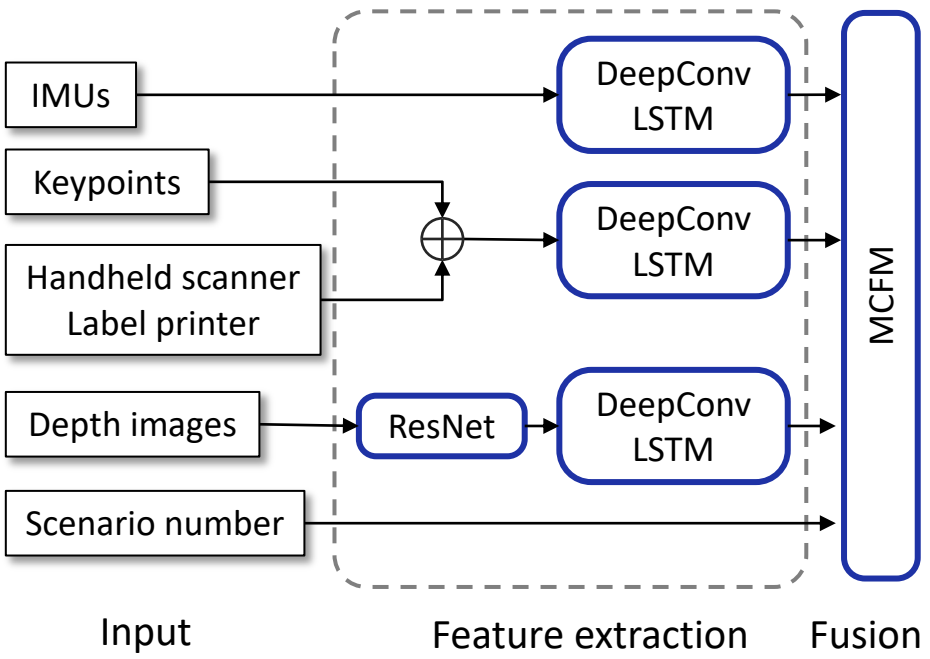
is a task to recognize 10 work operations along a timeline from sensor data (IMUs, vision, IoT-devices). It aims to solve the problem of "when and what kind of work" was performed.



* Source of photos (OpenPack Dataset's website) : https://open-pack.github.io/

Proposed Method

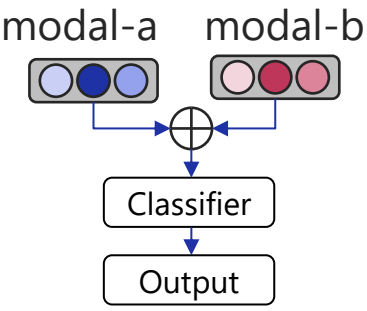
1. Overall framework



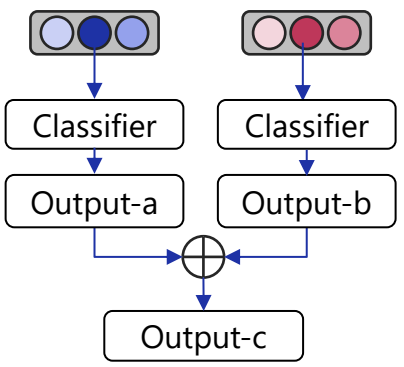
2. Fusion Strategies

a) Early fusion
b) Late fusion
c) Intermediate fusion

Cross Modality Feature Fusion

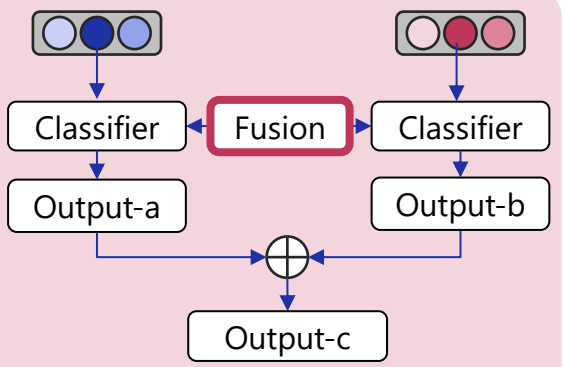


a) Early fusion



b) Late fusion

Our focus

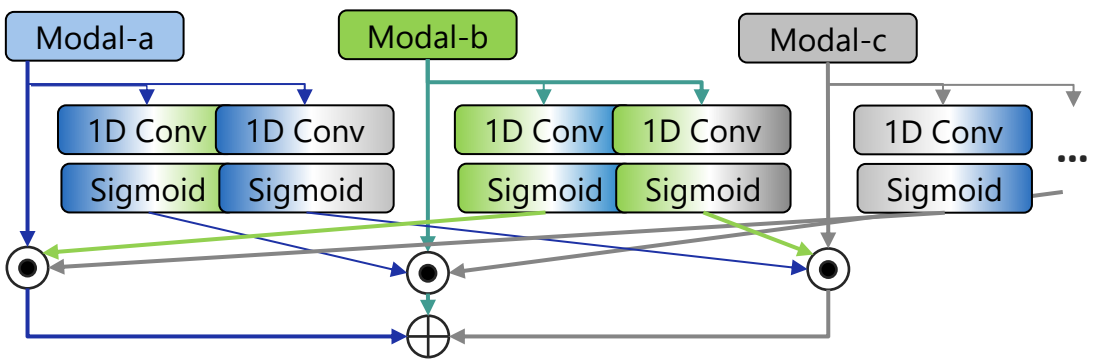


c) Intermediate fusion

3. Proposed fusion module (MCFM)

MCFM: Mutual Cross Fusion Module

- Recalibrate each feature from only other feature



Experiment

Modality	Test set	Submission set
IMU	0.940	-
Keypoints + IoT	0.964	0.945
Depth image	0.944	-
Fusion (Ours)	**0.975**	**0.959**

Scenario number improves the accuracy for S3 (Scenario.3: irregular situations/actions).

	Test set	(S1) U0102-	(S1) U0106-	(S3) U0202-	(S3) U0210-
Fusion (ours)	**0.975**	0.978	**0.978**	**0.968**	**0.976**
w/o Scenario num.	0.974	**0.981**	0.976	0.966	0.962