MAHASARAKHAM
UNIVERSITY

FOOD IMAGE CLASSIFICATION WITH IMPROVED MOBILENET ARCHITECTURE AND DATA AUGMENTATION

Sirawan Phiphiphatphaisit
Mahasarakham University, Thailand

OUTLINE

1 Introduction

2 Related work

3 Methodologies

4 Experimental and results

5 Conclusion

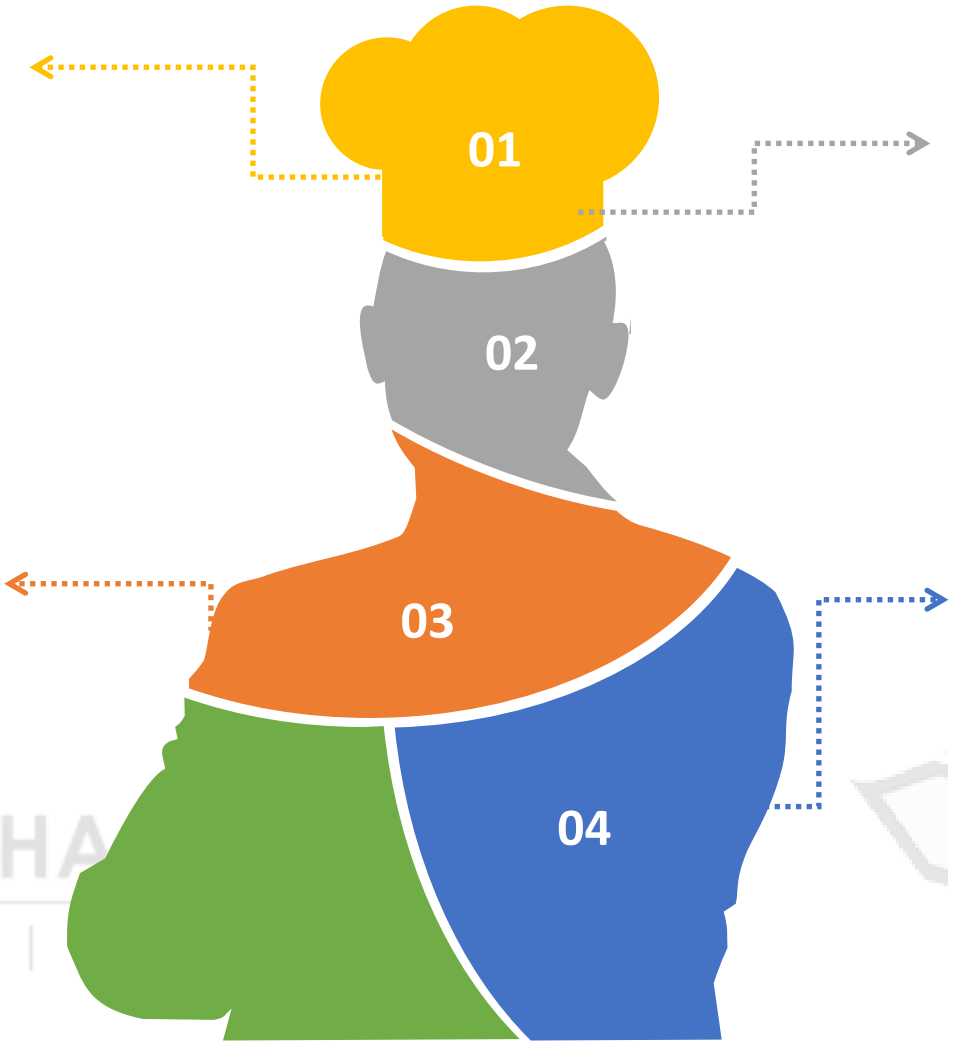
INTRODUCTION

What the problem?

- People are becoming obesity and overweight.

How to classification the images?

- Hand-craft feature
- Deep learning Algorithm



How to resolve the problem?

- The estimate calories system from food images.

What the challenge?

- Images are different
- Images are similar
- Images are many object

RELATED WORK

Reference	Method	Accuracy (%)
Yanai and Kawano [2015]	Deep Convolution Neural Network (DCNN)	70.41
Martinel et.al. [2016]	Supervised Extreme Learning Committee	55.89
Pandey et al. [2017]	FoodNet: Ensemble Net	72.10
Lin et al. [2018]	DeepFood	77.00

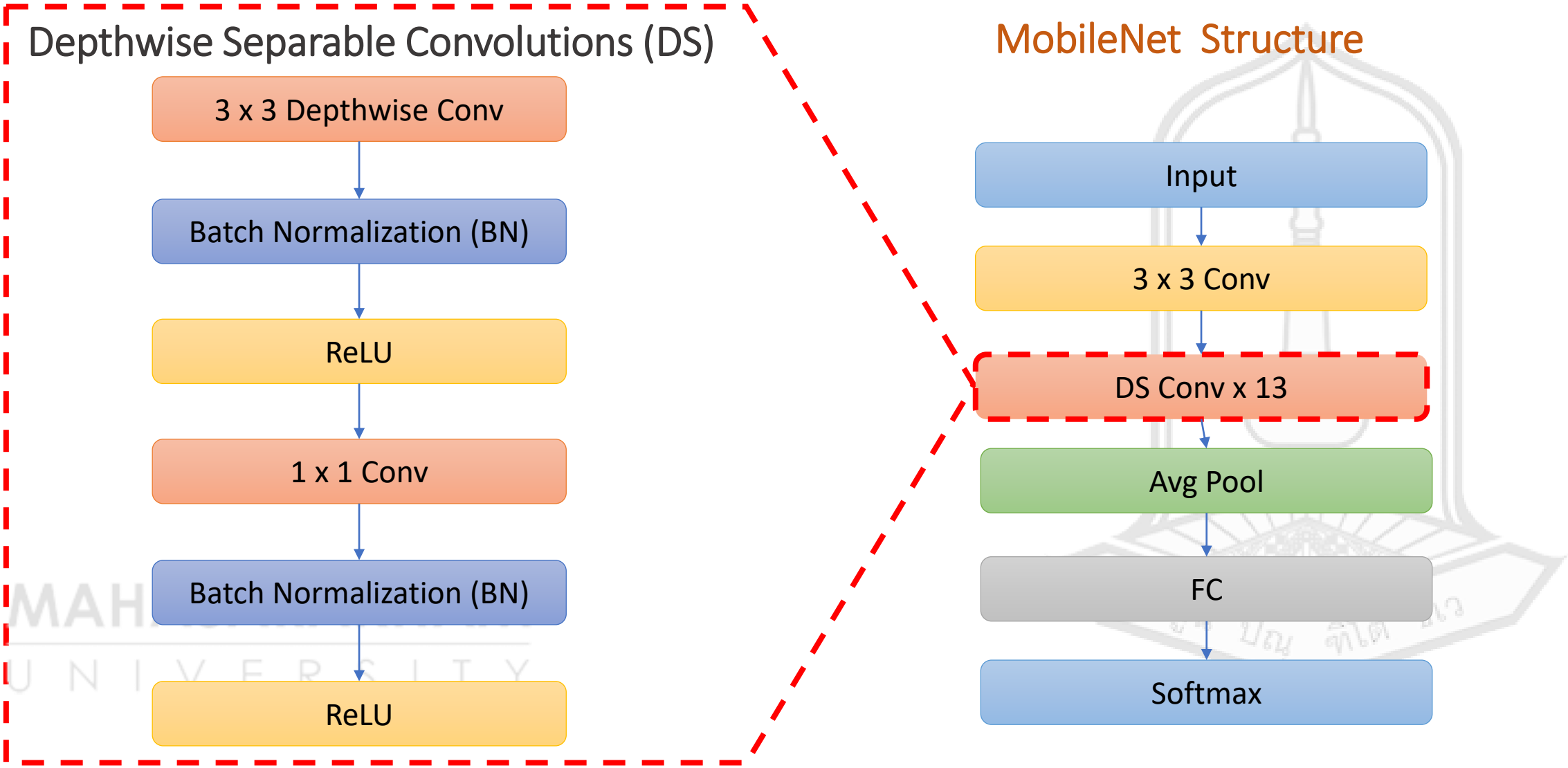
CONTRIBUTION

Improved MobileNet Architecture

Data Augmentation Technique

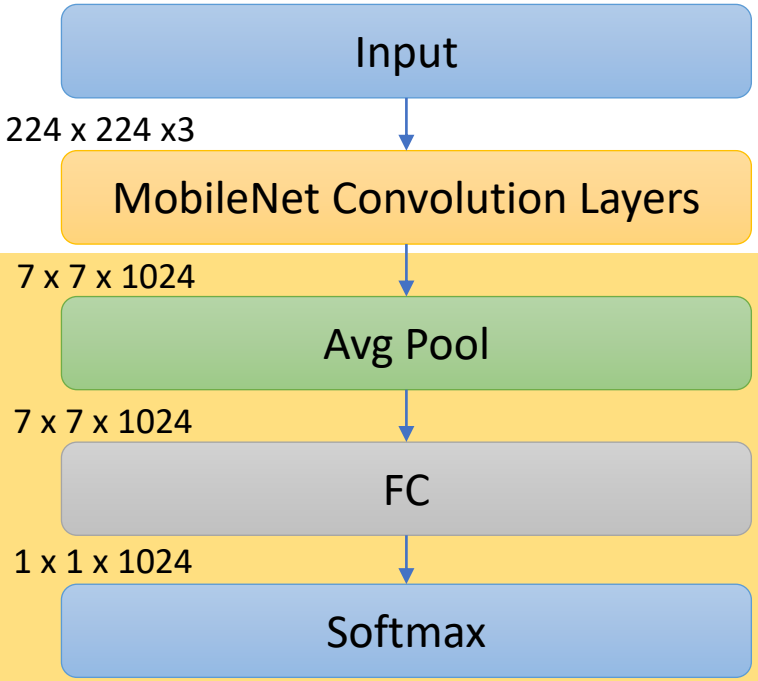
ETH Food-101 dataset

METODOLOGIES

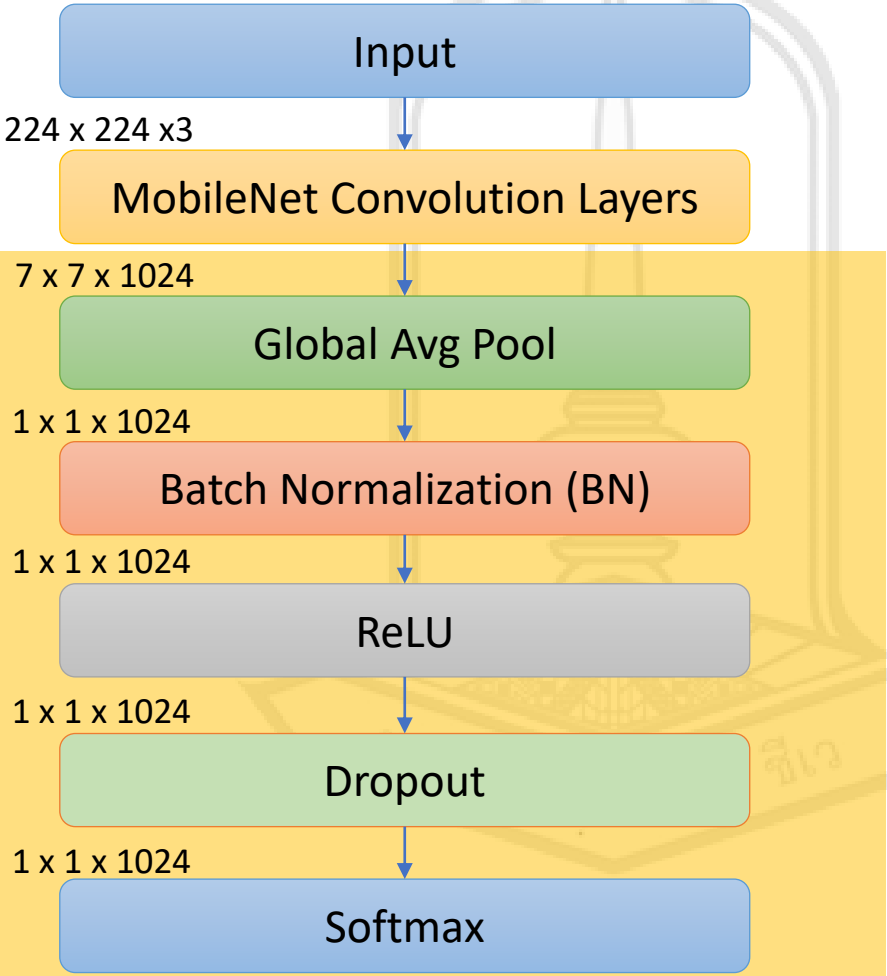


METODOLOGIES

MobileNet Architecture



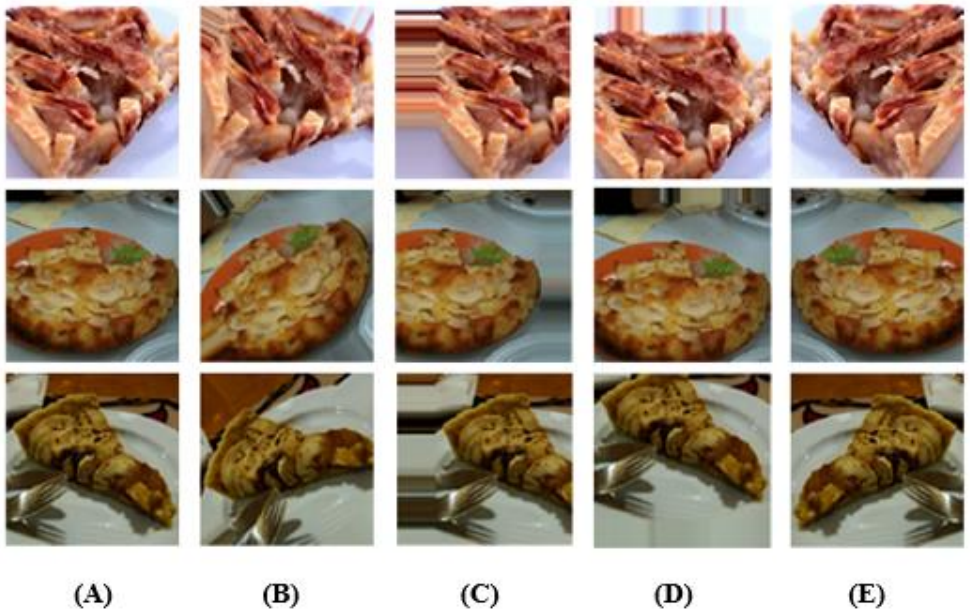
Proposed MobileNet Architecture



METODOLOGIES

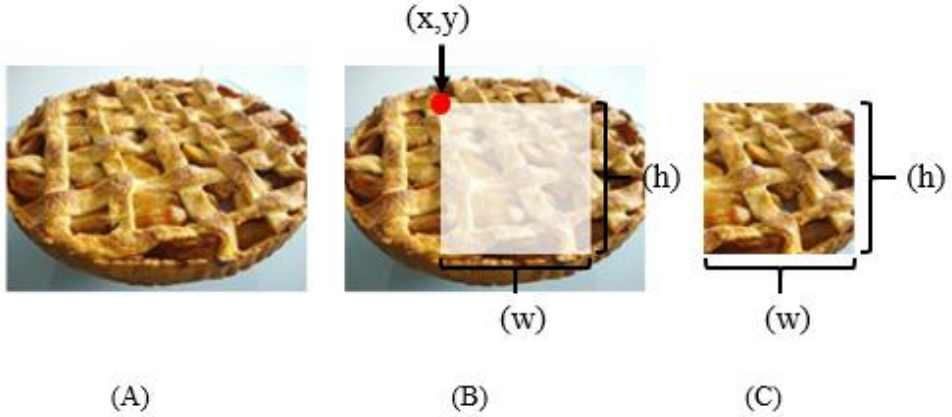
1

The data augmentation techniques ; rescaling, rotation, width shift, height shift, horizontal flip, shear, and zoom.



2

Random cropping, the position of points (x,y) are random, then it automatic cropping and resize to the target size.



Data Augmentation Techniques

EXPERIMENTAL & RESULTS

Dataset

ETH Food-101

- The real-world food images are collected by downloading from foodspotting.com website.
- The food images are a mix of eastern and western meals.
- The dataset consists of 101,000 images from 101 food categories.

EXPERIMENTAL & RESULTS

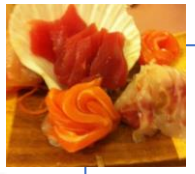
Experimental Setup

- Divided the dataset

Dataset
101,000 images

Set I
10,100 images

Set II
20,200 images



Set III
30,300 images

Set IV
40,400 images

Training
70%

Validation
10%

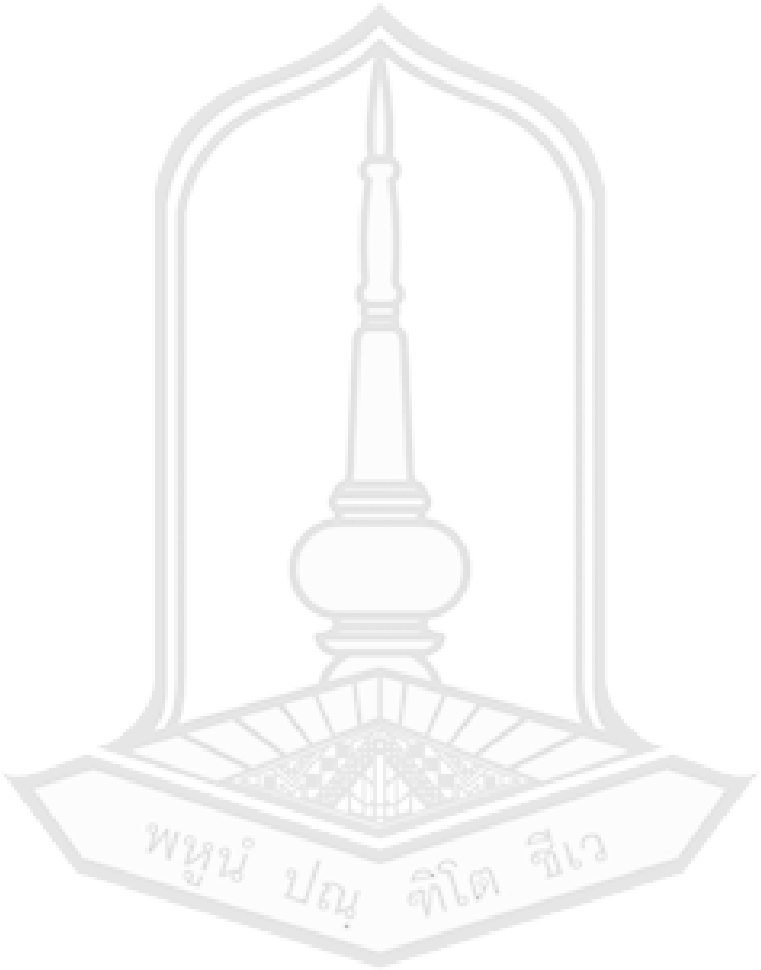
Testing
30%

EXPERIMENTAL & RESULTS

Experimental Setup

- The parameter setting

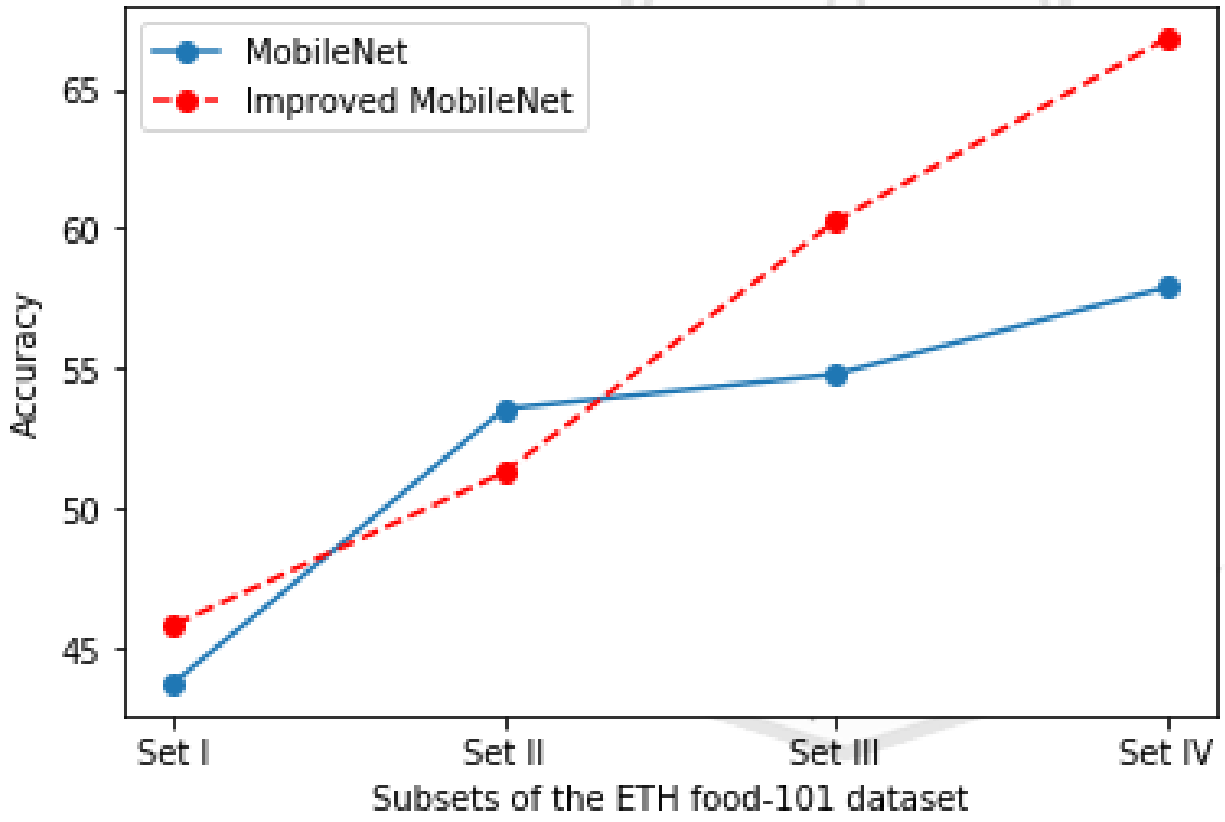
  - ✓ Stochastic gradient descent (SGD) solver
  - ✓ Batch size of 16
  - ✓ Learning rate at 0.0001
  - ✓ TensorFlow platform
  - ✓ Intel® Core™ i7-4790 CPU
  - ✓ 8GB RAM



EXPERIMENTAL & RESULTS

Experimental Results

The performance of the MobileNet and improved MobileNet architectures versus the different number of training samples (Set I - Set IV) on the ETH food-101 dataset.



EXPERIMENTAL & RESULTS

Experimental Results

The performance results of food image classification between MobileNet and improved MobileNet architectures with Data Augmentation.

Method	The number of image per class	Accuracy (%)
Data Augmentation + MobileNet	400	57.90
Data Augmentation + improved MobileNet	400	69.86

EXPERIMENTAL & RESULTS

Experimental Results

The performance results of food image classification on four subsets on ETH Food-101 dataset using the proposed MobileNet architecture.

Methods	Subsets of the EHT Food-101 dataset			
	I	II	III	IV
Without data augmentation	45.84	51.29	60.26	66.78
Random cropping	45.79	55.82	59.52	67.44
With data augmentation	48.71	56.71	62.49	69.86
With data augmentation + random cropping	51.39	59.68	65.97	**72.59**

EXPERIMENTAL & RESULTS

Experimental Results

The performance results of food image classification between the proposed MobileNet architecture and other previous methods

Method	The number of image per class	Accuracy (%)
Random Forest Discriminative Components	1,000	50.76
Supervised Extreme Learning Committee	1,000	55.89
Data Augmentation + Inception V3	1,000	70.41
FoodNet: Ensemble Net	1,000	72.10
Our proposed (Data Augmentation + MobileNet)	**400**	**72.59**

EXPERIMENTAL & RESULTS

Experimental Results

The performance results of food image classification between the proposed MobileNet architecture and other previous methods

Method	The number of image per class	Accuracy (%)
DeepFood	1,000	77.00
Our proposed (Data Augmentation + MobileNet)	**400**	**72.59**

CONCLUSION

✓The proposed MobileNet architecture is the best performance than the MobileNet architecture.

✓The data augmentation are impact to food image classification.

✓The best performance achieved an accuracy of 72.59% when the combination of the various data augmentation techniques and the proposed MobileNet architecture.

CONCLUSION

In future work

- We are interested in extracting the feature vector from the convolutional layers which may work better than individual deep CNN architecture.
- We plan to construct the deep ensemble convolutional neural network (CNN) architectures, which is the combination of the state-of-the-art deep CNN architectures.

MAHASARAKHAM
UNIVERSITY

THANK YOU FOR YOUR ATTENTION