FUJITSU

shaping tomorrow with you

NAS-EOD: an end-to-end Neural Architecture Search method for Efficient Object Detection

Huigang Zhang, Liuan Wang, Jun Sun, Li Sun

Fujitsu R&D Center, Beijing, China

Hiromichi Kobashi, Nobutaka Imamura

Fujitsu Laboratories Ltd., Kawasaki, Japan

Outline

FUJITSU

Introduction

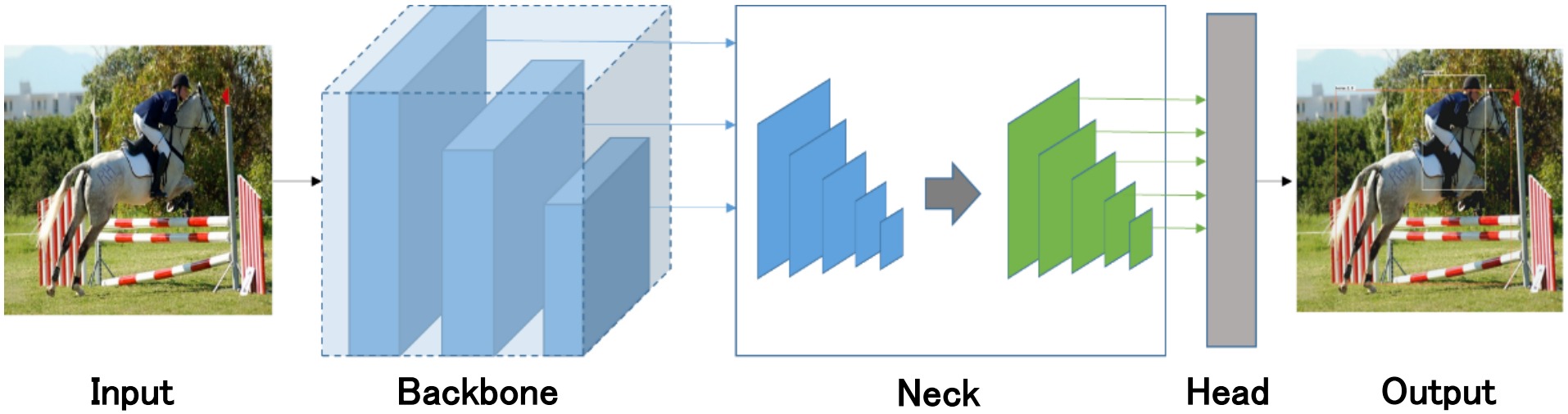
Proposed method

- Search space
- Search strategy

Experiment results

Introduction

Detection framework



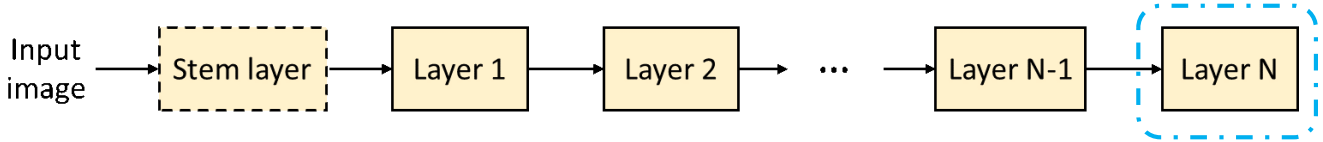
Our main contributions:

- We propose an end-to-end NAS method for detection task. The backbone and neck can be updated at the same time, which ensures an overall good performance.
- The search strategy can handle multi-task problems, it can balance accuracy and latency in the process of finding a better model.
- The discovered architecture (NAS-EOD) can be deployed on edge devices, and outperform the state-of-the-art models on the benchmark dataset.

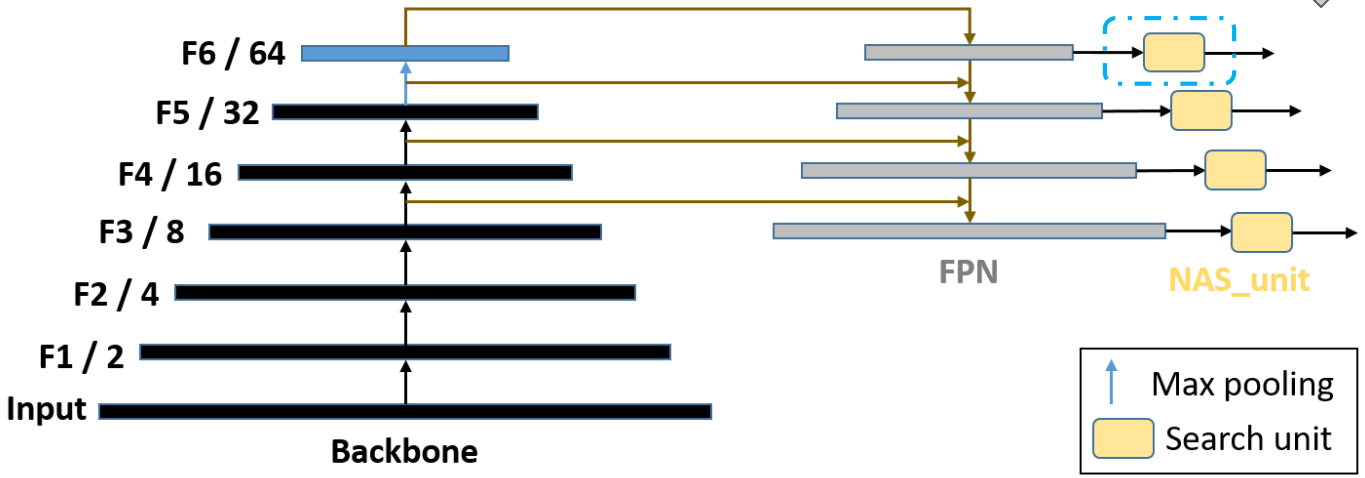
Search space

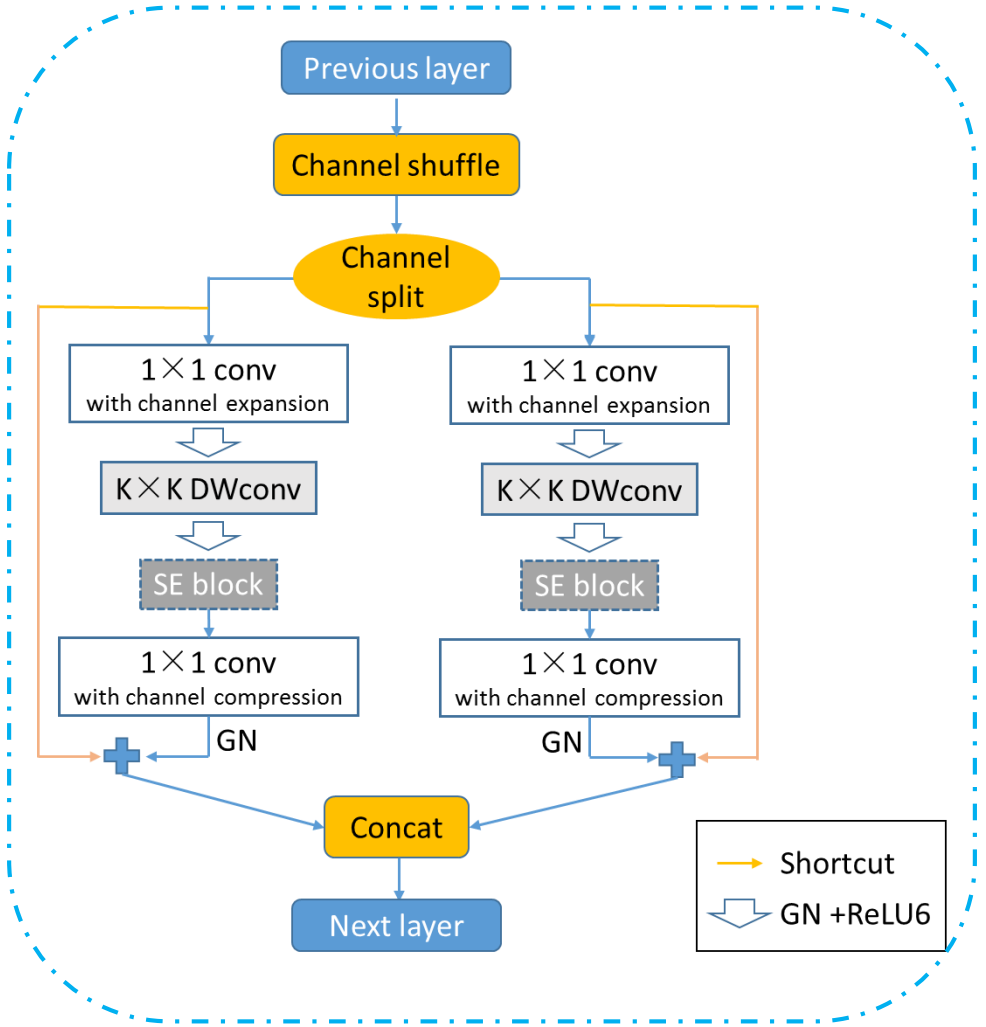
FUJITSU

Backbone:



Neck:





Search options:

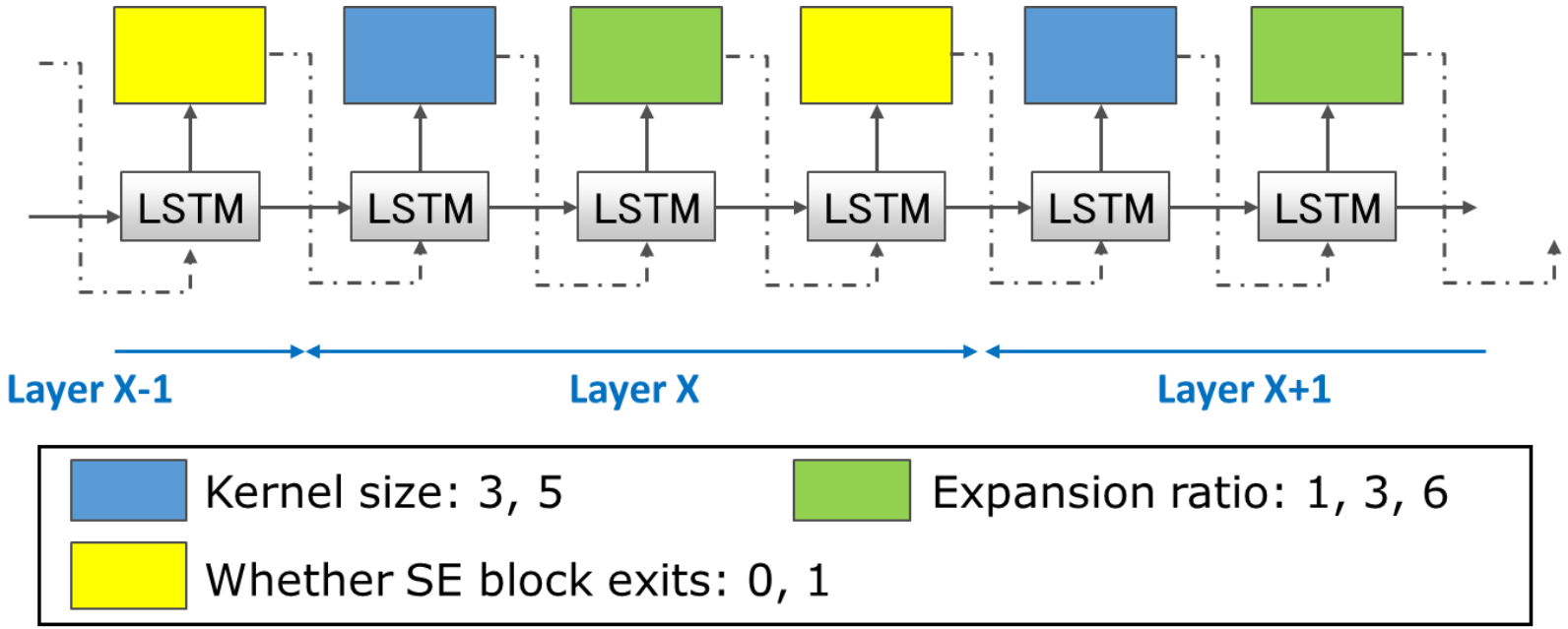
- The depth-wise convolution kernel size (K=3 or 5)
- Expansion ratio of the residual unit (1, 3, or 6)
- SE block* exists or not in each layer

* Jie Hu, Li Shen, and Gang Sun. “Squeeze-and-excitation networks”. In CVPR2018.

Search strategy

FUJITSU

- RNN design



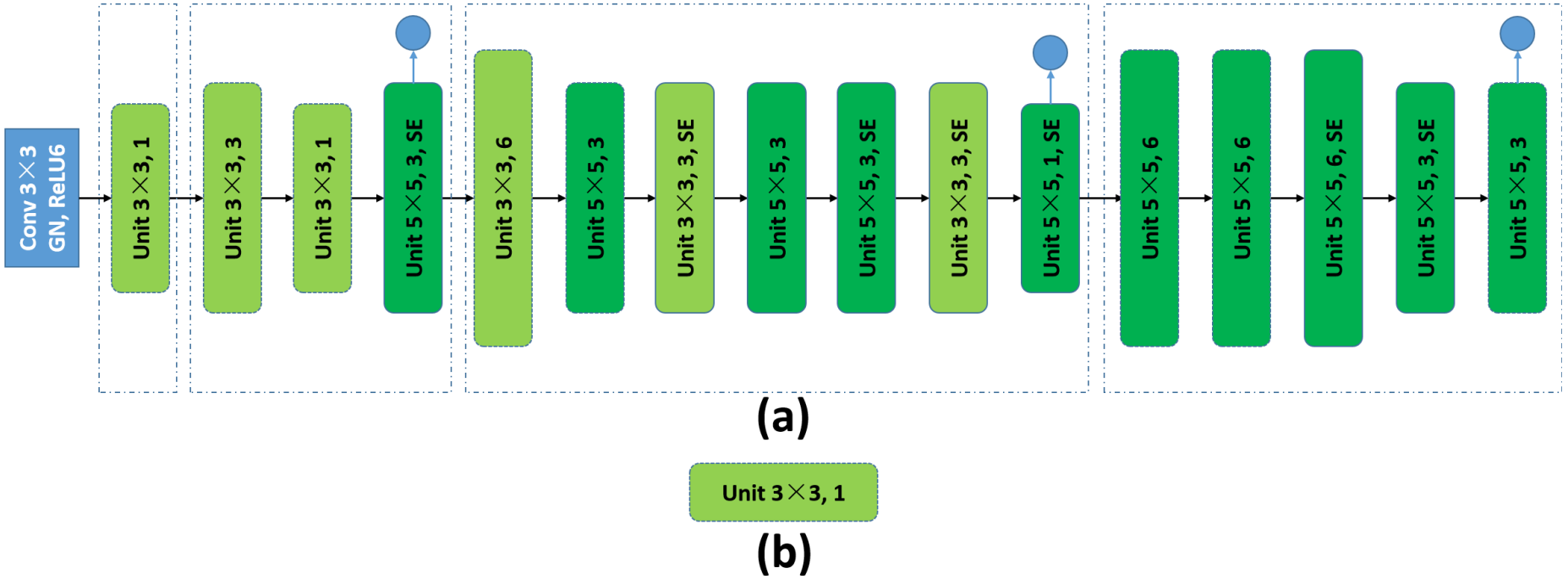
- Multi-objective loss

$$Loss(m) = \lambda_c Loss_c(m) + \lambda_r Loss_r(m) + Lat(m)$$

Experiment results

FUJITSU

- Architecture search result



- Some findings:
  - Large size kernels tend to be at later stages of the architecture.
  - SE blocks are chosen at the last layers of each stage.

Experiment results

FUJITSU

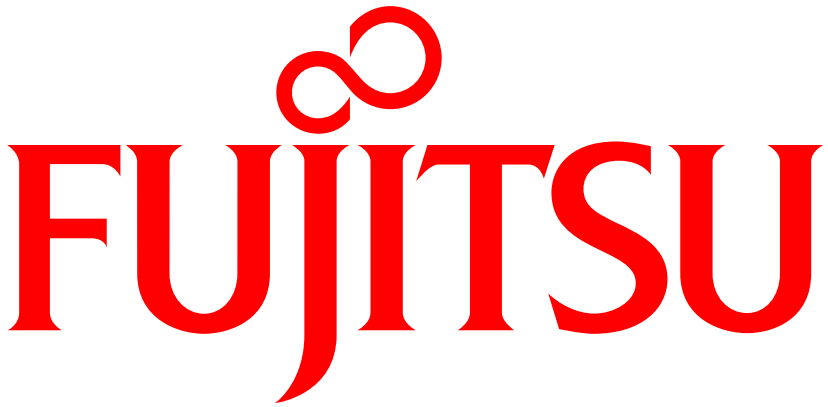
Comparison with State-of-the-arts

Comparisons to Training From-scratch Models

Model	Image size	# Params (M)	FLOPs (B)	mAP (%)
Tiny YOLOv2	416 × 416	15.9	7.0	57.1
Tiny YOLOv3	416 × 416	8.7	5.6	61.3
NAS-EOD @64	320 × 320	6.0	1.2	65.4
NAS-EOD @128	320 × 320	6.2	1.8	68.5

Comparisons to ImageNet Pre-trained Model

Model	Image size	# Params (M)	FLOPs (B)	mAP (%)
MobileNetV2-SSDLite	320 × 320	4.3	1.6	64.0
MobileNetV2-FPNLite @64	320 × 320	2.0	1.5	66.4
MobileNetV2-FPNLite @128	320 × 320	2.2	2.0	68.9
NAS-EOD @64	320 × 320	6.0	1.2	68.5
NAS-EOD @128	320 × 320	6.2	1.8	72.0



shaping tomorrow with you