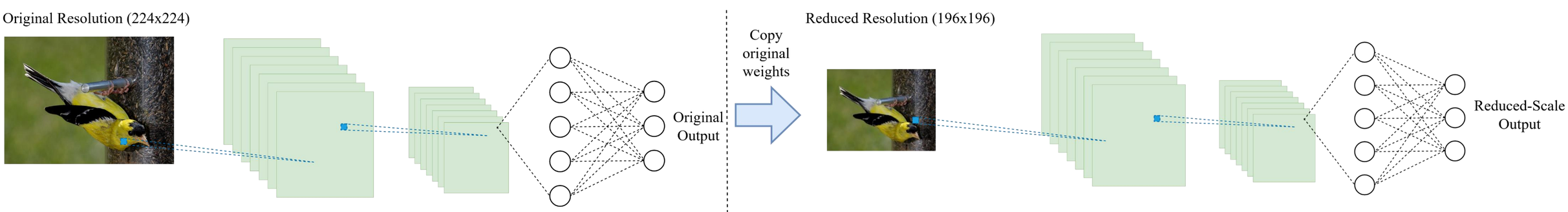
Projeto de Formatura – 2023

PCS - Departamento de Engenharia de Computação e Sistemas Digitais

Engenharia Elétrica – Ênfase Computação

Tema:

Resolution-wise Convolutional Neural Networks for Image Classification



Introduction

- Convolutional Neural Networks (CNNs) excel in computer vision tasks, especially image classification
- Modern CNN models and higher image resolutions increase computational costs, affecting portability and resource requirements
- In this work we explore how resolution affects CNNs, balancing between accuracy and increased memory and processing needs
- Our goal is to minimize network parameters, enhancing suitability for resource-constrained devices such as mobile phones.

Problem Statement

- **Research Questions**
  1. Is it possible to reduce the resolution of a trained Neural Network model and keep its predictive capacity?
  2. Is it possible to systematically encounter a resolution for which the model keeps its predictive ability?
- Answering these questions enables us to reduce the computational cost of Convolutional Neural Network-based solutions by simply reducing the resolution of pre-trained models

Experiments

- We use two well-known benchmarks, CIFAR-10 and ImageNet, and popular CNN architectures: NasNet, ResNet, and MobileNet
- **Main Experiments**
  1. Reduce the resolution of pre-trained CNNs and evaluate the floating points operations (FLOPs) and accuracy compared to the original model
  2. Select the resolution randomly from a set of pre-defined resolutions, and evaluate the FLOPs and accuracy compared to the original model

Results

- Resolution Reduction
  - NasNet on CIFAR-10

Resolution	FLOPs Drop (%)	Accuracy Drop (%)
28 x 28	23.43	1.92
24 x 24	43.74	4.34
20 x 20	60.93	13.32
16 x 16	74.99	29.58
12 x 12	85.93	52.40

- ResNet on ImageNet: Comparison with existing works

Resolution	FLOPs Drop (%)	Accuracy Drop (%)
He et al. [1]	20.00	1.70
Wang et al. [2]	20.00	2.00
He et al. [3]	20.00	1.40
Ours (210 x 210)	3.90	1.22
Ours (196 x 196)	14.72	2.90
Ours (182 x 182)	29.26	2.70
Ours (168 x 168)	38.49	5.13
Ours (154 x 154)	49.50	6.57
Ours (140 x 140)	57.26	10.72
Ours (Random Resolution)	17.30	2.42

Conclusions

- The experiments performed on ImageNet dataset suggest that **it is possible to reduce the CNN resolution with minor reduction on its predictive capability**
- These experiments achieved **significant reductions in FLOPs with a minor drop in accuracy**
- Random selection of model resolutions in experiments showed FLOPs reduction comparable to other methods in literature, with minimal accuracy loss
- There is potential for developing a more refined method for selecting the scale of Neural Networks

References

[1] HE, Y.; ZHANG, X.; SUN, J. Channel pruning for accelerating very deep neural networks. In: IEEE international conference on computer vision. [S.l.: s.n.], 2017

[2] WANG, X. et al. Skipnet: Learning dynamic routing in convolutional networks. In: European Conference on Computer Vision (ECCV) [S.l.: s.n.], 2018

[3] HE, Y. et al. Amc: Automl for model compression and acceleration on mobile devices. In: European Conference on Computer Vision (ECCV) [S.l.: s.n.], 2018.

Aluno: Jônatas de Souza Nascimento

Professor(a) Orientador(a): Prof. Dr. Artur Jordão