IEEE UEMCON 2021
Joint Disaster Classification and Victim Detection using Multi-Task Learning

Tham Mau Luen

Introduction

Background

- Every year, natural disasters generate substantial amount of:
  - damages
  - monetary costs
  - injuries and deaths
- The first 72 hours after a disaster are critical for rescuing survivors [1].
- Therefore, a disaster response system plays a vital role in facilitating search and rescue efforts.

ASEAN IVO

Introduction

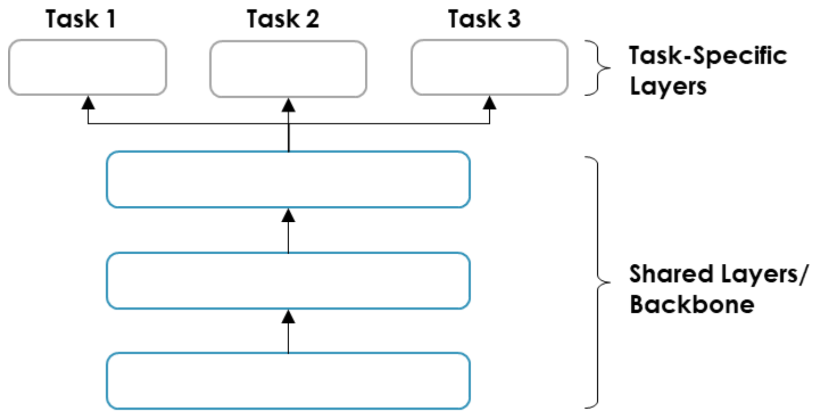
<u>Problem Statement</u>

- Emergency response unit can dispatch manpower more efficiently based on:
  - types of disasters
  - the reported number of victims
- Video analytics is regarded as one of the most promising candidates for detecting disaster [2].
- Modern detection methods automatically learn high-level feature through a Convolutional Neural Network (CNN), which is a form of deep learning.
- We aim to propose a unified multi-task CNN model that performs:
  - disaster classification
  - victim detection

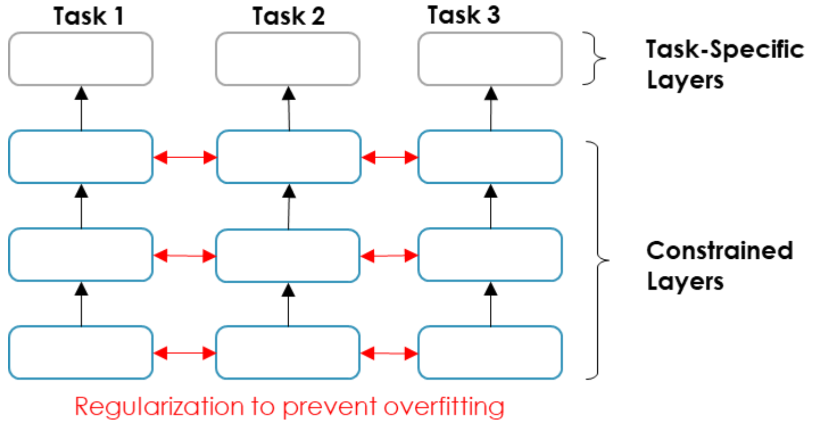
Literature Review

Multi-Task Learning (MTL)

- MTL is to perform more tasks using one model, without the need of using separate model for each task.
- Generally, MTL can be categorized into two classes.
- Hard parameter sharing
  - the most widely used approach to MTL [3]
  - multiple layers (backbone) are shared for all tasks
- Soft parameter sharing
  - each task has its own backbone, where the parameters of each backbone are regularized to encourage them to be similar



(a) Hard Parameter Sharing



(b) Soft Parameter Sharing

Literature Review

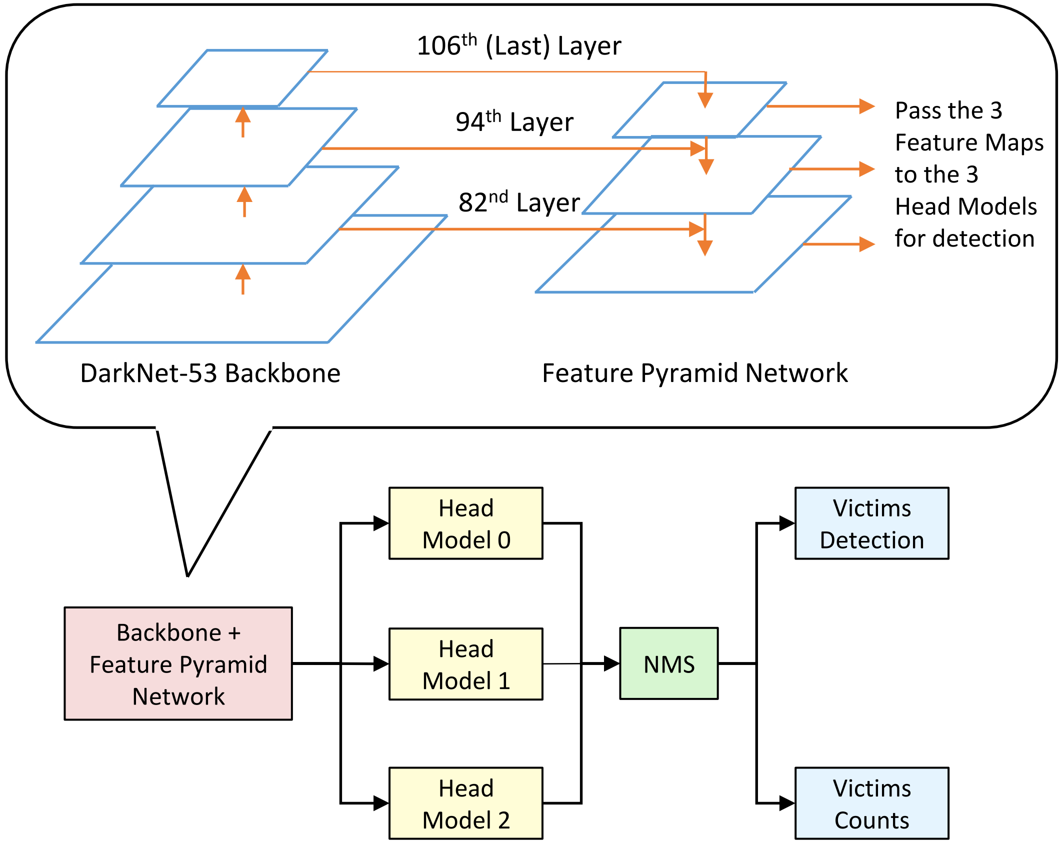
Disaster Detection

- "Crisis Image Benchmarks Dataset" is a dataset that consolidates several disaster-related datasets for benchmarking purpose [4], which is labeled for:
  - disaster types
  - informativeness
  - humanitarian
  - informativeness
- A multi-task classification model has been developed for the four classification tasks in "Crisis Image Benchmarks Dataset" [5].
- However, the solution is limited to classification tasks, without considering the additional requirement to locate instances in an image.

Proposed Solution

Victim Detection Model

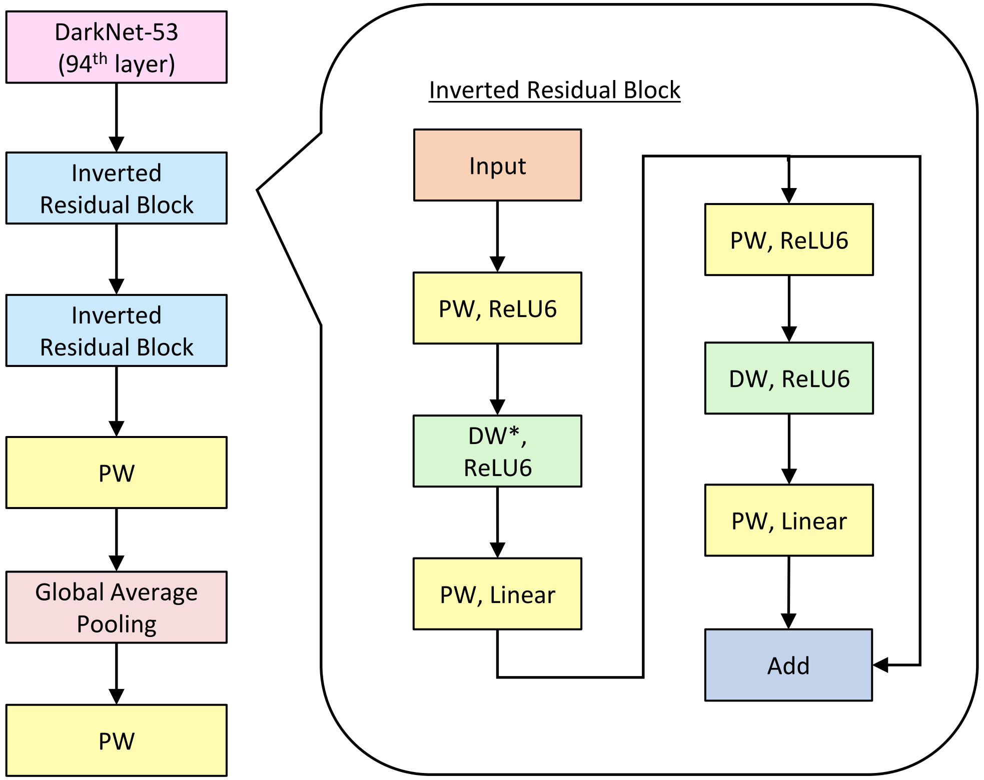
- YOLOv3 is selected because of its high speed and accuracy.
- YOLOv3
  = DarkNet-53 Backbone
  + Feature Pyramid Network
  + 3 Head Models (Decoders)
- In addition to bounding boxes, a victim counts will be returned as a Tensor.



Proposed Solution

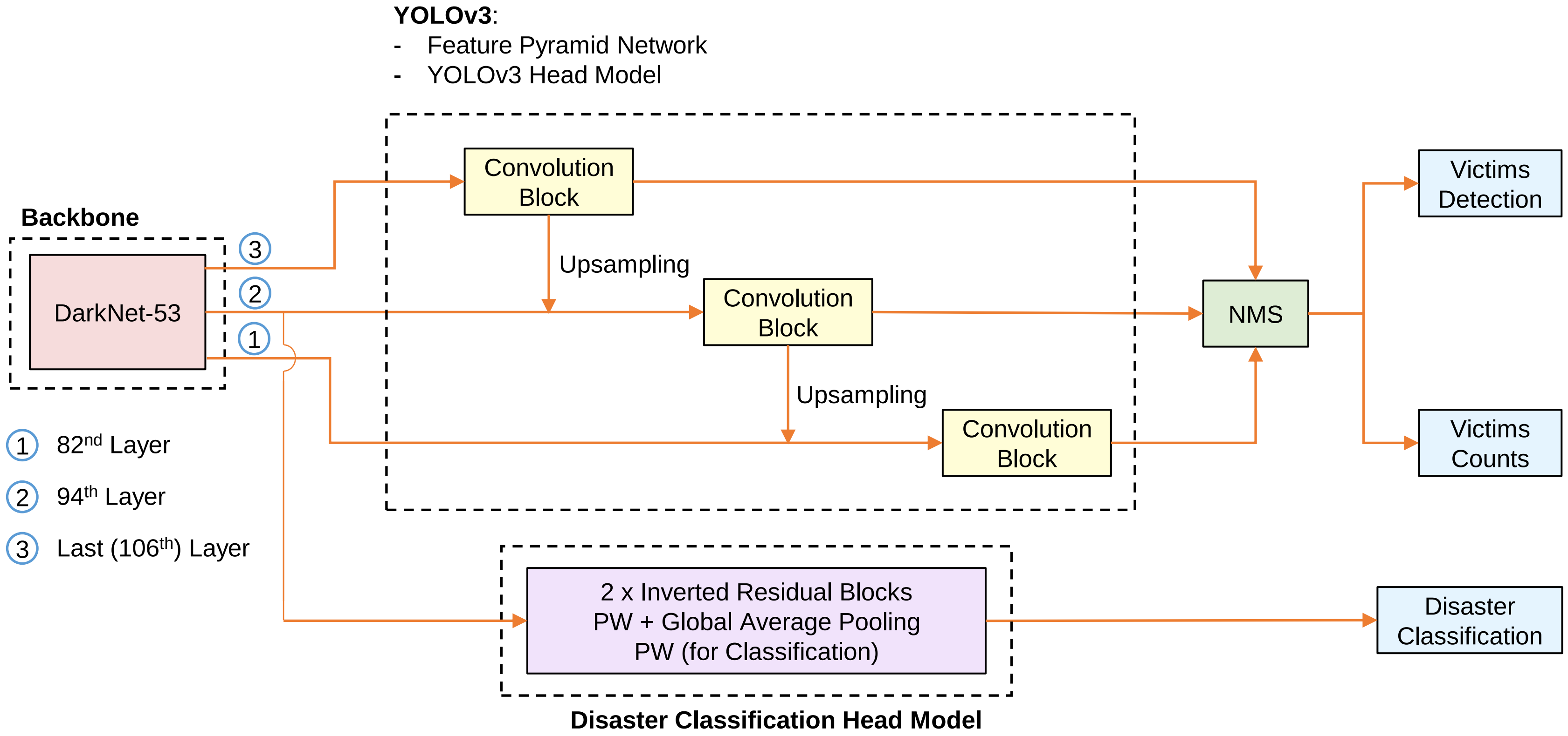
Disaster Classification Head Model

- Adopts the MobileNetv2 architecture.
- Consists of multiple depthwise-separable convolution layers.
- One concern: where should the head model be attached to the DarkNet-53.
- Attached to the 2nd output (94th layer) of the DarkNet-53 backbone (instead of 1st or last output layer).



All convolution layers have stride=1, except * (where stride=2 is used)

Proposed Solution



Datasets

Disaster Classification Dataset

- The images used for training are extracted from Crisis Image Benchmark Dataset [4].
- The sub-dataset for disaster classification task is used as the training images.
- There are seven labels in the dataset, as shown in the Table on the right.

Data Split for Disaster Types Dataset

Class Labels	Train	Validation	Test
Fire	1270	121	280
Hurricane	1444	175	352
Flood	2336	266	599
Earthquake	2058	207	404
Landslide	940	123	268
Other Disaster	1132	143	302
Not Disaster	3666	435	990
Total	12846	1470	3195

Datasets

Victim Detection Dataset

- There is a lack of publicly available victim detection dataset.
- Thus, the sub-dataset from Crisis Image Benchmark [4] for disaster classification is annotated for victim detection.
- In total, 8076 images are annotated for victim detection dataset.

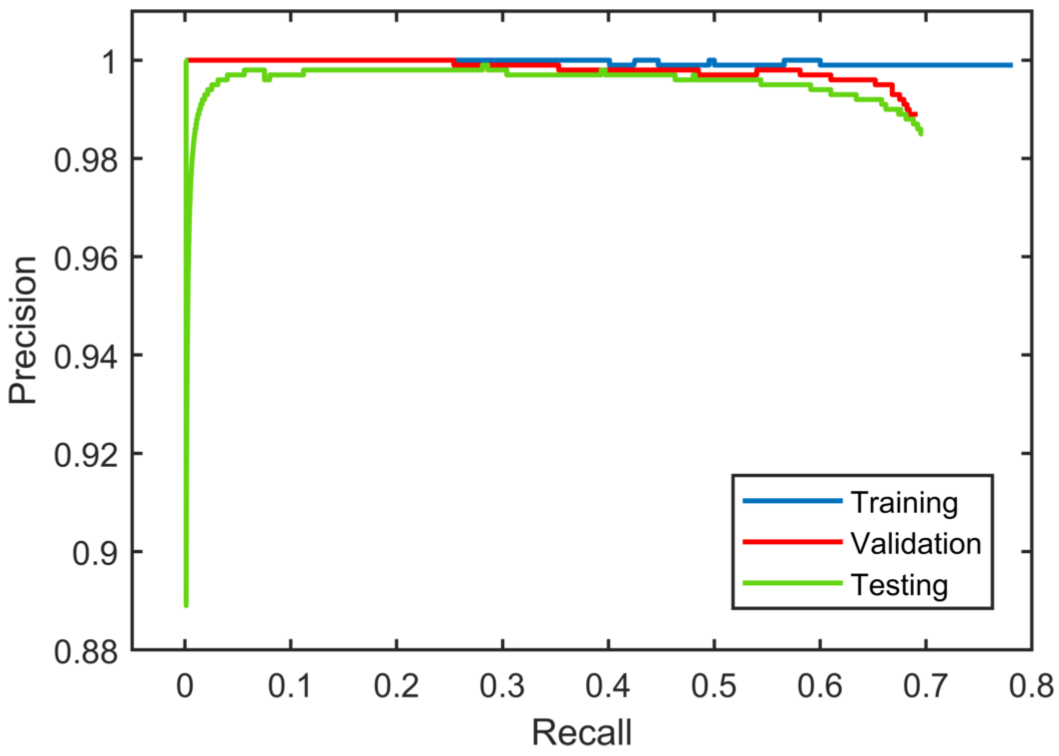
Data Split for Victim Detection Dataset

Class Labels	Count
Training	5994
Validation	634
Testing	1448
Total	8076

Results

Victim Detection

- For victim detection task, average precision (AP) is used as the evaluation metric.
- AP is equivalent to the area under the Precision-Recall curves.
- The train, validation and test AP of the model are 0.7814, 0.6907 and 0.6938, respectively.
- The average frame per second on NVIDIA GeForce RTX 2070 SUPER is 20.31 (suitable for real-time disaster monitoring)



Results

Disaster Classification

- Interestingly, the proposed solution is comparable with most CNN models that are specifically trained for disaster classification.
- The ability to detect victim on top of disaster classification comes at the cost of only 2% accuracy performance loss.
- Our model has the second highest precision and approximate the best model with 0.7 % gap.
- With regards to the recall and F1 score, the proposed method performs slightly worse than other models.

Backbone	Accuracy	Precision	Recall	F1 Score
ResNet18	0.812	0.807	0.809	0.809
ResNet50	0.817	0.810	0.812	0.812
ResNet101	0.819	0.815	0.816	0.816
AlexNet	0.755	0.753	0.753	0.753
VGG16	0.803	0.797	0.798	0.798
DenseNet (121)	0.817	0.811	0.813	0.813
SqueezeNet	0.726	0.719	0.717	0.717
InceptionNet (v2)	0.808	0.801	0.802	0.802
MobileNet (v2)	0.793	0.788	0.793	0.789
EfficientNet (b1)	**0.838**	**0.834**	**0.838**	**0.835**
Proposed Solution	0.792	0.827	0.769	0.766

Results

Examples

(a) Victim Detection in Flooded Area.

(b) Victim Detection in Landslide Area.

(c) Victim Detection in Earthquake Area.

Q&A

VISIT US HERE

UTAR
UNIVERSITI TUNKU ABDUL RAHMAN

IPSR Institute of Postgraduate Studies and Research
Intellectual Pursuits Through Sustainable Research

utar.edu.my

ipsr.utar.edu.my

research.utar.edu.my

References

- [1] United Nations Office for the Coordination of Humanitarian Affair, "Five essentials for the first 72 hours of disaster response," OCHA, 2017. https://www.unocha.org/story/five-essentials-first-72-hours-disaster-response (accessed Aug. 20, 2021).
- [2] S. A. Shah, D. Z. Seker, S. Hameed and D. Draheim, "The Rising Role of Big Data Analytics and IoT in Disaster Management: Recent Advances, Taxonomy and Prospects," in IEEE Access, vol. 7, pp. 54595-54614, 2019, doi: 10.1109/ACCESS.2019.2913340.
- [3] S. Ruder, "An Overview of Multi-Task Learning in Deep Neural Networks," CoRR, vol. abs/1706.0, 2017, [Online]. Available: http://arxiv.org/abs/1706.05098.
- [4] F. Alam, F. Ofli, M. Imran, T. Alam, and U. Qazi, "Deep Learning Benchmarks and Datasets for Social Media Image Classification for Disaster Response," in 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2020, pp. 151–158, doi: 10.1109/ASONAM49781.2020.9381294.
- [5] F. Alam, T. Alam, F. Ofli, and M. Imran, "Social Media Images Classification Models for Real-time Disaster Response," CoRR, vol. abs/2104.0, 2021, [Online]. Available: https://arxiv.org/abs/2104.04184.