An Efficient And Scalable Deep Learning Approach for Road Damage Detection

Sadra Naddaf, Mahdi Naddaf, Amir Kashani, Hassan Zargazadeh

RICS Lab, Lamar University
A.I. Lab, Stanley Black & Decker

2020 IEEE International Big Data Conference

Content

- Background
- Dataset
- Method
- Evaluation Metrics
- Augmentation
- Training
- results
- Conclusions

Background

- 16% of traffic crashes are result of poor pavement conditions according to NHTSA*.
- $25 billion per year, estimated costs of pavement maintenance. according to a study FHA in 1997**.
- Developing low cost and automated maintenance techniques is necessary.

- * National motor vehicle crash causation survey: Report to congress ," National Highway Traffic Safety Administration Technical Report DOT HS, vol. 811, p. 059, 2008
- ** summary of shrp research and economic benefits of pavement maintenance," Federal Highway Administration, Tech. Rep., 1997.

Dataset

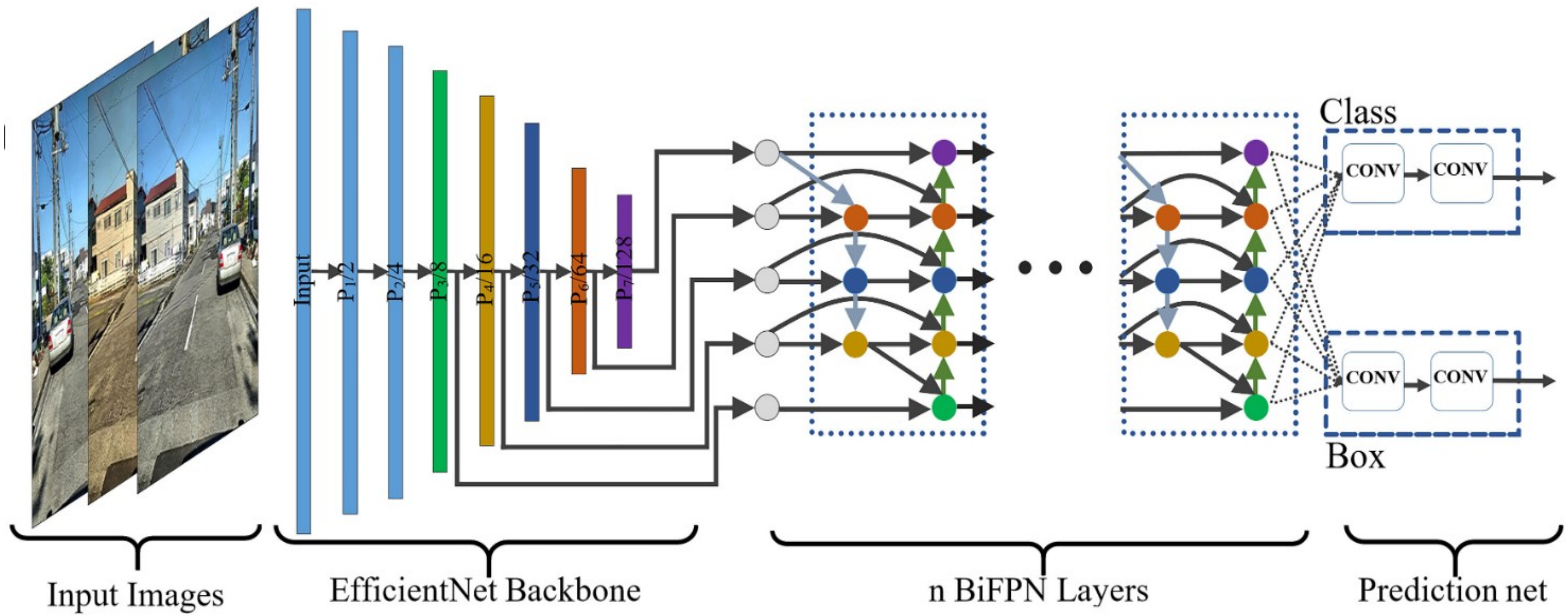
- Images from 3 different countries with 4 classes of damage types.
- Images gathered using a smartphone installed on dashboard of a car.

	Total	**Categories**			
		D00	***D10***	***D20***	***D40***
Train	18930	5918	4014	7535	5103
Validation	2111	674	432	846	524
Test1	2631	—[a]	—	—	—
Test2	2664	—	—	—	—
Total	26336	6592	4446	8381	5627

[a] Testset annotations were not released.

Method

- EfficientDet
- R_{input}= 512 +φ·128 W_{bifpn}= 64·(1.35φ) D_{bifpn}= 3 +φ D_{box}=D_{class}= 3 +[φ/3]



Evaluation Metrics

- F1 per Competition rules

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \qquad precision = \frac{tp}{tp + fp}; recall = \frac{tp}{tp + fn}$$

- AP metrics :
  - recall and precision of a robust object detector are not altering much, with varying confidence.
  - AP_{50}, AP_{75}, AP_s(objects with area<32^2) AP_m, AP_l (objects with area>96^2) and mAP.

Augmentation

- Transferred augmentation strategies learned through AutoAugment*.
- Trained A D0 network using 25% of the Training Set, for 150 epochs.
- Trained Networks using most-effective policy(policyV1)

policy Name	***Base (% of Improvement)***			
	AP_{50}	AP_s	AP_m	AP_l
NoAugment	31.7	3.8	7.3	12.7
policyV0[a]	31.9(**+0.6**)	7.0(**+84.2**)	7.8(**+6.8**)	11.9(**-6.2**)
policyV1[a]	34.1(**+7.5**)	5.9(**+55.2**)	8.2(**+12.3**)	12.9(**+1.5**)
policyV2[a]	33.0(**+4.1**)	6.4(**+68.4**)	8.6(**+17.8**)	12.2(**-4**)
policyV3[a]	33.4(**+5.3**)	5.1(**+34.2**)	7.7(**+5.4**)	13.4(**+5.5**)

[a] Rotation strategies are removed from these policies.

* “Learning data augmentation strategies for object detection” By Zoph et al.

Training

- Used K-means to find optimal anchor box ratios.
- For small EfficientDet Networks (D0-D4) Used 3 V100-16GB GPUs.
- Mixed precision Training and syncBN Normalization provided by apex.
- Exponential moving average used with weight decay of 0.9998
- initial weights are transferred from EffcientDet model trained on COCO dataset.
- Other parameters are similar to original EfficientDet Implementation.

Results

- Results

Model Name	*Input Image Resolution*	*Backbone Name*	*Test1 F1*	*Test2 F1*	*Validation*						
					AP	AP_{50}	AP_{75}	AP_s	AP_m	AP_l	$F1$
D0	512	B0	52.1	51.4	19.1	47.2	11.5	7.2	14.3	22.2	54.04
D0-AUG	512	B0	51.2	52.5	19.8	48.4	12.1	7.9	15.4	22.7	54.03
D1	640	B1	53.8	54.7	21.7	51.5	13.4	15.3	16.9	25.0	56.9
D1-AUG	640	B1	54.4	55.4	22.0	51.7	13.1	17.1	17.7	24.7	56.5
D2	768	B2	55.2	54.9	22.9	53.5	14.9	10.4	18.6	24.9	56.7
D2-AUG	768	B2	54.1	54.0	22.9	54.2	15.2	13.3	18.8	24.7	56.6
D3	896	B3	56.5	54.7	23.0	53.4	15.0	10.5	18.4	25.4	56.5
D3-AUG	896	B3	56.3	54.2	22.6	53.4	14.7	11.4	18.3	24.8	56.8
D4	1024	B4	54.53	54.6	22.8	53.3	15.1	15.5	18.1	25.7	57.2
D7-AUG	1536	B6	**56.5**	**54.9**	23.4	53.6	15.0	31.4	19.2	25.5	56.5

Results

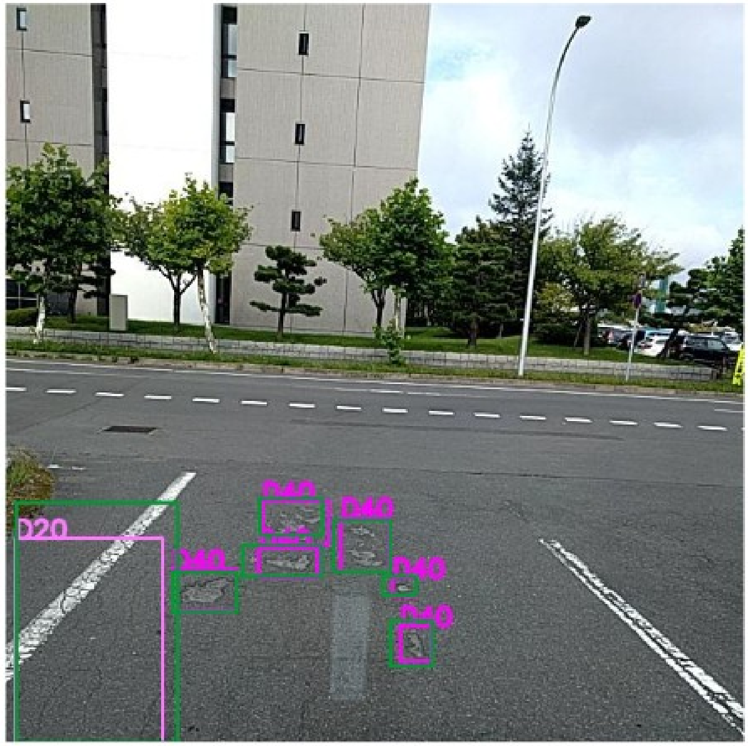
- Results
- Inference time
- Real-time
- Scalability

Model Name	*#Params*	*Batch size/ Learning rate*	*Inference Time(img/s) V100*		
			b^b=1	b^b=8	b^b=16
D0	3.8M	90/0.112	20	121	178
D1	6.5M	75/0.075	15	98	147
D2	8M	45/0.056	12	82	100
D3	11.9M	18/0.026	10	54	58
D4	20.5M	9/0.011	9	35	37
D7	51M	8/0.01	6	9	10

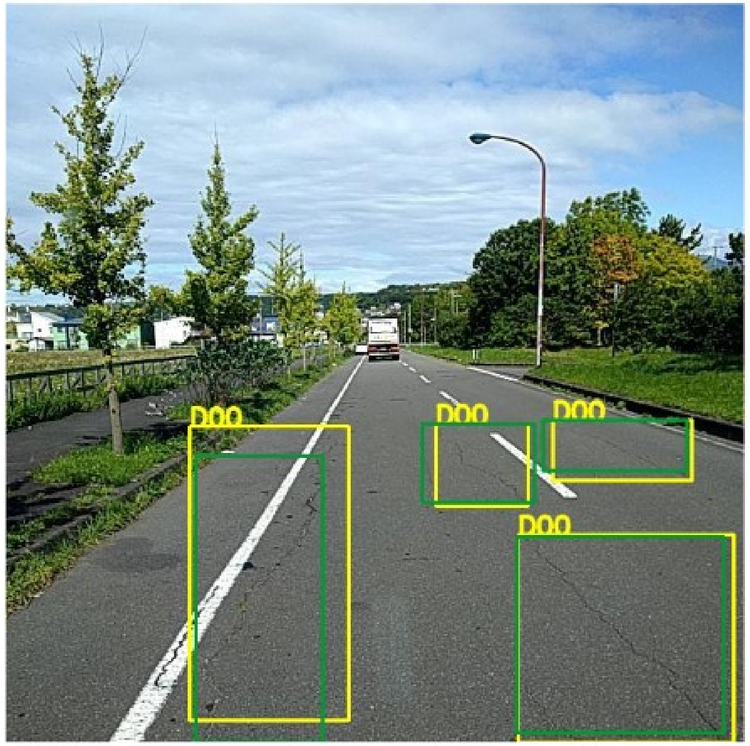
b denotes batch size for inference.

Results

- Discussion on results.



(1)

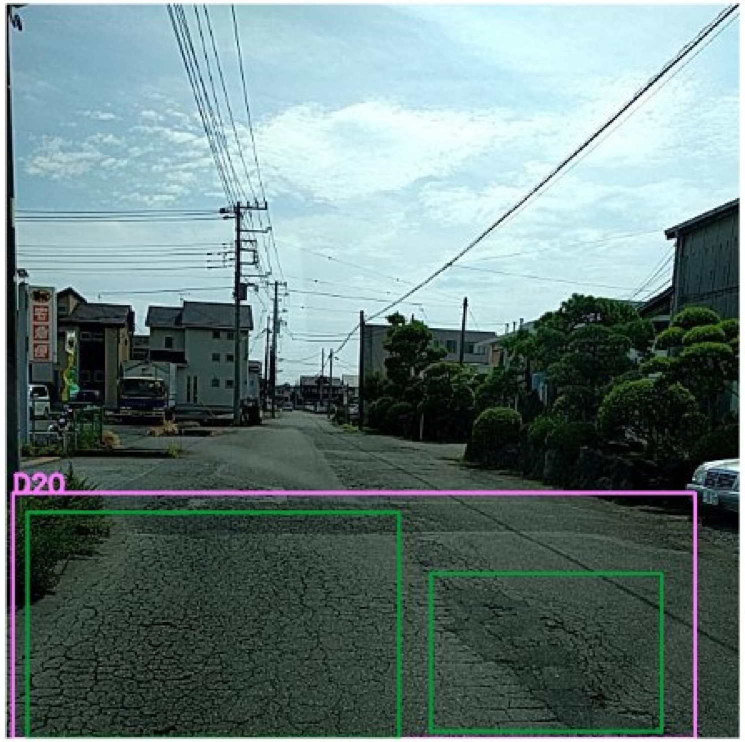


(2)

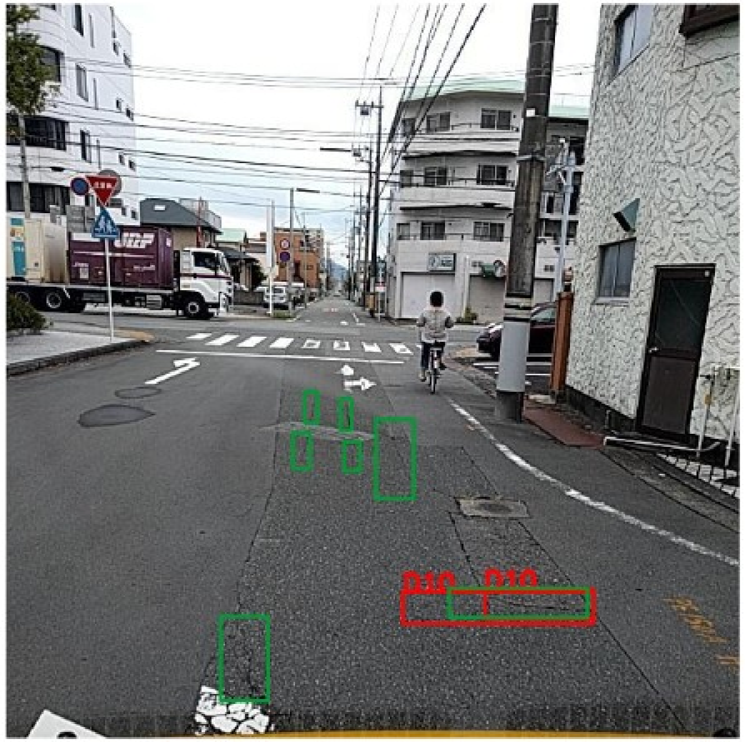
(3)

Results

(1)



(2)



(3)

Results

- Tested on local roads




Conclusions

- A family of efficient and scalable models for road damage detection is introduced.
- Model can be chosen based on with respect to hardware and time constraints.
- As time performance of models shows no ensemble or TTA is used.
- Transferred augmentation policies for other works.
- For Future It is suggested to revise dataset.
- Evaluate reliability of dataset with more Images from different countries.