Learning Image Statistics with CNNs?

Julian Biesheuvel, Supervisors: Tom Viering , Ziqi Wang , David Tax, Marco Loog Examiner: Klaus Hildebrandt

TU Delft, J.P.Biesheuvel@student.tudelft.nl,{T.J.Viering, Z.Wang-8, D.M.J.Tax, M.Loog}@tudelft.nl

CSE3000 | July 2nd 2021

1 BACKGROUND

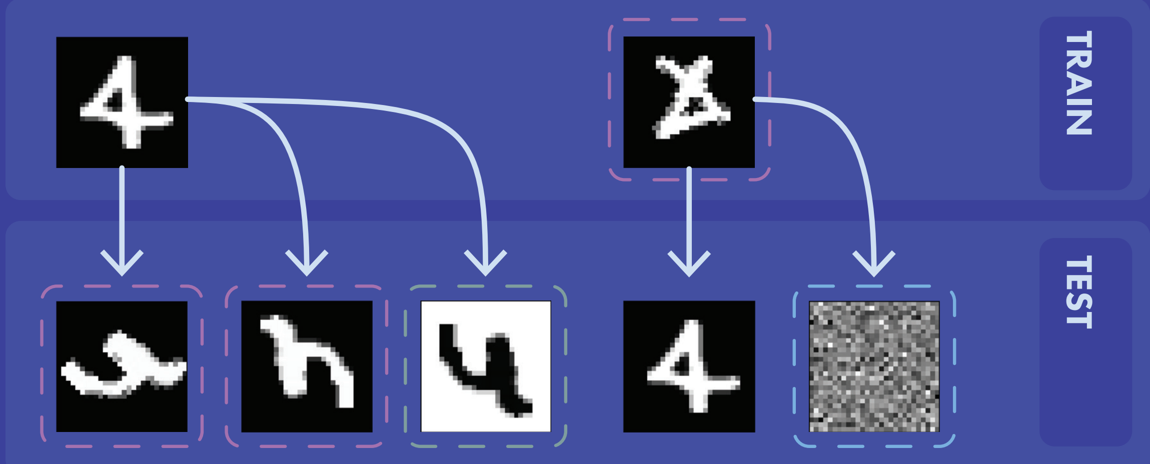
○ Deep neural networks are treated as omniscient.

- Improve our understanding of neural networks and what they can learn.

○ Domain shift is a recurring generalization problem in machine learning

- For example working with images: new lighting, angles, viewpoints, etc.

2 RESEARCH QUESTIONS & EXPERIMENTS



RESEARCH (SUB)QUESTION(S)

○ Are convolutional neural networks invariant to domain shifts, while their task is to learn to predict image statistics?

- What is the performance impact, if any, by domain shifts?

SYNTHETIC TASKS

○ Learn to predict the following image statistics:

- **Mean**, **median**, **standard deviation**, and **variance** pixel intensities.

BASELINES

○ The median value of targets in either the target domain or the source domain.

ROTATION

○ Experiment 1 & 2: Respectively 45 and 180 degrees rotated, **target** domain.

○ Experiment 3: Rotated between [0, 360] degrees, **source** domain.

- Does the order make a difference?

INVERSION

○ Experiment 4: Images are inverted to their negatives, **target** domain.

- Transformation: p′ = 255 - p
- What is the performance impact for color space transformation?

NOISY

○ Experiment 5: Source domain [0, 360] rotated images, Target domain noisy images.

- Does spatial structure, like shapes and objects, matter for these tasks?

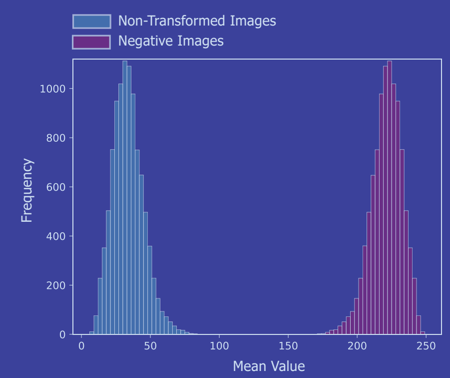
3 RESULTS

○ Experiments 1, 2 and 3:

- Median pixel intensity could not be learned, i.e. the test loss did not beat the baseline.
- No performance difference between experiment 1, 2, and 3. Order does not seem to matter here.

○ Experiment 4:

- All runs for the mean pixel intensity outperformed the baseline.
- Some runs for the median pixel intensity outperformed the baseline.
- What do these results tell us about the importance of spatial structure?



○ Experiment 5:

- Although the average test loss does not beat the baseline, in some runs a test loss lower than the baseline is registered.
- The network was less consistent in predicting the right image statistics.

Mean			Standard Deviation			Median			Variance		
$\hat{y}_{rw}$: 9.12			$\hat{y}_{rw}$: 9.83			$\hat{y}_{rw}$: 0.0			$\hat{y}_{rw}$: 1517.64		
Run	$\hat{y}_{train}$	#Epoch	Run	$\hat{y}_{train}$	#Epoch	Run	$\hat{y}_{train}$	#Epoch	Run	$\hat{y}_{train}$	#Epoch
1	12.61	40	1	10.34	32	1	1.39	37	1	813.64	26
2	14.53	40	2	19.21	40	2	0.93	39	2	2346.90	40
3	9.94	40	3	26.74	25	3	1.34	30	3	736.67	25
4	15.94	31	4	3.38	23	4	0.96	23	4	3106.85	40
5	6.92	20	5	4.75	40	5	1.29	15	5	792.86	26
6	6.93	19	6	3.65	16	6	2.21	15	6	2595.30	40
7	14.30	30	7	17.36	31	7	1.37	32	7	3376.52	40
8	7.92	28	8	3.60	21	8	1.02	27	8	985.66	18
9	7.71	40	9	6.32	39	9	1.66	27	9	778.08	25
10	5.37	26	10	7.17	38	10	0.91	22	10	672.69	19

5 DISCUSSION & CONCLUSIONS

○ The network seems to be able to learn to predict the mean, standard deviation and variance pixel intensity.

○ In the cases the network outperformed the baseline in experiment 5, the results were not significant.

- Does not tell us something for other regression tasks.
- More experiments are needed to investigate the importance of spatial structure for regression tasks.

○ The performance of the network, in this setting, is correlated with the type of transformation that was applied to the images.

- Color space transformations have more impact than geometric transformations.

TU Delft