THE UNIVERSITY OF TENNESSEE KNOXVILLE

374. Effects of Degradation on Face Detectors

Vijaysrinivas Rajagopal
The University of Tennessee, Knoxville

Abstract

In this research project, **we look into some of the standard image-based obstacles faced by state-of-the-art face detectors in the real world, such as degraded camera quality and transmission interruptions.** With these obstacles, we experiment and quantitatively describe how the CNNs react in these scenarios. Additionally, **we reconstruct the “degraded” image with various image processing methods in order to recover detection accuracy.** Furthermore, we quantitatively establish connections between certain recovery methods, their intensity, as well as the intensity of the degradation with the ability to recover the detection accuracy.

Methodology and Background

Face detectors (chosen based accuracy on WIDERFACE):

- **TinaFace** - 92% (Hard-set) **[1]**
- **RetinaFace** - 91% (Hard-set) **[1]**
- **Dual-shot Face Detector (Optimized)** - 81% (Hard-set) **[2]**

The following degradations were chosen based on their occurrence in real-life:

Figure 1. Comparisons of different noises. Gaussian noise is seen in most degraded digital photos. Gamma and Poisson noises occur in X-Rays and CT-Scans. Finally, Salt & Pepper occurs in sharp pulses during image transmission. **[3]**

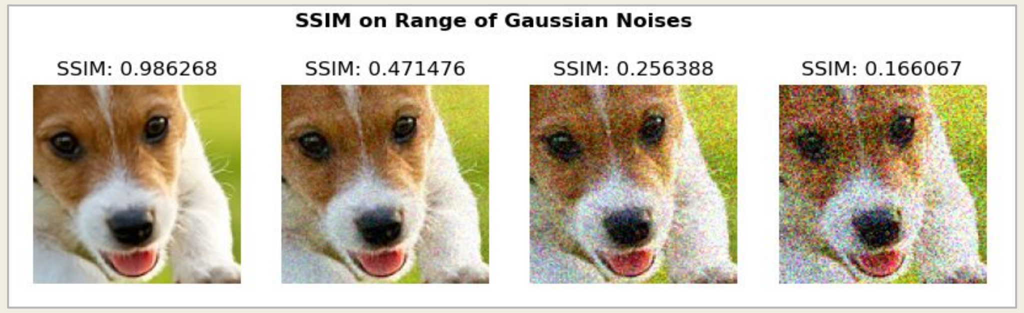

Figure 2. Demonstration of SSIM metric on increasing levels of gaussian noise. Note the SSIM value falling from 1 (same image) to 0 (recognizably different)

Correction Methods:

- **Median Filter** - Uses a *n* x *n* kernel, where n is a non-zero integer, and slides through image; smoothens out feature within the kernel.
- **YCbCr-Histogram Equalization** - Used on Gamma noise to balance overexposure. Applies histogram equalization on the luminance (Y) channel of the image after converting to YCbCr color space

Note: Both correction methods can have an optimized runtime of <10-20 ms. This makes them suitable for real-time corrections

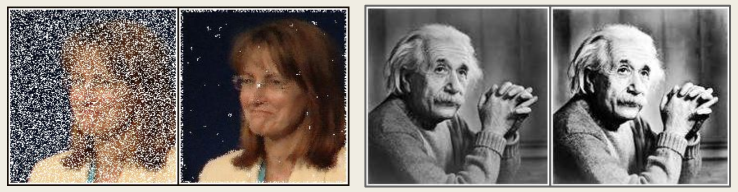

Figure 3. Examples of correction methods (median filter [4] on the right and histogram equalization [5] on the left)

Results

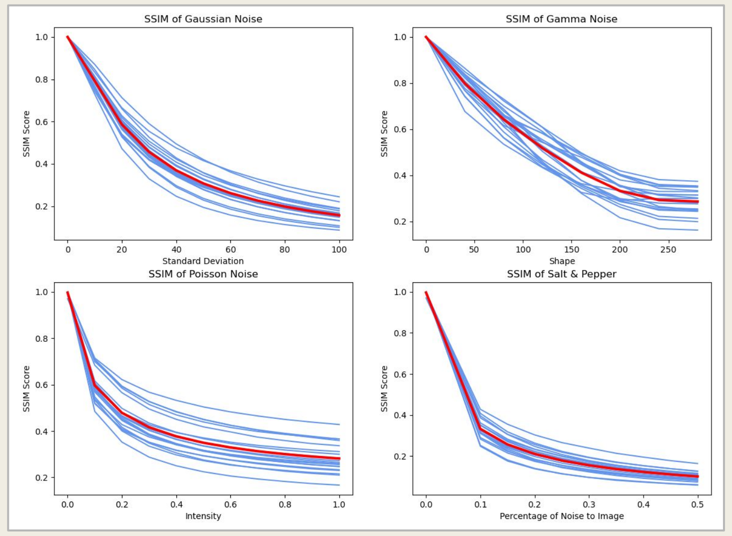

Figure 4. Comparisons of all image degradation methods against SSIM metric to determine the range of the chosen adjustable parameters. For Gaussian Noise, the standard deviation was the variable parameter. For Salt & Pepper noise, it was the percentage of the image covered in a given white or black pixel. For Poisson, it was the lambda parameter, and Gamma Noise utilized the “Gamma Shape”. This noise is unique because the parameter “washes” out the picture.

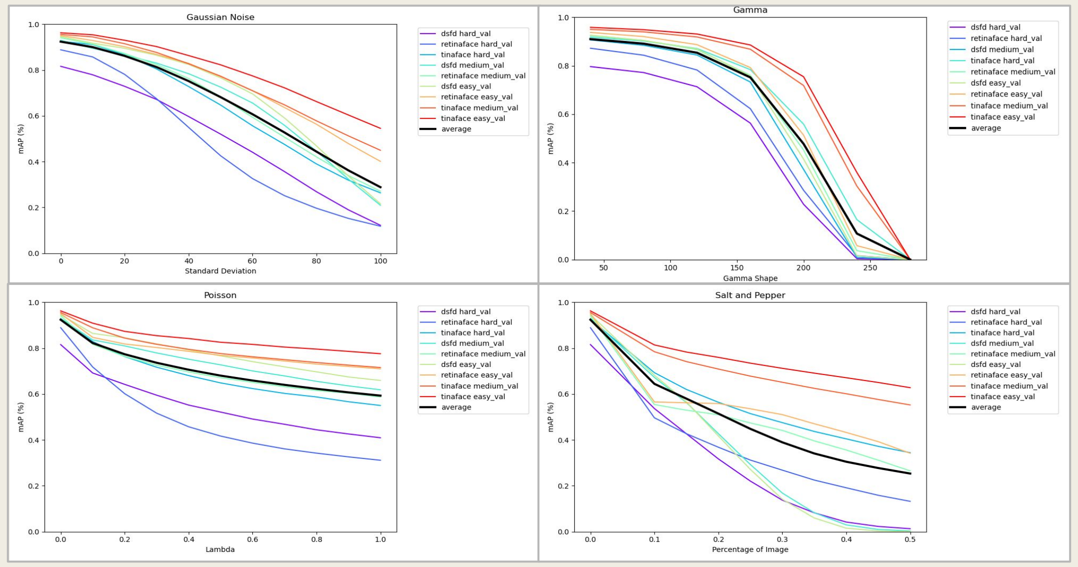

Figure 5. Shows the different models’ behavior through each of the noises at different noise-specific parameters. Each model’s performance is split further into the WIDERFACE categories: “easy”, “medium”,”hard”. These categories are determined by the area of the bounding box of a given face. All models experienced similar trend of accuracy decline, but TinaFace (SOTA method of WIDERFACE) consistently outperforms all other models.

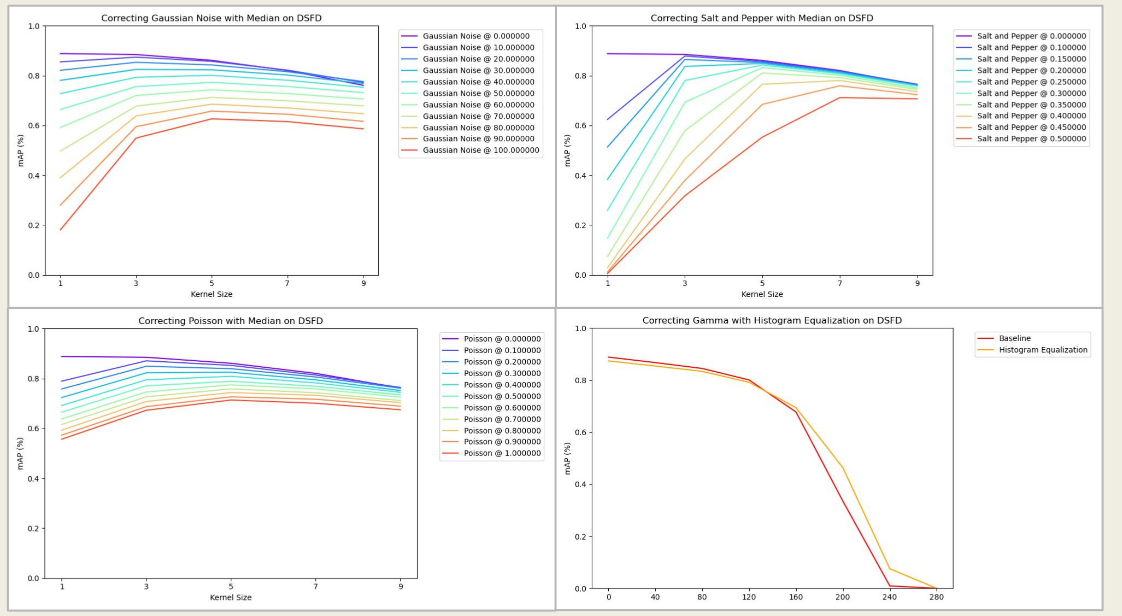

Figure 6. Graphs showing effects of the Median and Histogram Equalization corrections on the corresponding noises on DSFD (similar trends on the other models). Median Filter shows promising recovery results at a kernel size of 3 or 5. Histogram shows improvement only in the extreme noise ranges, but this is inconsistent with other models.

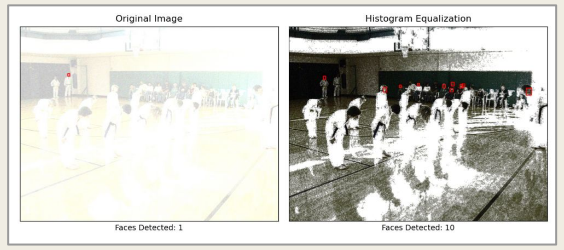

Figure 7. Example of Histogram Equalization being applied to Gamma Noise @ 240. It was effective at the higher levels of Gamma Noise. In this example, we improved detection from 1 face to 10 faces, but other examples were inconsistent in their ability to improve face detection rates.

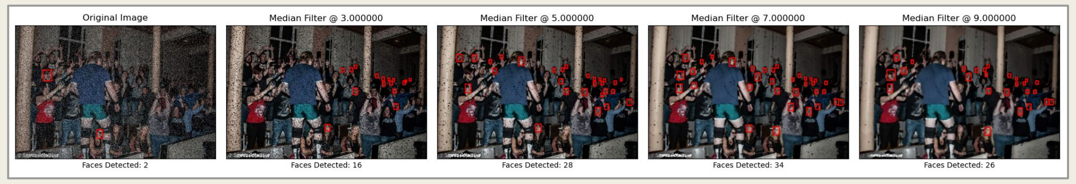

Figure 8. Example of median filter at different kernel sizes improving the detection rate of an image with Salt and Pepper @ 50. (4 faces to 34 faces at best Median Filter size)

Conclusions

- DSFD had the worst effect on all detectors
- TinaFace was most resilient against all noises (achieved highest mAP regardless of noise type and intensity)
- Median Filter worked effectively, but Histogram-Equalization not consistent
- When under noisy conditions, CNNS not able to infer the general shape and orientation of objects the same way humans can.
- Inconsistent relationship between SSIM and CNN degradation trends

Future Work

- For future, expand the noise types, correction types, face detectors
- Investigate any architecture changes that can help performance loss
- Adaptive correction methods

Citations

[1] - “Papers with Code - WIDER Face (Hard) Benchmark (Face Detection).” The Latest in Machine Learning, paperswithcode.com/sota/face-detection-on-wider-face-hard/.
[2] - Li, Jian, et al. “DSFD: Dual Shot Face Detector.” Http://Openaccess.thecvf.com/, 1 Jan. 1970, openaccess.thecvf.com/content_CVPR_2019/html/Li_DSFD_Dual_Shot_Face_Detector_CVPR_2019_paper.html.
[3] - Boyat, Ajay Kumar, and Brijendra Kumar Joshi. A Review Paper: Noise Models in Digital Image Processing. Signal & Image Processing : An International Journal, 2 Apr. 2015, arxiv.org/pdf/1505.03489.pdf.
[4] - By The original uploader was Anton at German Wikipedia. - Originally from de.wikipedia; description page is/was here., CC BY 2.5, https://commons.wikimedia.org/w/index.php?curid=2544834
[5] - “Histogram Equalization.” Tutorialspoint, www.tutorialspoint.com/dip/histogram_equalization.htm.

Acknowledgement

I would like to thank my advisor, Dr. Mongi Abidi, for his guidance on this project. Additionally, I would like to thank the authors of the networks that I used in this project for their contribution towards the face detection task.