Report 03

Yiming Bian

May 20th 2022

Major Professor

Arun Somani(arun@iastate.edu)

Committee Members

Henry Duwe (duwe@iastate.edu)
Aditya Ramamoorthy (adityar@iastate.edu)
Alexander Stoytchev (alexs@iastate.edu)
Cindy Yu (cindyyu@iastate.edu)

1 Introduction

In this report, I will present the results of two experiments: 1) compare the test accuracy of models trained on different levels of the same noise type 2) compare the test accuracy of models trained on different noise types. The detailed experiment set-ups and analysis are presented in the next section.

2 Experiment results and analysis

The basic set-ups of the following two experiments are:

- Network architecture: ResNet-18, VGG-16
- Noise type: Salt and pepper noise(SNP), Speckle noise(SPE)
- Noise level: For SNP, there are 5 noise levels. For SPE, there are 4 noise levels.
- Object category: vizsla, hamster, goldfinch, upright, pitcher

- Training set(*format*[**Index** *set name*(*# images for training per category*,*# images for validation per category*)]): **I** *Original*(*1040*,*130*), **II** *SNP_whole*(*5200*,*650*), **III** *SNP_0.1*(*1040*,*130*), **IV** *SNP_0.2*(*1040*,*130*), **V** *SNP_0.3*(*1040*,*130*), **VI** *SNP_0.4*(*1040*,*130*), **VII** *SNP_0.5*(*1040*,*130*), **VIII** *SPE_whole*(*4160*,*520*), **IX** *SPE_0.5*(*1040*,*130*), **X** *SPE_1.0*(*1040*,*130*), **XI** *SPE_5.0*(*1040*,*130*), **XII** *SPE_10.0*(*1040*,*130*)

- Testing set(*format*[**Index** *set name*(*# images for testing per category*)]): **I** *Original*(*130*), **II** *SNP_whole*(*650*), **III** *SNP_0.1*(*130*), **IV** *SNP_0.2*(*130*), **V** *SNP_0.3*(*130*), **VI** *SNP_0.4*(*130*), **VII** *SNP_0.5*(*130*), **VIII** *SPE_whole*(*520*), **IX** *SPE_0.5*(*130*), **X** *SPE_1.0*(*130*), **XI** *SPE_5.0*(*130*), **XII** *SPE_10.0*(*130*)

The baseline model of each network, **RtO_Res** and **RtO_VGG**, is trained on training set **I**, which contains $1,040$ images for training and 130 images for validation in each of five object categories.

2.1 Experiment 1: comparison among models trained on single level of noise data

The following models are trained only on a single level of noise data. For SNP noise, there are five re-trained models for each network architecture: **RtN_0.1**, **RtN_0.2**, **RtN_0.3**, **RtN_0.4**, and **RtN_0.5**. For SPE noise, there are four re-trained models for each network architecture: **RtN_0.5**, **RtN_1.0**, **RtN_5.0** and **RtN_10.0**. Each model is only trained on the data with corresponding noise level and all of these models are tested on the same type but all levels of noise data. Below we present the results regarding the network architectures.

2.1.1 ResNet-18

In Table 1, it shows the test accuracy on various SNP noise data sets of the baseline ResNet-18 model and five re-trained ResNet-18 models. The baseline model shows a performance as expected: very high accuracy on original images that has no noise and the accuracy steadily decrease as the noise level goes up. However, for other re-trained models, there is a very significant performance gap between original test set and noise test sets. Moreover, particular re-trained models have shows intriguing accuracy figures such as for SNP_0.4 test set, RtN_0.2 model has an accuracy of 90% and for SNP_0.5 test set, RtN_0.4 model has an accuracy of over 85%. An example of image with level 0.4 and 0.5 of SNP noise is shown in Fig. 1.

In Table 2, it shows the test accuracy on various SPE noise data sets of the baseline ResNet-18 model and four re-trained ResNet-18 models. The baseline model performs as expected. However, the re-trained models do not have much difference when processing images of SPE noise level 5.0 and 10.0. An example of such noise images are shown in Fig. 2. They look equally awful. It is questionable to choose such noise levels as they do not generate much difference.

Figure 1: **Left:** an image of vizsla with level 0.4 of SNP noise; **Right:** an image of goldfinch with level 0.5 of SNP noise

Table 1: Test accuracy of re-trained ResNet-18 models on original and various SNP noise testing sets

model / test set	**RtO**	**RtN_0.1**	**RtN_0.2**	**RtN_0.3**	**RtN_0.4**	**RtN_0.5**
I Original	98.615%	75.538%	52%	45.538%	37.231%	27.846%
III SNP_0.1	96%	97.385%	83.692%	75.846%	55.692%	36.462%
IV SNP_0.2	91.077%	95.077%	93.231%	88.308%	73.077%	51.692%
V SNP_0.3	80.923%	89.077%	91.846%	90%	82.769%	62%
VI SNP_0.4	67.846%	83.231%	90%	88.769%	85.077%	76%
VII SNP_0.5	56.769%	72.923%	84.615%	83.231%	85.385%	80.615%

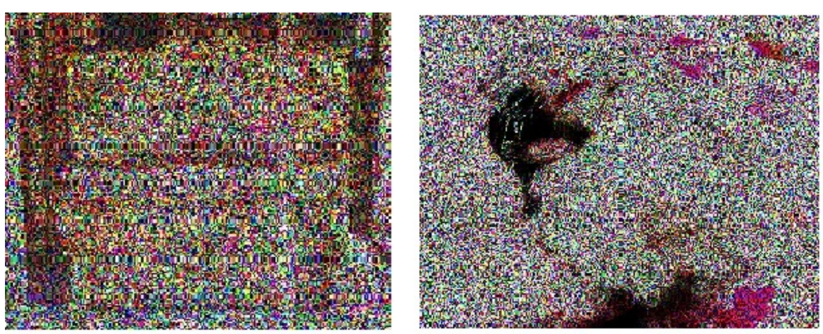

Figure 2: **Left:** an image of upright with level 5.0 of SPE noise; **Right:** an image of hamster with level 10.0 of SPE noise

Table 2: Test accuracy of re-trained ResNet-18 models on original and various SPE noise testing sets

model / test set	RtO	RtN_0.5	RtN_1.0	RtN_5.0	RtN_10.0
I Original	98.615%	71.385%	64.154%	51.077%	41.846%
IX SPE_0.5	84%	93.846%	91.231%	85.692%	67.846%
X SPE_1.0	73.077%	92.462%	91.385%	87.538%	71.385%
XI SPE_5.0	52%	81.385%	81.077%	80.923%	73.538%
XII SPE_10.0	47.846%	75.692%	78.769%	80.923%	74.154%

2.1.2 VGG-16

In Table 3, it shows the test accuracy on various SNP noise data sets of the baseline VGG-16 model and five re-trained VGG-16 models. As the baseline model behaves as expected, all the re-trained models show an astonishing robustness. There is little performance drop on original test set compared to the baseline model. Moreover, there is slight performance drop as the noise level goes up.

Table 3: Test accuracy of re-trained VGG-16 models on original and various SNP noise testing sets

model / test set	RtO	RtN_0.1	RtN_0.2	RtN_0.3	RtN_0.4	RtN_0.5
I Original	99.538%	99.385%	98.923%	99.077%	97.385%	80%
III SNP_0.1	96%	98.462%	98.154%	97.538%	95.077%	87.692%
IV SNP_0.2	91.231%	96%	95.692%	95.231%	91.692%	87.385%
V SNP_0.3	76.923%	88%	90.769%	91.077%	89.846%	86.615%
VI SNP_0.4	63.538%	75.846%	81.385%	83.077%	82.462%	79.385%
VII SNP_0.5	52.769%	62.462%	70.308%	74.615%	78.153%	75.538%

In Table 4, it shows the test accuracy on various SPE noise data sets of the baseline VGG-16 model and four re-trained ResNet-18 models. These re-trained model keep a consistent good performance as they did on the SNP noise sets. They show great robustness and a normal behavior, which is the decreasing accuracy on increasing noise data, unlike the random behaviors of ResNet-18 models. Therefore, compared to ResNet-18, VGG-16 benefits much more from training on noise data according to this experiment.

2.2 Experiment 2: comparison among models trained on same type of noise data

The following models are trained only on the same type, but multiple levels, of noise data. For SNP noise, the re-trained models **RtN_SNP** is trained on data set SNP_whole, which is the combination of set III, IV, V, VI and VII. For

Table 4: Test accuracy of re-trained VGG-16 models on original and various SPE noise testing sets

test set \ model	RtO	RtN_0.5	RtN_1.0	RtN_5.0	RtN_10.0
I Original	99.538%	99.231%	98.462%	97.077%	91.692%
IX SPE_0.5	84.923%	92.462%	93.538%	87.846%	86.308%
X SPE_1.0	72.154%	87.538%	86.923%	83.538%	80.923%
XI SPE_5.0	58%	72.769%	74.923%	75.077%	75.385%
XII SPE_10.0	54.462%	64.769%	69.538%	72.462%	70.308%

SPE noise, the re-trained models **RtN_SPE** is trained on data set SPE_whole, which is the combination of set IX, X, XI and XII. Below we present the results regarding the network architectures.

2.2.1 ResNet-18

In Table 5, the baseline model performs worse on SPE noise data compared to SNP noise data. To some sense, we can conclude that SPE noise data is more difficult to process than SNP noise data. On the other hand, **RtN_SNP** achieves the best accuracy of over 91% on SNP data and a descent 85% on SPE test data. **RtN_SNP** has a close performance on both SNP and SPE noise data. Both sacrifice the a prediction accuracy of around 30% on original images. One exciting potential conclusion to develop is that, training on one kind of noise data increase the robustness on processing other kinds of noise data. Due to the limitation of this experiment, we can only conclude on SNP and SPE noise.

Table 5: Test accuracy of re-trained ResNet-18 models on original and various noise testing sets

test set \ model	RtO	RtN_SNP	RtN_SPE
I Original	98.615%	68.308%	66.769%
II SNP_whole	80.215%	91.723%	84.585%
VIII SPE_whole	64.654%	85%	86.846%

2.2.2 VGG-16

Table 6 shows the performance of re-trained VGG-16 models and they are intriguing as there is very little accuracy loss on original image recognition ability compared to ResNet-18 models. Similar to the previous network architecture, **RtN_SNP** has a great performance on SNP noise data and descent performance on SPE noise data. **RtN_SPE** has a similar descent performance on both SNP and SPE nosie data. But again, this time, there is no accuracy sacrifice on original image processing ability .

Table 6: Test accuracy of re-trained VGG-16 models on original and various noise testing sets

test set \ model	**RtO**	**RtN_SNP**	**RtN_SPE**
I Original	99.538%	99.538%	99.231%
II SNP_whole	76%	90.062%	85.077%
VIII SPE_whole	67.231%	79.346%	83.154%

3 Plan for the next two weeks

I would like add more completeness and robustness to the conclusion that training on noise data improve the prediction accuracy on noise data. So here are a few steps I plan to do in the next two weeks:

- Test on more network architectures by adding newer models
- Add more noise types beyond pixel-level noises
- Investigate deeper on previous related research works, including re-establishing the implementations of previous methods
- Improve the analysis methods on experiment results using proper types of graphs and consolidate the conclusion

Apart from the above work, I also plan to start working on a survey paper on the topic of object recognition/detection in low-quality images. I will start collecting related papers and try to build a general frame of the survey paper.

4 Resource and Rule

Github Repo: Project Repo
Report Frequency: Every two weeks
Next Report: June 3, 2022