The Role of Convolutional Neural Networks In Diagnostic Imaging

Adam Salazar[1] Daniel Johnson, Ph.D.[2] Isaac Rutel, Ph.D.[1]
[1]The University of Oklahoma Health Sciences Center [2]University of Kansas Cancer Center

Introduction

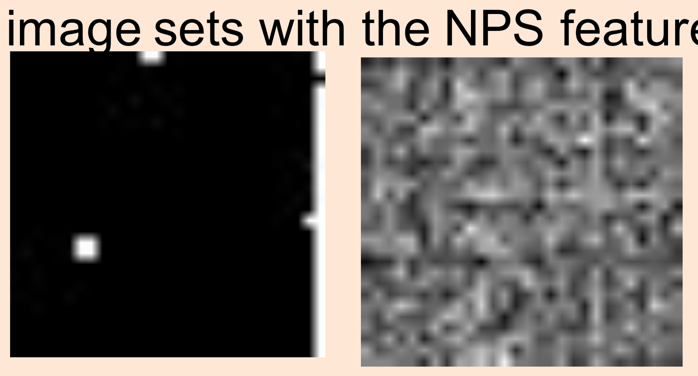
Digital Detection Systems are constructed from arrays of thousands of pixels. It is common that some pixels will become defective over time. Issues arise when clusters of dead pixels are in vital areas of anatomy or pathology. Compounding this problem is that the vendor specific criteria testing and information for replacing and investigating the detector condition tends to be proprietary. This project aims to provide vendor independent evaluation methods for dead detector analysis.

Methods and Procedures

Using the on-board imaging system of a Varian LINAC, 30 images were acquired using a range of common diagnostic kVp and mAs combinations. A range of Copper filtration was also applied (0-2 mm) to build a training/testing image library. A convolutional neural network (CNN) was then created using Google's open source software Tensorflow and Keras. The CNN was tasked with categorizing samples into four distinct categories: 1) No dead detectors were present in the image 2) A single dead detector was present in the image 3) Two dead detectors were present 4) Up to 5% of the detectors were dead (potential line defect)

Two data sets were created utilizing sub-sampled 28x28 images that were labeled for each dead detector categories by comparing to a vendor supplied dead element map. The two different training sets consisted of 78,838 images and another consisting of 53,230 images. To give the CNN a specific feature from which to learn, a local Noise Power Spectrum (NPS) calculation[1] (utilizing a 3 x 3 pixel neighborhood) was determined for each test pixel, for each image creating a 28 x 28 resultant matrix.

Four tests were performed on the data sets: 1) Training on the image sets alone 2) Training on the image sets concatenated with the local NPS matrix feature map 3) Image sets where minority categories are weighted (based on prevalence in the data set, and 4) Weighted image sets with the NPS feature maps.

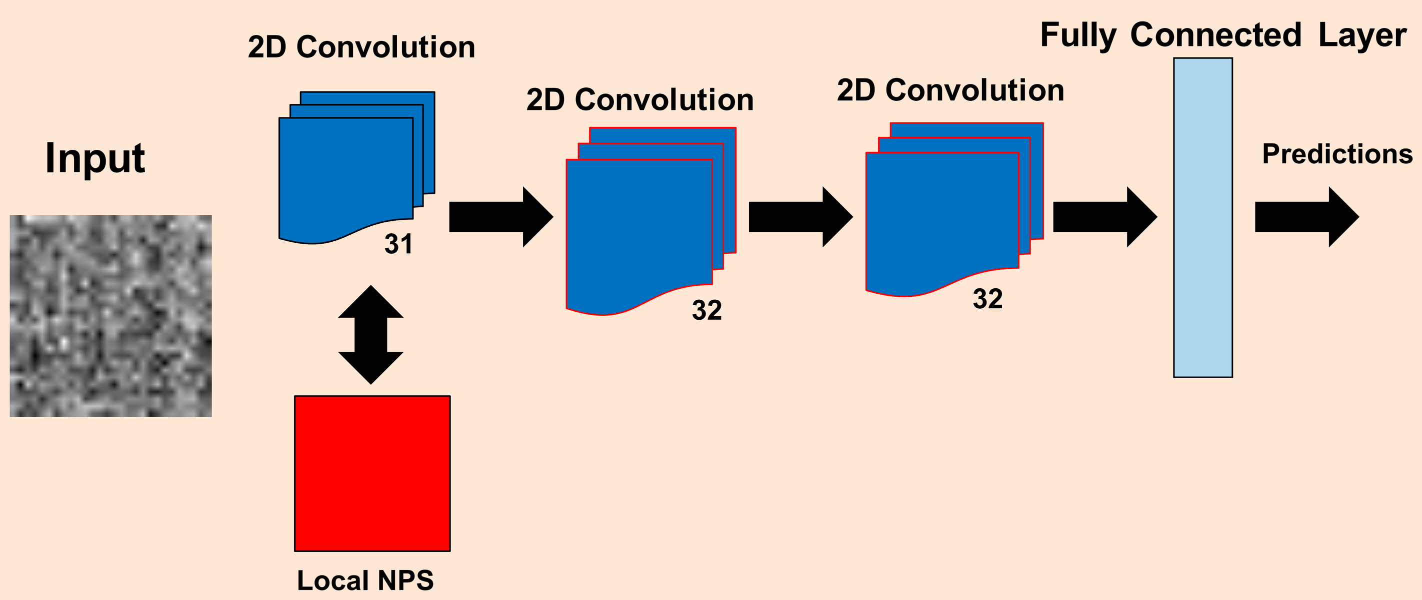


Sample from the Dead Detector Element Map extracted from the Varian LINAC and corresponding sample image.

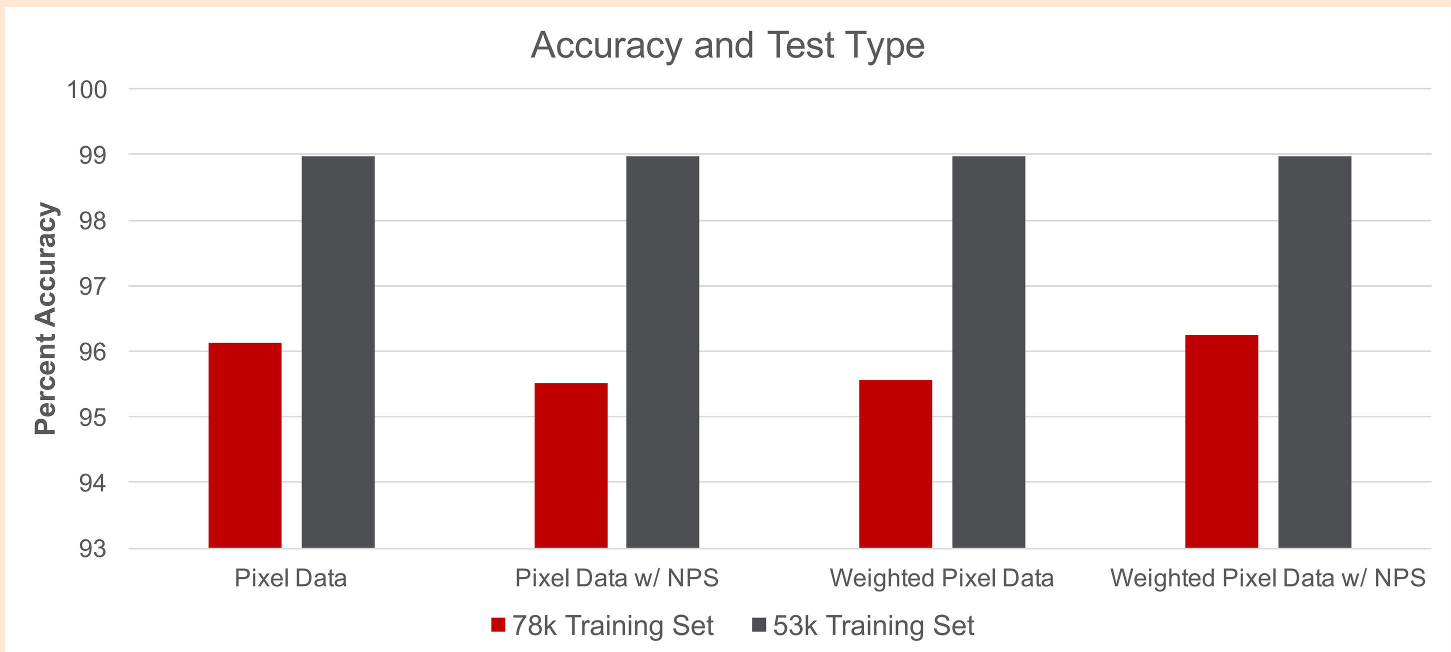
The black corresponds to functioning detector elements, while the white spots are detector elements that are being corrected by the vendor algorithm.

Custom Feature Maps and CNN Structure

Insertion of the custom (NPS) feature map occurs after the first convolution performed by the CNN. To maintain the number of total feature maps, only 31 (instead of 32) initial feature maps are created in the first layer when inserting the NPS layer. NPS slopes (the result of the NPS calculation) may indicate pixel correlation where a correction algorithm has been implemented. The following is the CNN structure.



Results



The data set with fewer samples outperformed the larger data set. The 53k set consistently achieved an accuracy of 98.97% for each case. To test whether it was simply guessing the category of majority, precision and recall tests were performed for both training set to determine which categories the system was succeeding or failing with during training.

The 78k data set achieved a highest accuracy utilizing weighted categories without the NPS. The NPS appears to make no noticeable difference for these configurations.

78k Training Set

Scenario	Precision (TP/(TP+FP))				Recall(TP/((TP+FN))			
	0 Dead	1 Dead	2 Dead	5% Dead	0 Dead	1 Dead	2 Dead	5% Dead
Pixel Data Only	1	1	1	0	0	0	0	1
Pixel Data with NPS	0.99	1	1	0.04	0	0	0	1
Weighted Pixel Data	1	1	1	0.04	0	0	0	1
Weighted Pixel Data with NPS	0.95	1	1	1	1	0	0	0

53k Training Set

Scenario	Precision (TP/(TP+FP))				Recall(TP/((TP+FN))			
	0 Dead	1 Dead	2 Dead	5% Dead	0 Dead	1 Dead	2 Dead	5% Dead
Pixel Data Only	0.99	1	1	1	1	0	0	0
Pixel Data with NPS	0.99	1	1	1	1	0	0	0
Weighted Pixel Data	0.99	1	1	1	1	0	0	0
Weighted Pixel Data with NPS	0.99	1	1	1	1	0	0	0

Overall, the model appears capable of avoiding False Positives but is weak at avoiding False Negatives, as seen in the high Precision values and low Recall values. While fewer False Negatives would be desirable, it is encouraging that the testing indicates the model is training in all categories. The high Precision values indicate that the model can pick up on subtle differences and predict with high accuracy from the training data.

Conclusion and Future Research

The goal of this research project is to create a tool that allows a Diagnostic Medical Physicist to test the integrity a digital detection system independently from the vendor. Due to the proprietary nature of the detection systems by the vendors, this would also increase the Physicist's ability to perform vendor independent checks. A CNN was created to find and quantify numbers of dead detectors within a digital detection system. Weighting the categories in proportion to how many dead elements appear in each category increased the precision of the CNN response, preventing the CNN from naively guessing all images contain no dead elements, naively improving the calculated accuracy. The current implementation of the NPS feature map is ineffective or slightly deleterious to the accuracy of the CNN categorization.

Future improvements include training the CNN on different modalities and vendors to see the effect from different systems and correction algorithms.

References

1)Hanson, K. M. (1998). Simplified method of estimating noise-power spectra. *Medical Imaging 1998: Physics of Medical Imaging*. doi: 10.1117/12.317023