Detecting Weapons in X-Ray Image Data using Machine Learning

Atticus Nafziger

Table of Contents

Introduction

The Transportation Security Administration (TSA) has an alarmingly low success rate at detecting security threats such as guns or explosives. In a 2015 test done by the Department of Homeland Security, TSA failed to detect and stop sixty-seven out of seventy fake weapons the Department of Homeland Security sent through security: a failure rate of over 95 percent. The security of every flight in the United States is at risk due to the shortcomings of TSA and it is time to implant some forms of image recognition software to help increase the detection rate of weapons. Machine learning can help increase airport security dramatically. Some work in image recognition by Battelle to classify luggage at different risk levels for weapons, but very limited research has been done into implementing machine learning image recognition to identify weapons in luggage. In this paper, the creation of three different models (multilayer neural network, support vector machine, and convolutional neural network) to accurately detect weapons in X-rays will be constructed in order to create safer airports and airplanes in the United States using machine learning.

Data Discussion

The X-ray data consists of 20,978 different images from Domingo Mery and the Department of Computer Science at the Universidad Catolica de Chile. The images depict X-rays of baggage with revolvers, razorblades, and knives in it; weapons outside of baggage; and an assortment of X-rays of castings, wooden parts, and various other items. The data for the most part is 850 x 850 pixels but fluctuations in image size existed. To make the data useable in a machine learning process, it is standardized and resized to a greyscale 64 x 64 image as described in *image 1* below:

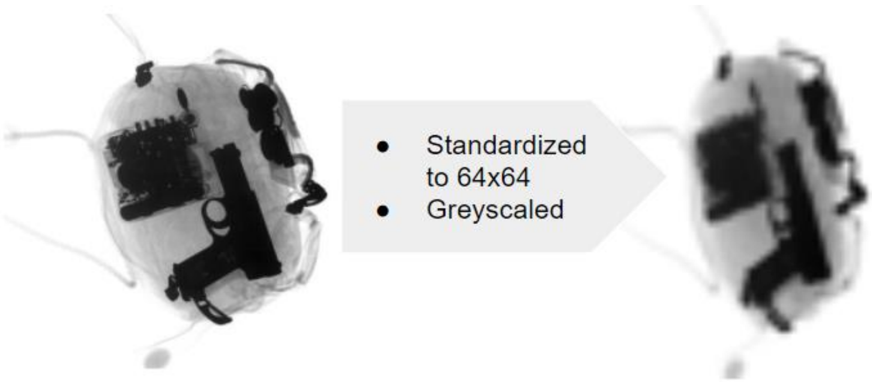

Image 1: Example of data modification

The folders of images with weapons and images without weapons are uploaded to Google Drive and then imported into Colab to be used to create the machine learning algorithms. Examples images are included below in *Image 2, Image 3,* and *Image 4*:



Image 2: An image of a casting, Image 3: A casting image, Image 4: An image of a razor blade

Different Model Types & Building the Models

Three different model types will be constructed to find the optimal model for detecting weapons in luggage. The three models are a multilayer neural network, a support vector classifier, and a convolutional neural network. First a simple multilayer neural network is created as a baseline model with a test size of 25 percent. Second, a support vector classifier is generated using a radial basis function and a test size of 20%. Finally, a convolutional neural network is generated with seven separate layers. Convolutional neural networks are widely regarded as the best machine model for working with image data, so using a convolutional neural network along with other models should guarantee the best results.

Model Performances

Multilayer Neural Network

As the baseline model, without adjusting parameters this model performed at a 93.6% testing accuracy over 10 epochs which can be seen below in *Image 5*. The model has three layers, two sigmoid layers with densities of 4096 and one SoftMax layer to slim down to the classification (either weapon or no-weapon). With some more adjustments such as increasing the number of epochs and adding more layers, a better performance could be reached. This model ran in roughly 10 minutes. Looking at the confusion matrix of the model in *Image 6* shows that the model did not detect 243 weapons and incorrectly identified 82 non-weapons as weapons. Ideally, the safest model would not fail to detect any weapons and only misclassify non-weapons.

Image 5: Training and Validation for Multilayer Neural Network over 10 epochs

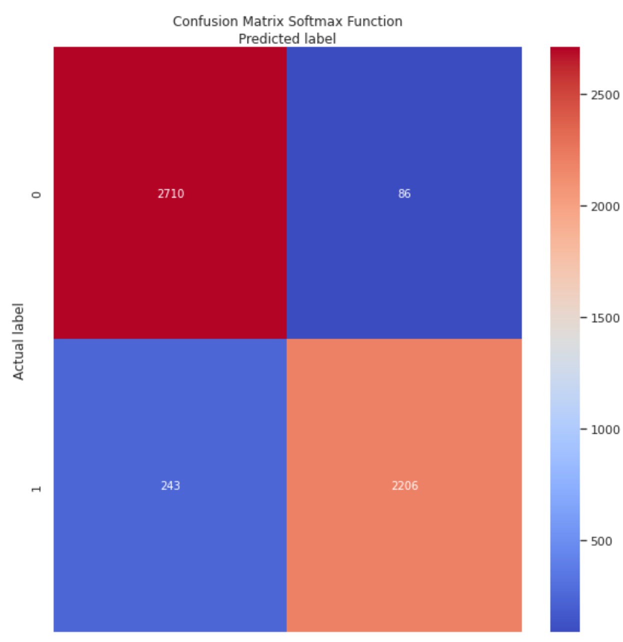

Image 6: Confusion Matrix of the Multilayer Neural Network

Support Vector Classifier

The radial based support vector classifier with a C value of 1 performed very well with an accuracy of 96%. The support vector classifier also trained the quickest out of the three models with a runtime of roughly three minutes. Looking at the confusion matrix of the support vector classifier, shows that it did not detect 94 weapons and misclassified 62 non-weapons. This is an improvement from the multilayer neural network, but again it would be more ideal if less weapons were being undetected.

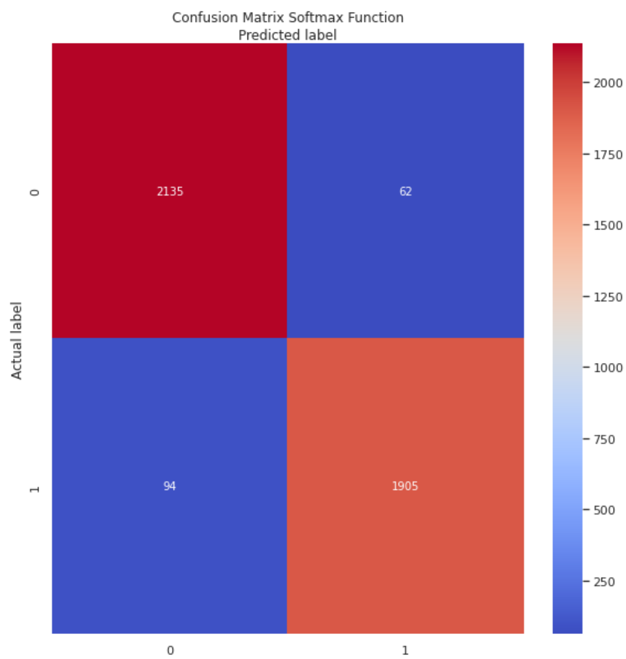

Image 7: Confusion matrix for Support Vector Classifier

Convolutional Neural Network

The final model, a convolutional neural network with seven layers performed with an accuracy of 97.3%, a slight improvement from the previous two models. This model took 83 minutes to run, much longer than the other two models. The convolutional neural network is made of seven layers. The first layer is a rectified linear unit function with 3x3 filter. The second layer is another rectified linear unit function, this time with a pooling size of 2x2. The third layer is a repeat of layer one. Layer four is a repeat of layer 2 with the addition of being flattened. Layers five and six are 512 dense layers. Finally, layer seven is a 10% dropout layer with a density of two because it is the final layer. The resulting confusion matrix, seen in *Image 8*, shows that the model only missed 31 weapons (8 times less than the multilayer neural network and 3 times less than the support vector classifier). The model also misclassified 103 non-weapons. This model is clearly the best for airport security out of the three generated. With more tweaking to reduce the number of missed weapons, this model could reach a standard of safety needed in airports across the country.

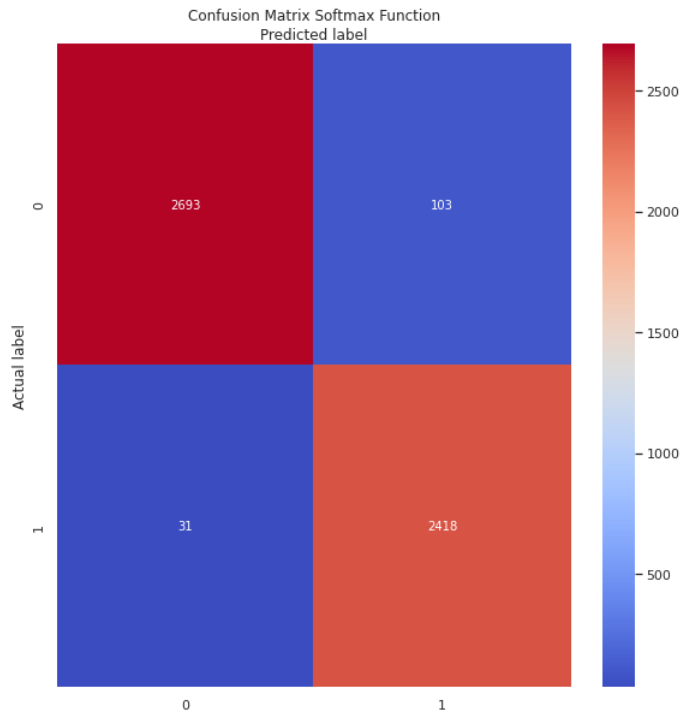

Image 8: Confusion matrix for convolutional neural network

Discussion and Conclusion

The resulting models vastly outperform the TSA's current weapon detection rate (TSA detecting 5% of weapons; Generated convolutional neural network detecting 97.3%), but practically the models may not. Due to the data sets used, the machine learning models did not have a data set that includes images of baggage without weapons, just baggage with weapons and non-baggage items. A more useful dataset to this project would have included only baggage data with and without weapons. Additionally, weapons being smuggled through airport security are most likely disassembled and hidden, making it harder to detect them. To increase the accuracy of the resulting models, higher definition images could be used instead of the 64x64 scaled-down grayscale images used. With the three models (multilayer neural network, support vector classifier, convolutional neural network) all performing with high accuracies (93.6%, 96%, 97.3%) the implementation of a machine learning detection system into airport security is clearly beneficial. To further increase safety, reducing the number of false negatives (not detecting weapons) these models produce is crucial, even at the cost of lack of accuracy and an increase in false negatives (detecting non-weapons). The convolutional neural network is currently performing at a false negative rate of 31/5245 or 0.59% which is already extremely safe when compared to the 95 percent weapon detection of the TSA. The introduction of machine learning into airport security can vastly increase the safety of airport passengers.

Citations

Battelle.org. (n.d.). Retrieved April 24, 2022, from https://www.battelle.org/insights/case-studies/case-study-details/improving-x-ray-imaging-for-the-transportation-security-administration

CNN Wire Service (2019, December 21). *Whistleblower says TSA is trading speed for security*. The Mercury News. Retrieved April 18, 2022, from www.mercurynews.com/2019/12/20/whistleblower-says-tsa-is-trading-speed-for-security/

Mery, D.; Riffo, V.; Zscherpel, U.; Mondragón, G.; Lillo, I.; Zuccar, I.; Lobel, H.; Carrasco, M. (2015): GDXray: The database of X-ray images for nondestructive testing. Journal of Nondestructive Evaluation, 34.4:1-12. [PDF]