Preprocessing Techniques' Effect On Overfitting for VGG16 Fast-RCNN Pistol Detection

Jiahao Li* Charles Ablan* Rui Wu† Shanyue Guan* Jason Yao*
East Carolina University, Greenville, NC 27858 USA

Abstract

Within the past two decades, gun detection became an increasingly popular research topic as gun violence continued to threaten public safety. Of all machine learning algorithms employed to identify weapons, Convolution Neural Networks (CNN) stood out as the most robust method for identifying guns in images. Although CNN had outstanding image classification performances, it is not without limitations. A CNN without large quantities of data suffers from overfitting. While complex architectures reduce overfitting, it also results in slower detection speed and increased memory usage. This study analyzed three image preprocessing techniques' effect on reducing overfitting in VGG16 Fast Regional Convolutional Neural Network (F-RCNN) without modifying network architectures. The base VGG16 was trained with transfer learning in MATLAB on a dataset of 1500 selected images to artificially induce overfitting. The average testing precision of the base VGG16 detector was then compared with the results of other VGG16 detectors supplemented with image processing techniques. The three image processing techniques used are color contrast enhancement, principle component analysis (PCA), and combined preprocessing methods. The study concluded that color contrast enhancement had the greatest impact on reducing the effects of overfitting. It was found that with proper levels of color contrast enhancements, the average testing precision went up noticeably. The PCA supplemented model failed to reduce the number of irrelevant features and did not retain the important features. The PCA method proved to be ineffective in reducing overfitting and resulted in an overall loss of average precision. The combined preprocessing methods combined the images of both PCA and color contrast enhancements into two different training datasets. The first dataset combined PCA with color enhancements and the second only combined color enhancement results. Both combined preprocessing methods did not increase the average precision potentially due to conflicting features.

Key Words: VGG16; convolution neural networks; image preprocessing techniques; gun detection.

*Department of Engineering. Email: lij17@students.ecu.edu, ablanc17@students.ecu.edu, guans18@ecu.edu, yaoj@ecu.edu
†Department of Computer Science. Email: wur18@ecu.edu

1 Introduction

Past breakthroughs in Convolutional Neural Networks (CNN) using VGG16 and VGG19 architectures achieved an incredible 90 percent accuracy in image classification [3, 23]. As the search for better neural networks continues, increasing complexity becomes unavoidable [3]. A CNN's classification accuracy often increases with layer complexity, however, increasing complexity also causes slower detection speeds and increased memory usage [6]. The new 1000+ layered Inception V networks are strong evidence that CNN is getting increasingly complex at the expense of memory and speed [26]. Maintaining the simplicity of the network architecture while achieving high identification precision becomes a growing concern in recent years [24]. Of the existing neural networks, VGG16 remained one of the simplest yet robust neural networks ever created. This study applied a Fast-Regional Convolutional Neural Network (F-RCNN) model that modified a VGG16 net into an F-RCNN object detector for pistol detection [7]. A base VGG16 net was trained using transfer learning with an original blend of 1500 ground truth images. The base VGG16 was trained within MATLAB to achieve an overall true positive percentage of over 98 percent on pistol detection for the training dataset. Such high levels of training accuracy are often a sign of overfitting [19]. By running the precision versus recall test [8], it was revealed that the base detector had a low average precision. This study focused on determining viable image preprocessing techniques that can address the overfitting problem and increase network performance without modifying the internal architecture.

This experiment utilized MATLAB to both train the F-RCNN object detector and implement various image preprocessing techniques. MATLAB was used because the Deep Learning toolbox is easy to implement with robust support for various neural networks [11, 25]. The built-in conversion function of images to matrix form made image processing easy and efficient due to various image processing techniques requiring matrix transform and matrix transpose [16, 17]. Three image pre-processing trials were applied to the training dataset for the VGG16 F-RCNN pistol detector in hopes of raising average testing precision. A neural network performs best if trained with great variance in the pose and lighting of an object [29]. The image preprocessing techniques applied in this study operated under the assumption that pose and lighting are key features that greatly affect the precision of any object recognition system [29]. The neural network's tendency of relying on color and lighting makes it challenging to distinguish dark simple objects from lowlighting backgrounds [18]. To resolve the problem of low color contrast with the background, the first trial used color enhancement techniques at set intervals to widen the gap between dark and light-colored regions. The second trial used PCA analysis to reduce the background while retaining most of the pistol's features [10, 22]. After the results of the first two methods were obtained, a fusion of the two modified image datasets along with the original dataset was used to train two combined VGG16 F-RCNN detectors.

In the rest of this paper, Section 2 describes the related works in the area of pistol detection and image preprocessing. Section 3 describes the dataset used in the study as well as the source images used. Section 4 describes the performance of the VGG16 F-RCNN detector on the base dataset. Section 5 describes the effects of various color enhancement trials on detector precision. Section 6 describes the effect of PCA analysis on detector performance. Section 7 describes the effect of combining the different image preprocessing techniques into one dataset. Section 8 discusses the results of the previous experimental trials. The conclusion is given in Section 9.

2 Related Works

Pistol detection has been a widely researched topic ever since Neural Networks first gained popularity with Alex-net in 2012 [12]. Many studies in the past have tackled the problem of handgun detection using Regional Proposal Networks and transfer learning from established neural networks such as VGG16 [2, 18]. Two methods are generally used to improve the performance of specific object detectors. The first method is modifying the neural network architectures to become deeper and more robust [26], and the second method is applying image preprocessing techniques to help detectors better separate the important features from background noise [21]. This paper tackles the problem of using image preprocessing techniques to increase the performance of a VGG16 F-RCNN pistol detector.

2.1 Pistol Detection

Pistol detection differs from pistol classification in that object detectors must identify the part of the image with the weapon [5, 13-14, 18]. Various pistol detectors have been trained in the past using both neural networks and traditional machine learning techniques such as SVM(Support Vector Machine) and Histogram of Oriented Gradients (HOG) [18, 27]. However, the best performances came from using Regional Proposal Networks(RCNN) and its variants the FAST-RCNN and Faster-RCNN [5, 18].

Akcay et al. [2] used a variety of neural networks with different types of object detectors to detect pistols in x-ray images. Five different object detection models were used in the study including Sliding Windows Convolutional Neural Network(SW-CNN), Regional Convolutional Neural Network (RCNN), Faster Regional Convolutional Neural Network (Faster-RCNN), Region-based Fully Convolutional Networks (R-FCN) and You Only Look Once object detectors (YOLOv2). The object detection models were trained with Alexnet, VGG16, VGG19, and residual neural networks (ResNets). The resultant data shows that both RCNN and Faster-RCNN outperformed both traditional handcrafted Bag of visual words(BoVW) features and fellow SW-CNN detectors. Akcay et al. [2] have proved with their study that it is possible to achieve high object detection precision with R-CNN models and their variants.

Olmos, Tabik, and Herra [18] analyzed a VGG16 based Faster- RCNN detector for video pistol detection with limited success. The Faster-RCNN's large false positives rates drastically lowered the overall detector precision. The researchers contributed the high false positives ratio to low contrast and luminosity of certain video frames. The pistols that are not clearly distinguished from the backgrounds are often missed and other objects are falsely identified.

2.2 Neural Network Image Preprocessing

Within the field of image analysis, image preprocessing techniques are frequently used in combination with a variety of image classification algorithms [5, 13, 28, 30]. Many image preprocessing techniques aimed to either reduce noise or enhance desired features. Rehman et al [21] discussed a variety of image preprocessing methods for character recognition using neural networks. Within the various preprocessing techniques discussed, the researchers highlighted the importance of thresholding in image processing. Thresholding sets a boundary on the original color scheme from which the image is converted to binary or grayscale. Threshold simplifies the image and highlights the desired characters making it an important preliminary image processing technique. The other image preprocessing techniques included the elimination of unwanted features and extraction of key features. The importance of image preprocessing techniques was highlighted by Rehman et al [21] and its effect in increasing neural network performances cannot be overlooked.

3 Dataset

This study utilized 2,000 open source images from various image datasets under an academic license as well as original images taken by the research team. Out of a training dataset of 2,000 images, 1,500 were randomly selected for training and 500 were used for testing. The first dataset used was the IMFDb online firearm dataset, a free gun image dataset that contains a variety of pistols, rifles, shotguns, and other firearms [9]. The second dataset used was the Sci2s weapons dataset from the University of Granada [5, 18]. For this experiment, only pistols were selected from a variety of movie images. The selection criteria for the combined image dataset included a variety of background and contextual information at various backgrounds, angles, and distances for a balanced selection of various pistol images. The combined image dataset was biased towards images with a variety of background and contextual information for accurate object recognition [20]. The limited number of training images meant a larger chance for the network to overfit, while a diverse spread of gun sample images helped the detector to maintain an acceptable testing precision.

4 Base VGG16 F-RCNN Performance

This study was conducted on the basis that large false positives are generated often due to the overfitting of neural networks [19]. Where a model trained to recognize features in the training dataset might fail to generalize features onto the testing dataset resulting in low testing identification rates. To fully express the influences of various color enhancement methods on overfitting, this research focused on simple VGG-16 nets trained with limited images. This was meant to introduce overfitting in the model and compare the effect of various preprocessing methods on reducing overfitting. Within MATLAB, a pre-trained VGG16 net was trained with the 1,500 pistol images. The VGG16 object detector reached a training accuracy of 98.8 percent after 5 epochs. The trained detector was then tested on the 500 testing images with an average precision of 0.2138. The low precision likely resulted from boxing errors where either a part of the gun or too much of the background was selected. However, upon physical examination, it was found that although the detection boxes did not precisely match the ground truth labeling, over 90 percent of the highest confidence detection box correctly identified most of the weapons and did not mistake other objects for pistols. Thus, the overall accuracy of the base F-RCNN object detector was found to be in an acceptable range. Figure 1 shows the MATLAB's precision evaluation of the VGG16 during testing. Figure 1 shows the performance graph plotted for precision versus recall, where recall and precision are ratios of true positive instances to the sum of true positives and false negatives in the detector, based on the ground truth [2, 18]. By running through each image and visually judging the accuracy of the neural net, three different category metrics were used to determine true the performance of a neural net performance. i) True Positive means that the network has correctly identified the weapon with an appropriate bounding box. ii) False Positive means that the network has failed to select most of the weapon or has mistaken something else for the pistol. iii) Negative means the Neural net has failed to recognize guns in the image with high confidence. The results of the visual analysis will change slightly from person to person and trial to trial, but five successive repetitions on the testing dataset prove the general effectiveness of the trained F-RCNN object detector. Table 1 shows the classification criteria for accuracy evaluation and the average performance of the VGG16 detector over the five repetitions.

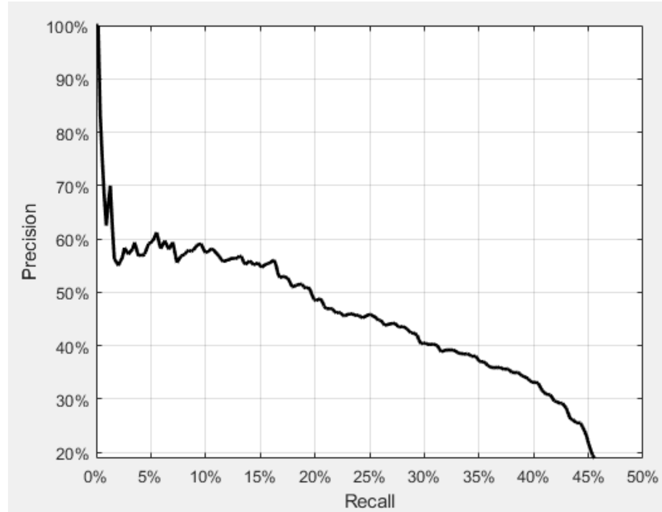
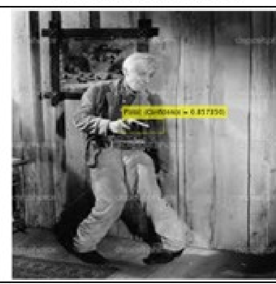


Figure 1: VGG16 precision vs recall on 500 images

Table 1: VGG16 judgment criteria and percentage number

True Positive	False Positive
	
453/500	47/500

Negative

0/500

5 Fixed Ratio Color Enhancement Trial

The color enhancement preprocessing techniques operate under the assumption that neural nets rely heavily on color and texture for target identification [1, 18]. Since most pistols have a darker color as opposed to their surroundings, by increasing the color contrast the neural network should be able to better distinguish dark pistols from light backgrounds to avoid false positives. A flowchart of the color enhancement process is shown in Figure 2.

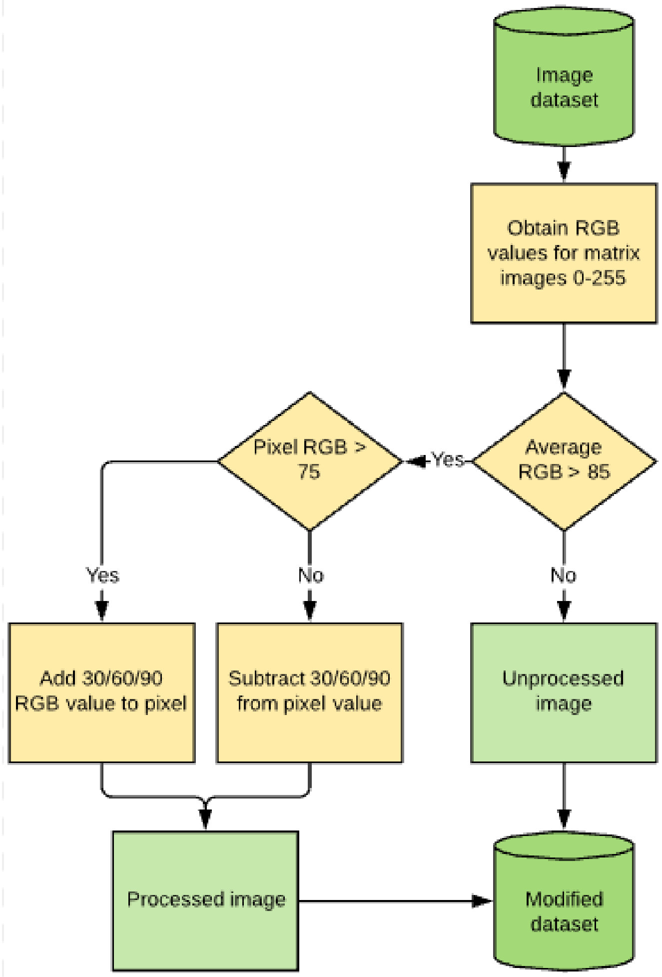


Figure 2: Color enhancement trial flowchart

The first preprocessing method separated light and dark colors at 75 out of 255 for all three RGB color schemes. The 75 separation value is selected because it separates the darker gun color scheme from the lighter background. The parts of the image with an RGB value less than 75 are darkened by -30 and the parts with a number larger than 75 are lightened by +30. The 30 enhancement value was chosen because initial trials indicated that any value less than 30 produced had no significant impact on the testing precision. If a particular pixel has an RGB value of less than 30 it will simplify to zero after preprocessing and the same concept applies to values greater than 225. The images with a mean RGB value of 85 or below were deemed too dark for the pistol to be separated from the background and thus the color enhancement was not performed. Only the training dataset underwent this preprocessing technique and the testing results on the original testing dataset are shown in Figure 3. In the rest of this paper, the “x” of “enhanced-x” refers to the specific RGB value that is modified in the image.

From Figure 3 it can be observed that the enhanced-30 dataset did not result in significant improvements in MATLAB recorded average precision from 0.2138 to 0.2208. Since initial trials with the color enhanced-30 showed only slight improvements in precision, the second trial doubled the modification ratio to +60 and -60. From Figure 3 it is observed that the enhanced-60 trial noticeably increased the average precision from 0.2138 to 0.2944 for a 38 percent increase in precision. Based on the success of the enhanced-60 trial, the enhanced-90 trial further increased the contrast to +90 and -90 in an attempt to achieve higher precision. As it can be observed from Figure 3, overly increasing the color enhancement ratio to +90 and -90 negatively impacted detector performance decreasing the average precision to only 0.0759.

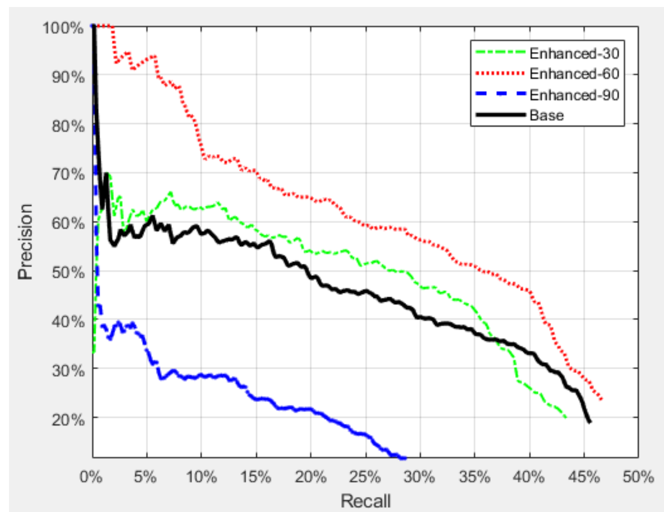


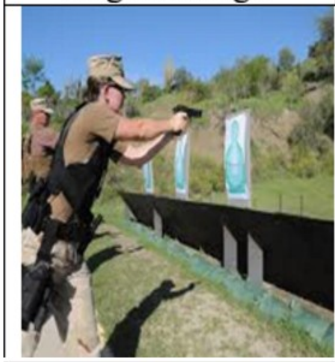
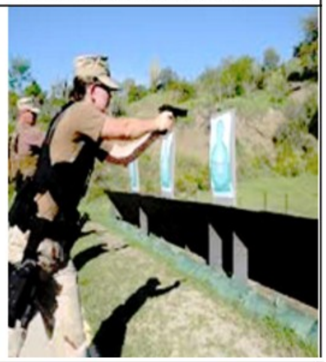
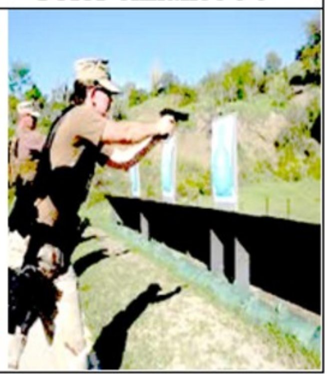
Figure 3: Color enhancement trial performance

Table 2 shows the full comparison of the effect of different color enhancement ratios on training images. It can be seen from the color-enhanced images that the pistol has a drastically higher color contrast as opposed to the background. By using color enhancement techniques, the background noise is also whitewashed resulting in an overall simpler image.

5.1 Varied Ratio Color Enhancement Trial

After initial promising results with the enhanced-60 dataset, new color enhancement trials were conducted. The new color enhancement trials focused on analyzing the effect of changing threshold values and enhancement ratios had on average precision. The result of the extended tests is shown in Table 3. FR stands for Fast R-CNN model trained and the number after the ” - ” symbol stands for various color enhancement ratios used with a threshold of 75. For FR-60/50, FR-60/100, and FR-60/125 the number after the ”/” symbol indicates the threshold used to separate light and dark regions. Pixel values below the low threshold are considered light regions. Pixel values above the high threshold are considered dark regions. Low ratios are applied to light regions and high ratios are applied to dark

Table 2: Original and color-enhanced images

Original image	Color enhance 30
	
Color enhance 60	Color enhance 90
	

regions; where integer values indicate subtraction and decimal values indicate multiplication. The average precision shows the model performance.

The previous enhanced-60 model used a fixed threshold of 75 to differentiate between light and dark colors. To test for the effect on changing the threshold, new tests were conducted for thresholds at 50, 100, and 125 with the same enhancement ratio of +60 and -60. From Table 3, it can be observed that both decreasing and increasing the threshold has a detrimental effect on model performance. However, by comparing the three new models, FR-60/125 had the most detrimental effect on accuracy with a -0.1743 or 81 percent decrease in average precision. Overall, the new trials indicate that changing the thresholds too far up or down can have a detrimental effect onperformance.

The second variable that changed was the enhancement ratio. While enhancement-30, 60, and 90 had fixed changes to pixel values, new tests incorporated varying changes based on the old pixel value. The new trials multiplied the original pixel values with a ratio between 0.1 - 0.9 for value reduction, and 1.1 to 1.9 for value increase. Out of the eight new models, FR-0.1 and FR-0.5 tested for the effect of only enhancing the light portions of the image. FR-1.5 and FR-1.9 tested for the effect of enhancing the darker sections of the image. Trials FR-0.1-1.9 to FR-0.9-1.1 tested for the effect of increasing both the dark and light contrasts. Out of the eight new trials, FR-0.1 had the

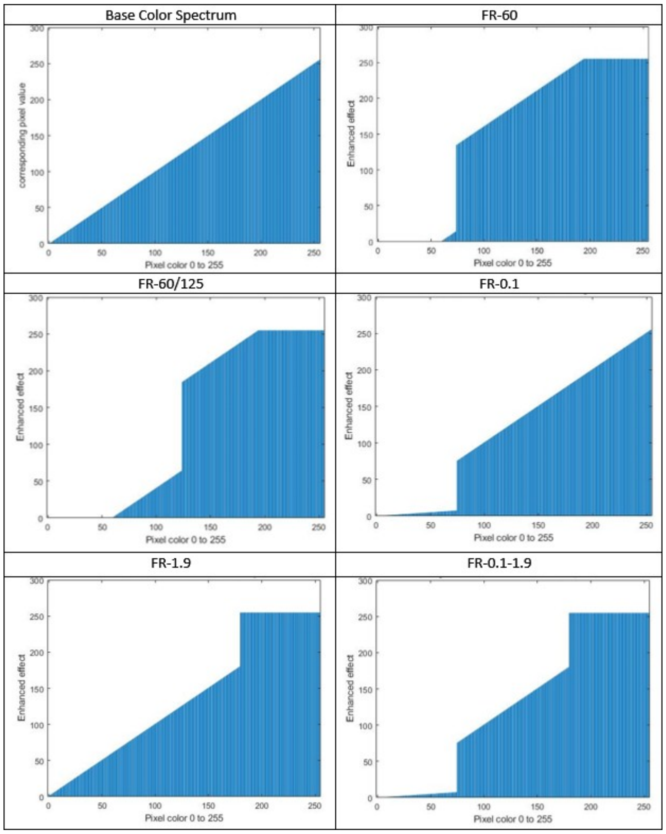
Table 3: Color-enhancement Trials.

Trials	Low Threshold	High Threshold	Low Ratio	High Ratio	Average Precision
Base	N/A	N/A	N/A	N/A	0.2138
FR-30	75	75	30	30	0.2208
FR-60	75	75	60	60	0.2944
FR-90	75	75	90	90	0.0759
FR-60/50	50	50	60	60	0.1158
FR-60/100	100	100	60	60	0.1569
FR-60/125	125	125	60	60	0.0395
FR-0.1	75	N/A	0.1	N/A	0.2852
FR-0.5	75	N/A	0.5	N/A	0.2386
FR-1.5	N/A	180	N/A	1.5	0.1299
FR-1.9	N/A	180	N/A	1.9	0.1725
FR-0.1-1.9	75	180	0.1	1.9	0.0015
FR-0.3-1.7	75	180	0.3	1.7	0.1077
FR-0.5-1.5	75	180	0.5	1.5	0.1503
FR-0.7-1.3	75	180	0.7	1.3	0.1755
FR-0.9-1.1	75	180	0.9	1.1	0.1822

highest increase in precision of 34 percent to 0.2852 followed by an 11 percent increase for FR-0.5 at 0.2386. In contrast, FR-1.5 and FR-1.9 both had detrimental effects on precision at 0.1299 and 0.1725. From the single ratio trials, it can be judged that that enhancing the lighter part of the image is more beneficial than enhancing the darker parts. The last four rows of 3 show the result of the multiple ratio trials. The FR-0.9-1.1 had the highest precision and FR-0.1-1.9 had the lowest precision. The average precision of the "FR-X-X" trials followed a distinct trend, where larger enhancement ratios are related to decreases in performance. The combined result of new color enhancement ratio trials suggested that lighter region enhancements have the best result, where dark value enhancements likely resulted in a loss of important pistol features.

Table 4 shows the effect of various color enhancement trials on pixel values. The X-axis is the original pixel values from 0 to 255. The Y-axis is the modified pixel values. The base color spectrum shows the unmodified pixel graph where the X and Y-axis follows a linear one-to-one correlation. Both FR-60 and FR-0.1 both outperformed the base model, and from Table 4 it can be observed that both had a noticeable drop in values less than 75. FR-60/125 had the most dramatic effect

Table 4: Color spectrum comparison chart



on the color spectrum, which could correlate to loss of important features resulting in poor performance. Both FR-1.9 and FR-0.1-1.9 underperformed against the base model and both had a noticeable rise in darker color regions. Overall,it can be inferred from Table 4 that lighter region enhance-ments without sharp rises in darker regions offer the best enhancement results.

6 PCA Feature Reduction

Each layer of a Convolutional Neural Network retains a specific feature of the image [6, 26]. Because of the curse of dimensionality, some researchers showed that fewer features can be less misleading for a machine learning model. PCA feature reduction aimed to reduce the number of features the neural net was exposed to. PCA compresses the data and retains principle features in the original image [10, 22]. The PCA works substantially better for binary images that have only have two dimensions as opposed to the three dimensions of RGB images [15]. While there are ways to perform PCA for RGB images using multi-linear subspace learning algorithms [15], performing PCA on binary images often produces cleaner results. However, binary PCA reduced images cannot take on color as the principle components are analyzed as 1 and 0 inputs. To restore the original color to the PCA image, MATLAB was used to fuse the original image with the PCA binary reduction image. The resultant color scheme differs from the original image because of the merging process. However, most of the color features are retained through this method.

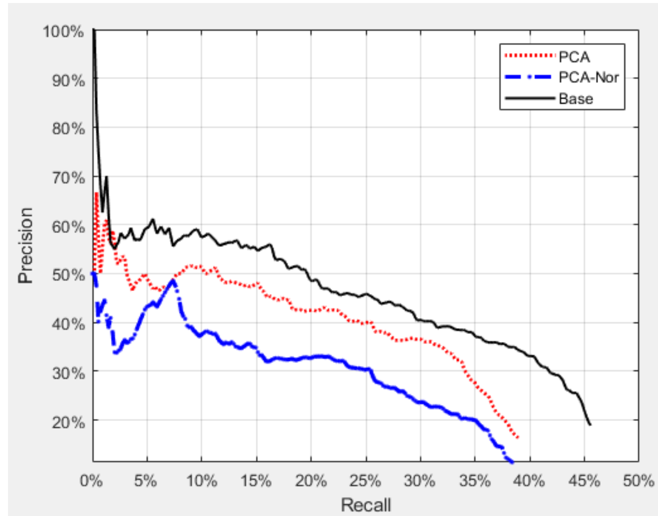
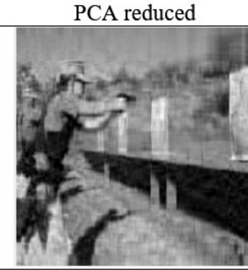
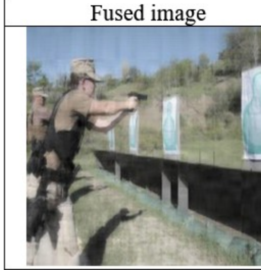


Figure 4: PCA detector performance

Table 5 shows the comparison between the original image, PCA reduced image, and fused image used for analysis. Figure 4 shows the performance of the VGG16 net trained with PCA reduced images on unmodified test images. Figure 4 also shows the performance of the PCA trained net on PCA reduced test images. From Figure 4, it can be observed that the PCA reduced detector tested with the original testing dataset resulted in a significant decrease in average precision from 0.2138 to 0.1223. In comparison, the PCA reduced detector tested with PCA reduced testing dataset resulted in a less significant decrease in average precision from 0.2138 to0.1636. This disparity between the test results can be

Table 5: PCA sample images

Original image	PCA reduced
	

Fused image


attributed to the loss of features between original and PCA datasets.

7 Combined Dataset

For a neural network to retain information about an object from various angles and lighting conditions, a great variance in the training dataset is required [4]. While individual approaches can yield no conclusive results towards increasing accuracy, combining the various approaches into a single dataset can increase data variations without requiring additional images. The first combined dataset used a combination of the color enhance-30, color enhance-60, PCA analysis, and original images to form the 1500 training dataset. The second combined dataset used a combination of the three-color enhance trials and original images. Since color enhance-60 yielded the best results, the combined dataset uses 750 images from the color enhance-60 dataset, 250 images each from color enhance-90, color enhance-30, and the original images. Figure 5 shows the combined F-RCNN performance.

From Figure 5 it can be observed that combining PCA reduced images with color-enhanced images had no statistically significant changes in performance as the average precision raised from 0.2138 to 0.2398. From Figure 5 it can also be observed that mixing different color enhancement ratios significantly decreases the average precision to 0.1524.

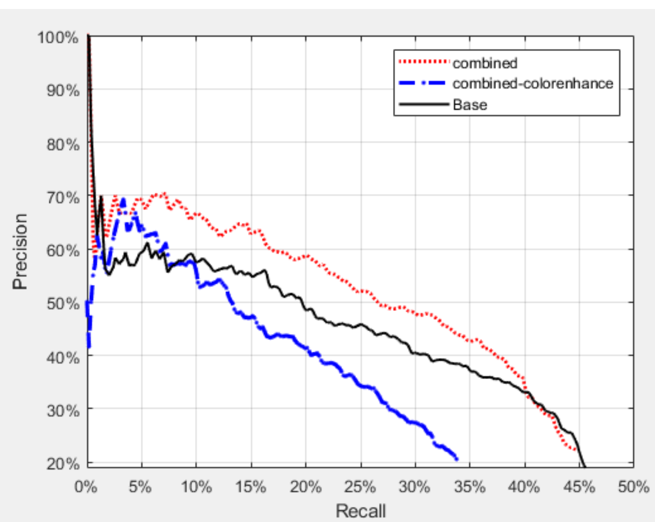


Figure 5: Combined detector performance

8 Discussion

Of all the common image preprocessing techniques used in machine vision, three methods were chosen to address the overfitting problem of the base detector. Color enhancement was chosen to address target and background separation. Its flexible nature with many parameters makes it easily transferable from detecting pistols to other objects of interest. Secondly, PCA was introduced to reduce the number of overlapping features thus lowering the chance for overfitting. Finally, the combined approach focused on merging the two techniques to produce high feature variations.

Three different fixed color enhancement ratios were used at intervals of 30, 60, and 90 with a 75 threshold. Figure 3 shows color enhancement-30 slightly increased the average precision by 0.007 from 0.2138 to 0.2208, which was not statistically significant. The second trial increased the enhancement ratio to 60 which resulted in a noticeable rise in average precision by 0.0806, from 0.2138 to 0.2944. Although this result is not a significant change in precision, it shows that basic color enhancement can mitigate the background noises associated with low lighting environments. The third trial increased the enhancement ratio to 90 and resulted in a significant drop in average precision of -0.1379 from 0.2138 to 0.0759.

To further study the extensive effect of color enhancement, new trials were conducted using various thresholds and enhancement ratios. From the results of Table 3, it can be seen that the varying ratio trials indicated a strong correlation between color contrast and network performance. Since the color enhancement-60 trials yielded the greatest increase in precision, the new trials first aimed at adjusting the 75 percent threshold. FR-60/50 lowered the threshold to 50 and saw a 46 percent loss of precision which is surprising due to the success of the color enhancement-60 trials. FR-60/100 and FR-60/125 then had a more detrimental effect on average precision with FR- 60/125 having an 82 percent loss of precision. The failure to raise average precision despite using a tested enhancement ratio suggests that there is an optimal threshold that separates light and dark regions for each database, and deviations from the threshold can lead to detrimental results. FR-60/50 enhanced made 25 pixels in the light region darker when compared to the original color enhancement-60 trials, and this difference contributed to the loss of information. By adjusting the threshold to 100 and 125, more regions are classified as light regions and had their values decreased, also leading to a loss of information. At this point it is unclear why 75 presumably works for this dataset, however, it can be hypothesized that only regions with color values less than 75 are lighter regions that require light enhancement to reduce background noise. The rest of the varying ratio trials will continue to adopt 75 as the low threshold since it provided the best results. The high threshold was chosen at 180 because it was 75-pixel values lower than 255. The varied ratio trials multiplied each pixel value by a ratio instead of adjusting every pixel by the same amount. The difference between the fixed and varying ratio adjustments can be observed in Table 4.

Of all the varied ratio trials, FR-0.1 showed a 34 percent increase in average precision to 0.2852. A follow-up experiment with FR-0.5 showed a less significant increase of 11 percent in average precision to 0.2386. The two trails suggest that large lighter region enhancements can have positive effects on testing precision. FR-1.5 and FR-1.9 were then tested for the effect of enhancements on darker regions. However, FR-1.5 resulted in a 39 percent decrease in average precision and FR-1.9 resulted in a 19 percent decrease in average precision. The two dark enhancement results suggest that enhancement of dark regions correlates with a small loss of precision. Trials FR-0.1-1.9 to FR-0.9-1.1 tested the combined effect of both enhancing the lighter and darker regions. Although all four trials resulted in a loss of average precision, there is a clear trend linking the level of enhancement with the amount of precision loss. FR- 0.1-1.9 had the most loss in precision of 99 percent. It can then be observed that as the enhancement ratios went down to FR-0.5-1.5, the precision loss also went down to 30 percent. While the lowest enhancement ratio of FR-0.9-1.1 only had a 15 percent loss in precision. The exponential loss of precision following larger enhancement ratios can be attributed to the loss of critical information during the double enhancement process. By enhancing the color spectrum from 0 to 75 and 180 to 255, 150 different pixel values were enhanced. Enhancing over 58 percent of the original color scheme could lead to conflicting features and loss of useful features.

From the color enhancement trials, it can be determined that the ratios of enhancement greatly affect the detector performance; too little color enhancement results in no significant changes but a huge enhancement ratio can have negative impacts. The rise in precision with the color enhancement-60 trial can be attributed to correctly separating key features of the pistol from the background while retaining most of the context. The color enhancement-90 trial likely overly enhanced the image so that key features of the pistol might be lost while the background context became too mono-toned. Although the current color enhancement trials do not show significant increases in precision, the preliminary results obtained using minimalist preprocessing algorithms open the way for more advanced enhancement techniques. For future works, potential extensions of this algorithm are directly amplifying a range of pistol color spectrum with machine learning algorithms.

Applying PCA to the training dataset did not result in any increases in average precision. PCA reduced the number of features by half in the training data images, which resulted in a loss of distinguishing features for the VGG16 to use. Figure 4 showed that the PCA trained detector performed poorly when tested on the original testing dataset as the average precision dropped from 0.2138 to 0.1223. In comparison, the detector performed noticeably better when tested on the PCA reduced testing dataset with the average precision only dropping from 0.2138 to 0.1636. This change in performance can be attributed to the VGG16 only retaining features pertinent to the PCA images. While too many features can confuse the neural network, too little feature will cause the learned PCA features to not transfer to original images.

The first combined dataset with an evenly distributed color enhanced, PCA and original images only raised the average precision from 0.2138 to 0.2398. The insignificant change in average precision indicates that blind variations in the training dataset do not necessarily improve performance due to conflicting features. The PCA reduced features likely interfered with the positive effects of the color-enhanced trials. The second dataset focused entirely on the effect of varying color enhancements. The second combined dataset hoped to retain thepositive effects from the color enhance-60 trials while adding beneficial contrast variations. Contradicting initial predictions, the second detector trained with the three color-enhanced image datasets performed drastically worse than the original detector. The combined color-enhanced detector lowered the average precision from 0.2138 to 0.1524. The significant drop in average precision can be attributed to the color enhance-30 and 90 dataset offering conflicting features that overruled the positive effects of the color enhance-60 dataset.

9 Conclusions

In conclusion, this paper analyzed three different image preprocessing techniques that can be applied to reduce overfitting in a VGG16 F-RCNN detector. The color enhancement method proved to be most impactful in increasing the average testing precision. A positive correlation was found between increasing the lighter region enhancement ratios and increases in average testing precision. While enhancing darker regions were observed to have detrimental impacts on precision. It was found that on average a small amount of color enhancement is unlikely to result in noticeable changes in detector performance. However, over enhancement can also have negative impacts on performance. The optimal color enhancement ratio and threshold will depend on the target objects and their relative backgrounds. It is recommended that different object databases go through extensive trials to find the correct enhancement ratios and thresholds to optimize performance.

Applying PCA to reduce the number of features in the image proved to be an ineffective method to increase average testing precision for pistol detection. The VGG16 F-RCNN detector trained with PCA reduced images had an overall detrimental effect on testing precision. The loss of precision was largely attributed to a loss of critical information during the feature reduction phase. The first combined dataset of PCA and color enhancement images resulted in no significant increases in performance. The second dataset of only color enhancement trials resulted in a detriment to performance. Overall, the combined methods had no significant impact on performance. The underwhelming performance of the combined method can be contributed to conflicting features. Too large of a color variation can result in conflicting features and had a negative impact on performance.

Acknowledgments

The following people of ECU's Innovation Design Lab contributed greatly in funding and supporting the project: Dr. Todd Fraley, Dr. Ted Moris, Director Wayne Godwin, lab assistants Marco Agostini and Elliot Paul. This material is based upon work supported by the National Science Foundation under grant number IUSE/PFE: RED award #1730568. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the

authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] A. Akbarinia and K. R. Gegenfurtner, "How is Contrast Encoded in Deep Neural Networks?", *arXiv preprint arXiv:1809.01438*, 2018. https://arxiv.org/abs/1809.01438, Last Accessed 17 Feb 2021.

[2] S. Akcay, M. E. Kundegorski, C. G. Willcocks, and T. P. Breckon, "Using Deep Convolutional Neural Network Architectures for Object Classification and Detection Within X-ray Baggage Security Imagery," *IEEE Transactions on Information Forensics and Security*, 13(9):2203–2215, 2018.

[3] M. Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M. S. Nasrin, B. C. Van Esesn, A. A. S. Awwal, and V. K. Asari, "The History Began from Alexnet: A Comprehensive Survey on Deep Learning Approaches," *arXiv preprint arXiv:1803.01164*, 2018.

[4] C. M. Bishop, *Neural Networks for Pattern Recognition*.Oxford University Press, 1995.

[5] "Brightness Guided Preprocessing for Automatic Cold Steel Weapon Detection in Surveillance Videos with Deep Learning," *Neurocomputing*, 330:151–161, 2019.

[6] A. Dhingra, "Model Complexity-Accuracy Trade-off for a Convolutional Neural Network," *arXiv preprint arXiv:1705.03338*, 2017. https://arxiv.org/abs/1705.03338, Last Accessed 17 Feb 2021.

[7] R. Girshick, "Fast R-CNN," *Proceedings of the IEEE International Conference On Computer Vision*, pp. 1440–1448, 2015.

[8] J. Grau, I. Grosse, and J. Keilwagen, "PRROC: Computing and Visualizing Precision-Recall and Receiver Operating Characteristic Curves in R," *Bioinformatics*, 31(15):2595–2597, 2015.

[9] N. A, "IMFDB: Internet Movie Firearms Database," 2020. http://www.imfdb.org/wiki/Main Page [Accessed on 5 May 2020].

[10] I. T. Jolliffe and J. Cadima, "Principal Component Analysis: a Review and Recent Developments," *Philosophical Transactions of the Royal Society A: Mathematical,Physical and Engineering Sciences*, 374(2065):20150202, 2016.

[11] P. Kim, "Matlab Deep Learning," *With Machine Learning, Neural Networks and Artificial Intelligence*, vol. 130, 2017.

[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks," *Advances in Neural Information Processing Systems*, pp. 1097–1105, 2012.

[13] J. Li, X. Liang, S. Shen, T. Xu, J. Feng, and S. Yan, "Scale-Aware Fast R-CNN For Pedestrian Detection," *IEEE Transactions on Multimedia*, 20(4):985–996, 2017.

[14] M. Lokanath, K. Sai Kumar, and E. Sanath Keerthi, "Accurate Object Classification and Detection by Faster-RCNN," *Materials Science and Engineering Conference Series*, vol. 263, 2017.

[15] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "A Survey of Multilinear Subspace Learning for Tensor Data," *Pattern Recognition*, 44(7):1540–1551, 2011.

[16] A. McAndrew, "An Introduction to Digital Image Processing with Matlab Notes for SCM2511 Image Processing," *School of Computer Science and Mathematics, Victoria University of Technology*, 264(1): 1–264, 2004.

[17] J. G. Nagy, K. Palmer, and L. Perrone, "Iterative Methods for Image Deblurring: a Matlab Object-Oriented Approach," *Numerical Algorithms*, 36(1):73– 93, 2004.

[18] R. Olmos, S. Tabik, and F. Herrera, "Automatic Handgun Detection Alarm in Videos Using Deep Learning," *Neurocomputing*, 275:66–72, 2018.

[19] A. P. Piotrowski and J. J. Napiorkowski, "A Comparison of Methods to Avoid Overfitting in Neural Networks Training in the Case of Catchment Runoff Modelling," *Journal of Hydrology*, 476:97–111, 2013.

[20] J. Ponce, T. L. Berg, M. Everingham, D. A. Forsyth, M. Hebert, S. Lazebnik, M. Marszalek, C. Schmid, B. C. Russell, A. Torralba, *et al.*, "Dataset Issues in Object Recognition," *Toward Category-Level Object Recognition*, pp. 29–48, 2006.

[21] A. Rehman and T. Saba, "Neural Networks for Document Image Preprocessing: State of The Art," *Artificial Intelligence Review*, 42(2):253–273, 2014.

[22] G. Shakhnarovich and B. Moghaddam, "Face Recognition in Subspaces," in *Handbook of Face Recognition*, pp. 141–168, 2005.

[23] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv preprint arXiv:1409.1556*, 2014. https://arxiv.org/abs/1409.1556 Last Accessed 17 Feb 2021.

[24] J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller, "Striving for Simplicity: The All Convolutional Net," *arXiv preprint arXiv:1412.6806*, 2014. https://arxiv.org/abs/1412.6806 Last Accessed 17 Feb 2021.

[25] D. P. Strik, A. M. Domnanovich, L. Zani, R. Braun, and P. Holubar, "Prediction of Trace Compounds in Biogas from Anaerobic Digestion using the MATLAB Neural Network Toolbox," *Environmental Modelling & Software*, 20(6):803–810, 2005.

[26] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going Deeper with Convolutions," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, 2015.

[27] G. K. Verma and A. Dhillon, "A Handheld Gun Detectionusing Faster R-CNN Deep Learning," *Proceedings of the 7th International Conference on Computer and Communication Technology*, pp. 84–88,

2017.

[28] Y. Xu, Z. Zhang, G. Lu, and J. Yang, “Approximately Symmetrical Face Images for Image Preprocessing in Face Recognition and Sparse Representation based Classification,” *Pattern Recognition*, 54:68–82, 2016.

[29] L. Yann, H. Fu, and B. Leon, “Learning Methods for Generic Object Recognition with Invariance to Pose and Lighting: Computer Vision and Pattern Recognition.” CVPR 2004, *Proceedings of the 2004 IEEE Computer Society Conference*, vol. 2, 2004.

[30] Y. Zhu, H. Yuan, C. Zhang, and C. Lee, “Image-Preprocessing Method for Near-Wall Particle Image Velocimetry (PIV) Image Interrogation with Very Large In-Plane Displacement,” *Measurement Science and Technology*, 24(12):125302, 2013.

Jiahao Li is a machine learning researcher at East Carolina University. His primary research interest includes the main areas of machine learning from deep neural networks to evolutionary computing systems. His recent research is aimed at developing intelligent threat response systems with robust auto-regressive models. Outside of technical research pursuits, Li is also pursuing sensible A.I. ethics and studying the impact of automation on society.

Charles Ablan is an undergraduate researcher at East Carolina University pursuing his BS in Engineering. His main areas of research interest include machine learning and engineering technological innovation. He is a member of the ECU underground water level research team developing autoregressive models.

Wu Rui received a bachelor’s degree in Computer Science and Technology from Jilin University, China in 2013. He received his Master and Ph.D. degrees in Computer Science and Engineering from the University of Nevada, Reno in 2015 and 2018, respectively. Dr. Wu is now working as an assistant professor in the Department of Computer Science at East Carolina University and collaborates with geological and hydrological scientists to protect the ecological system. His main research interests are data imputation, machine learning, and data visualization using AR/VR devices.

Guan Shanyue received his B.S. degree in civil engineering from Tongji University, Shanghai, China, in 2011, and the Ph.D. degree in civil engineering from the University of Florida, Gainesville, FL, USA, in 2017. From 2017 to 2018, he was a Post-Doctoral Fellow with the University of Florida. He joined East Carolina University, Greenville, NC, USA, in 2018, where he is currently an Assistant Professor of Engineering.

His research interests include smart and resilient infrastructure, wireless sensor networks, UAV-based monitoring, image processing, and data analysis.

Jason Yao has research interests in the areas of wireless/wearable medical sensors, sensor networks for home environments, telemedicine, and industrial process monitoring and control. Dr. Yao received his Ph.D. degree in electrical engineering from Kansas State University. He is a senior member and an active volunteer of IEEE.