



**IDENTIFYING IMAGE MANIPULATION
SOFTWARE FROM IMAGE FEATURES**

THESIS

Devlin T. Boyter, CPT, USA

AFIT-ENG-MS-15-M-051

**DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY**

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

The views expressed in this document are those of the author and do not reflect the official policy or position of the United States Air Force, the United States Department of Defense or the United States Government. This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

AFIT-ENG-MS-15-M-051

IDENTIFYING IMAGE MANIPULATION
SOFTWARE FROM IMAGE FEATURES

THESIS

Presented to the Faculty
Department of Electrical and Computer Engineering
Graduate School of Engineering and Management
Air Force Institute of Technology
Air University
Air Education and Training Command
in Partial Fulfillment of the Requirements for the
Degree of Master of Science in Cyber Operations

Devlin T. Boyter, B.S.

CPT, USA

March 26, 2015

DISTRIBUTION STATEMENT A
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AFIT-ENG-MS-15-M-051

IDENTIFYING IMAGE MANIPULATION
SOFTWARE FROM IMAGE FEATURES

THESIS

Devlin T. Boyter, B.S.
CPT, USA

Committee Membership:

Dr. G. L. Peterson, PhD
Chair

Dr. R. K. Martin, PhD
Member

Maj B. G. Woolley, PhD
Member

Abstract

As technology steadily increases in the field of image manipulation, determining which software was used to manipulate an image becomes increasingly complex for law enforcement and intelligence agencies. When only the image is available, the task becomes much more difficult. However, this should be surmountable because different image manipulation programs will have implemented the manipulation algorithms differently. This research examines the use of four preexisting image manipulation detection techniques: Two-Dimensional Second Derivative, One-Dimensional Zero Crossings, Quantization Matrices Identification, and File Metadata analysis applied to the task of identifying the image manipulation program used. The determination is based on each image manipulation software program having implemented the manipulation algorithms differently. These differences in the implementation will leave behind different artifacts in the resultant image. Experimental results demonstrate the framework's ability to identify from the 48 combinations of image manipulation software programs, scaling, and the algorithm used with a true positive rate of 0.54, false positive rate of 0.01, and a Kappa statistic of 0.53 for Joint Photographic Experts Group (JPEG). The results for Tagged Image File Format (TIFF) images were a true positive rate of 0.53, false positive rate of 0.01, and a Kappa statistic of 0.52.

Acknowledgements

I would like to express my appreciation to Dr. Peterson for the superb guidance and extraordinary patience you showed me. Your advice allowed me to exceed my own expectations. To my loving family, thank you for the support and encouragement throughout this endeavor. I would not have accomplished anything without you.

Devlin T. Boyter

Table of Contents

	Page
Abstract	iv
Acknowledgements	v
List of Figures	viii
List of Tables	ix
List of Abbreviations	xii
I. Introduction	1
1.1 Problem Statement	1
1.2 Hypothesis, Goals, and Assumptions	2
1.3 Implementation	2
1.4 Testing	3
1.5 Success Criteria	4
1.6 Paper Organization	5
II. Related Work	6
2.1 Mathematical Image Definition, Image Formats, and Interpolation Algorithms	6
2.1.1 Mathematical Image Definition	6
2.1.2 Lossy and Lossless Image Formats	8
2.1.3 Interpolation Algorithms	11
2.2 Image Manipulation and Detection	19
2.2.1 Pixel Based	19
2.2.2 Format Based	19
2.2.3 Camera Based	20
2.2.4 Physical Based	20
2.2.5 Geometry Based	21
2.3 Feature Detection	21
2.4 Implemented Techniques	22
2.4.1 Header Format	23
2.4.2 Resampling	23
2.4.3 Statistical	28
2.4.4 Double JPEG Compression	29
2.5 Classification Overview	31
2.6 Summary	31

	Page
III. Methodology	33
3.1 Detection Framework	33
3.2 Feature Generation	35
3.2.1 File Metadata	35
3.2.2 Quantization Matrices Identification	35
3.2.3 Two-Dimensional Second Difference	37
3.2.4 One-Dimensional Zero Crossings	38
3.3 Classification	39
3.4 Summary	41
IV. Experiments and Results	42
4.1 Experiments	42
4.2 Results and Analysis	45
4.2.1 Group 1: Individual Modules	47
4.2.2 Group 2: Pairing of modules	48
4.2.3 Group 3: Three Modules	49
4.2.4 Group 4: All Modules	50
4.2.5 Results Conclusion	50
4.3 Summary	54
V. Conclusion	56
5.1 Hypothesis and Success Criteria	56
5.2 Results Synopsis	57
5.3 Significance in the Area of Research	58
5.4 Future Work	58
Appendix A. Results	60
A.1 Group 1	61
A.2 Group 2	74
A.3 Group 3	93
A.4 Group 4	100
Bibliography	104

List of Figures

Figure		Page
2.1	DCT Based Encoder Simplified Diagram.	8
2.2	Preparation of Quantized Coefficients for Entropy Encoding [5].	9
2.3	Superposition Interpolation.	12
2.4	Nearest Neighbor Interpolation.	12
2.5	Bilinear Interpolation.	14
2.6	Bicubic Interpolation.	15
2.7	Interpolation Algorithms in Black and White.	17
2.8	Interpolation Algorithms in color.	18
2.9	Single and Double Quantization [25].	30
3.10	Framework Process Flow.	34

List of Tables

Table		Page
2.1	TIFF Compression Dictionary 1.....	11
2.2	TIFF Compression Dictionary 2.....	11
4.3	Image Modification Settings.	43
4.4	Group Overall Results.	46
4.5	Group By Rate and Algorithm Results.	46
A.6	2DSD Only By Class (JPEG).....	62
A.7	2DSD Only By Rate (JPEG).	63
A.8	2DSD Only By Algorithm (JPEG).	63
A.9	2DSD Confusion Matrix (JPEG).	64
A.10	1DZC Only By Class (JPEG).....	65
A.11	1DZC Only By Rate (JPEG).	66
A.12	1DZC Only By Algorithm (JPEG).	66
A.13	1DZC Confusion Matrix (JPEG).	67
A.14	2DSD Only By Class (TIFF).	68
A.15	2DSD Only By Rate (TIFF).....	69
A.16	2DSD Only By Algorithm (TIFF).	69
A.17	2DSD Confusion Matrix (TIFF).....	70
A.18	1DZC Only By Class (TIFF).	71
A.19	1DZC Only By Rate (TIFF).....	72
A.20	1DZC Only By Algorithm (TIFF).	72
A.21	1DZC Confusion Matrix (TIFF).....	73
A.22	2DSD and 1DZC By Class (JPEG).	75

Table		Page
A.23	2DSD and 1DZC By Rate (JPEG).	76
A.24	2DSD and 1DZC By Algorithm (JPEG).	76
A.25	2DSD and 1DZC Confusion Matrix (JPEG).	77
A.26	1DZC and QMI By Class (JPEG).	78
A.27	1DZC and QMI By Rate (JPEG).	79
A.28	1DZC and QMI By Algorithm (JPEG).	79
A.29	1DZC and QMI Confusion Matrix (JPEG).	80
A.30	2DSD and QMI By Class (JPEG).	81
A.31	2DSD and QMI By Rate (JPEG).	82
A.32	2DSD and QMI By Algorithm (JPEG).	82
A.33	2DSD and QMI Confusion Matrix (JPEG).	83
A.34	2DSD and 1DZC By Class (TIFF).	84
A.35	2DSD and 1DZC By Rate (TIFF).	85
A.36	2DSD and 1DZC By Algorithm (TIFF).	85
A.37	2DSD and 1DZC Confusion Matrix (TIFF).	86
A.38	1DZC and FMD By Class (TIFF).	87
A.39	1DZC and FMD By Rate (TIFF).	88
A.40	1DZC and FMD By Algorithm (TIFF).	88
A.41	1DZC and FMD Confusion Matrix (TIFF).	89
A.42	2DSD and FMD By Class (TIFF).	90
A.43	2DSD and FMD By Rate (TIFF).	91
A.44	2DSD and FMD By Algorithm (TIFF).	91
A.45	2DSD and FMD Confusion Matrix (TIFF).	92
A.46	2DSD, 1DZC, and QMI By Class (JPEG).	94

Table		Page
A.47	2DSD, 1DZC, and QMI By Rate (JPEG).....	95
A.48	2DSD, 1DZC, and QMI By Algorithm (JPEG).....	95
A.49	2DSD, 1DZC, and QMI Confusion Matrix (JPEG).....	96
A.50	2DSD, 1DZC, and QMI By Class (TIFF).....	97
A.51	2DSD, 1DZC, and QMI By Rate (TIFF).	98
A.52	2DSD, 1DZC, and QMI By Algorithm (TIFF).	98
A.53	2DSD, 1DZC, and FMD (TIFF).	99
A.54	All Modules By Class (TIFF).....	101
A.55	All Modules By Rate (TIFF).	102
A.56	All Modules By Algorithm (TIFF).	102
A.57	All Modules (JPEG).....	103

List of Abbreviations

Abbreviation		Page
IMSP	Image Manipulation Software Programs	1
JPEG	Joint Photographic Experts Group	2
DFT	Discrete Fourier Transform	2
APS	Adobe Photoshop Creative Cloud Release 2014.0.0	5
MSPM	Microsoft Office Picture Manager 14.0.7010.1000	5
GIMP	GNU Image Manipulation Program 2.8.10	5
JPEG	Joint Photographic Experts Group	6
TIFF	Tagged Image File Format	6
RGB	Red-Green-Blue	6
JFIF	JPEG File Interchange Format	7
DCT	Discrete Cosine Transform	8
FDCT	Forward Discrete Cosine Transform	8
DC	Discrete Cosine	9
AC	Alternate Coefficients	9
IJG	Independent JPEG Group	9
LZW	Lempel-Ziv and Welch	10
IQM	Image Quality Measures	22
HOWS	Higher Order Wavelet Statistics	22
BSM	Binary Similarity Measures	22
JFS	Joint Feature Set	22
CFS	Core Feature Set	22
SFFS	Sequential Forward Feature Selection	22

Abbreviation		Page
EM	Expectation Maximization	24
DFT	Discrete Fourier Transform	25
PSNR	Peak Signal to Noise Ratio	28
MSE	Mean Square Error	28
SIMM	Structural Similarity	28
WEKA	Waikato Environment for Knowledge Analysis	31
QMI	Quantization Matrices Identification	35
2DSD	Two-Dimensional Second Difference	37
1DZC	One-Dimensional Zero Crossings	38
ARFF	Attribute-Relation File Format	44
IMSP	Image Manipulation Software Programs	56
GUI	Graphical User Interface	59

IDENTIFYING IMAGE MANIPULATION SOFTWARE FROM IMAGE FEATURES

I. Introduction

The progression in technology since the invention of digital photography has led to the creation of many high quality Image Manipulation Software Programs (IMSP) being readily available to the general public. IMSPs can make a manipulated image appear as though no alteration has occurred. Numerous techniques have been developed in the area of digital image authentication to detect image manipulation when it is not obvious or visible to the viewer [9, 11, 16, 17, 26]. However, there is little research focusing on identification of the IMSP that created the manipulated image when only in possession of the altered image.

Law enforcement and intelligence agencies have a need to identify the utilized image manipulation software as part of the investigative and evidence gathering process. By detecting the super/sub-position algorithms used in IMSP, an image suspected of undergoing an alteration can be associated with an IMSP. Knowledge of the IMPS aids in identifying the computer system used in altering the image. This assists examiners in reconstructing how and by whom an image manipulation occurred.

1.1 Problem Statement

This research demonstrates the development of techniques that can be employed to identify the software and interpolation algorithm used to resize an image. Although different IMSPs make use of the same algorithms for resizing, we hypothesize that differences in implementation of the algorithms leave detectable traces in the modified

image. The techniques for identifying a specific IMSP entail building a framework of modules used to detect artifacts present after an image manipulation occurred employing existing image authentication methods.

1.2 Hypothesis, Goals, and Assumptions

It is hypothesized that the proposed framework improves an examiner's ability to classify an image to which IMSP created the manipulated image beyond the probability of a random guess. To evaluate the hypothesis a framework is created that classifies a manipulated image to the correct IMSP.

The research's hypothesis contains two main assumptions. First, the only manipulation addressed is resizing an image using specific rates, interpolation algorithms, and IMSPs. Additionally, only one manipulation occurs per image and no cross contamination occurs by using multiple IMSPs on a single image.

1.3 Implementation

This research implements unique Python modules using four previously developed image manipulation techniques to create a framework to detect specific IMSPs [11, 16, 17, 26]. The first module reads in the file metadata and conducts string matching against known signatures IMSPs embedded in the image [17]. The second module extracts the quantization tables for Joint Photographic Experts Group (JPEG) images only and compares them to tables known to be used by specific IMSPs [16]. Another module examines the image's statistics by calculating the derivative of a row of pixels and processing that data with a Discrete Fourier Transform (DFT) function in order to find spikes that correspond to an IMSP [11]. The final module examines the zero crossings of an image to create a binary sequence that is processed by a DFT function to find spikes that correspond to an IMSP [26]. A Bayesian Network

classifier processes the data from the modules and classifies an image to an IMSP.

1.4 Testing

The experiments in this research were used to determine the framework's accuracy. Two datasets were built for the purpose of training and testing the two image format specific classifiers: one consists of only JPEG images and the other of only Tagged Image File Format (TIFF) images. One hundred images were taken using the Nikon D5100 digital camera in the Nikon RAW format and then converted to either JPEG or TIFF.

After conversion to the two image formats, the two sets of 100 images were resized 48 different ways for a total of 4,800 images per set. The classes are sorted by software used, Adobe Photoshop Creative Cloud Release 2014.0.0 (APS), Microsoft Office Picture Manager 14.0.7010.1000 (MSPM), or GNU Image Manipulation Program 2.8.10 (GIMP), algorithm used (bilinear, bicubic, nearest-neighbor, or undetermined), and the interpolation rate (0.50, 0.66, 0.75, 0.90, 1.10, 1.25, 1.33, or 2.00). For example, images 1-100 were resized with GIMP using the bicubic algorithm at the interpolation rate of 1.10 would be grouped into the same class. The interpolation rates were chosen to include both superposition and subposition rates. None of the rates are instances in which one rate is implemented more than once to get the other rate. For example, 0.25 can be implemented by using the rate 0.50 twice. See Table 4.3 for a graphical representation of the classes.

The framework was then combined in thirteen different variations of the original four modules: Two-Dimensional Second Derivative, One-Dimensional Zero Crossings, Quantization Matrices Identification, and File Metadata analysis. These thirteen variations are tested by classifying of the manipulated images using the Bayesian Network classifier. In order to create a variation in the TPR results, the number of

images to be processed by the classifier were scaled down to 80% of the original size of the dataset by randomly selecting images 10 times to create 10 different unbalanced sub-datasets. The Bayesian Network classifier processed each sub-dataset 10 times using a 10 fold cross-validation method with different random number seeds.

Each test was performed ten times using the cross-validation method implemented in Waikato Environment for Knowledge Analysis (WEKA) with ten different random seeds, 1-10. The mean true positive rates (TPR), false positive rates (FPR), and F-Measure for each class and Kappa Statistic for the overall framework were determined for each grouping of modules.

1.5 Success Criteria

The defined success criteria requires the framework to have a statistically significant Kappa statistic that compares the classification results to a random guess. This comparison was conducted using a t-statistic test with a 95% confidence interval. The size of the tested data set and number of classes defines a random guess to be 0.02084. The data set contains 48 distinct classes. The results show that the framework can distinguish between many of the classes but previous research shows several of the classes will not have a statistically significant improvement in the accuracy rate when compared to a random guess.

The results also examined the framework's ability classify an image to a specific interpolation rate and algorithm. A comparison between the true positive rate and a random guess using a t-statistic test was conducted with a 95% confidence interval. The random guess probability for the interpolation rate is 0.125 and for the algorithm is 0.25.

1.6 Paper Organization

The remainder of the thesis follows this outline. Chapter 2 discusses previous research conducted in the area of image authentication and image manipulation detection. The related work in image authentication and manipulation detection spans five different categories of image forensic detection and techniques as described by Farid, [9]. Farid's categories assist in the organization of Chapter 2 with the associated techniques and details the literature on the specific methods implemented in this paper. Chapter 3 presents the proposed framework. Each of the proposed modules described in Section 1.4 are detailed in this chapter.

Chapter 4 discusses the results and analyzes the findings. Testing focused on Nearest Neighbor, Bilinear and Bicubic interpolation algorithms used in Adobe Photoshop Creative Cloud Release 2014.0.0 (APS), Microsoft Office Picture Manager 14.0.7010.1000 (MSPM), and GNU Image Manipulation Program 2.8.10 (GIMP) with the following resizing rates; 0.50, 0.66, 0.75, 0.90, 1.10, 1.25, 1.33, and 2.00. The Chapter begins with discussion on the overall experiment process. Then the individual tests and their results are presented. Finally, the chapter discusses the analysis of the results and observations made. The results of the four main experiments show with statistical significance the framework's ability to determine a specific IMSP that manipulated an image when compared to a random guess.

Chapter 5 concludes the paper and discusses further areas of research. The next chapter discusses previous work in the area of image authentication and manipulation detection.

II. Related Work

The ability to detect an image that has undergone a common image processing operation, such as scaling, rotation, brightness adjustment, compression, etc., is important for use by law enforcement and intelligence agencies in identifying which Image Manipulation Software Program (IMSP) conducted the image manipulation. The research discussed in this thesis builds on the concepts and algorithms for detecting image manipulations.

This chapter presents the literature on the superposition/subpostion algorithms and the implementation of the algorithms within various IMSPs. The presented literature covers five different categories of image forensic detection and associated techniques as developed by Farid [9]. The following section defines the mathematical definition of an image, discusses two image formats, and three interpolation algorithms.

2.1 Mathematical Image Definition, Image Formats, and Interpolation Algorithms

This section defines an image mathematically, discusses the Joint Photographic Experts Group (JPEG) [14] lossy and Tagged Image File Format (TIFF) [1] lossless image formats, and discusses the interpolation algorithms related to the IMSPs.

2.1.1 Mathematical Image Definition.

An image (I) can be defined as a row (i) x column (j) matrix of pixels values, Equation 2.1 [14]. A single pixel value for digital images relevant to this research consists of three primary color channels, Red-Green-Blue (RGB) [23] where each channel ranges between 0 and 255. A digital camera stores digital images in this

color space format.

JPEG File Interchange Format (JFIF) images can also be represented in YCbCr format [14]. The conversions from RGB to YCbCr and YCbCr to RGB are illustrated in Equation 2.2 and Equation 2.3 [14].

$$I = \begin{bmatrix} 0,0 & 0,1 & \cdots & 0,j-1 & 0,j \\ 1,0 & 1,1 & \cdots & 1,j-1 & 1,j \\ \vdots & \vdots & & \vdots & \vdots \\ i-1,0 & i-1,1 & \cdots & i-1,j-1 & i-1,j \\ i,0 & i,1 & \cdots & i,j-1 & i,j \end{bmatrix} \quad (2.1)$$

$$RGB \Rightarrow YCbCr = \begin{bmatrix} Y = 0.299R + 0.587G + 0.114B \\ Cb = -0.1687R - 0.3313G + 0.5B + 128 \\ Cr = 0.5R - 0.4187G - 0.0813B + 128 \end{bmatrix} \quad (2.2)$$

$$YCbCr \Rightarrow RGB = \begin{bmatrix} R = Y + 1.402(Cr - 128) \\ G = Y - 0.34414(Cb - 128) - 0.71414(Cr - 128) \\ B = Y + 1.772(Cb - 128) \end{bmatrix} \quad (2.3)$$

Many image feature detection techniques use images that have been converted to grayscale. When an image is converted to grayscale, the pixel values are intensity values of black and white. Two generic intensity value formulas are presented in Equation 2.4 and Equation 2.5, where R , G , B represent the respective channel values and $w1$, $w2$, and $w3$ represent a value to weight each channel representation [23]. After the intensity value has been calculated the value replaces each of the channel values in order to get a grayscale value for each pixel. The research in this thesis uses the Equation 2.5, where $w1 = 0.299$, $w2 = 0.587$, and $w3 = 0.114$ [23].

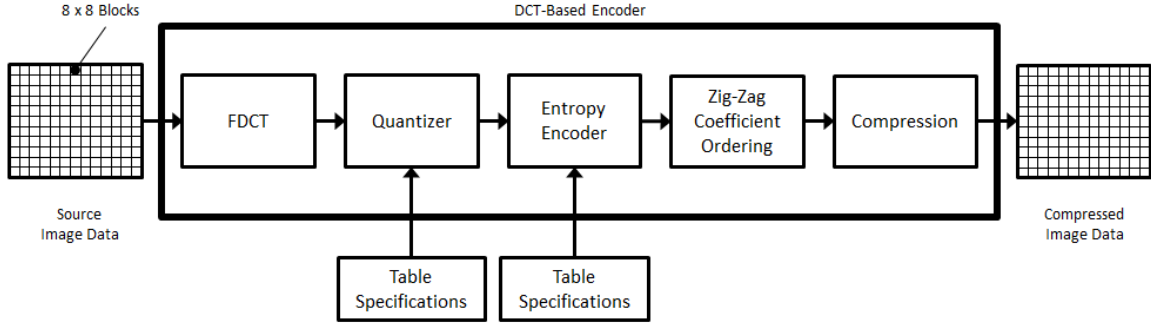


Figure 2.1. DCT Based Encoder Simplified Diagram.

Lossy and lossless image formats are discussed next.

$$intensity\ value = \frac{R + G + B}{3} \quad (2.4)$$

$$intensity\ value = R \cdot w1 + G \cdot w2 + B \cdot w3 \quad (2.5)$$

2.1.2 Lossy and Lossless Image Formats.

There are two main types of image file compression, lossy and lossless [5, 1, 14]. Lossy file compression reduces image quality in order to attain a smaller file size. The JPEG lossy format uses Discrete Cosine Transform (DCT) encoding for compression while still producing a high quality replica of a source image [5]. Figure 2.1 shows an overview of the DCT based encoding process [5]. When an image is processed by lossless compression, a file's size is reduced while still maintaining a pixel value replica of the original file.

The CCITT, a branch of the International Telecommunication Union in the United Nations, developed the JPEG format and describes the format in the publication T.81 [5]. The CCITT describes the processes of formatting an image to a JPEG in Steps. Step 1 decomposes the source image into 8 x 8 blocks. Each block undergoes a transformation by the Forward Discrete Cosine Transform (FDCT) into an 8 x 8

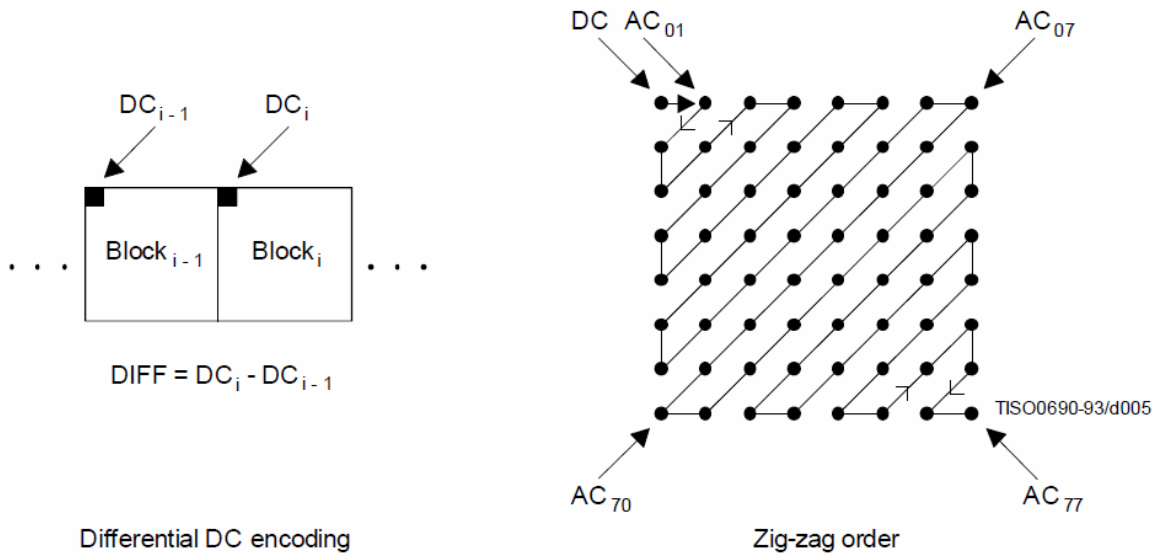


Figure 2.2. Preparation of Quantized Coefficients for Entropy Encoding [5].

matrix composed of 64 DCT coefficient values. The upper left most coefficient in the matrix is called the Discrete Cosine (DC) coefficient. The other 63 coefficients are the AC coefficients. These 64 coefficients are quantized using the values from an 8 x 8 quantization table for color channels and sometimes a 4 x 4 table for non-color channels, Figure 2.1. After quantization, the DC coefficient and the 63 Alternate Coefficients (AC) undergo entropy encoding. The entropy encoding process uses the previous quantized DC coefficient to predict the current quantized DC coefficient and encode the difference. Figure 2.1 illustrates this process. The 63 quantized AC coefficients are converted into a one-dimensional zig-zag sequence, Figure 2.2. After quantization is complete, the entropy encoding process is performed using one of two compression methods, Huffman or arithmetic [5].

The quantization stage uses between one and four tables [16]. An image normally contains two tables. The third table becomes a duplicate of the second when three tables are used. Each block is transformed by the source image into 8 x 8 blocks. The most commonly used tables are those published by the Independent JPEG Group (IJG) [16, 12]. As noted by Kornblum, [16], several of the IMSPs use their own

custom non-IJG tables based on the JPEG quality desired.

The pseudocode in Algorithm 1, illustrates the TIFF lossless Lempel-Ziv and Welch (LZW) compression algorithm [1]. The characters in the LZW strings are bytes containing TIFF uncompressed image data. Omega, Ω , represents the prefix string. InitializeStringTable() initializes a string table of all possible character strings numbered 0 to 255. WriteCode() writes a code to the output stream. GetNextCharacter() retrieves the next character value from the input stream. AddTableEntry() adds a table entry [1].

Algorithm 1 Lempel-Ziv and Welch Compression Algorithm.

```
1: InitializeStringTable()
2: WriteCode(ClearCode)
3:  $\Omega$  = the empty string
4: for each character in the strip do
5:   K = GetNextCharacter()
6:   if  $\Omega+K$  is in the string table then
7:      $\Omega = \Omega+K$  /* string concatenation */
8:   else
9:     WriteCode (CodeFromString( $\Omega$ ))
10:    AddTableEntry( $\Omega+K$ )
11:     $\Omega = K$ 
12:   end if
13: end for
14: WriteCode (CodeFromString( $\Omega$ ))
15: WriteCode (EndOfInformation)
```

The algorithm reduces the file size by concatenating sequences together. For example, there are two dictionaries, Tables 2.1 and 2.2, used to encode the message “weartearbear#”, where the “#” tells the encoder it has arrived at the end of the message. The first dictionary is all letters in the alphabet and the “#” symbol. The second dictionary concatenates the letters together to create an extended dictionary. Using this method reduces the encoding space from the original size of 13 symbols \cdot 5 bits/symbol = 65 bits to 6 codes \cdot 5 bits/code + 4 codes \cdot 6 bits/code = 54 bits.

Table 2.1. TIFF Compression Dictionary 1.

Symbol	Binary	Decimal
#	00000	0
A	00001	1
B	00010	2
C	00011	3
D	00100	4
E	00101	5
F	00110	6
G	00111	7
H	01000	8
I	01001	9
J	01010	10
K	01011	11
L	01100	12
M	01101	13
N	01110	14
O	01111	15
P	10000	16
Q	10001	17
R	10010	18
S	10011	19
T	10100	20
U	10101	21
V	10110	22
W	10111	23
X	11000	24
Y	11001	25
Z	11010	26
	11100	
	⋮	
	11111	

Table 2.2. TIFF Compression Dictionary 2.

Current Sequence	Next Char	Output		Extended Dictionary	Comments
		Code	Bits		
NULL	W				
W	E	23	10111	27: WE	27 = first available code after 0 through 26
E	A	5	00101	28: EA	
A	R	1	00001	29: AR	
R	T	18	10010	30: RT	
T	E	20	10100	31: TE	
EA	R	28	11100	32: EAR	32 requires 6 bits, so for next output use 6 bits
R	B	18	010010	33: RB	
B	E	2	000002	34: BE	
EAR	#	32	100000		# stops the algorithm; send the cur seq and the stop code
		0	000000		

2.1.3 Interpolation Algorithms.

The objective of image interpolation is to “produce acceptable images at different resolutions,” where the image variations are derived from one original [29]. Interpolation refers to a process of approximating new data points within the range of discrete known data points [15]. This research focuses on two positional interpolation algorithm categories, superposition and subposition. Superposition refers to resizing a picture to a larger pixel grid and subposition refers to resizing a picture to a smaller pixel grid. An example of superposition appears in Figure 2.3.

Commercial applications implement three common interpolation algorithms: nearest-neighbor [21], bilinear [21], and bicubic [21]. Each IMSP can implement details of the three algorithms differently resulting in slight variations in the manipulated image. The following explanations are just a few examples of possible implementation.

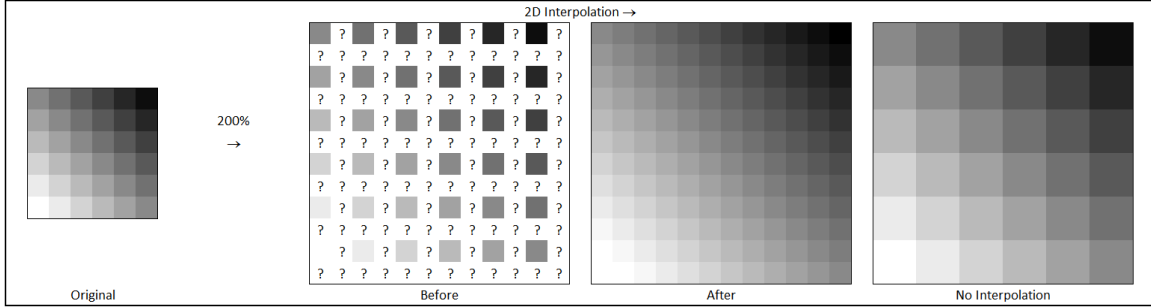


Figure 2.3. Superposition Interpolation.

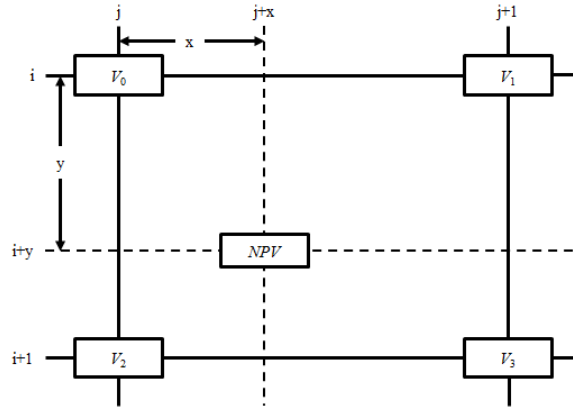


Figure 2.4. Nearest Neighbor Interpolation.

Nearest-Neighbor.

The nearest-neighbor algorithm [21] takes the value of the nearest old pixel coordinate point and sets it to the value of the closest new pixel coordinate point. Applications typically do not implement this algorithm when large images are resized, therefore, it is not a popular choice for IMSPs. Figure 2.4 represents an example of the implementation of the nearest neighbor algorithm for a grayscale image, where the new position value box is at the intersection where the new pixel value is being interpolated. Algorithm 2 represents a method for computing the new pixel value, where x is the distance from the j column, y is the distance from the i row, and NPV is the new pixel value.

Figure 2.4 and Algorithm 2 are one representation of the implementation of the algorithm. The IMSPs can alter the implementation so that a different NPV can be

calculated. For instance, instead of looking at the four surrounding pixels, the IMSP implementation might only examine the horizontal values (i, j) and $i, j + 1$. In this instance, whichever location is closer to the $(i + y, j + x)$ location would set the *NPV* as that location's pixel value.

Algorithm 2 Nearest-Neighbor Interpolation Algorithm.

```

1:  $V_0$  = pixel value at  $(i, j)$ 
2:  $V_1$  = pixel value at  $(i+1, j)$ 
3:  $V_2$  = pixel value at  $(i, j+1)$ 
4:  $V_3$  = pixel value at  $(i+1, j+1)$ 
5:  $NPV$  = pixel value at  $(i+y, j+x)$ 
6: if  $((j + 1) - (j + x)) > x$  and  $((i + 1) - (i + y)) > y$  then
7:    $NPV = V_0$ 
8: else if  $((j + 1) - (j + x)) < x$  and  $((i + 1) - (i + y)) > y$  then
9:    $NPV = V_1$ 
10: else if  $((j + 1) - (j + x)) > x$  and  $((i + 1) - (i + y)) < y$  then
11:    $NPV = V_2$ 
12: else
13:    $NPV = V_3$ 
14: end if

```

Bilinear.

The bilinear interpolation algorithm [21] is a commonly implemented algorithm in IMSPs. However, it has a slightly higher computational cost than the nearest-neighbor. The bilinear interpolation algorithm implementation process begins by interpolating the new pixel point linearly between the old pixel points and then calculating the weighted average value of the four surrounding pixels. Figure 2.5 is an example of the bilinear algorithm for a grayscale image, where the value at the intersection of the dotted lines is the new pixel value. Equation 2.6 is the set of equations for computing the new pixel value, where x is the distance from the j column, y is the distance from the i row, and NPV is the new pixel value.

The bilinear algorithm can be implemented differently across the IMSPs. An

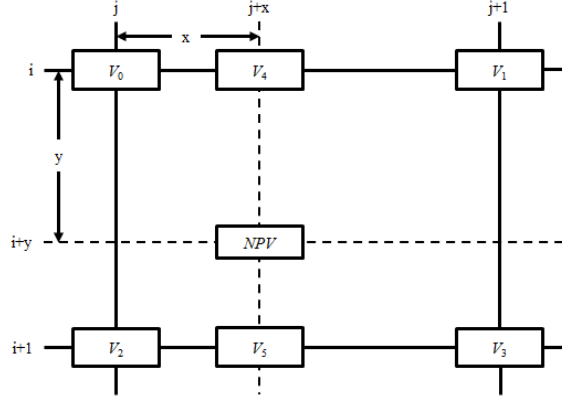


Figure 2.5. Bilinear Interpolation.

alternate implementation could be that all four known pixel values V_0 , V_1 , V_2 , and V_3 are summed together and then divided by four to create the average value and set as the new NPV . Alternatively, a weighting function such as an inverse Gaussian filter could be used to result in a crisper modified image.

$$\left[\begin{array}{l} V_4 = \frac{(j+1)-(j+x)}{(j+1)-j} \cdot V_0 + \frac{(j+x)-j}{(j+1)-j} \cdot V_1 \\ V_5 = \frac{(j+1)-(j+x)}{(j+1)-j} \cdot V_2 + \frac{(j+x)-j}{(j+1)-j} \cdot V_3 \\ NPV = \frac{(i+1)-(i+y)}{(i+1)-i} \cdot V_4 + \frac{(i+y)-i}{(i+1)-i} \cdot V_5 \end{array} \right] \quad (2.6)$$

Bicubic.

Adobe Photoshop (APS) and GNU Image Manipulation Program (GIMP) implement the bicubic algorithm as the default interpolation algorithm. The bicubic algorithm [21] begins by taking the weighted average of the nearest sixteen pixels to set the new pixel value. Figure 2.6 is an example of an implementation of the bicubic algorithm for a grayscale image, where the value at the intersection of the dotted lines is the new pixel value. Equation 2.7 is the set of equations for computing the new pixel value, where x_0 is the distance from the j column, x_1 is the distance from the $j + 1$ column, y_0 is the distance from the i row, y_1 is the distance from the $i + 1$ row, and NPV is the new pixel value.

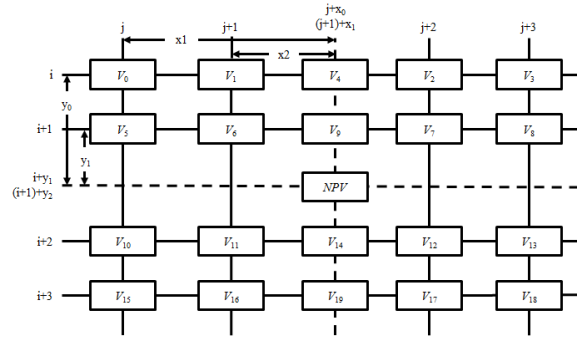


Figure 2.6. Bicubic Interpolation.

The bicubic algorithm can also be implemented differently across the IMSPs. An alternate implementation could be that all sixteen known pixel values are summed together and then divided by sixteen to create the average value and set as the *NPV*.

$$\left[\begin{aligned}
V_5 &= \frac{\|(j+3)-(j+x_0)|-|(j+2)-((j+1)+x_1)|-|((j+1)+x_1)-(j+1)||}{(j+3)-j+(j+2)-(j+1)} \cdot V_0 + \\
&\frac{\|(j+3)-(j+x_0)|-|(j+2)-((j+1)+x_1)|-|(i+x_0)-j||}{(j+3)-j+(j+2)-(j+1)} \cdot V_1 + \\
&\frac{\|(j+3)-(j+x_0)|-|(j+x_0)-j|-|((j+1)+x_1)-(j+1)||}{(j+3)-j+(j+2)-(j+1)} \cdot V_2 + \\
&\frac{\|(j+x_0)-j|-|((j+1)+x_1)-(j+1)|-|((j+1)+x_1)-(j+2)||}{(j+3)-i+(j+2)-(j+1)} \cdot V_3 \\
V_9 &= \frac{\|(j+3)-(j+x_0)|-|(j+2)-((j+1)+x_1)|-|((j+1)+x_1)-(j+1)||}{(j+3)-j+(j+2)-(j+1)} \cdot V_5 + \\
&\frac{\|(j+3)-(j+x_0)|-|(j+2)-((j+1)+x_1)|-|(j+x_0)-j||}{(j+3)-j+(j+2)-(j+1)} \cdot V_6 + \\
&\frac{\|(j+3)-(j+x_0)|-|(j+x_0)-j|-|((j+1)+x_1)-(j+1)||}{(j+3)-j+(j+2)-(j+1)} \cdot V_7 + \\
&\frac{\|(j+x_0)-j|-|((j+1)+x_1)-(j+1)|-|((j+1)+x_1)-(j+2)||}{(j+3)-j+(j+2)-(j+1)} \cdot V_8 \\
V_{14} &= \frac{\|(j+3)-(j+x_0)|-|(j+2)-((j+1)+x_1)|-|((j+1)+x_1)-(j+1)||}{(j+3)-j+(j+2)-(j+1)} \cdot V_{10} + \\
&\frac{\|(j+3)-(j+x_0)|-|(j+2)-((j+1)+x_1)|-|(j+x_0)-j||}{(j+3)-j+(j+2)-(j+1)} \cdot V_{11} + \\
&\frac{\|(j+3)-(j+x_0)|-|(j+x_0)-j|-|((j+1)+x_1)-(j+1)||}{(j+3)-j+(j+2)-(j+1)} \cdot V_{12} + \\
&\frac{\|(j+x_0)-j|-|((j+1)+x_1)-(j+1)|-|((j+1)+x_1)-(j+2)||}{(j+3)-j+(j+2)-(j+1)} \cdot V_{13} \\
V_{19} &= \frac{\|(j+3)-(j+x_0)|-|(j+2)-((j+1)+x_1)|-|((j+1)+x_1)-(j+1)||}{(j+3)-j+(j+2)-(j+1)} \cdot V_{15} + \\
&\frac{\|(j+3)-(j+x_0)|-|(j+2)-((j+1)+x_1)|-|(j+x_0)-j||}{(j+3)-j+(j+2)-(j+1)} \cdot V_{16} + \\
&\frac{\|(j+3)-(j+x_0)|-|(j+x_0)-j|-|((j+1)+x_1)-(j+1)||}{(j+3)-j+(j+2)-(j+1)} \cdot V_{17} + \\
&\frac{\|(j+x_0)-j|-|((j+1)+x_1)-(j+1)|-|((j+1)+x_1)-(j+2)||}{(j+3)-j+(j+2)-(j+1)} \cdot V_{18} \\
NPV &= \frac{\|(i+3)-(i+y_0)|-|(i+2)-((i+1)+y_1)|-|((i+1)+y_1)-(i+1)||}{(i+3)-i+(i+2)-(i+1)} \cdot V_4 + \\
&\frac{\|(i+3)-(i+y_0)|-|(i+2)-((i+1)+y_1)|-|(i+y_0)-i||}{(i+3)-i+(i+2)-(i+1)} \cdot V_9 + \\
&\frac{\|(i+3)-(i+y_0)|-|(i+y_0)-i|-|((i+1)+y_1)-(i+1)||}{(i+3)-i+(i+2)-(i+1)} \cdot V_{14} + \\
&\frac{\|(i+y_0)-i|-|((i+1)+y_1)-(i+1)|-|((i+1)+y_1)-(i+2)||}{(i+3)-i+(i+2)-(i+1)} \cdot V_{19}
\end{aligned} \right] \tag{2.7}$$

Figure 2.7 and Figure 2.8 illustrate the resulting images after undergoing processing by the three IMSPs and their respective interpolation algorithms at a 0.75 interpolation rate. This means that there are 75% of the original number of pixels, 12×12 pixels to 9×9 pixels. Figure 2.7 displays an original image black and white checkerboard with a group of 3×3 pixel squares and the images after the resizing

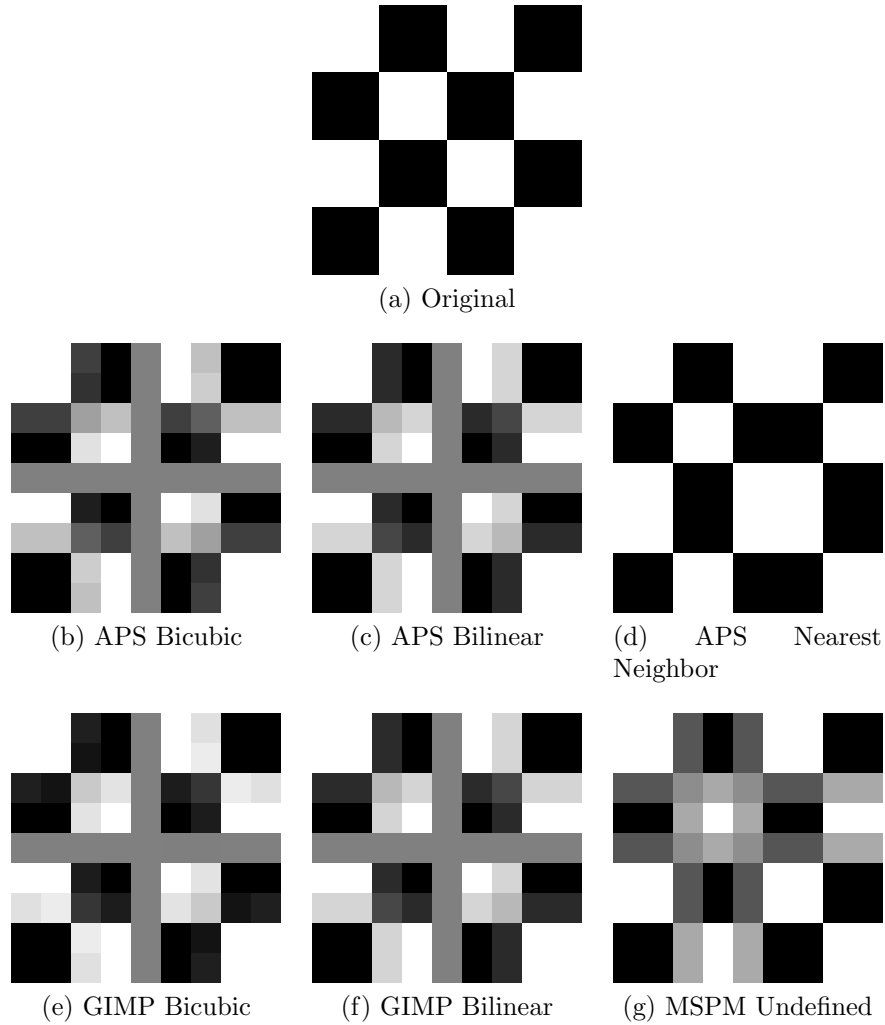
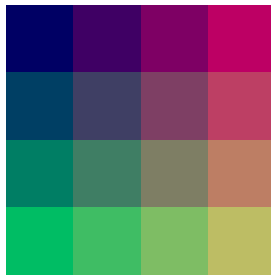


Figure 2.7. Interpolation Algorithms in Black and White.

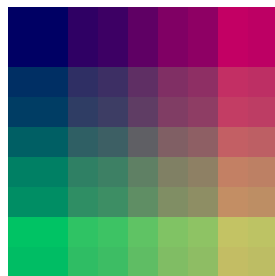
process in grayscale. In order to achieve this all three channels for each pixel were set to either black (0) or white (255).

Figure 2.8 displays an original image with different colored checkerboards with a group of 3 x 3 pixel squares and the images after the resizing process. In order to achieve this, all three channels for each pixel were set to different values to attain different colors for each 3 x 3 pixel square.

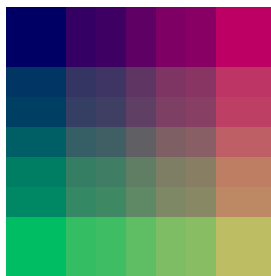
The resulting images result in noticeable differences. APS bicubic appears to smooth the image while the GIMP bicubic is more crisp. Both bilinear images are



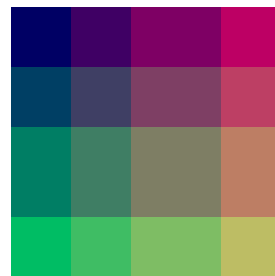
(a) Original



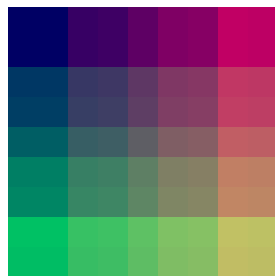
(b) APS Bicubic



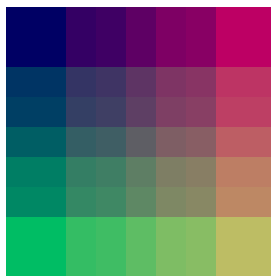
(c) APS Bilinear



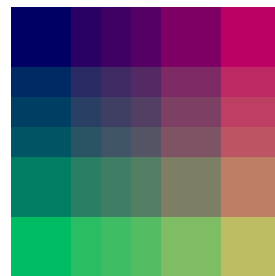
(d) APS Nearest Neighbor



(e) GIMP Bicubic



(f) GIMP Bilinear



(g) MSPM Undefined

Figure 2.8. Interpolation Algorithms in color.

too similar to determine any differences and when examining the RGB channel values the pixels are only off by one or two digits. The APS nearest neighbor and the MSPM undefined algorithms behave like none of the other algorithms.

2.2 Image Manipulation and Detection

Alterations to images occur in many ways: rotation, resize, crop, sharpen, soften, contrast changes, etc. Farid [9] described categories and techniques used to detect image manipulations by looking at how image processing alters the underlying statistics of the images. The two categories (pixel and format) and four techniques (statistical, resampling, header, and double JPEG compression) that are pertinent to this research are discussed in the following sections.

2.2.1 Pixel Based.

In the pixel based category, four different techniques to detect image manipulation by analyzing correlation between pixels are discussed [9]. In order to detect this manipulation, an analysis on the Fourier statistics of the derivative of an image is conducted by looking at the frequency composition of a signal. The main techniques used to detect an altered image within the pixel based category are: cloning, resampling, splicing, and statistical.

2.2.2 Format Based.

The format based category focuses on the JPEG lossy compression scheme and image headers [9]. Cameras and image manipulation software programs can use different quantization tables when compressing images. By extracting the table from the image, a forensic examiner can compare the extracted table from tables known to be used by a suspect's cameras or image manipulation software to determine what

device took the photo or what software altered the image. As mentioned previously, JPEG compression is based on block Discrete Cosign Transform (DCT) where each 8 x 8 pixel image block is individually transformed and quantized. Artifacts are created as edges at the border of these blocks. JPEG blocking occurs when an image is altered and then re-compressed leaving behind a new set of blocking artifacts that may not align with the original image artifacts. These misaligned blocks can be detected by the examiners [9, 18, 22]. Farid discusses three techniques that utilize the JPEG lossy compression scheme: JPEG quantization, double compression, and blocking. A simpler technique analyzes the JPEG headers, specifically the data within the header and the format of the headers [17, 3].

2.2.3 Camera Based.

Farid equates the camera based category to bullet ballistics analysis in which a bullet fired by a specific weapon can be traced back to that weapon by unique marks created on the bullet by grooves within the barrel [9]. When modeling and estimating camera artifacts, inconsistencies can be found during the analysis of the suspected altered image. The four techniques described in [9] are chromatic aberration, color filter array, camera response, and sensor noise.

2.2.4 Physical Based.

The physical based category includes three techniques that are based on determining the lighting environments in which photographs were taken [9]. Differences in lighting on parts of an image can show that an image has been altered. The techniques relevant to this category are 2-dimensional light direction, 3-dimensional light direction, and the light environment.

2.2.5 Geometry Based.

In his discussion of the geometry based category, Farid includes examining an image and how geometric shapes should appear in the images. Light principle point and metric measurement are techniques relevant to the geometric based category [9]. These techniques use geometrics of surrounding objects in the image as well as the knowledge that the projection of the camera's center is near the image's center. The next section discusses the individual feature detection.

2.3 Feature Detection

Fontani, et al. [10] describes a method of detecting image splicing manipulations using existing detection techniques. The authors created a framework that utilizes existing image manipulation detection techniques and fuses the results of each algorithm together to return an easily readable output to an image forensic analyst. Fontani, et al. treat each technique as an expert in its focused area and used a decision fusion engine to determine trace relationships between the techniques [10].

Avcibas, et al. [2] discuss the use of detectable features to design a classifier to discriminate between an original and manipulated image. Avcibas, et al. focused on the idea that most manipulated images have undergone some sort of standard image processing operation such as scaling, rotation, and brightness adjustment. The authors developed a group of classifiers that were able to discern between an original image and an altered image by determining whether one or more of these image processing operations were conducted on the altered image.

Bayram, et al., [4] builds on Avcibas, et al.'s, [2] work and discusses how to detect image manipulations by looking at different operations in image processing and detected image feature alterations. By developing tools that individually examined features common to these operations and then fusing the results of the individual

analysis together, the authors created over 100 features to detect when images had been manipulated. Bayram, et al. then grouped the features into three categories: Image Quality Measures (IQM), Higher Order Wavelet Statistics (HOWS), and Binary Similarity Measures (BSM). IQM features analyze the differences between the original and the altered image. HOWS features decompose images by filters and compute the mean, variance, skewness, and kurtosis of the sub-band coefficients at each orientation and scale. The BSM features analyze the correlation and texture properties between and within the lower significant bit planes.

The authors in [4] grouped these categories into two sets, the Joint Feature Set (JFS) and Core Feature Set (CFS). An examiner can then choose specific features. The JFS includes 188 features: 108 BSM, 72 HOWS, and 8 IQM. The CFS is a smaller population of features that were selected using the Sequential Forward Feature Selection (SFFS). Pudil, et al. [27] discussed the construction of the best feature set by adding to and/or removing from the current set of features until no more performance improvement is possible.

The research presented in this thesis combines parts of these authors' research and the techniques outlined in the next section to develop a framework of modules that have the capability of being used to determine which interpolation algorithm, rate, and image manipulation software program was used.

2.4 Implemented Techniques

The techniques used in this research are grouped within Farid's [9] categories. Resampling and statistical are in the Pixel category, while header format and double JPEG compression are in the Format category. Each technique is discussed in one of the next four sections.

2.4.1 Header Format.

When analyzing the data within the header and the format of the header, the Exchange image file format (Exif), [7], includes information that can reveal which software manipulated a suspect image. For the purposes of this research, file metadata includes the information found in the Exif data. As first noted by Levine and Liberatore [17], an IMSP can add its name to the file metadata or it could rearrange the file metadata to a format of its choosing. Levine and Liberatore [17] and Ball and Keefer [3] note that while this technique is simple in its execution, it is easy to manipulate the header to provide false information about the suspect software. Therefore, the results of analyzing the file metadata should not be the only source of data in determining image manipulation software.

As noted by the authors in [3, 17], file metadata in the header is normally added by the device (camera most often) that took the picture [7]. Some IMSPs alter the file metadata by signing its name in the “software” tag. Ball and Keefer [3] developed a method of extracting that data and analyzing it for the signatures added by the image manipulation software programs. As noted prior, this method is not 100% accurate as the file metadata can be manipulated by using a HEX editor or one of the many file metadata manipulation tools available online.

Ball and Keefer [3] developed a method similar to the previously described method to examine the header format of the image files. Some of the IMSPs add various APP headers that are unique to specific programs. The authors’ method detects these APP headers and can analyze to determine a probable IMSP [3].

2.4.2 Resampling.

Resampling is the resizing, rotation, or stretching of part of an image. Parker, et al. [21] defines re-sampling as “the process of transforming a discrete image which is

defined at one set of coordinate locations to a new set of coordinate locations.” This alteration occurs when the original image is laid onto a new sampling lattice. The key element in detection of this manipulation is the unnatural periodic correlations between neighboring pixels occurring on this new lattice [9]. Note that detection of resampling of images focuses on the pixel domain [26] and the frequency domain [25]. This paper’s research focuses on the pixel domain. Popescu and Farid [24, 25] describe the resampling process of a one dimensional signal, $x[t]$, with m samples by a factor of p/q to n samples in three steps. Popescu and Farid’s steps are outlined below:

1. *Up-sample*: create a new signal $x_u[t]$ with pm samples, where $x_u[pt] = x[t]$, $t = 1, 2, \dots, m$, and $x_u[t] = 0$.
2. *Interpolate*: convolve $x_u[t]$ with a low-pass filter: $x_i[t] = x_u[t] \star h[t]$.
3. *Down-sample*: create a new signal $x_d[t]$ with n samples, where $x_d[t] = x_i[qt]$, $t = 1, 2, \dots, n$. Denote the resampled signal as $y[t] \equiv x_d[t]$.

Where, x_u is the upsampled signal, x_i is the interpolated signal, and x_d is the downsampled signal. Depending on the re-sampling algorithm, linear or cubic, the interpolation filter $h[t]$ will differ. These steps can be re-written in vector form $\vec{y} = A_{p/q}\vec{x}$ where \vec{x} is the original image, \vec{y} is the re-sampled signal, and $A_{p/q}$ is an $m \times n$ matrix representing the process outlined above in a two dimensional space [24].

Popescu and Farid [24] discuss a method of detecting resampling using an implementation of the Expectation Maximization (EM) algorithm to estimate a set of periodic samples that are correlated to their neighbors and the form of these correlations. The EM Algorithm 3 used in the authors’ research is outlined below, where N is the neighborhood size, $\vec{\alpha}$ is the set of weights that satisfy Equation 2.8.

$$\vec{a}_i = \sum_{k=-N}^N \alpha_k y_{i+k} \quad (2.8)$$

In Equation 2.8, a_i is the i th row of the resampling matrix, and $i = 3, 7, 11, \dots$ etc, if the correlations, $\vec{\alpha}$, are known then solve for y_i using the Equation 2.9 and σ is the variance.

$$y_i = \sum_{k=-N}^N \alpha_k y_{i+k} \quad (2.9)$$

$$Y = \begin{bmatrix} y_1 & \cdots & y_N & y_{N+2} & \cdots & y_{2N+1} \\ y_2 & \cdots & y_{N+1} & y_{N+3} & \cdots & y_{2N+2} \\ \vdots & & \vdots & \vdots & & \vdots \\ y_i & \cdots & y_{N+i-1} & y_{N+i+1} & \cdots & y_{2N+i} \\ \vdots & & \vdots & \vdots & & \vdots \end{bmatrix} \quad (2.10)$$

After processing the original and scaled one-dimensional signals with the EM algorithm, Popescu and Farid [24] show how the probability of a sample is related to its neighbors. The probability a sample is related to its neighbors becomes periodic when interpolated. While the authors focus on one-dimensional signals that have been interpolated with a linear formula, they note in their work that this method can be modified to include the two-dimensional space and additional interpolation algorithms. The EM algorithm creates a two-dimensional probability map that represents the spatial correlations in the image. The Discrete Fourier Transform (DFT) is then computed for the probability map to show the frequency representation of the spatial correlations.

Prasad and Ramakrishnan [26] discuss a technique in the pixel domain that is relevant to the research presented in this paper. In this technique, the authors analyzed

Algorithm 3 Expectation-Maximization Algorithm.

```
1: /* Initialize */
2: choose a random  $\vec{\alpha}_0$ 
3: choose  $N$  and  $\sigma_0$ 
4: set  $p_0$  to the reciprocal of the range of the signal  $\vec{y}$ 
5: set matrix  $Y$  /* equation (2.10) */
6: set  $h$  to be a binomial low-pass filter of size  $(N_h \times N_h)$ 
7:  $n = 0$ 
8: repeat
9:   /* Expectation Step */
10:  for each sample  $i$  do
11:     $R(i) = |y(i) - \sum_{k=-N}^N \alpha_N(k)y(i+k)|$  /* residual */
12:  end for
13:   $R = R \star h$  /* spatially average the residual error */
14:  for each sample  $i$  do
15:     $P(i) = (\frac{1}{\sigma_n \sqrt{2\pi}}) e^{-\frac{R(i)^2}{2\sigma_n^2}}$  /* conditional probability */
16:     $w(i) = \frac{P(i)}{P(i)+p_0}$  /* posterior probability */
17:  end for
18:  /* Maximization Step */
19:   $W = 0$ 
20:  for each sample  $i$  do
21:     $W(i, i) = W(i)$  /* weighting matrix */
22:  end for
23:   $\sigma_{n+1} = (\frac{\sum_i w(i)R^2}{\sum_i w(i)})^{\frac{1}{2}}$  /*
24:  new variance estimate */
25:   $\vec{\alpha}_{n+1} = (Y^T W Y)^{-1} Y^T W \vec{y}$  /* new estimate */
26:   $n = n + 1$ 
27: until ( $\|\alpha_n - \alpha_{n-1}\| < \epsilon$ )
```

the zero-crossings of the second difference of a resampled image. Prasad and Ramakrishnan first construct a second difference sequence of the grayscale of an image and then find the zero crossings along a row. The binary sequence, $p[j]$, is constructed using Equation 2.11, where I'' is the second derivative of a row of the image and j is the pixel location in the image's row. The binary sequence, $p[j]$, is used to compute the DFT in order to convert it into the frequency domain. The magnitude of the resulting DFT is analyzed for spikes that indicate if a resizing manipulation has occurred. A candidate spike is the local maximum and is t times larger than its local average magnitude. This method can also be altered to analyze the two-dimensional space of the entire image. A periodicity will occur if the image has been resampled.

$$p[j] = \begin{cases} 1, & \text{if } I''[j] = 0 \\ 0, & \text{otherwise} \end{cases} \quad (2.11)$$

Mahdain and Saic [19] present a method of examining the affine transformation. The authors conduct this examination in four steps: region of interest selection, second derivative computation, radon transformation, and search for periodicity. For the region of interest selection, Mahdain and Saic select a block of pixels, $b(i, j)$ (block of $i \times j$ pixels), suspected of having undergone resampling. If a selection cannot be made then the image can be composed of overlapping blocks, $b_i(i, j)$. Each block, $b(i, j)$, then has its second derivative $D''b(i, j)$ taken. Mahdain and Saic note that similar results can be achieved using other derivative orders. To compute the radon transform, Mahdain and Saic used Equation 2.12 [19].

$$p_\theta = \int_{-\infty}^{\infty} |D''b(i, j)| \cdot (i' \cos \theta - j' \sin \theta, i' \sin \theta + j' \cos \theta) dj' \quad (2.12)$$

Theta, $p(\theta)$ is 180 one-dimensional vectors and θ denotes the orientation of the x' axis counterclockwise from the x axis [19]. In order to determine if resampling has

occurred, $p(\theta)$ is searched for periodicity in the autocovariance of these vectors using Equation 2.13:

$$R_{p_\theta}(k) = \sum_i (p_\theta(i+k) - \bar{p}_\theta)(p_\theta(i) - \bar{p}_\theta) \quad (2.13)$$

Mahdian and Saic applied a derivative filter of order one to emphasize and detect periodicity in p_θ [19]. The DFT is computed on this sequence and the magnitude of the resulting DFT is analyzed for spikes that indicate a resizing manipulation has occurred. A candidate spike is the local maximum and is t times larger than its local average magnitude.

Ouwerkerk examined several super-resolution algorithms to determine differences in their implementation and resulting images [20]. The author performed objective measuring tests using various error measures and subjective tests looking at performance in edge blurring, edge blocking, and generation of detail. Each test image was first decimated using the same filter and down sampling step. Then, each of the decimated images were up sampled and then processed by a super-resolution algorithm. The resulting images are examined using objective and subjective tests. The objective tests included evaluating the Peak Signal to Noise Ratio (PSNR) based on the Mean Square Error (MSE), Structural Similarity (SIMM) error measure, and edge stability error measures using edge detectors. The subjective tests involve the author examining a portion of the super-resolution images to determine how much edge blurring and blocking were introduced. The author also determined which image appeared to keep the most detail in the images [20].

2.4.3 Statistical.

Researchers also analyze the statistical properties of an image to determine if an image has been manipulated, in particular, the higher-order statistics from a wavelet

decomposition [9]. This technique can be used to detect several different manipulation methods. Farid [9] discusses how resizing, filtering, discerning between photographic and computer generated images, and steganography can all be detected by this method.

Gallagher [11] demonstrated a two-dimensional second derivative method of determining the interpolation factor with which an image had been manipulated. The two-dimensional second derivative method calculates the second derivative of an image for each row and then averages each row to create a one-dimensional frequency histogram. The resulting sequence is then used to compute the DFT. Similar to the One-Dimensional Zero Crossing method developed by Prasad and Ramakrishnan [26], the magnitude of the resulting DFT is analyzed for spikes that indicate a resizing manipulation has occurred. A candidate spike is the local maximum and is t times larger than its local average magnitude.

2.4.4 Double JPEG Compression.

The steps to manipulate images by IMSPs are: import an original image in the software, manipulate it, and then resave the image. In the process of resaving the image, a second compression of the JPEG occurs, hence the name of the detection technique: double JPEG compression. If an image was not cropped, detectable artifacts are embedded in the image [9, 25].

In order to understand double JPEG compression, it is important to first understand how quantization and dequantization work. Section 1.1 presented the JPEG lossy image format, the process of compression by the quantization of the DCT coefficients, c . Farid describes the process as: $q_a(c) = \lfloor c/a \rfloor$, where a is the quantization step (positive integer), [9]. Dequantization is converting the new values back to the original values: $q_a^{-1}(c) = ac$. Double compression is the double quantization using

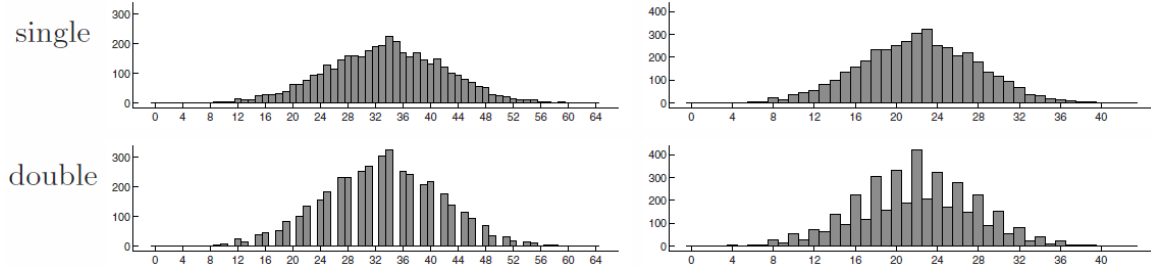


Figure 2.9. Single and Double Quantization [25].

the equation $q_{ab}(c) = \lfloor \lfloor c/b \rfloor b/a \rfloor$, where a and b are the quantization steps [9].

Farid represents double quantization in a sequence of three steps [9]:

1. Quantization with step b .
2. Dequantization with step b .
3. Quantization with step a .

Popescu and Farid examined histograms of an image's DCT coefficients [25]. Periodicity occurs within the histograms when an image has undergone double quantization. The authors note that the presence of double quantization artifacts does not mean an image has been altered, only that it has been saved more than once. Figure 2.9 displays both single and double quantization of a one-dimensional signal. The left column is quantized with Step 3 then 2 and the right column uses Step 2 then 3.

Kornblum [16] developed a software library, Calvin, that extracts quantization tables from images and matches the tables to standard tables, extended tables, and tables from APS. Kornblum, used this library to determine if an image was taken by a specific camera. The author's process can also be used to determine what IMSP was used to alter an image. By matching the tables used in a image and tables known to be used by an IMSP, the correlation can be made between the IMSP and the image.

2.5 Classification Overview

Whitten, et al. [28] define data mining as “discovering patterns in data” to solve problems. This research used the open source University of Waikato software data mining program Waikato Environment for Knowledge Analysis (WEKA) [13]. The classifier examined in this research is the WEKA Bayesian Network algorithm. The Bayesian network implemented in this research uses the K2 learning algorithm, first introduced by Cooper and Herskovits [8]. The K2 algorithm orders the features, then processes each feature and considers adding edges from previously processed features to the current one. The added edge must maximize the networks score until there is no additional improvement. Whitten, et al. state that since the results depend on the initial ordering, it is advisable to run the algorithm with different random orderings.

Whitten, et al. [28] discuss a ten fold cross-validation as a concept where the entire dataset is randomly partitioned into ten parts. Each part should hold approximately the same proportion of each class as the whole dataset. The machine learning algorithm processes the dataset ten times using nine of the parts for training and one part for testing. Each part should be tested once and then the statistics for each test combined to create the overall test statistics. Whitten, et al. discuss how the ten fold cross-validation should be performed ten times on the dataset with different random number seeds for partitioning of the dataset. This will randomize how the parts are distributed. The statistics for the overall experiment can be extracted from these ten runs.

2.6 Summary

This chapter first defined an image as a two dimensional array and then lossy and lossless compression implementation was discussed. The three relevant interpolation algorithms were discussed next, nearest-neighbor, bilinear, and bicubic. The

colorspace of a pixel in RGB and the black and white conversion were outlined. Also, this chapter discussed the literature on the four image manipulation detection techniques implemented in this research: the Two-Dimensional Second Derivative, One-Dimensional Zero Crossings, Quantization Matrices Identification, and File Metadata analysis. Finally, the classifier used in this research was briefly given an overview. The next chapter discusses the focus of this research, the framework of modules.

III. Methodology

To identify artifacts present in images after alteration by an Image Manipulation Software Program (IMSP) requires identification using preexisting techniques: Two-Dimensional Second Derivative, One-Dimensional Zero Crossings, Quantization Matrices Identification, and File Metadata analysis [11, 16, 17, 26]. These four modules are used to generate features using the data within an image file. Then, these features are sent to a Bayesian Network classifier.

This chapter first discusses the basic flow of the framework. Then each module and the features generated from them are described in detail. Finally, the implemented classifier and the correlation of the classifier's output are discussed. The detection framework is discussed in the next section.

3.1 Detection Framework

To evaluate the hypothesis that the proposed framework improves an examiners' ability to classifying an image to which IMSP created the manipulated image beyond the probability of a random guess a framework is created that classifies a manipulated image to what IMSP, interpolation algorithm, and rate was used to alter an image. Subsequently, a digital image forensic analyst can reconstruct how an image was manipulated and by what IMSP by utilizing this framework. This is an important step in reconstructing a sequence of events in an investigation by law enforcement agencies.

The developed framework leverages existing image manipulation detection techniques used to detect specific manipulations conducted on digital images. An image suspected of undergoing manipulation by an IMSP is processed by three to four modules. Three modules are used for the Tagged Image File Format (TIFF) images and

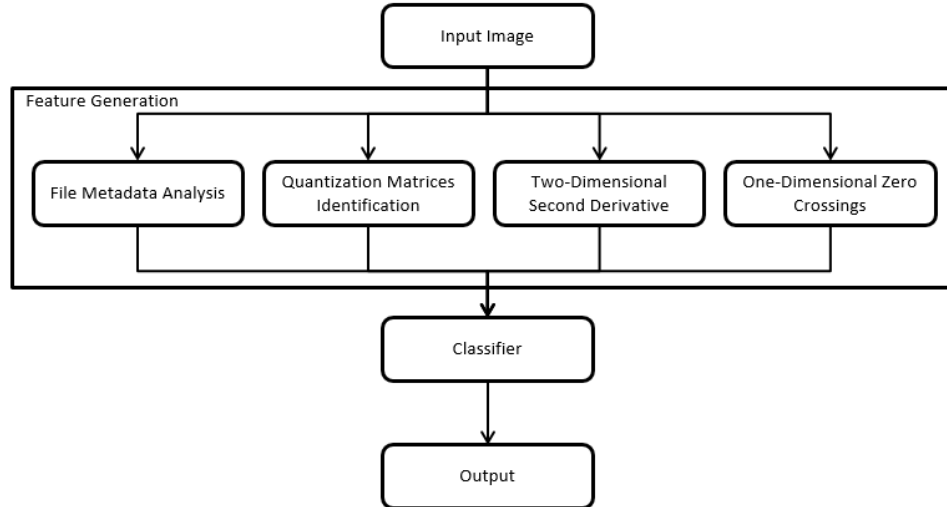


Figure 3.10. Framework Process Flow.

four modules are used for the Joint Photographic Experts Group (JPEG) images. The first module detects the artifacts present in the file metadata left behind by an IMSP [17]. The second module is used on only the JPEG images and it determines the Quantization matrix used on the image [16]. The third module examines the averages across the columns of the second derivative of an image [11]. Finally, the fourth module examines the zero crossings of the second derivative of an image [26]. Figure 3.10 depicts the process flow for the framework.

An assumption of this framework is that the analyst does not have any prior knowledge of which IMSP was used. A suspect image is first processed individually by each of the four modules. The modules identify features present after an image has been manipulated. The features identified are then processed by a classifier. The classifier’s output is a prediction as to which IMSP was used to manipulate an image. The next section discusses in detail the modules and what features are identified by each.

3.2 Feature Generation

This section describes the modules and provides the pseudocode for each. Each module returns features identified from the suspect image. These features are artifacts left behind in the file metadata or image statistics after a manipulation has occurred. Different types of image alterations leave different artifacts within an image. As mentioned in Chapter 2, this research focuses on interpolating an image and, therefore, the modules presented here focus on identifying the artifacts left behind by interpolation operations. Each of the modules are detailed next.

3.2.1 File Metadata.

During the processing of a image by an IMSP, the file metadata in an image is often altered to leave a signature of the IMSP. The file metadata module inspects the file metadata of an image and performs string matching against a predefined known signature of an IMSP. See Algorithm 4 for the file metadata module pseudocode. An image's file metadata is extracted and then each string from an array of a predefined signature is compared to the extracted data. If a match is found, the module returns the IMSP that corresponds to the matching signature.

As noted in Chapter 2, the file metadata is easily manipulated with a HEX editor or other software available for download from the internet. The next sections describe modules that inspect the statistics of an image that are more difficult to alter.

3.2.2 Quantization Matrices Identification.

Many IMSPs use different quantization matrices. For example, Adobe PhotoShop (APS) uses 12 proprietary quantization matrices, one for each JPEG compression and quality level. The Quantization Matrices Identification (QMI) module extracts the quantization matrix from the image and matches it to known quantization matrices.

Algorithm 4 File Metadata.

```
1: /* Initialize and populate signature arrays */
2: Signature1FileMetadata[]
3: Signature2FileMetadata[]
4: Signature3FileMetadata[]
5: /* Extract File Metadata */
6: ImageFileMetadata = GetFileMetadata(Image)
7: /* Perform string matching for each signature */
8: for each element(i) in Signature1FileMetadata do
9:   if Signature1FileMetadata[i] is in ImageFileMetadata then return Signa-
   ture1
10:  end if
11: end for
12: for each element(j) in Signature2FileMetadata do
13:   if Signature2FileMetadata[j] is in ImageFileMetadata then return Signa-
   ture2
14:  end if
15: end for
16: for each element(k) in Signature3FileMetadata do
17:   if Signature3FileMetadata[k] is in ImageFileMetadata then return Signa-
   ture3
18:  end if
19: end for
```

There are currently only 14 signatures included in the module. See Algorithm 5 for the QMI module pseudocode.

Algorithm 5 Quantization Matrices.

```
1: /* Initialize and populate signature arrays */
2: Signature1Qmatrix[]
3: Signature2Qmatrix[]
4: Signature3Qmatrix[]
5: /* Extract qmatrix data */
6: ImageQmatrix = string(Image.quantization)
7: /* Perform string matching for each signature */
8: if Signature1Qmatrix[i] is in ImageQmatrix then return Signature1
9: end if
10: if Signature1Qmatrix[j] is in ImageQmatrix then return Signature2
11: end if
12: if Signature1Qmatrix[k] is in ImageQmatrix then return Signature3
13: end if
```

The quantization matrices embedded in an image can be easily altered by saving the altered image with a different IMSP than the IMSP used to manipulate the image.

3.2.3 Two-Dimensional Second Difference.

The Two-Dimensional Second Difference (2DSD) module examines the averages across the columns of the second derivative of an image. See Algorithm 6 for the 2DSD module pseudocode. The grayscale of an image is first read into memory as a matrix, $I[i, j]$, of the pixel values. The Algorithm then computes the second derivative of an image, $I''[i, j]$, across the rows. Next, the mean of the magnitude of the column values are computed and placed into a one-dimensional array, $avg[k]$, where k is the index value of the array. This sequence is then used to compute a Discrete Fourier Transform (DFT) representation, dft , of the data in the frequency domain. After converting the absolute values of the dft to integers, the sequence is searched for the 250 largest spikes while ignoring the lower and upper two percent of DFT frequencies. The frequencies are normalized to values between 0 and 1. The frequencies are then

classified as candidate peaks if the frequency's magnitude is a local maximum and is T times larger than its local average magnitude. For the 2DSD module the local maximum area is the surrounding one percent of frequencies and $T = 5$. The 2DSD and 1DZC modules can each generate 1,000 possible features from an image when the DFT frequencies are normalized from 0 to 1 and when the accuracy is limited to the thousandths decimal place. The frequencies that occur from JPEG compression are ignored: $\frac{1}{8}$, $\frac{1}{4}$, $\frac{3}{8}$, $\frac{5}{8}$, $\frac{3}{4}$, and $\frac{7}{8}$.

Algorithm 6 Two-Dimensional Second Difference.

```

1: Read grayscale of image  $I[i, j]$  into memory
2: for  $i$  in  $I[i, j]$  do
3:   for  $j$  in  $I[i, j]$  do
4:      $I'[i, j] = |I[i, j] - I[i, j + 1]|$ 
5:   end for
6: end for
7: for  $i$  in  $I'[i, j]$  do
8:   for  $j$  in  $I'[i, j]$  do
9:      $I''[i, j] = |I'[i, j] - I'[i, j + 1]|$ 
10:  end for
11: end for
12: for  $j$  in  $I''[i, j]$  do
13:    $avg[k] = \langle \sum_{i=0}^{\text{number of rows}} |I''[i, j]| \rangle$ 
14: end for
15:  $dft = \text{DFT}(avg[k])$ 
16:  $dft = dft.\text{real}$ 
17:  $dft = \text{np.absolute}(dft.\text{astype}(\text{int}))$ 
18:  $lowlimit = dft.\text{size} * .02$ 
19:  $highlimit = dft.\text{size} * .98$ 
20:  $peaks = -\text{bn.partsort}(-dft[lowlimit : highlimit], 250)[: 250]$ 
21:  $peaks = \text{np.sort}(peaks)$ 
22:  $candiditepeaks = \text{findPeaks}(dft, peaks, dftsize, lowlimit, highlimit)$ 

```

3.2.4 One-Dimensional Zero Crossings.

The One-Dimensional Zero Crossings (1DZC) module inspects where an image's second derivative pixel value is zero across a single row. See Algorithm 7 for the

1DZC module pseudocode. Similar to the 2DSD module, the 1DZC module starts by reading the grayscale of an image into memory as a matrix, $I[i, j]$, of the pixel values. Next, the algorithm computes the second derivative of the middle row of an image, $I''[\frac{i}{2}, \text{all } j]$. Then the row, $I''[\frac{i}{2}, \text{all } j]$, is searched for zero crossings and a new sequence, $Z[k]$, is created using Equation 3.14, where the location k in the sequence $Z[k]$ corresponds to the location j in the row $I''[\frac{i}{2}, j]$.

$$Z[k] = \begin{cases} 1, & \text{if } I''[\frac{i}{2}, j] = 0 \\ 0, & \text{otherwise} \end{cases} \quad (3.14)$$

This sequence, $Z[k]$, is then used to compute a DFT representation, dft , of the data in the frequency domain. After converting the absolute values of the dft to integers, it is searched for the 250 largest spikes while ignoring the lower and upper two percent of DFT frequencies. The frequencies are normalized to values between 0 and 1. The frequencies are then classified as candidate peaks if the frequency's magnitude is a local maximum and is T times larger than its local average magnitude. For the 1DZC module, the local maximum area is the surrounding three percent of frequencies and $T = 4$. Similar to the 2DSD module, the frequencies that occur from JPEG compression are ignored and total possible 1,000 features can be generated.

3.3 Classification

After the modules have generated the features of an image, the features are processed by a classifier. The output of the classifier is the determination of which IMSP manipulated the image.

This research used the open source University of Waikato software data mining program Waikato Environment for Knowledge Analysis (WEKA) [13]. The classifier examined in this research is the WEKA Bayesian Network algorithm. The WEKA

Algorithm 7 One-Dimensional Zero Crossings.

```
1: Read grayscale of image  $I[i, j]$  into memory
2: for  $j$  in  $I[i/2, j]$  do
3:    $I'[j] = |I[i, j] - I[i, j + 1]|$ 
4: end for
5: for  $j$  in  $I'[j]$  do
6:    $I''[j] = |I'[j] - I'[j + 1]|$ 
7: end for
8: for  $j$  in  $I''[j]$  do
9:    $k = j$ 
10:  if  $I''[j] == 0$  then
11:     $Z[k] = 1$ 
12:  else
13:     $Z[k] = 0$ 
14:  end if
15: end for
16:  $dft = \text{DFT}(Z[k])$ 
17:  $dft = dft.\text{real}$ 
18:  $dft = \text{np.absolute}(dft.\text{astype(int)})$ 
19:  $lowlimit = dft.\text{size} * .02$ 
20:  $highlimit = dft.\text{size} * .98$ 
21:  $peaks = -\text{bn.partsort}(-dft[lowlimit : highlimit], 250)[: 250]$ 
22:  $peaks = \text{np.sort}(peaks)$ 
23:  $candiditepeaks = \text{findPeaks}(dft, peaks, dftsize, lowlimit, highlimit)$ 
```

Bayesian Network algorithm was chosen for implementation because it returned the best results when higher valued features were not present in an image, specifically features extracted by the file metadata and QMI modules. The J4.8 classifier was a contender for possible implementation, but when the file metadata was not present the classifier classified all images as APS. The Bayesian Network algorithm did not do this. The next chapter will discuss the experiments and results of this research.

3.4 Summary

The Two-Dimensional Second Derivative, One-Dimensional Zero Crossings, Quantization Matrices Identification, and File Metadata analysis modules facilitate the identification of artifacts present in images after alteration by an IMSP [11, 16, 17, 26]. These four modules are used to generate features using the data within an image file. After feature generation the Bayesian Network classifier classifies an image using the features. The next chapter discusses experimentation on the methods presented in this chapter.

IV. Experiments and Results

To evaluate the framework presented in Chapter 3, a series of tests were performed to determine if the framework is able to classify an image better than a random guess. In addition to classifying an image into the correct class, the classification of an image to the correct algorithm and rate is also examined through the results of the experiment.

The overall experiment process is discussed first. The first step in the experiment process begins with building the databases of images. The features are then generated from the images by the framework of modules. Classification of the images into 48 classes occurs next and statistics of the accuracy of the classification. This is followed by hypothesis testing to determine the statistical significance of the results. The analysis of the results are discussed after the classification. Finally, the results are concluded and a summary of the findings are discussed. The next section discusses the experiments.

4.1 Experiments

The experiments in this research were used to determine the framework's accuracy. Two datasets were built for the purpose of training and testing the two image format specific classifiers: one consists of only JPEG images and the other of only TIFF images. One hundred images were taken using the Nikon D5100 digital camera in the Nikon RAW format where the image's original dimensions are 3264×4928 pixels. The images were a combination of both outdoor scenes and indoor scenes. The intent in selection of each image was to have a variety of images. The RAW format was used to allow for conversion of the images to both JPEG or TIFF images.

After conversion to the two image formats using the Nikon Picture Image Control

Table 4.3. Image Modification Settings.

GIMP-0.50-bilinear	GIMP-0.50-bicubic	MSPM-0.50-undefined	APS-0.50-bilinear	APS-0.50-bicubic	APS-0.50-nearest neighbor
GIMP-0.66-bilinear	GIMP-0.66-bicubic	MSPM-0.66-undefined	APS-0.66-bilinear	APS-0.66-bicubic	APS-0.66-nearest neighbor
GIMP-0.75-bilinear	GIMP-0.75-bicubic	MSPM-0.75-undefined	APS-0.75-bilinear	APS-0.75-bicubic	APS-0.75-nearest neighbor
GIMP-0.90-bilinear	GIMP-0.90-bicubic	MSPM-0.90-undefined	APS-0.90-bilinear	APS-0.90-bicubic	APS-0.90-nearest neighbor
GIMP-1.10-bilinear	GIMP-1.10-bicubic	MSPM-1.10-undefined	APS-1.10-bilinear	APS-1.10-bicubic	APS-1.10-nearest neighbor
GIMP-1.25-bilinear	GIMP-1.25-bicubic	MSPM-1.25-undefined	APS-1.25-bilinear	APS-1.25-bicubic	APS-1.25-nearest neighbor
GIMP-1.33-bilinear	GIMP-1.33-bicubic	MSPM-1.33-undefined	APS-1.33-bilinear	APS-1.33-bicubic	APS-1.33-nearest neighbor
GIMP-2.00-bilinear	GIMP-2.00-bicubic	MSPM-2.00-undefined	APS-2.00-bilinear	APS-2.00-bicubic	APS-2.00-nearest neighbor

Utility 2, the two sets of 100 images were resized 48 different ways for a total of 4,800 images per set. The classes are sorted by software used, Adobe Photoshop Creative Cloud Release 2014.0.0 (APS), Microsoft Office Picture Manager 14.0.7010.1000 (MSPM), or GNU Image Manipulation Program 2.8.10 (GIMP), algorithm used (bilinear, bicubic, nearest-neighbor, or undefined), and the interpolation rate (0.50, 0.66, 0.75, 0.90, 1.10, 1.25, 1.33, or 2.00). For example, the images resized with GIMP using the bicubic algorithm at the interpolation rate of 1.10 would be grouped into the same class. The interpolation rates were chosen to include both superposition and subposition rates. None of the rates are instances in which one rate is implemented more than once to get the other rate. For example, 0.25 could be implemented by using the 0.50 rate twice. See Table 4.3 for a list of the 48 classes.

Four collections of different variations in the collection of modules were tested in order to examine each module’s performance and their combined performance. The TIFF dataset was tested on only the file metadata (FMD), Two-Dimensional Second Derivative (2DSD) and the One-Dimensional Zero Crossings (1DZC) modules because TIFF images do not contain quantization matrices.

Group 1 contains the individual modules, 2DSD and 1DZC (JPEG and TIFF), tested on each dataset separately. The Quantization Matrices Identification (QMI) and FMD modules were excluded from this test because it was determined that the signature for each image altered by a specific IMSP were the same across the classes associated with the IMSP. For example, the FMD and QMI modules tested on all 24 classes associated with APS returned only that it was altered by APS. The two

modules are unable to determine, with any statistically significant accuracy, a specific class an image belongs to. These variations also indicate anticipated results if the header had been modified.

Group 2 contains different combinations of pairings of the modules: 2DSD with 1DZC (JPEG and TIFF), 2DSD with QMI (JPEG), 1DZC with QMI (JPEG), 2DSD with FMD (TIFF), and 1DZC with FMD (TIFF).

Group 3 contained only the QMI, 2DSD, 1DZC modules (JPEG) and the FMD, 2DSD, and 1DZC modules (TIFF).

Group 4 included four modules: FMD, QMI, 2DSD, and 1DZC modules. This group of modules did not process the TIFF dataset because TIFF images do not contain quantization matrices.

After the images were processed by the feature generation modules, the classification of the images was performed. The results for each variation were converted to an Attribute-Relation File Format (ARFF) file to be processed by the Bayesian Network classifier. In order to create a variation in the true positive rates (TPR) results, the number of images to be processed by the classifier were scaled down to 80% of the original size of the dataset by randomly selecting images 10 times to create 10 different unbalanced sub-datasets. The Bayesian Network classifier processed each sub-dataset 10 times using a 10 fold cross-validation method with different random number seeds.

Each test was performed ten times using the cross-validation method implemented in Waikato Environment for Knowledge Analysis (WEKA) with ten different random seeds, 1-10. The mean TPR, false positive rates (FPR), and F-Measure for each class and Kappa statistic for the overall framework were determined for each grouping of modules. Each group's results were also examined for the framework's ability to classify an image to the correct algorithm and rate.

For the group results, the Kappa statistic [6], k , is a chance corrected measure of accuracy. If $k > 0$ then the classifier is performing better than a random guess. This statistic was used to evaluate the classifier’s ability to perform better than random guessing. For the rate and algorithm results the TPR is evaluated against a random guess using a t-statistic test. The individual class analysis examined the TPR, FPR, and F-measure. The F-measure is a measure of the test’s accuracy by considering the precision Equation 4.15 and recall Equation 4.16 of each class [28].

$$t = \frac{\text{test rate} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad (4.15)$$

$$t = \frac{\text{test rate} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad (4.16)$$

The one tailed t -statistic test was determined for each class and using Equation 4.17. Where $\mu = 0.0208$ is the probability expected for a random guess and the sample size $n = 10$. Using the 95% confidence interval with the degrees of freedom, $df = 9$, the critical t value is 1.833. Therefore, if the t-statistic is greater than 1.833 and the computed P-Value is less than 0.05 then H_0 is rejected and H_1 is accepted.

$$t = \frac{\text{test rate} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad (4.17)$$

4.2 Results and Analysis

This section discusses the results of the tests conducted on each variation: Groups 1-4. Appendix A contains in-depth results for each variation and a representative confusion matrix. Table 4.4 gives the mean TPR, FPR, Kappa statistic, and each of their variances for each of the 13 variations of modules when classifying an image to one of the 48 classes, “IMSP-RATE-ALGORITHM”. Each individual class was

tested to determine if its F-Measure was statistically significant when compared to a random guess. Table 4.5 examined the framework’s ability to classify an image to the rate and algorithm used during manipulation. When conducting hypothesis testing on the overall rate compared to the random probability of guessing the rate for algorithm, the resulting P-Value was extremely small. This indicates that each of the variations was a statistically significant improvement when compared to a random guess. The random guess probability of 0.125 was used for hypothesis testing on the rate classification and 0.25 for hypothesis testing on the algorithm classification.

Table 4.4. Group Overall Results.

Framework	TPR		FPR		Kappa Statistic	
	Mean	STDEV	Mean	STDEV	Mean	STDEV
2DSD Only (JPEG)	0.381	0.005	0.014	0.000	0.367	0.005
1DZC Only (JPEG)	0.092	0.007	0.023	0.001	0.069	0.007
2DSD Only (TIFF)	0.343	0.005	0.015	0.000	0.328	0.005
1DZC Only (TIFF)	0.112	0.006	0.022	0.001	0.090	0.006
2DSD & 1DZC (JPEG)	0.390	0.003	0.014	0.001	0.377	0.003
2DSD & QMI (JPEG)	0.534	0.008	0.010	0.000	0.524	0.008
1DZC & QMI (JPEG)	0.143	0.008	0.022	0.001	0.122	0.008
2DSD & 1DZC (TIFF)	0.384	0.008	0.014	0.000	0.370	0.008
2DSD & FMD (TIFF)	0.506	0.011	0.011	0.000	0.495	0.011
1DZC & FMD (TIFF)	0.190	0.009	0.020	0.000	0.171	0.009
2DSD & 1DZC & QMI (JPEG)	0.544	0.006	0.010	0.000	0.533	0.006
2DSD & 1DZC & FMD (TIFF)	0.532	0.008	0.010	0.000	0.522	0.008
2DSD & 1DZC & QMI & FMD (JPEG)	0.542	0.013	0.011	0.000	0.532	0.013

Table 4.5. Group By Rate and Algorithm Results.

Framework	Rate Classification					Algorithm Classification				
	TPR			FPR		TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV	Mean	STDEV	P-Value	Mean	STDEV
2DSD Only (JPEG)	0.706	0.010	0.000	0.042	0.001	0.557	0.028	0.000	0.190	0.014
1DZC Only (JPEG)	0.194	0.009	0.000	0.116	0.001	0.356	0.046	0.000	0.260	0.036
2DSD Only (TIFF)	0.884	0.003	0.000	0.017	0.000	0.487	0.012	0.000	0.208	0.011
1DZC Only (TIFF)	0.370	0.021	0.000	0.090	0.003	0.320	0.057	0.005	0.253	0.039
2DSD & 1DZC (JPEG)	0.697	0.009	0.000	0.043	0.001	0.571	0.022	0.000	0.183	0.013
2DSD & QMI (JPEG)	0.713	0.010	0.000	0.041	0.001	0.659	0.013	0.000	0.151	0.014
1DZC & QMI (JPEG)	0.204	0.009	0.000	0.114	0.001	0.508	0.006	0.000	0.191	0.043
2DSD & 1DZC (TIFF)	0.886	0.007	0.000	0.016	0.001	0.523	0.023	0.000	0.186	0.011
2DSD & FMD (TIFF)	0.886	0.004	0.000	0.016	0.001	0.541	0.012	0.000	0.187	0.011
1DZC & FMD (TIFF)	0.382	0.016	0.000	0.089	0.002	0.397	0.013	0.000	0.246	0.024
2DSD & 1DZC & QMI (JPEG)	0.719	0.008	0.000	0.040	0.001	0.668	0.014	0.000	0.150	0.010
2DSD & 1DZC & FMD (TIFF)	0.883	0.005	0.000	0.017	0.001	0.572	0.011	0.000	0.178	0.006
2DSD & 1DZC & QMI & FMD (JPEG)	0.717	0.014	0.000	0.041	0.002	0.662	0.010	0.000	0.150	0.015

In addition to Tables 4.4 and 4.5, the by class, rate, and algorithm results are presented in Appendix A. The next four sections discuss results and observations

for the different variations in the module combinations. Results for same modules performed on different image types had similar results, therefore, unless specifically stated all observations discussed are referring to the modules on both image types. The next section discusses the Group 1 Results.

4.2.1 Group 1: Individual Modules.

This experiment focused on testing the individual modules for both JPEG and TIFF images to determine their accuracy. In this experiment, the 2DSD Only and 1DZC Only modules generated the features and the previously outlined process was conducted to determine the results.

Both modules, 2DSD and 1DZC, were able to classify an image to its correct class, “IMSP-RATE-ALGORITHM”, with a statistically significant improvement over a random guess. The Kappa Statistics for both modules are both greater than zero. Both modules are also able to classify the image to the rate and algorithm with a statistically significant improvement over a random guess. This determination was made by examining the TPR for each module. When comparing the 2DSD and 1DZC modules against each other it was noted that the 2DSD module performed significantly better than the 1DZC in classifying an image based the rate, algorithm, and class.

The 2DSD modules for both JPEG and TIFF images generated features that relate to the interpolation rate used by an IMSP with high statistical significance. This is expected because the 2DSD module is based on Gallagher’s work on detecting interpolation rate [11]. The 1DZC module was also able to identify the interpolation rate with which an image was manipulated with statistical significance when compared to a random guess. It was noted that the 2DSD module was able to classify the interpolation rate best when the images were manipulated at the 0.75, 0.90, 1.10, 1.25, and

2.00. The 2DSD module’s mean TPR when identifying the images manipulated at these rates was above 0.90. It was noted that the majority of the FPs were classified at the 0.50 rate. The 1DZC module performed best at 1.10, 1.25, 1.33, and 2.00 with its mean TPR for these interpolation rates above 0.35. More detailed results are in Appendix A.1.

4.2.2 Group 2: Pairing of modules.

This experiment focused on testing pairings of modules for both JPEG and TIFF images to determine their accuracy. The results were similar when the variation was applied to both JPEG and TIFF images. In this experiment, the pairings of modules generated the features. The process outlined in the previous section was conducted to determine the results. More detailed results are in Appendix A.2.

The first part of this experiment examined the combination of the 2DSD and 1DZC modules. The main observation from this pairing was the slightly improved TPR and Kappa statistic from the 2DSD Only variation and significant improvements from the 1DZC Only variation. The rate and algorithm classification had slight improvements in the FPR rate but decline in the TPR from the 2DSD Only variation and significant improvements from the 1DZC Only variation.

Both the QMI and FMD modules only determine which IMSP altered an image for JPEG images. The 2DSD and QMI pairing was used on the JPEG images. Since the TIFF images do not contain quantization matrices, the 2DSD and FMD pairing was used on these images.

The 2DSD and QMI modules (JPEG) and the 2DSD and FMD (TIFF) were the next pairing of modules to be examined. These results were a significant improvement to the overall results for the 2DSD and 1DZC pairing. The mean TPR reached over 0.50 with this pairing. The reason for this increase is the inclusion of the QMI module

data. With the inclusion of this data, the classifier is now able to classify the image, using the features generated from the 2DSD module, to a more limited number of classes because the QMI module is able to specify which IMSP manipulated the image. The rate and algorithm classification only had a slight increase. This is expected because the QMI module only gives information on which IMSP manipulated the image; it does not generate features based on the statistics of an image that can be used to determine the rate or algorithm used during the manipulation.

The final pairing was between the 1DZC and QMI modules. As expected, this pairing performed similar to the 2DSD and QMI pairing. The QMI module increased the 1DZC module's ability to classify the image, using the features generated from the 1DZC module, to a more limited number of classes. This pairing also did not significantly improve the rate and algorithm classification.

In summary all pairings showed an increase in the TPR and Kappa statistic while decreasing the FPR for the overall variation results. The TPR for the results by algorithm and by rate also increased for all pairings except when comparing the 2DSD Only and 2DSD and 1DZC modules. There was a slight decline in the TPR but the TPRs for both are close together. Therefore the decline is not statistically significant. The increase in TPRs and Kappa statistics while decreasing FPRs illustrates how grouping different modules focused on specific features can be used to assist in classifying an image to a specific IMSP, rate, and algorithm. The best performing combination of modules was the 2DSD and QMI for the JPEG images and the 2DSD FMD for the TIFF images.

4.2.3 Group 3: Three Modules.

This experiment focused on a variation of the modules containing three modules. The variation examined is 2DSD, 1DZC, and QMI for the JPEG images and 2DSD,

1DZC, and FMD for the TIFF images. It was noted in this experiment that the addition of the 1DZC module does not significantly increase the accuracy of the framework when compared to the 2DSD and QMI pairing. However, it displays a significant improvement when compared to the 1DZC and QMI pairing. This is because of the 2DSD module and its ability to classify an image into one of the 48 classes. More detailed results are in Appendix A.3.

4.2.4 Group 4: All Modules.

The next experiment examined the implementation of all modules. Only the JPEG images were examined in this experiment because the TIFF images do not contain quantization matrices and therefore no information can be gained by processing TIFF images through the QMI module. The results of this experiment show a slight decline in the results from the variation of three modules, 2DSD, 1DZC, and QMI. More detailed results are in Appendix A.4.

4.2.5 Results Conclusion.

This section discusses the observations made during the testing of the 13 variations of the module groupings, the rankings of the modules using specific quantitative and qualitative measures, and identification of the best group of modules to keep in the framework. Observations made during the 13 tests will be discussed first.

The interpolation rate features are significant to the classifier. For tests conducted when the 2DSD module was present, the interpolation rate was correctly classified between 69.698% – 71.852% of the time on JPEG images and 88.352% – 88.598% of the time on TIFF images. When the 1DZC module was present, the interpolation rate was correctly detected between 19.392% – 71.852% of the time on JPEG images and 37.025% – 88.272% of the time on TIFF images. For the variations where either

module was paired with QMI or FMD the TPR increased and the FPR decreased. For reference, a random guess of the interpolation rate is expected to yield a correct guess 12.5% of the time. The grouping of modules that returned the best by algorithm results were the 2DSD, 1DZC, and QMI for JPEG images 71.852% and the best results were returned from the 2DSD and FMD grouping for TIFF images.

The algorithm only results show trends similar to the rate only results. For tests conducted when the 2DSD module was present, the algorithm was correctly classified between 55.670% – 66.758% of the time on JPEG images and 48.746% – 57.241% of the time on TIFF images. When the 1DZC module was present, the algorithm was correctly classified between 35.621% – 66.758% of the time on JPEG images and 37.025% – 57.241% of the time on TIFF images. For the variations where the either module was paired with QMI or FMD the TPR increased and the FPR decreased. For reference, a random guess of the algorithm is expected to yield a correct guess 25% of the time. The grouping of modules that returned the best by algorithm results were the 2DSD, 1DZC, and QMI for JPEG images 71.852% and the best results were returned from the 2DSD, 1DZC, and FMD grouping for TIFF images.

The next observation is the large number of false positives appearing consistently in several classes when the 2DSD and/or the 1DZC modules are used in conjunction with the QMI or the FMD modules. The ReliefFAttributeEval attribute evaluator implementing the Ranker search method within WEKA was implemented to determine the highest ranking features. The consistent number of false positives are most likely images that do not contain any of the high ranking features of the classifier. This is especially true when only the false positives that are not classified to the correct interpolation rate are examined. A simple way to rectify this is to decrease the sensitivity for spike detection in the 2DSD and 1DZC modules. In other words accept smaller less distinct spikes as candidate spikes. This will increase the features

the classifier uses to make its decision. This can also have the opposite desired effect and generate too many features for the classifier to correctly classify an image. If this occurs then the FPR will increase and the TPR will decrease.

Another observation is the improved results when the modules are grouped together into a framework instead of an independent implementation. The Kappa statistic for the individual modules when processing JPEG images was the highest with the 2DSD module at 0.36738. When 2DSD, 1DZC, and QMI modules were implemented together, the Kappa statistic was 0.53342. The results after processing TIFF images showed similar trends. The 2DSD Only test for TIFF images resulted in a Kappa statistic of 0.32815 and 0.52186 when the 2DSD, 1DZC, and FMD modules were grouped together. The 1DZC module's results show similar trends but the Kappa statistic was not as high as the 2DSD module.

When ranking the modules for implementation in the framework, both qualitative and quantitative comparisons were conducted. The quantitative comparison examined the ability of the individual module only and grouping the module with others to classify an image with an improvement over a random guess. The qualitative comparison looked at how easily an image could be manipulated to hide the features the module extracted from them.

The module with the highest Kappa statistic when implemented individually was the 2DSD module for both JPEG and TIFF images. This module had a Kappa statistic of 0.36738 for JPEG images and 0.32815 for TIFF images when implemented alone and showed the largest increase when implemented in conjunction with the QMI or FMD module a 0.52404 for JPEG images and 0.49512 for TIFF images rate. Both results are a statistically significant improvement over a random guess. The 1DZC module also performed better than a random guess with its Kappa statistic at 0.06944 for JPEG images and 0.09025 for TIFF images when implemented alone and a 0.12224

for JPEG images and 0.17064 for TIFF images rate when implemented in conjunction with the QMI or FMD module. However, this is significantly less than 2DSD module with and without the QMI or FMD module. The QMI and FMD modules were not rated quantitatively because both modules are strictly signature based. This means the signatures are built into the program and no alteration was made to the QMI or FMD in the images, they would both be able to detect which IMSP altered an image with a 1.0 TP rate. As noted previously the QMI and FMD modules do not detect the interpolation rate or the algorithm used in the manipulations. Therefore, the modules cannot be compared to the 2DSD and 1DZC modules.

The qualitative comparison between the modules examines how easy it is to manipulate an image to hide features the modules generates within the constraints of this research. In other words, this research only looks at what happens when an image is manipulated by only one IMSP and only one resizing occurs. It does not look at what happens to an image when it has been altered by more than one IMSP because multiple IMSPs can leave conflicting features. For example, the FMD is altered by each IMSP, it is possible that an image can contain FMD related to each IMSP that altered it. This would give conflicting evidence on which IMSP created the most recent manipulations.

The easiest data to manipulate is the FMD. Erasing the data does not change any characteristics of the actual image but it can leave the module without any detectable signatures. The QMI data is more difficult to alter given confines of the experiment, only one IMSP can manipulate the image. However, if the IMSP used to manipulate an image then altered the quantization tables then the module would fail. This is harder to accomplish in a proprietary software like APS or MSPM compared to an open source software like GIMP. It is possible to change the tables of an image if the image is opened and altered within a simple HEX editor. The 2DSD and 1DZC data

are the most difficult to alter because the features they detect are within the statistics of the images and cannot be changed without additional image manipulations such as resizing the images multiple times or resaving with different quantization table settings. The next section discusses a summary of the results and a recommendation for the best combination of modules for the framework.

4.3 Summary

In summary, thirteen different combinations of the four modules that compose the framework were tested to attain results and determine the best combination of modules for the framework. Testing started on the individual modules using both JPEG and TIFF images. Both the 2DSD and 1DZC modules improvement over a random guess probability were statistically significant. The 2DSD module's Kappa statistic was greater than the 1DZC module with statistical significance. The next test examined the results of grouping the modules into pairs. During this test, pairing the modules together significantly improved the Kappa statistic. Also, the pairing of the 2DSD and 1DZC modules only yielded a slight improvement when compared to the 2DSD only. Testing of groups of three modules occurred next. 2DSD, 1DZC, and QMI were tested with JPEG images and 2DSD, 1DZC, and FMD were tested with TIFF images. These groupings showed a slight improvement over the pairings of modules. Finally, all four modules were tested with JPEG images only. This grouping was a slightly less accurate when compared to the 2DSD, 1DZC, and QMI grouping.

With the current settings on the modules, 2DSD, 1DZC, and QMI is the best combination of modules to comprise the framework for JPEG images. This combination contained the best TPR of 0.54219, the lowest FPR of 0.01030, and best Kappa statistic of 0.53342. 2DSD, 1DZC and FMD is the best combination for TIFF images. The TPR was 0.53220, lowest FPR of 0.01042, and best Kappa statistic of 0.52186.

The next chapter concludes the research and gives possible future work.

V. Conclusion

The invention of Image Manipulation Software Programs (IMSP) within the field of digital photography has made it so a manipulated image appears as though no alteration has occurred. Numerous techniques have been developed in the area of digital image authentication to detect image manipulation when it is not obvious or visible to the viewer [9, 11, 16, 17, 26]. This research has adapted these preexisting techniques to be able to determine which IMSP manipulated an image.

Law enforcement and intelligence agencies have a need to identify the utilized image manipulation software as part of the investigative and evidence gathering process. By detecting the super/sub-position algorithms used in IMSP, an image suspected of undergoing an alteration can be associated with an IMSP. Knowledge of the IMPS aids in identifying the computer system used in altering the image. This will assist examiners in reconstructing how and by whom an image manipulation occurred. The results of this research show an ability of the framework to classify an image better than a random guess. The final chapter discusses the hypothesis with the success criteria and whether the framework produced a true positive accuracy rate that is a statistically significant improvement over the probability of random guess using a t-statistic test with a 95% confidence interval.

5.1 Hypothesis and Success Criteria

This research demonstrated the development of techniques that can be employed to identify the software and interpolation algorithm used to resize an image. Although different IMSPs make use of the same algorithms for resizing, we hypothesized that differences in implementation of the algorithms leave detectable traces in the modified image. The techniques for identifying a specific IMSP entail building a framework

of modules used to detect artifacts present after an image manipulation occurred employing existing image authentication methods.

It is hypothesized that the proposed framework improves an examiner's ability to classify an image to which IMSP created the manipulated image beyond the probability of a random guess. To evaluate the hypothesis a framework was created that classifies a manipulated image to the correct IMSP. The research's hypothesis contained two main assumptions. First, the only manipulation addressed is resizing an image using specific rates, interpolation algorithms, and IMSPs. Additionally, only one manipulation occurs per image and no cross contamination occurs by using multiple IMSPs on a single image.

The research is implemented by creating unique Python modules using four previously developed image manipulation detection techniques to generate features then a Bayesian Network classifier would use the features to classify an image to a specific IMSP. The modules are the Two-Dimensional Second Derivative, One-Dimensional Zero Crossings, Quantization Matrices Identification, and File Metadata analysis [11, 16, 17, 26]. Testing was conducted using a dataset of Joint Photographic Experts Group (JPEG) images and a dataset of Tagged Image File Format (TIFF) images.

5.2 Results Synopsis

The framework's ability to correctly classify an image is a statistically significant improvement when compared to a random guess probability. The author acknowledges that the dataset used in testing is relatively small when compared to the large number of IMSPs. However, the three IMSPs used have widespread applications in digital imagery manipulations. The main observations made after the conclusion of the tests are discussed next.

The Two-Dimensional Second Difference (2DSD) module focused on the interpo-

lation rate used in the manipulations and the features extracted by the module were significant to the classifier. The Quantization Matrices Identification (QMI) and File Metadata (FMD) modules return only which IMSP was used to manipulate an image. When used separately, the two modules could only differentiate between the three IMSPs. However, when one of the modules is used in conjunction with the 2DSD module the Kappa statistic of the 2DSD module improved significantly. The One-Dimensional Zero Crossings (1DZC) module also showed an improvement, but the improvement was not as significant as with the 2DSD and QMI pairing. After analyzing different combinations of the modules, the ideal framework composition when inspecting a JPEG image consists of the 2DSD, 1DZC, and QMI modules which had a Kappa statistic $k = 0.53342$. For TIFF images, the 2DSD, 1DZC, and FMD modules were selected for inclusion in the framework with $k = 0.52186$. These groups contained the best Kappa statistic k . Because $k > 0$ the hypothesis is confirmed at the framework is an improvement when compared to a random guess probability.

5.3 Significance in the Area of Research

The findings are significant to this area of research because it shows that previously developed techniques to find specific image manipulations can also be used to determine which specific IMSP was used to manipulate the image. The findings show that if the correct techniques are selected, then a framework of several techniques can be built that can improve the ability of law enforcement agencies to classify manipulated images based on which IMSP conducted the manipulations.

5.4 Future Work

This research is intended to be a progression towards creating a framework that is able to classify an image to a specific IMSP. In order to improve accuracy of the

proposed framework, several additional areas of research are recommended. The first proposed area is adding a capability to process other image types. The current framework was only tested on JPEG and TIFF image formats. While the results showed the implementation worked on lossy and lossless, expanding the capability and testing to other image formats is desired.

The second recommended improvement is the addition of modules focused on features not specifically associated with the rate. The current framework can correctly classify 71.852% of the images JPEG images and 88.272% of the TIFF images to the correct interpolation rate. The framework's focus on detecting rate was a first step in creating a framework that can detect multiple manipulation types. Modules that detect skewing, copy-paste forgeries, rotating, etc. manipulations and how each IMSP implements the manipulation should be a focus of additional research. Chapter 2 discussed several of these recommended image manipulation detection techniques.

The next recommended improvement to the framework is with the 1DZC module. The 1DZC module showed promise for several classes. Currently, the module uses a one-dimensional implementation of the technique outlined in the literature. It can also be implemented in a two-dimensional technique. This two-dimensional technique should be investigated for future development of the framework.

The final recommended improvement to the framework is a front end Graphical User Interface (GUI). This will give a user with no programming experience the ability to interact with the framework. The GUI should allow the user to increase the framework's classification ability by inputting a group of images that have been manipulated in the exact same manner by a single IMSP. This will assist the framework in the detection of manipulations by additional IMSPs and the ability to classify the images with an IMSP.

Appendix A. Results

The tables in Appendix A are detailed results from the thirteen tests of the variations of modules. Each test's results are displayed in four tables. The first of the four tables is the by class results. This table details the True Positive Rate (TPR), False Positive Rate (FPR) and F-Measure for each class. The next two tables contain the by rate and by algorithm results. Each table contains the TPR and the FPR. The final table is the confusion matrix for the classifier.

A.1 Group 1

Table A.6. 2DSD Only By Class (JPEG).

Class	TPR		FPR		F-Measure		
	Mean	STDEV	Mean	STDEV	Mean	STDEV	P-Value
GIMP-0.50-bicubic	0.101	0.101	0.024	0.029	0.076	0.048	0.007
GIMP-0.50-bilinear	0.121	0.177	0.035	0.066	0.073	0.052	0.015
GIMP-0.66-bicubic	0.094	0.045	0.003	0.001	0.148	0.061	0.000
GIMP-0.66-bilinear	0.098	0.081	0.002	0.002	0.154	0.096	0.002
GIMP-0.75-bicubic	0.224	0.087	0.006	0.003	0.284	0.095	0.000
GIMP-0.75-bilinear	0.058	0.088	0.003	0.003	0.078	0.108	0.146
GIMP-0.90-bicubic	0.381	0.091	0.014	0.003	0.368	0.068	0.000
GIMP-0.90-bilinear	0.087	0.090	0.005	0.005	0.111	0.097	0.021
GIMP-1.10-bicubic	0.348	0.074	0.016	0.004	0.332	0.058	0.000
GIMP-1.10-bilinear	0.527	0.049	0.012	0.002	0.505	0.045	0.000
GIMP-1.25-bicubic	0.297	0.051	0.008	0.002	0.352	0.046	0.000
GIMP-1.25-bilinear	0.419	0.112	0.012	0.003	0.424	0.078	0.000
GIMP-1.33-bicubic	0.195	0.205	0.023	0.034	0.150	0.069	0.000
GIMP-1.33-bilinear	0.211	0.185	0.041	0.034	0.117	0.056	0.001
GIMP-2.00-bicubic	0.330	0.305	0.023	0.020	0.202	0.142	0.004
GIMP-2.00-bilinear	0.050	0.027	0.002	0.003	0.083	0.036	0.001
MSPM-0.50-undefined	0.044	0.029	0.002	0.001	0.077	0.050	0.008
MSPM-0.66-undefined	0.989	0.013	0.001	0.001	0.965	0.018	0.000
MSPM-0.75-undefined	0.666	0.091	0.020	0.002	0.519	0.064	0.000
MSPM-0.90-undefined	0.885	0.029	0.011	0.001	0.738	0.027	0.000
MSPM-1.10-undefined	0.885	0.023	0.000	0.000	0.931	0.020	0.000
MSPM-1.25-undefined	0.941	0.020	0.007	0.001	0.838	0.030	0.000
MSPM-1.33-undefined	0.870	0.020	0.000	0.000	0.929	0.011	0.000
MSPM-2.00-undefined	0.517	0.060	0.001	0.001	0.667	0.049	0.000
APS-0.50-bicubic	0.400	0.378	0.115	0.126	0.119	0.039	0.000
APS-0.50-bilinear	0.179	0.329	0.053	0.104	0.046	0.057	0.221
APS-0.50-nearest	0.204	0.310	0.052	0.098	0.095	0.040	0.000
APS-0.66-bicubic	0.094	0.241	0.029	0.080	0.032	0.039	0.430
APS-0.66-bilinear	0.263	0.083	0.007	0.002	0.326	0.085	0.000
APS-0.66-nearest	0.235	0.075	0.003	0.001	0.331	0.085	0.000
APS-0.75-bicubic	0.055	0.034	0.003	0.001	0.091	0.055	0.004
APS-0.75-bilinear	0.709	0.030	0.004	0.001	0.742	0.033	0.000
APS-0.75-nearest	0.395	0.046	0.003	0.001	0.512	0.051	0.000
APS-0.90-bicubic	0.272	0.128	0.010	0.004	0.290	0.108	0.000
APS-0.90-bilinear	0.061	0.036	0.002	0.001	0.102	0.057	0.002
APS-0.90-nearest	0.624	0.032	0.001	0.001	0.749	0.031	0.000
APS-1.10-bicubic	0.679	0.050	0.008	0.001	0.667	0.048	0.000
APS-1.10-bilinear	0.545	0.072	0.008	0.001	0.567	0.063	0.000
APS-1.10-nearest	0.632	0.044	0.002	0.000	0.740	0.035	0.000
APS-1.25-bicubic	0.498	0.063	0.007	0.001	0.544	0.056	0.000
APS-1.25-bilinear	0.972	0.011	0.003	0.001	0.920	0.023	0.000
APS-1.25-nearest	0.566	0.078	0.005	0.001	0.622	0.059	0.000
APS-1.33-bicubic	0.062	0.053	0.005	0.003	0.089	0.062	0.009
APS-1.33-bilinear	0.239	0.087	0.003	0.001	0.337	0.092	0.000
APS-1.33-nearest	0.485	0.042	0.005	0.001	0.564	0.037	0.000
APS-2.00-bicubic	0.322	0.302	0.024	0.023	0.216	0.133	0.002
APS-2.00-bilinear	0.156	0.201	0.011	0.018	0.134	0.107	0.011
APS-2.00-nearest	0.023	0.023	0.002	0.003	0.040	0.040	0.172
Framework	0.381	0.005	0.014	0.000	0.378	0.007	0.000

Table A.7. 2DSD Only By Rate (JPEG).

Rate	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
0.50	0.322	0.023	0.000	0.230	0.056
0.66	0.837	0.163	0.000	0.031	0.074
0.75	0.944	0.011	0.000	0.005	0.001
0.90	0.977	0.009	0.000	0.002	0.001
1.10	0.982	0.002	0.000	0.003	0.000
1.25	0.998	0.003	0.000	0.000	0.000
1.33	0.575	0.107	0.000	0.062	0.029
2.00	0.976	0.021	0.000	0.003	0.003
Overall	0.706	0.010	0.000	0.042	0.001

Table A.8. 2DSD Only By Algorithm (JPEG).

Algorithm	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
Bicubic	0.506	0.026	0.000	0.302	0.122
Bilinear	0.571	0.079	0.000	0.205	0.121
Nearest	0.667	0.178	0.000	0.078	0.105
Undefined	0.753	0.016	0.000	0.049	0.004
Overall	0.557	0.028	0.000	0.190	0.014

Table A.10. 1DZC Only By Class (JPEG).

Class	TPR		FPR		F-Measure		
	Mean	STDEV	Mean	STDEV	Mean	STDEV	P-Value
GIMP-0.50-bicubic	0.022	0.055	0.018	0.051	0.008	0.017	0.056
GIMP-0.50-bilinear	0.083	0.169	0.077	0.155	0.010	0.019	0.119
GIMP-0.66-bicubic	0.057	0.170	0.049	0.148	0.005	0.016	0.015
GIMP-0.66-bilinear	0.087	0.260	0.080	0.238	0.006	0.017	0.028
GIMP-0.75-bicubic	0.002	0.005	0.002	0.005	0.001	0.002	0.000
GIMP-0.75-bilinear	0.108	0.262	0.091	0.243	0.027	0.029	0.537
GIMP-0.90-bicubic	0.000	0.000	0.000	0.000	0.000	0.001	0.000
GIMP-0.90-bilinear	0.004	0.009	0.003	0.005	0.004	0.010	0.001
GIMP-1.10-bicubic	0.008	0.018	0.001	0.002	0.013	0.029	0.413
GIMP-1.10-bilinear	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GIMP-1.25-bicubic	0.041	0.112	0.038	0.103	0.006	0.015	0.020
GIMP-1.25-bilinear	0.103	0.240	0.078	0.219	0.027	0.044	0.706
GIMP-1.33-bicubic	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GIMP-1.33-bilinear	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GIMP-2.00-bicubic	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GIMP-2.00-bilinear	0.003	0.007	0.002	0.003	0.004	0.010	0.001
MSPM-0.50-undefined	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MSPM-0.66-undefined	0.000	0.000	0.000	0.001	0.000	0.000	0.000
MSPM-0.75-undefined	0.015	0.044	0.014	0.042	0.005	0.010	0.001
MSPM-0.90-undefined	0.001	0.003	0.000	0.001	0.002	0.005	0.000
MSPM-1.10-undefined	0.406	0.214	0.015	0.008	0.362	0.155	0.000
MSPM-1.25-undefined	0.912	0.047	0.017	0.002	0.685	0.049	0.000
MSPM-1.33-undefined	0.945	0.030	0.015	0.003	0.715	0.044	0.000
MSPM-2.00-undefined	0.935	0.021	0.034	0.002	0.534	0.027	0.000
APS-0.50-bicubic	0.007	0.022	0.000	0.001	0.012	0.037	0.497
APS-0.50-bilinear	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APS-0.50-nearest	0.070	0.196	0.063	0.187	0.013	0.020	0.256
APS-0.66-bicubic	0.000	0.000	0.000	0.001	0.000	0.000	0.000
APS-0.66-bilinear	0.003	0.010	0.003	0.010	0.001	0.003	0.000
APS-0.66-nearest	0.016	0.031	0.017	0.034	0.005	0.011	0.002
APS-0.75-bicubic	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APS-0.75-bilinear	0.006	0.012	0.006	0.012	0.003	0.005	0.000
APS-0.75-nearest	0.003	0.008	0.003	0.010	0.001	0.004	0.000
APS-0.90-bicubic	0.001	0.002	0.000	0.001	0.001	0.004	0.000
APS-0.90-bilinear	0.115	0.261	0.107	0.240	0.010	0.020	0.144
APS-0.90-nearest	0.079	0.236	0.068	0.204	0.005	0.015	0.013
APS-1.10-bicubic	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APS-1.10-bilinear	0.057	0.145	0.055	0.137	0.008	0.017	0.045
APS-1.10-nearest	0.001	0.004	0.002	0.005	0.001	0.002	0.000
APS-1.25-bicubic	0.000	0.001	0.001	0.001	0.001	0.002	0.000
APS-1.25-bilinear	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APS-1.25-nearest	0.000	0.000	0.000	0.001	0.000	0.000	0.000
APS-1.33-bicubic	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APS-1.33-bilinear	0.026	0.044	0.024	0.046	0.013	0.018	0.226
APS-1.33-nearest	0.001	0.004	0.002	0.005	0.001	0.002	0.000
APS-2.00-bicubic	0.006	0.009	0.002	0.002	0.010	0.015	0.061
APS-2.00-bilinear	0.045	0.107	0.045	0.107	0.007	0.015	0.027
APS-2.00-nearest	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Framework	0.092	0.007	0.023	0.001	0.054	0.006	0.000

Table A.11. 1DZC Only By Rate (JPEG).

Rate	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
0.50	0.165	0.184	0.534	0.157	0.259
0.66	0.041	0.057	0.002	0.148	0.271
0.75	0.215	0.280	0.357	0.115	0.234
0.90	0.071	0.072	0.052	0.176	0.287
1.10	0.426	0.133	0.000	0.070	0.133
1.25	0.454	0.159	0.000	0.135	0.233
1.33	0.523	0.140	0.000	0.041	0.047
2.00	0.388	0.088	0.000	0.083	0.108
Overall	0.194	0.009	0.000	0.116	0.001

Table A.12. 1DZC Only By Algorithm (JPEG).

Algorithm	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
Bicubic	0.355	0.060	0.001	0.108	0.166
Bilinear	0.346	0.070	0.003	0.546	0.281
Nearest	0.104	0.076	0.000	0.153	0.248
Undefined	0.517	0.055	0.000	0.098	0.051
Overall	0.356	0.046	0.000	0.260	0.036

Table A.14. 2DSD Only By Class (TIFF).

Class	TPR		FPR		F-Measure		
	Mean	STDEV	Mean	STDEV	Mean	STDEV	P-Value
GIMP-0.50-bicubic	0.306	0.374	0.063	0.078	0.104	0.066	0.004
GIMP-0.50-bilinear	0.130	0.237	0.022	0.055	0.093	0.045	0.001
GIMP-0.66-bicubic	0.140	0.087	0.005	0.004	0.192	0.092	0.000
GIMP-0.66-bilinear	0.108	0.089	0.004	0.002	0.149	0.114	0.008
GIMP-0.75-bicubic	0.043	0.046	0.004	0.004	0.062	0.061	0.076
GIMP-0.75-bilinear	0.173	0.095	0.011	0.007	0.201	0.083	0.000
GIMP-0.90-bicubic	0.465	0.037	0.003	0.001	0.570	0.033	0.000
GIMP-0.90-bilinear	0.423	0.142	0.026	0.009	0.319	0.066	0.000
GIMP-1.10-bicubic	0.457	0.220	0.022	0.010	0.342	0.092	0.000
GIMP-1.10-bilinear	0.201	0.076	0.010	0.005	0.230	0.057	0.000
GIMP-1.25-bicubic	0.388	0.282	0.032	0.023	0.222	0.103	0.000
GIMP-1.25-bilinear	0.186	0.199	0.014	0.017	0.158	0.086	0.001
GIMP-1.33-bicubic	0.395	0.347	0.022	0.017	0.264	0.171	0.002
GIMP-1.33-bilinear	0.227	0.262	0.012	0.014	0.193	0.131	0.003
GIMP-2.00-bicubic	0.260	0.163	0.012	0.011	0.263	0.092	0.000
GIMP-2.00-bilinear	0.144	0.182	0.008	0.011	0.145	0.109	0.008
MSPM-0.50-undefined	0.069	0.055	0.007	0.014	0.092	0.039	0.000
MSPM-0.66-undefined	0.994	0.008	0.002	0.001	0.945	0.017	0.000
MSPM-0.75-undefined	0.125	0.085	0.006	0.004	0.163	0.096	0.002
MSPM-0.90-undefined	0.314	0.058	0.004	0.001	0.415	0.059	0.000
MSPM-1.10-undefined	0.785	0.054	0.000	0.000	0.876	0.035	0.000
MSPM-1.25-undefined	0.322	0.235	0.024	0.019	0.238	0.090	0.000
MSPM-1.33-undefined	0.862	0.028	0.000	0.000	0.921	0.016	0.000
MSPM-2.00-undefined	0.538	0.088	0.004	0.002	0.622	0.048	0.000
APS-0.50-bicubic	0.269	0.308	0.059	0.072	0.101	0.062	0.004
APS-0.50-bilinear	0.186	0.319	0.038	0.070	0.073	0.060	0.029
APS-0.50-nearest	0.120	0.246	0.024	0.055	0.062	0.049	0.034
APS-0.66-bicubic	0.023	0.021	0.002	0.001	0.041	0.038	0.137
APS-0.66-bilinear	0.256	0.133	0.010	0.004	0.286	0.127	0.000
APS-0.66-nearest	0.268	0.067	0.004	0.001	0.367	0.076	0.000
APS-0.75-bicubic	0.021	0.018	0.001	0.001	0.038	0.033	0.159
APS-0.75-bilinear	0.051	0.053	0.003	0.003	0.078	0.072	0.042
APS-0.75-nearest	0.731	0.107	0.035	0.008	0.438	0.028	0.000
APS-0.90-bicubic	0.276	0.111	0.017	0.008	0.257	0.054	0.000
APS-0.90-bilinear	0.233	0.196	0.016	0.014	0.202	0.114	0.001
APS-0.90-nearest	0.827	0.023	0.005	0.001	0.796	0.021	0.000
APS-1.10-bicubic	0.423	0.130	0.011	0.009	0.440	0.069	0.000
APS-1.10-bilinear	0.224	0.147	0.014	0.008	0.221	0.096	0.000
APS-1.10-nearest	0.649	0.022	0.009	0.002	0.616	0.027	0.000
APS-1.25-bicubic	0.262	0.044	0.002	0.001	0.382	0.046	0.000
APS-1.25-bilinear	0.049	0.038	0.002	0.002	0.083	0.061	0.014
APS-1.25-nearest	0.415	0.168	0.017	0.004	0.359	0.109	0.000
APS-1.33-bicubic	0.646	0.039	0.002	0.001	0.737	0.039	0.000
APS-1.33-bilinear	0.385	0.327	0.020	0.017	0.269	0.147	0.001
APS-1.33-nearest	0.480	0.059	0.004	0.001	0.573	0.062	0.000
APS-2.00-bicubic	0.460	0.255	0.028	0.014	0.305	0.110	0.000
APS-2.00-bilinear	0.094	0.112	0.007	0.007	0.108	0.093	0.020
APS-2.00-nearest	0.596	0.041	0.021	0.001	0.468	0.025	0.000
Framework	0.343	0.005	0.015	0.000	0.319	0.005	0.000

Table A.15. 2DSD Only By Rate (TIFF).

Rate	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
0.50	0.531	0.007	0.000	0.126	0.002
0.66	0.948	0.021	0.000	0.004	0.001
0.75	0.991	0.005	0.000	0.001	0.001
0.90	0.996	0.003	0.000	0.001	0.000
1.10	0.997	0.002	0.000	0.000	0.000
1.25	0.995	0.003	0.000	0.001	0.000
1.33	0.998	0.002	0.000	0.000	0.000
2.00	0.992	0.003	0.000	0.001	0.000
Overall	0.884	0.003	0.000	0.017	0.000

Table A.16. 2DSD Only By Algorithm (TIFF).

Algorithm	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
Bicubic	0.470	0.025	0.000	0.309	0.092
Bilinear	0.493	0.035	0.000	0.217	0.100
Nearest	0.451	0.050	0.000	0.138	0.063
Undefined	0.665	0.090	0.000	0.057	0.026
Overall	0.487	0.012	0.000	0.208	0.011

Table A.18. 1DZC Only By Class (TIFF).

Class	TPR		FPR		F-Measure		
	Mean	STDEV	Mean	STDEV	Mean	STDEV	P-Value
GIMP-0.50-bicubic	0.039	0.092	0.033	0.079	0.009	0.020	0.113
GIMP-0.50-bilinear	0.057	0.167	0.041	0.117	0.008	0.020	0.078
GIMP-0.66-bicubic	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GIMP-0.66-bilinear	0.018	0.054	0.016	0.047	0.005	0.014	0.006
GIMP-0.75-bicubic	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GIMP-0.75-bilinear	0.080	0.197	0.058	0.140	0.011	0.022	0.196
GIMP-0.90-bicubic	0.003	0.008	0.000	0.001	0.005	0.015	0.011
GIMP-0.90-bilinear	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GIMP-1.10-bicubic	0.064	0.106	0.007	0.007	0.064	0.093	0.200
GIMP-1.10-bilinear	0.157	0.089	0.015	0.008	0.159	0.066	0.000
GIMP-1.25-bicubic	0.051	0.123	0.008	0.010	0.042	0.080	0.454
GIMP-1.25-bilinear	0.385	0.258	0.031	0.018	0.234	0.114	0.000
GIMP-1.33-bicubic	0.002	0.004	0.002	0.003	0.003	0.005	0.000
GIMP-1.33-bilinear	0.047	0.041	0.003	0.003	0.075	0.058	0.020
GIMP-2.00-bicubic	0.245	0.197	0.063	0.145	0.169	0.118	0.004
GIMP-2.00-bilinear	0.010	0.022	0.001	0.004	0.014	0.030	0.502
MSPM-0.50-undefined	0.096	0.268	0.068	0.187	0.010	0.022	0.171
MSPM-0.66-undefined	0.003	0.008	0.002	0.007	0.001	0.004	0.000
MSPM-0.75-undefined	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MSPM-0.90-undefined	0.097	0.277	0.065	0.183	0.009	0.021	0.130
MSPM-1.10-undefined	0.119	0.178	0.012	0.015	0.092	0.108	0.078
MSPM-1.25-undefined	0.413	0.208	0.022	0.011	0.310	0.154	0.000
MSPM-1.33-undefined	0.918	0.020	0.045	0.004	0.452	0.040	0.000
MSPM-2.00-undefined	0.931	0.025	0.024	0.002	0.595	0.026	0.000
APS-0.50-bicubic	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APS-0.50-bilinear	0.118	0.265	0.086	0.184	0.017	0.028	0.659
APS-0.50-nearest	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APS-0.66-bicubic	0.073	0.219	0.050	0.151	0.007	0.021	0.083
APS-0.66-bilinear	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APS-0.66-nearest	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APS-0.75-bicubic	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APS-0.75-bilinear	0.001	0.003	0.001	0.002	0.001	0.002	0.000
APS-0.75-nearest	0.005	0.014	0.005	0.013	0.002	0.006	0.000
APS-0.90-bicubic	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APS-0.90-bilinear	0.074	0.221	0.048	0.145	0.007	0.020	0.069
APS-0.90-nearest	0.003	0.008	0.000	0.001	0.005	0.013	0.005
APS-1.10-bicubic	0.036	0.074	0.008	0.010	0.032	0.053	0.530
APS-1.10-bilinear	0.272	0.192	0.023	0.013	0.200	0.098	0.000
APS-1.10-nearest	0.231	0.157	0.023	0.012	0.166	0.107	0.003
APS-1.25-bicubic	0.030	0.053	0.008	0.003	0.037	0.056	0.403
APS-1.25-bilinear	0.151	0.116	0.009	0.009	0.167	0.087	0.001
APS-1.25-nearest	0.232	0.084	0.016	0.004	0.226	0.071	0.000
APS-1.33-bicubic	0.080	0.194	0.064	0.156	0.011	0.023	0.248
APS-1.33-bilinear	0.018	0.023	0.004	0.002	0.028	0.035	0.555
APS-1.33-nearest	0.012	0.013	0.001	0.001	0.022	0.024	0.859
APS-2.00-bicubic	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APS-2.00-bilinear	0.029	0.062	0.003	0.005	0.031	0.063	0.650
APS-2.00-nearest	0.065	0.189	0.046	0.132	0.008	0.019	0.066
Framework	0.112	0.006	0.022	0.001	0.069	0.006	0.000

Table A.19. 1DZC Only By Rate (TIFF).

Rate	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
0.50	0.122	0.075	0.900	0.218	0.259
0.66	0.050	0.069	0.010	0.066	0.145
0.75	0.037	0.067	0.003	0.061	0.146
0.90	0.132	0.116	0.866	0.108	0.209
1.10	0.756	0.007	0.000	0.029	0.002
1.25	0.804	0.015	0.000	0.027	0.002
1.33	0.686	0.207	0.000	0.085	0.163
2.00	0.452	0.139	0.000	0.130	0.187
Overall	0.370	0.021	0.000	0.090	0.003

Table A.20. 1DZC Only By Algorithm (TIFF).

Algorithm	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
Bicubic	0.387	0.025	0.000	0.235	0.225
Bilinear	0.394	0.060	0.000	0.346	0.248
Nearest	0.232	0.033	0.128	0.095	0.125
Undefined	0.314	0.074	0.028	0.258	0.253
Overall	0.320	0.057	0.005	0.253	0.039

A.2 Group 2

Table A.22. 2DSD and 1DZC By Class (JPEG).

Class	TPR		FPR		F-Measure		
	Mean	STDEV	Mean	STDEV	Mean	STDEV	P-Value
GIMP-0.50-bicubic	0.217	0.163	0.046	0.035	0.113	0.043	0.000
GIMP-0.50-bilinear	0.053	0.040	0.004	0.003	0.080	0.060	0.017
GIMP-0.66-bicubic	0.162	0.215	0.043	0.079	0.096	0.060	0.004
GIMP-0.66-bilinear	0.119	0.086	0.005	0.003	0.166	0.107	0.003
GIMP-0.75-bicubic	0.254	0.138	0.007	0.002	0.307	0.139	0.000
GIMP-0.75-bilinear	0.049	0.044	0.003	0.001	0.080	0.066	0.025
GIMP-0.90-bicubic	0.397	0.083	0.014	0.002	0.385	0.065	0.000
GIMP-0.90-bilinear	0.057	0.038	0.004	0.002	0.087	0.050	0.003
GIMP-1.10-bicubic	0.376	0.098	0.015	0.002	0.351	0.068	0.000
GIMP-1.10-bilinear	0.532	0.085	0.013	0.003	0.505	0.045	0.000
GIMP-1.25-bicubic	0.308	0.083	0.009	0.002	0.349	0.069	0.000
GIMP-1.25-bilinear	0.433	0.059	0.010	0.003	0.451	0.047	0.000
GIMP-1.33-bicubic	0.192	0.114	0.024	0.022	0.163	0.081	0.001
GIMP-1.33-bilinear	0.148	0.131	0.025	0.025	0.113	0.056	0.001
GIMP-2.00-bicubic	0.436	0.295	0.026	0.017	0.278	0.151	0.001
GIMP-2.00-bilinear	0.097	0.082	0.004	0.004	0.134	0.101	0.008
MSPM-0.50-undefined	0.057	0.030	0.006	0.008	0.087	0.050	0.003
MSPM-0.66-undefined	0.991	0.010	0.001	0.001	0.973	0.017	0.000
MSPM-0.75-undefined	0.629	0.057	0.020	0.002	0.490	0.052	0.000
MSPM-0.90-undefined	0.861	0.052	0.009	0.002	0.743	0.041	0.000
MSPM-1.10-undefined	0.938	0.037	0.000	0.000	0.965	0.022	0.000
MSPM-1.25-undefined	0.962	0.014	0.002	0.001	0.933	0.013	0.000
MSPM-1.33-undefined	0.929	0.023	0.000	0.000	0.963	0.012	0.000
MSPM-2.00-undefined	0.748	0.056	0.003	0.001	0.782	0.035	0.000
APS-0.50-bicubic	0.373	0.323	0.116	0.115	0.113	0.024	0.000
APS-0.50-bilinear	0.112	0.230	0.030	0.073	0.056	0.048	0.057
APS-0.50-nearest	0.200	0.254	0.055	0.087	0.098	0.036	0.000
APS-0.66-bicubic	0.098	0.105	0.028	0.040	0.071	0.037	0.003
APS-0.66-bilinear	0.213	0.099	0.006	0.002	0.276	0.111	0.000
APS-0.66-nearest	0.294	0.076	0.004	0.002	0.391	0.075	0.000
APS-0.75-bicubic	0.054	0.026	0.002	0.000	0.091	0.042	0.001
APS-0.75-bilinear	0.715	0.045	0.004	0.002	0.749	0.036	0.000
APS-0.75-nearest	0.369	0.066	0.003	0.001	0.490	0.071	0.000
APS-0.90-bicubic	0.291	0.099	0.011	0.003	0.315	0.089	0.000
APS-0.90-bilinear	0.099	0.075	0.004	0.003	0.143	0.091	0.003
APS-0.90-nearest	0.594	0.053	0.001	0.000	0.722	0.039	0.000
APS-1.10-bicubic	0.655	0.091	0.007	0.001	0.649	0.070	0.000
APS-1.10-bilinear	0.483	0.072	0.008	0.001	0.518	0.056	0.000
APS-1.10-nearest	0.633	0.042	0.002	0.001	0.742	0.036	0.000
APS-1.25-bicubic	0.600	0.072	0.007	0.002	0.628	0.043	0.000
APS-1.25-bilinear	0.977	0.015	0.003	0.001	0.914	0.024	0.000
APS-1.25-nearest	0.656	0.076	0.004	0.001	0.706	0.063	0.000
APS-1.33-bicubic	0.061	0.045	0.005	0.003	0.086	0.057	0.007
APS-1.33-bilinear	0.188	0.068	0.004	0.001	0.270	0.091	0.000
APS-1.33-nearest	0.544	0.061	0.003	0.001	0.639	0.052	0.000
APS-2.00-bicubic	0.198	0.207	0.010	0.015	0.201	0.113	0.001
APS-2.00-bilinear	0.196	0.227	0.012	0.013	0.178	0.110	0.002
APS-2.00-nearest	0.020	0.018	0.001	0.001	0.037	0.031	0.152
Framework	0.390	0.003	0.014	0.001	0.390	0.004	0.000

Table A.23. 2DSD and 1DZC By Rate (JPEG).

Rate	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
0.50	0.307	0.023	0.000	0.216	0.064
0.66	0.660	0.196	0.000	0.069	0.079
0.75	0.954	0.009	0.000	0.004	0.001
0.90	0.978	0.010	0.000	0.002	0.001
1.10	0.981	0.006	0.000	0.003	0.001
1.25	0.997	0.003	0.000	0.000	0.000
1.33	0.627	0.127	0.000	0.049	0.030
2.00	0.977	0.011	0.000	0.002	0.001
Overall	0.697	0.009	0.000	0.043	0.001

Table A.24. 2DSD and 1DZC By Algorithm (JPEG).

Algorithm	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
Bicubic	0.507	0.028	0.000	0.348	0.103
Bilinear	0.620	0.061	0.000	0.137	0.088
Nearest	0.646	0.170	0.000	0.078	0.089
Undefined	0.756	0.035	0.000	0.050	0.010
Overall	0.571	0.022	0.000	0.183	0.013

Table A.26. 1DZC and QMI By Class (JPEG).

Class	TPR		FPR		F-Measure		
	Mean	STDEV	Mean	STDEV	Mean	STDEV	P-Value
GIMP-0.50-bicubic	0.036	0.051	0.009	0.013	0.038	0.046	0.299
GIMP-0.50-bilinear	0.089	0.265	0.027	0.081	0.015	0.042	0.670
GIMP-0.66-bicubic	0.004	0.012	0.002	0.006	0.003	0.010	0.001
GIMP-0.66-bilinear	0.102	0.273	0.031	0.083	0.026	0.047	0.729
GIMP-0.75-bicubic	0.049	0.146	0.016	0.046	0.014	0.037	0.566
GIMP-0.75-bilinear	0.088	0.256	0.029	0.086	0.018	0.041	0.854
GIMP-0.90-bicubic	0.103	0.286	0.031	0.085	0.022	0.043	0.930
GIMP-0.90-bilinear	0.003	0.007	0.000	0.000	0.005	0.013	0.005
GIMP-1.10-bicubic	0.057	0.043	0.004	0.002	0.086	0.065	0.015
GIMP-1.10-bilinear	0.036	0.079	0.009	0.025	0.028	0.044	0.650
GIMP-1.25-bicubic	0.101	0.266	0.030	0.082	0.029	0.050	0.643
GIMP-1.25-bilinear	0.182	0.303	0.055	0.094	0.053	0.050	0.084
GIMP-1.33-bicubic	0.020	0.033	0.003	0.004	0.026	0.040	0.686
GIMP-1.33-bilinear	0.195	0.365	0.057	0.105	0.036	0.055	0.418
GIMP-2.00-bicubic	0.011	0.016	0.003	0.004	0.015	0.022	0.448
GIMP-2.00-bilinear	0.059	0.039	0.007	0.004	0.078	0.045	0.004
MSPM-0.50-undefined	0.208	0.320	0.016	0.024	0.124	0.141	0.057
MSPM-0.66-undefined	0.110	0.249	0.009	0.019	0.061	0.124	0.353
MSPM-0.75-undefined	0.399	0.369	0.029	0.030	0.218	0.137	0.002
MSPM-0.90-undefined	0.210	0.337	0.015	0.024	0.120	0.147	0.073
MSPM-1.10-undefined	0.487	0.131	0.002	0.000	0.613	0.123	0.000
MSPM-1.25-undefined	0.897	0.032	0.004	0.001	0.855	0.026	0.000
MSPM-1.33-undefined	0.918	0.032	0.003	0.001	0.901	0.027	0.000
MSPM-2.00-undefined	0.924	0.039	0.002	0.001	0.903	0.023	0.000
APS-0.50-bicubic	0.002	0.005	0.002	0.002	0.004	0.008	0.000
APS-0.50-bilinear	0.010	0.026	0.003	0.004	0.011	0.029	0.323
APS-0.50-nearest	0.061	0.043	0.005	0.005	0.087	0.059	0.008
APS-0.66-bicubic	0.001	0.002	0.001	0.002	0.002	0.004	0.000
APS-0.66-bilinear	0.001	0.003	0.001	0.002	0.002	0.005	0.000
APS-0.66-nearest	0.090	0.234	0.044	0.112	0.016	0.030	0.643
APS-0.75-bicubic	0.002	0.003	0.001	0.001	0.002	0.003	0.000
APS-0.75-bilinear	0.000	0.001	0.000	0.000	0.001	0.002	0.000
APS-0.75-nearest	0.098	0.273	0.047	0.132	0.017	0.032	0.727
APS-0.90-bicubic	0.048	0.142	0.023	0.067	0.009	0.025	0.209
APS-0.90-bilinear	0.003	0.008	0.002	0.004	0.003	0.007	0.000
APS-0.90-nearest	0.195	0.384	0.085	0.167	0.022	0.041	0.925
APS-1.10-bicubic	0.154	0.313	0.072	0.146	0.021	0.041	0.990
APS-1.10-bilinear	0.033	0.037	0.006	0.004	0.042	0.046	0.193
APS-1.10-nearest	0.051	0.150	0.025	0.071	0.010	0.026	0.245
APS-1.25-bicubic	0.094	0.273	0.045	0.131	0.014	0.030	0.542
APS-1.25-bilinear	0.022	0.035	0.005	0.008	0.025	0.041	0.747
APS-1.25-nearest	0.022	0.040	0.005	0.006	0.025	0.043	0.780
APS-1.33-bicubic	0.034	0.073	0.016	0.039	0.021	0.025	0.970
APS-1.33-bilinear	0.051	0.087	0.020	0.042	0.029	0.042	0.586
APS-1.33-nearest	0.092	0.275	0.044	0.130	0.010	0.031	0.348
APS-2.00-bicubic	0.035	0.040	0.007	0.003	0.046	0.052	0.179
APS-2.00-bilinear	0.020	0.022	0.005	0.005	0.025	0.027	0.621
APS-2.00-nearest	0.039	0.085	0.020	0.043	0.013	0.027	0.423
Framework	0.143	0.008	0.021	0.001	0.101	0.009	0.000

Table A.27. 1DZC and QMI By Rate (JPEG).

Rate	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
0.50	0.186	0.040	0.001	0.065	0.086
0.66	0.098	0.072	0.290	0.092	0.136
0.75	0.218	0.087	0.011	0.130	0.161
0.90	0.149	0.061	0.269	0.162	0.176
1.10	0.387	0.188	0.002	0.120	0.151
1.25	0.425	0.187	0.001	0.149	0.211
1.33	0.424	0.215	0.002	0.148	0.170
2.00	0.451	0.109	0.000	0.046	0.042
Overall	0.204	0.009	0.000	0.114	0.001

Table A.28. 1DZC and QMI By Algorithm (JPEG).

Algorithm	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
Bicubic	0.451	0.067	0.000	0.248	0.243
Bilinear	0.476	0.030	0.000	0.212	0.103
Nearest	0.336	0.015	0.000	0.225	0.165
Undefined	1.000	0.000	0.000	0.000	0.000
Overall	0.508	0.006	0.000	0.191	0.043

Table A.30. 2DSD and QMI By Class (JPEG).

Class	TPR		FPR		F-Measure		
	Mean	STDEV	Mean	STDEV	Mean	STDEV	P-Value
GIMP-0.50-bicubic	0.214	0.227	0.024	0.031	0.147	0.067	0.000
GIMP-0.50-bilinear	0.434	0.274	0.051	0.032	0.202	0.058	0.000
GIMP-0.66-bicubic	0.302	0.270	0.030	0.032	0.189	0.084	0.000
GIMP-0.66-bilinear	0.288	0.068	0.004	0.010	0.406	0.068	0.000
GIMP-0.75-bicubic	0.328	0.102	0.005	0.002	0.410	0.093	0.000
GIMP-0.75-bilinear	0.291	0.129	0.006	0.001	0.356	0.139	0.000
GIMP-0.90-bicubic	0.422	0.055	0.005	0.001	0.507	0.058	0.000
GIMP-0.90-bilinear	0.325	0.025	0.004	0.001	0.436	0.027	0.000
GIMP-1.10-bicubic	0.647	0.080	0.006	0.001	0.663	0.051	0.000
GIMP-1.10-bilinear	0.733	0.063	0.005	0.002	0.736	0.037	0.000
GIMP-1.25-bicubic	0.574	0.061	0.006	0.002	0.614	0.040	0.000
GIMP-1.25-bilinear	0.651	0.081	0.007	0.001	0.661	0.063	0.000
GIMP-1.33-bicubic	0.335	0.169	0.022	0.016	0.274	0.062	0.000
GIMP-1.33-bilinear	0.087	0.082	0.007	0.011	0.102	0.057	0.002
GIMP-2.00-bicubic	0.540	0.243	0.010	0.004	0.497	0.192	0.000
GIMP-2.00-bilinear	0.283	0.238	0.005	0.005	0.324	0.162	0.000
MSPM-0.50-undefined	1.000	0.000	0.013	0.002	0.778	0.041	0.000
MSPM-0.66-undefined	1.000	0.000	0.000	0.000	0.807	0.055	0.000
MSPM-0.75-undefined	0.679	0.074	0.000	0.000	0.984	0.017	0.000
MSPM-0.90-undefined	1.000	0.000	0.000	0.000	0.977	0.010	0.000
MSPM-1.10-undefined	0.986	0.010	0.000	0.000	0.901	0.061	0.000
MSPM-1.25-undefined	0.990	0.013	0.000	0.000	0.463	0.315	0.002
MSPM-1.33-undefined	0.947	0.018	0.000	0.000	0.133	0.088	0.004
MSPM-2.00-undefined	0.735	0.036	0.000	0.000	0.112	0.064	0.002
APS-0.50-bicubic	0.166	0.258	0.017	0.045	0.112	0.072	0.004
APS-0.50-bilinear	0.273	0.369	0.046	0.068	0.254	0.143	0.001
APS-0.50-nearest	0.271	0.338	0.044	0.061	0.429	0.053	0.000
APS-0.66-bicubic	0.288	0.358	0.051	0.067	0.361	0.123	0.000
APS-0.66-bilinear	0.292	0.043	0.004	0.002	0.549	0.278	0.000
APS-0.66-nearest	0.356	0.092	0.003	0.001	0.680	0.109	0.000
APS-0.75-bicubic	0.150	0.029	0.003	0.002	0.560	0.081	0.000
APS-0.75-bilinear	0.789	0.068	0.003	0.001	0.482	0.077	0.000
APS-0.75-nearest	0.579	0.091	0.007	0.002	0.645	0.168	0.000
APS-0.90-bicubic	0.415	0.044	0.006	0.002	0.783	0.021	0.000
APS-0.90-bilinear	0.404	0.091	0.007	0.001	0.791	0.037	0.000
APS-0.90-nearest	0.716	0.035	0.001	0.001	0.763	0.051	0.000
APS-1.10-bicubic	0.698	0.034	0.002	0.001	0.789	0.029	0.000
APS-1.10-bilinear	0.840	0.046	0.005	0.001	0.883	0.054	0.000
APS-1.10-nearest	0.687	0.057	0.003	0.001	0.906	0.039	0.000
APS-1.25-bicubic	0.771	0.038	0.002	0.001	0.499	0.270	0.000
APS-1.25-bilinear	0.969	0.016	0.002	0.000	0.304	0.068	0.000
APS-1.25-nearest	0.814	0.051	0.002	0.000	0.499	0.175	0.000
APS-1.33-bicubic	0.216	0.123	0.013	0.012	0.468	0.213	0.000
APS-1.33-bilinear	0.212	0.074	0.002	0.001	0.409	0.199	0.000
APS-1.33-nearest	0.555	0.043	0.003	0.000	0.230	0.111	0.000
APS-2.00-bicubic	0.481	0.281	0.015	0.009	0.404	0.138	0.000
APS-2.00-bilinear	0.333	0.334	0.009	0.008	0.241	0.234	0.020
APS-2.00-nearest	0.249	0.129	0.016	0.010	0.005	0.016	0.020
Framework	0.534	0.008	0.010	0.000	0.000	0.000	0.000

Table A.31. 2DSD and QMI By Rate (JPEG).

Rate	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
0.50	0.354	0.019	0.000	0.180	0.061
0.66	0.566	0.170	0.000	0.085	0.067
0.75	0.954	0.016	0.000	0.004	0.001
0.90	0.986	0.007	0.000	0.002	0.001
1.10	0.986	0.005	0.000	0.002	0.001
1.25	0.996	0.003	0.000	0.001	0.000
1.33	0.669	0.126	0.000	0.040	0.025
2.00	0.891	0.057	0.000	0.014	0.008
Overall	0.713	0.010	0.000	0.041	0.001

Table A.32. 2DSD and QMI By Algorithm (JPEG).

Algorithm	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
Bicubic	0.574	0.026	0.000	0.228	0.051
Bilinear	0.619	0.075	0.000	0.188	0.088
Nearest	0.659	0.116	0.000	0.077	0.057
Undefined	1.000	0.001	0.000	0.000	0.000
Overall	0.659	0.013	0.000	0.151	0.014

Table A.34. 2DSD and 1DZC By Class (TIFF).

Class	TPR		FPR		F-Measure		
	Mean	STDEV	Mean	STDEV	Mean	STDEV	P-Value
GIMP-0.50-bicubic	0.253	0.292	0.038	0.064	0.161	0.031	0.000
GIMP-0.50-bilinear	0.258	0.271	0.047	0.060	0.138	0.041	0.000
GIMP-0.66-bicubic	0.147	0.076	0.005	0.004	0.205	0.076	0.000
GIMP-0.66-bilinear	0.240	0.122	0.008	0.004	0.279	0.112	0.000
GIMP-0.75-bicubic	0.082	0.076	0.006	0.003	0.111	0.093	0.018
GIMP-0.75-bilinear	0.129	0.061	0.007	0.004	0.168	0.062	0.000
GIMP-0.90-bicubic	0.488	0.043	0.004	0.001	0.583	0.047	0.000
GIMP-0.90-bilinear	0.407	0.195	0.023	0.011	0.310	0.116	0.000
GIMP-1.10-bicubic	0.291	0.130	0.012	0.004	0.292	0.087	0.000
GIMP-1.10-bilinear	0.304	0.109	0.012	0.005	0.317	0.086	0.000
GIMP-1.25-bicubic	0.080	0.137	0.005	0.006	0.084	0.126	0.165
GIMP-1.25-bilinear	0.350	0.128	0.019	0.007	0.302	0.080	0.000
GIMP-1.33-bicubic	0.512	0.082	0.024	0.003	0.381	0.048	0.000
GIMP-1.33-bilinear	0.338	0.189	0.012	0.005	0.324	0.128	0.000
GIMP-2.00-bicubic	0.205	0.104	0.009	0.004	0.242	0.091	0.000
GIMP-2.00-bilinear	0.186	0.178	0.010	0.012	0.190	0.114	0.002
MSPM-0.50-undefined	0.262	0.297	0.048	0.066	0.135	0.044	0.000
MSPM-0.66-undefined	0.993	0.008	0.002	0.001	0.946	0.016	0.000
MSPM-0.75-undefined	0.162	0.075	0.008	0.004	0.204	0.087	0.000
MSPM-0.90-undefined	0.352	0.074	0.005	0.001	0.446	0.075	0.000
MSPM-1.10-undefined	0.747	0.053	0.000	0.000	0.843	0.032	0.000
MSPM-1.25-undefined	0.430	0.131	0.018	0.006	0.362	0.084	0.000
MSPM-1.33-undefined	0.897	0.032	0.000	0.000	0.938	0.020	0.000
MSPM-2.00-undefined	0.990	0.018	0.001	0.000	0.844	0.271	0.000
APS-0.50-bicubic	0.167	0.196	0.024	0.048	0.120	0.050	0.000
APS-0.50-bilinear	0.144	0.221	0.022	0.047	0.109	0.055	0.001
APS-0.50-nearest	0.148	0.227	0.023	0.049	0.104	0.054	0.001
APS-0.66-bicubic	0.069	0.038	0.005	0.004	0.136	0.094	0.005
APS-0.66-bilinear	0.184	0.141	0.006	0.004	0.252	0.113	0.000
APS-0.66-nearest	0.281	0.049	0.004	0.001	0.325	0.119	0.000
APS-0.75-bicubic	0.047	0.032	0.002	0.003	0.069	0.051	0.019
APS-0.75-bilinear	0.056	0.086	0.005	0.007	0.133	0.133	0.032
APS-0.75-nearest	0.706	0.085	0.031	0.009	0.422	0.077	0.000
APS-0.90-bicubic	0.369	0.173	0.020	0.011	0.291	0.090	0.000
APS-0.90-bilinear	0.206	0.139	0.014	0.008	0.304	0.203	0.002
APS-0.90-nearest	0.831	0.035	0.005	0.001	0.767	0.089	0.000
APS-1.10-bicubic	0.722	0.063	0.019	0.007	0.529	0.058	0.000
APS-1.10-bilinear	0.284	0.144	0.011	0.004	0.345	0.171	0.000
APS-1.10-nearest	0.651	0.048	0.007	0.002	0.625	0.051	0.000
APS-1.25-bicubic	0.730	0.052	0.017	0.003	0.515	0.153	0.000
APS-1.25-bilinear	0.193	0.096	0.010	0.008	0.284	0.110	0.000
APS-1.25-nearest	0.461	0.084	0.010	0.004	0.517	0.101	0.000
APS-1.33-bicubic	0.645	0.042	0.003	0.001	0.665	0.136	0.000
APS-1.33-bilinear	0.340	0.142	0.013	0.006	0.386	0.135	0.000
APS-1.33-nearest	0.518	0.059	0.005	0.002	0.530	0.120	0.000
APS-2.00-bicubic	0.387	0.226	0.017	0.012	0.347	0.089	0.000
APS-2.00-bilinear	0.294	0.217	0.020	0.013	0.243	0.140	0.001
APS-2.00-nearest	0.573	0.083	0.013	0.002	0.507	0.089	0.000
Framework	0.384	0.008	0.014	0.000	0.306	0.127	0.000

Table A.35. 2DSD and 1DZC By Rate (TIFF).

Rate	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
0.50	0.537	0.020	0.000	0.120	0.006
0.66	0.898	0.039	0.000	0.008	0.004
0.75	0.992	0.006	0.000	0.001	0.001
0.90	0.997	0.003	0.000	0.000	0.000
1.10	0.997	0.002	0.000	0.000	0.000
1.25	0.998	0.002	0.000	0.000	0.000
1.33	0.998	0.002	0.000	0.000	0.000
2.00	0.994	0.003	0.000	0.001	0.000
Overall	0.886	0.007	0.000	0.016	0.001

Table A.36. 2DSD and 1DZC By Algorithm (TIFF).

Algorithm	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
Bicubic	0.530	0.030	0.000	0.225	0.079
Bilinear	0.529	0.044	0.000	0.230	0.080
Nearest	0.503	0.065	0.000	0.114	0.055
Undefined	0.622	0.143	0.000	0.095	0.073
Overall	0.523	0.023	0.000	0.186	0.011

Table A.38. 1DZC and FMD By Class (TIFF).

Class	TPR		FPR		F-Measure		
	Mean	STDEV	Mean	STDEV	Mean	STDEV	P-Value
GIMP-0.50-bicubic	0.008	0.013	0.002	0.005	0.011	0.017	0.132
GIMP-0.50-bilinear	0.129	0.261	0.041	0.078	0.035	0.051	0.422
GIMP-0.66-bicubic	0.085	0.256	0.024	0.072	0.015	0.045	0.699
GIMP-0.66-bilinear	0.038	0.114	0.014	0.041	0.011	0.033	0.387
GIMP-0.75-bicubic	0.082	0.244	0.025	0.074	0.015	0.042	0.692
GIMP-0.75-bilinear	0.066	0.162	0.018	0.051	0.033	0.047	0.452
GIMP-0.90-bicubic	0.034	0.052	0.006	0.015	0.040	0.063	0.376
GIMP-0.90-bilinear	0.113	0.278	0.030	0.084	0.039	0.080	0.516
GIMP-1.10-bicubic	0.156	0.114	0.012	0.006	0.161	0.104	0.003
GIMP-1.10-bilinear	0.412	0.145	0.018	0.005	0.351	0.089	0.000
GIMP-1.25-bicubic	0.234	0.167	0.012	0.006	0.230	0.103	0.000
GIMP-1.25-bilinear	0.370	0.145	0.017	0.006	0.324	0.107	0.000
GIMP-1.33-bicubic	0.001	0.004	0.003	0.001	0.002	0.007	0.000
GIMP-1.33-bilinear	0.051	0.033	0.002	0.001	0.089	0.056	0.005
GIMP-2.00-bicubic	0.217	0.100	0.012	0.009	0.247	0.122	0.000
GIMP-2.00-bilinear	0.039	0.079	0.010	0.025	0.031	0.036	0.399
MSPM-0.50-undefined	0.085	0.232	0.026	0.070	0.019	0.039	0.894
MSPM-0.66-undefined	0.080	0.234	0.026	0.075	0.014	0.038	0.624
MSPM-0.75-undefined	0.064	0.191	0.021	0.062	0.013	0.037	0.563
MSPM-0.90-undefined	0.107	0.298	0.027	0.079	0.027	0.050	0.697
MSPM-1.10-undefined	0.202	0.151	0.012	0.008	0.203	0.132	0.003
MSPM-1.25-undefined	0.523	0.066	0.013	0.003	0.488	0.042	0.000
MSPM-1.33-undefined	0.928	0.023	0.024	0.003	0.613	0.032	0.000
MSPM-2.00-undefined	0.942	0.023	0.012	0.001	0.753	0.032	0.000
APS-0.50-bicubic	0.188	0.348	0.069	0.134	0.040	0.052	0.311
APS-0.50-bilinear	0.010	0.025	0.002	0.003	0.012	0.032	0.424
APS-0.50-nearest	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APS-0.66-bicubic	0.175	0.330	0.066	0.123	0.032	0.052	0.546
APS-0.66-bilinear	0.067	0.080	0.011	0.021	0.071	0.072	0.066
APS-0.66-nearest	0.001	0.002	0.001	0.001	0.001	0.003	0.000
APS-0.75-bicubic	0.091	0.268	0.034	0.099	0.013	0.035	0.518
APS-0.75-bilinear	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APS-0.75-nearest	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APS-0.90-bicubic	0.000	0.001	0.000	0.000	0.001	0.002	0.000
APS-0.90-bilinear	0.113	0.262	0.043	0.100	0.022	0.045	0.915
APS-0.90-nearest	0.131	0.259	0.032	0.091	0.070	0.113	0.221
APS-1.10-bicubic	0.060	0.065	0.006	0.004	0.076	0.072	0.048
APS-1.10-bilinear	0.639	0.061	0.012	0.002	0.590	0.046	0.000
APS-1.10-nearest	0.215	0.142	0.010	0.005	0.231	0.129	0.001
APS-1.25-bicubic	0.018	0.016	0.004	0.001	0.029	0.025	0.328
APS-1.25-bilinear	0.808	0.035	0.014	0.002	0.662	0.031	0.000
APS-1.25-nearest	0.245	0.033	0.011	0.002	0.285	0.039	0.000
APS-1.33-bicubic	0.173	0.348	0.065	0.130	0.023	0.047	0.873
APS-1.33-bilinear	0.548	0.056	0.014	0.002	0.497	0.044	0.000
APS-1.33-nearest	0.009	0.011	0.001	0.000	0.017	0.020	0.543
APS-2.00-bicubic	0.009	0.021	0.002	0.003	0.013	0.027	0.389
APS-2.00-bilinear	0.166	0.096	0.008	0.007	0.203	0.118	0.001
APS-2.00-nearest	0.068	0.193	0.025	0.070	0.015	0.038	0.648
Framework	0.190	0.009	0.020	0.000	0.142	0.010	0.000

Table A.39. 1DZC and FMD By Rate (TIFF).

Rate	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
0.50	0.230	0.088	0.006	0.136	0.138
0.66	0.194	0.110	0.091	0.139	0.167
0.75	0.256	0.265	0.173	0.095	0.160
0.90	0.244	0.208	0.120	0.133	0.186
1.10	0.762	0.020	0.000	0.028	0.003
1.25	0.813	0.013	0.000	0.025	0.002
1.33	0.670	0.213	0.000	0.087	0.138
2.00	0.483	0.151	0.000	0.066	0.069
Overall	0.382	0.016	0.000	0.089	0.002

Table A.40. 1DZC and FMD By Algorithm (TIFF).

Algorithm	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
Bicubic	0.386	0.013	0.000	0.340	0.158
Bilinear	0.441	0.051	0.000	0.287	0.196
Nearest	0.399	0.049	0.000	0.070	0.093
Undefined	0.458	0.086	0.000	0.153	0.106
Overall	0.397	0.013	0.000	0.246	0.024

Table A.42. 2DSD and FMD By Class (TIFF).

Class	TPR		FPR		F-Measure		
	Mean	STDEV	Mean	STDEV	Mean	STDEV	P-Value
GIMP-0.50-bicubic	0.437	0.375	0.038	0.034	0.232	0.103	0.000
GIMP-0.50-bilinear	0.256	0.260	0.020	0.027	0.188	0.091	0.000
GIMP-0.66-bicubic	0.128	0.053	0.002	0.001	0.205	0.074	0.000
GIMP-0.66-bilinear	0.424	0.094	0.004	0.001	0.517	0.078	0.000
GIMP-0.75-bicubic	0.183	0.168	0.007	0.006	0.214	0.141	0.003
GIMP-0.75-bilinear	0.382	0.213	0.011	0.005	0.367	0.161	0.000
GIMP-0.90-bicubic	0.521	0.053	0.002	0.002	0.645	0.034	0.000
GIMP-0.90-bilinear	0.819	0.076	0.017	0.003	0.623	0.045	0.000
GIMP-1.10-bicubic	0.751	0.069	0.014	0.002	0.623	0.041	0.000
GIMP-1.10-bilinear	0.438	0.091	0.007	0.002	0.496	0.076	0.000
GIMP-1.25-bicubic	0.385	0.305	0.014	0.012	0.311	0.181	0.001
GIMP-1.25-bilinear	0.332	0.262	0.013	0.011	0.292	0.142	0.000
GIMP-1.33-bicubic	0.452	0.283	0.010	0.006	0.427	0.159	0.000
GIMP-1.33-bilinear	0.566	0.322	0.011	0.005	0.484	0.198	0.000
GIMP-2.00-bicubic	0.528	0.213	0.011	0.004	0.504	0.118	0.000
GIMP-2.00-bilinear	0.400	0.211	0.010	0.005	0.400	0.151	0.000
MSPM-0.50-undefined	0.348	0.356	0.030	0.031	0.182	0.116	0.002
MSPM-0.66-undefined	1.000	0.000	0.001	0.000	0.969	0.009	0.000
MSPM-0.75-undefined	0.218	0.154	0.006	0.004	0.265	0.142	0.001
MSPM-0.90-undefined	0.474	0.053	0.004	0.002	0.569	0.043	0.000
MSPM-1.10-undefined	0.765	0.040	0.000	0.000	0.861	0.025	0.000
MSPM-1.25-undefined	0.364	0.303	0.012	0.013	0.328	0.139	0.000
MSPM-1.33-undefined	0.902	0.024	0.000	0.000	0.942	0.016	0.000
MSPM-2.00-undefined	0.974	0.048	0.003	0.001	0.925	0.028	0.000
APS-0.50-bicubic	0.140	0.239	0.011	0.025	0.118	0.087	0.008
APS-0.50-bilinear	0.387	0.376	0.035	0.039	0.204	0.105	0.001
APS-0.50-nearest	0.506	0.410	0.054	0.042	0.201	0.120	0.001
APS-0.66-bicubic	0.054	0.026	0.001	0.001	0.097	0.047	0.001
APS-0.66-bilinear	0.400	0.054	0.004	0.002	0.501	0.034	0.000
APS-0.66-nearest	0.394	0.106	0.004	0.002	0.497	0.099	0.000
APS-0.75-bicubic	0.074	0.069	0.003	0.002	0.116	0.095	0.015
APS-0.75-bilinear	0.223	0.101	0.004	0.001	0.298	0.117	0.000
APS-0.75-nearest	0.758	0.066	0.014	0.001	0.625	0.048	0.000
APS-0.90-bicubic	0.565	0.219	0.010	0.003	0.535	0.131	0.000
APS-0.90-bilinear	0.474	0.205	0.009	0.005	0.471	0.151	0.000
APS-0.90-nearest	0.932	0.034	0.003	0.001	0.904	0.033	0.000
APS-1.10-bicubic	0.649	0.117	0.011	0.003	0.598	0.055	0.000
APS-1.10-bilinear	0.466	0.149	0.008	0.002	0.487	0.097	0.000
APS-1.10-nearest	0.681	0.037	0.006	0.001	0.694	0.030	0.000
APS-1.25-bicubic	0.408	0.230	0.007	0.008	0.446	0.112	0.000
APS-1.25-bilinear	0.567	0.294	0.015	0.008	0.451	0.175	0.000
APS-1.25-nearest	0.544	0.119	0.009	0.002	0.551	0.097	0.000
APS-1.33-bicubic	0.626	0.051	0.003	0.001	0.713	0.035	0.000
APS-1.33-bilinear	0.937	0.027	0.014	0.002	0.723	0.023	0.000
APS-1.33-nearest	0.487	0.049	0.003	0.001	0.598	0.045	0.000
APS-2.00-bicubic	0.502	0.200	0.011	0.005	0.480	0.101	0.000
APS-2.00-bilinear	0.429	0.234	0.010	0.005	0.412	0.158	0.000
APS-2.00-nearest	0.582	0.039	0.009	0.001	0.586	0.032	0.000
Framework	0.506	0.011	0.011	0.000	0.483	0.011	0.000

Table A.43. 2DSD and FMD By Rate (TIFF).

Rate	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
0.50	0.528	0.014	0.000	0.124	0.005
0.66	0.956	0.012	0.000	0.003	0.001
0.75	0.995	0.005	0.000	0.001	0.000
0.90	0.996	0.003	0.000	0.000	0.000
1.10	0.998	0.002	0.000	0.000	0.000
1.25	0.994	0.003	0.000	0.001	0.000
1.33	0.998	0.002	0.000	0.000	0.000
2.00	0.993	0.003	0.000	0.001	0.000
Overall	0.886	0.004	0.000	0.016	0.001

Table A.44. 2DSD and FMD By Algorithm (TIFF).

Algorithm	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
Bicubic	0.529	0.028	0.000	0.207	0.060
Bilinear	0.487	0.025	0.000	0.268	0.061
Nearest	0.573	0.062	0.000	0.108	0.041
Undefined	0.699	0.103	0.000	0.063	0.034
Overall	0.541	0.012	0.000	0.187	0.011

A.3 Group 3

Table A.46. 2DSD, 1DZC, and QMI By Class (JPEG).

Class	TPR		FPR		F-Measure		
	Mean	STDEV	Mean	STDEV	Mean	STDEV	P-Value
GIMP-0.50-bicubic	0.318	0.227	0.031	0.030	0.221	0.064	0.000
GIMP-0.50-bilinear	0.346	0.344	0.039	0.042	0.166	0.104	0.002
GIMP-0.66-bicubic	0.434	0.365	0.041	0.038	0.216	0.123	0.001
GIMP-0.66-bilinear	0.307	0.055	0.002	0.001	0.433	0.063	0.000
GIMP-0.75-bicubic	0.384	0.127	0.005	0.002	0.458	0.109	0.000
GIMP-0.75-bilinear	0.250	0.120	0.005	0.003	0.323	0.126	0.000
GIMP-0.90-bicubic	0.425	0.049	0.005	0.001	0.507	0.052	0.000
GIMP-0.90-bilinear	0.308	0.056	0.003	0.001	0.425	0.058	0.000
GIMP-1.10-bicubic	0.655	0.063	0.006	0.002	0.671	0.043	0.000
GIMP-1.10-bilinear	0.729	0.069	0.005	0.001	0.735	0.046	0.000
GIMP-1.25-bicubic	0.582	0.063	0.007	0.002	0.618	0.058	0.000
GIMP-1.25-bilinear	0.639	0.057	0.007	0.001	0.644	0.051	0.000
GIMP-1.33-bicubic	0.301	0.165	0.015	0.017	0.303	0.071	0.000
GIMP-1.33-bilinear	0.146	0.153	0.008	0.011	0.161	0.100	0.002
GIMP-2.00-bicubic	0.441	0.275	0.009	0.005	0.417	0.220	0.000
GIMP-2.00-bilinear	0.371	0.274	0.006	0.005	0.390	0.195	0.000
MSPM-0.50-undefined	0.971	0.019	0.007	0.002	0.838	0.026	0.000
MSPM-0.66-undefined	1.000	0.000	0.000	0.000	0.853	0.078	0.000
MSPM-0.75-undefined	0.701	0.052	0.000	0.000	0.955	0.064	0.000
MSPM-0.90-undefined	1.000	0.000	0.000	0.000	0.850	0.261	0.000
MSPM-1.10-undefined	0.993	0.009	0.000	0.000	0.539	0.351	0.002
MSPM-1.25-undefined	1.000	0.000	0.000	0.000	0.253	0.262	0.026
MSPM-1.33-undefined	0.973	0.023	0.000	0.000	0.156	0.158	0.031
MSPM-2.00-undefined	0.945	0.073	0.001	0.000	0.171	0.087	0.001
APS-0.50-bicubic	0.237	0.223	0.032	0.041	0.276	0.158	0.001
APS-0.50-bilinear	0.158	0.207	0.026	0.036	0.380	0.123	0.000
APS-0.50-nearest	0.230	0.296	0.034	0.055	0.394	0.179	0.000
APS-0.66-bicubic	0.378	0.375	0.064	0.064	0.511	0.218	0.000
APS-0.66-bilinear	0.317	0.041	0.003	0.001	0.600	0.133	0.000
APS-0.66-nearest	0.348	0.054	0.003	0.001	0.520	0.093	0.000
APS-0.75-bicubic	0.121	0.040	0.003	0.001	0.546	0.103	0.000
APS-0.75-bilinear	0.770	0.080	0.003	0.002	0.653	0.154	0.000
APS-0.75-nearest	0.568	0.097	0.008	0.003	0.772	0.112	0.000
APS-0.90-bicubic	0.403	0.051	0.006	0.001	0.786	0.070	0.000
APS-0.90-bilinear	0.409	0.074	0.007	0.001	0.794	0.028	0.000
APS-0.90-nearest	0.732	0.035	0.001	0.001	0.806	0.069	0.000
APS-1.10-bicubic	0.681	0.050	0.003	0.001	0.861	0.050	0.000
APS-1.10-bilinear	0.883	0.062	0.004	0.001	0.770	0.221	0.000
APS-1.10-nearest	0.717	0.063	0.002	0.001	0.560	0.279	0.000
APS-1.25-bicubic	0.738	0.061	0.002	0.001	0.441	0.200	0.000
APS-1.25-bilinear	0.984	0.013	0.003	0.001	0.450	0.149	0.000
APS-1.25-nearest	0.808	0.036	0.001	0.001	0.490	0.171	0.000
APS-1.33-bicubic	0.205	0.095	0.016	0.012	0.427	0.122	0.000
APS-1.33-bilinear	0.248	0.048	0.003	0.001	0.343	0.147	0.000
APS-1.33-nearest	0.573	0.049	0.002	0.001	0.349	0.144	0.000
APS-2.00-bicubic	0.425	0.318	0.011	0.009	0.216	0.210	0.021
APS-2.00-bilinear	0.427	0.324	0.010	0.009	0.055	0.114	0.385
APS-2.00-nearest	0.246	0.111	0.016	0.010	0.033	0.099	0.720
Framework	0.544	0.006	0.010	0.000	0.000	0.000	0.000

Table A.47. 2DSD, 1DZC, and QMI By Rate (JPEG).

Rate	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
0.50	0.401	0.083	0.000	0.156	0.070
0.66	0.530	0.170	0.000	0.106	0.067
0.75	0.945	0.014	0.000	0.005	0.001
0.90	0.985	0.004	0.000	0.002	0.000
1.10	0.986	0.006	0.000	0.002	0.001
1.25	0.999	0.002	0.000	0.000	0.000
1.33	0.675	0.091	0.000	0.038	0.017
2.00	0.892	0.059	0.000	0.015	0.009
Overall	0.719	0.008	0.000	0.040	0.001

Table A.48. 2DSD, 1DZC, and QMI By Algorithm (JPEG).

Algorithm	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
Bicubic	0.572	0.021	0.000	0.258	0.058
Bilinear	0.639	0.047	0.000	0.159	0.064
Nearest	0.682	0.117	0.000	0.066	0.055
Undefined	0.999	0.001	0.000	0.000	0.000
Overall	0.668	0.014	0.000	0.150	0.010

Table A.50. 2DSD, 1DZC, and QMI By Class (TIFF).

Class	TPR		FPR		F-Measure		
	Mean	STDEV	Mean	STDEV	Mean	STDEV	P-Value
GIMP-0.50-bicubic	0.396	0.289	0.032	0.029	0.254	0.078	0.000
GIMP-0.50-bilinear	0.324	0.308	0.028	0.030	0.208	0.091	0.000
GIMP-0.66-bicubic	0.203	0.072	0.004	0.001	0.288	0.086	0.000
GIMP-0.66-bilinear	0.360	0.086	0.004	0.001	0.469	0.073	0.000
GIMP-0.75-bicubic	0.118	0.069	0.003	0.002	0.175	0.084	0.000
GIMP-0.75-bilinear	0.509	0.107	0.013	0.004	0.480	0.061	0.000
GIMP-0.90-bicubic	0.492	0.045	0.002	0.001	0.626	0.044	0.000
GIMP-0.90-bilinear	0.783	0.086	0.017	0.002	0.605	0.052	0.000
GIMP-1.10-bicubic	0.681	0.053	0.013	0.002	0.591	0.035	0.000
GIMP-1.10-bilinear	0.477	0.074	0.008	0.002	0.518	0.063	0.000
GIMP-1.25-bicubic	0.507	0.168	0.012	0.005	0.485	0.090	0.000
GIMP-1.25-bilinear	0.361	0.173	0.011	0.007	0.360	0.100	0.000
GIMP-1.33-bicubic	0.656	0.106	0.010	0.002	0.614	0.069	0.000
GIMP-1.33-bilinear	0.525	0.109	0.008	0.002	0.544	0.076	0.000
GIMP-2.00-bicubic	0.449	0.144	0.009	0.004	0.480	0.083	0.000
GIMP-2.00-bilinear	0.608	0.185	0.012	0.003	0.544	0.105	0.000
MSPM-0.50-undefined	0.347	0.315	0.028	0.027	0.227	0.082	0.000
MSPM-0.66-undefined	1.000	0.001	0.001	0.000	0.974	0.009	0.000
MSPM-0.75-undefined	0.304	0.155	0.006	0.004	0.361	0.122	0.000
MSPM-0.90-undefined	0.449	0.044	0.006	0.002	0.522	0.041	0.000
MSPM-1.10-undefined	0.746	0.047	0.001	0.000	0.842	0.036	0.000
MSPM-1.25-undefined	0.483	0.117	0.010	0.003	0.472	0.072	0.000
MSPM-1.33-undefined	0.912	0.027	0.000	0.000	0.946	0.013	0.000
MSPM-2.00-undefined	1.000	0.001	0.000	0.000	0.423	0.327	0.005
APS-0.50-bicubic	0.374	0.290	0.038	0.032	0.167	0.091	0.001
APS-0.50-bilinear	0.206	0.278	0.020	0.027	0.203	0.082	0.000
APS-0.50-nearest	0.395	0.329	0.035	0.036	0.147	0.099	0.004
APS-0.66-bicubic	0.124	0.176	0.009	0.019	0.387	0.133	0.000
APS-0.66-bilinear	0.365	0.068	0.004	0.001	0.466	0.089	0.000
APS-0.66-nearest	0.380	0.091	0.004	0.002	0.295	0.143	0.000
APS-0.75-bicubic	0.142	0.085	0.003	0.001	0.250	0.122	0.000
APS-0.75-bilinear	0.184	0.110	0.005	0.002	0.499	0.192	0.000
APS-0.75-nearest	0.730	0.063	0.014	0.004	0.573	0.099	0.000
APS-0.90-bicubic	0.608	0.195	0.011	0.004	0.549	0.075	0.000
APS-0.90-bilinear	0.479	0.200	0.009	0.004	0.772	0.230	0.000
APS-0.90-nearest	0.942	0.027	0.002	0.001	0.756	0.095	0.000
APS-1.10-bicubic	0.816	0.038	0.010	0.002	0.695	0.028	0.000
APS-1.10-bilinear	0.602	0.046	0.003	0.001	0.704	0.033	0.000
APS-1.10-nearest	0.672	0.054	0.005	0.001	0.688	0.033	0.000
APS-1.25-bicubic	0.691	0.068	0.007	0.002	0.659	0.067	0.000
APS-1.25-bilinear	0.711	0.089	0.009	0.001	0.580	0.063	0.000
APS-1.25-nearest	0.495	0.058	0.007	0.002	0.685	0.095	0.000
APS-1.33-bicubic	0.720	0.080	0.004	0.002	0.737	0.044	0.000
APS-1.33-bilinear	0.865	0.092	0.011	0.003	0.696	0.065	0.000
APS-1.33-nearest	0.558	0.055	0.003	0.001	0.516	0.121	0.000
APS-2.00-bicubic	0.527	0.182	0.012	0.005	0.442	0.116	0.000
APS-2.00-bilinear	0.392	0.202	0.009	0.005	0.543	0.116	0.000
APS-2.00-nearest	0.588	0.041	0.008	0.001	0.543	0.034	0.000
Framework	0.532	0.008	0.010	0.000	0.145	0.207	0.104

Table A.51. 2DSD, 1DZC, and QMI By Rate (TIFF).

Rate	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
0.50	0.537	0.018	0.000	0.118	0.015
0.66	0.885	0.082	0.000	0.012	0.016
0.75	0.990	0.006	0.000	0.001	0.001
0.90	0.995	0.002	0.000	0.001	0.000
1.10	0.999	0.001	0.000	0.000	0.000
1.25	0.995	0.003	0.000	0.001	0.000
1.33	0.998	0.002	0.000	0.000	0.000
2.00	0.992	0.002	0.000	0.001	0.000
Overall	0.883	0.005	0.000	0.017	0.001

Table A.52. 2DSD, 1DZC, and QMI By Algorithm (TIFF).

Algorithm	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
Bicubic	0.543	0.028	0.000	0.229	0.050
Bilinear	0.531	0.030	0.000	0.233	0.038
Nearest	0.623	0.067	0.000	0.085	0.037
Undefined	0.710	0.082	0.000	0.058	0.028
Overall	0.572	0.011	0.000	0.178	0.006

A.4 Group 4

Table A.54. All Modules By Class (TIFF).

Class	TPR		FPR		F-Measure		
	Mean	STDEV	Mean	STDEV	Mean	STDEV	P-Value
GIMP-0.50-bicubic	0.320	0.245	0.039	0.042	0.197	0.037	0.000
GIMP-0.50-bilinear	0.509	0.313	0.057	0.035	0.216	0.085	0.000
GIMP-0.66-bicubic	0.154	0.168	0.011	0.019	0.154	0.076	0.001
GIMP-0.66-bilinear	0.280	0.068	0.001	0.001	0.411	0.086	0.000
GIMP-0.75-bicubic	0.470	0.091	0.007	0.002	0.532	0.070	0.000
GIMP-0.75-bilinear	0.185	0.121	0.003	0.002	0.260	0.149	0.001
GIMP-0.90-bicubic	0.467	0.032	0.005	0.001	0.548	0.019	0.000
GIMP-0.90-bilinear	0.294	0.041	0.003	0.001	0.406	0.050	0.000
GIMP-1.10-bicubic	0.652	0.043	0.007	0.001	0.664	0.029	0.000
GIMP-1.10-bilinear	0.702	0.055	0.006	0.001	0.714	0.040	0.000
GIMP-1.25-bicubic	0.576	0.049	0.007	0.002	0.615	0.040	0.000
GIMP-1.25-bilinear	0.645	0.086	0.007	0.001	0.653	0.067	0.000
GIMP-1.33-bicubic	0.224	0.088	0.009	0.008	0.268	0.083	0.000
GIMP-1.33-bilinear	0.272	0.180	0.023	0.016	0.217	0.087	0.000
GIMP-2.00-bicubic	0.508	0.237	0.009	0.004	0.485	0.187	0.000
GIMP-2.00-bilinear	0.333	0.175	0.006	0.004	0.396	0.137	0.000
MSPM-0.50-undefined	0.983	0.009	0.008	0.002	0.840	0.025	0.000
MSPM-0.66-undefined	1.000	0.000	0.000	0.000	0.802	0.036	0.000
MSPM-0.75-undefined	0.677	0.047	0.000	0.000	0.970	0.034	0.000
MSPM-0.90-undefined	1.000	0.000	0.000	0.000	0.749	0.306	0.000
MSPM-1.10-undefined	0.994	0.006	0.000	0.000	0.210	0.139	0.003
MSPM-1.25-undefined	1.000	0.000	0.000	0.000	0.138	0.057	0.000
MSPM-1.33-undefined	0.988	0.017	0.000	0.000	0.123	0.039	0.000
MSPM-2.00-undefined	0.932	0.086	0.000	0.000	0.185	0.091	0.000
APS-0.50-bicubic	0.242	0.269	0.033	0.054	0.379	0.065	0.000
APS-0.50-bilinear	0.202	0.302	0.031	0.052	0.388	0.130	0.000
APS-0.50-nearest	0.296	0.331	0.045	0.062	0.371	0.245	0.002
APS-0.66-bicubic	0.218	0.308	0.035	0.053	0.700	0.122	0.000
APS-0.66-bilinear	0.297	0.055	0.004	0.001	0.608	0.078	0.000
APS-0.66-nearest	0.315	0.083	0.002	0.001	0.479	0.065	0.000
APS-0.75-bicubic	0.159	0.167	0.016	0.037	0.571	0.118	0.000
APS-0.75-bilinear	0.807	0.056	0.002	0.001	0.766	0.062	0.000
APS-0.75-nearest	0.599	0.086	0.007	0.002	0.795	0.034	0.000
APS-0.90-bicubic	0.409	0.065	0.006	0.001	0.828	0.036	0.000
APS-0.90-bilinear	0.409	0.057	0.006	0.002	0.803	0.021	0.000
APS-0.90-nearest	0.706	0.065	0.001	0.001	0.851	0.063	0.000
APS-1.10-bicubic	0.718	0.029	0.003	0.001	0.905	0.023	0.000
APS-1.10-bilinear	0.875	0.065	0.004	0.001	0.703	0.253	0.000
APS-1.10-nearest	0.726	0.025	0.003	0.001	0.283	0.151	0.001
APS-1.25-bicubic	0.768	0.072	0.002	0.001	0.430	0.123	0.000
APS-1.25-bilinear	0.970	0.012	0.002	0.001	0.538	0.156	0.000
APS-1.25-nearest	0.834	0.036	0.001	0.001	0.455	0.140	0.000
APS-1.33-bicubic	0.168	0.106	0.011	0.010	0.319	0.109	0.000
APS-1.33-bilinear	0.245	0.054	0.002	0.001	0.332	0.119	0.000
APS-1.33-nearest	0.539	0.053	0.002	0.001	0.374	0.190	0.000
APS-2.00-bicubic	0.426	0.264	0.011	0.007	0.050	0.088	0.346
APS-2.00-bilinear	0.383	0.252	0.011	0.008	0.000	0.000	0.000
APS-2.00-nearest	0.261	0.112	0.018	0.010	0.000	0.000	0.000
Framework	0.542	0.013	0.011	0.000	0.000	0.000	0.000

Table A.55. All Modules By Rate (TIFF).

Rate	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
0.50	0.358	0.022	0.000	0.196	0.056
0.66	0.697	0.180	0.000	0.048	0.055
0.75	0.895	0.148	0.000	0.019	0.040
0.90	0.982	0.008	0.000	0.002	0.001
1.10	0.984	0.004	0.000	0.002	0.001
1.25	0.998	0.002	0.000	0.000	0.000
1.33	0.652	0.056	0.000	0.041	0.016
2.00	0.883	0.062	0.000	0.016	0.009
Overall	0.717	0.014	0.000	0.041	0.002

Table A.56. All Modules By Algorithm (TIFF).

Algorithm	TPR			FPR	
	Mean	STDEV	P-Value	Mean	STDEV
Bicubic	0.585	0.046	0.000	0.222	0.083
Bilinear	0.623	0.051	0.000	0.188	0.073
Nearest	0.658	0.115	0.000	0.079	0.060
Undefined	1.000	0.000	0.000	0.000	0.000
Overall	0.662	0.010	0.000	0.150	0.015

Bibliography

1. Adobe. *TIFF*. Technical Report 6.0, Adobe Systems Incorporated, 1992. URL <https://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf>.
2. Avcibas, Ismail, Sevinc Bayram, Nasir Memon, Mahalingam Ramkumar, and Bulent Sankur. “A classifier design for detecting image manipulations”. *Image Processing, 2004 International Conference on, 2004. ICIP'04.*, volume 4, 2645–2648. IEEE, 2004.
3. Ball, Justin and Jordan Keefer. “Detecting Resizing Software”, 2013.
4. Bayram, Sevinc, Bulent Sankur, Nasir Memon, and Ismail Avcibas. “Image manipulation detection”. *Journal of Electronic Imaging*, 15(4):041102–041102–17, 2006.
5. CCITT. *Terminal Equipment and Protocols for Telematic Services: Information Technology - Digital Compression and Coding of Continuous-tone Still Images - Requirements and Guidelines*. Recommendation T.81, United Nations: International Telecommunications Union, 1992. URL <http://www.w3.org/Graphics/JPEG/itu-t81.pdf>.
6. Cohen, Jacob. “A coefficient of agreement for nominal scales”. *Educational and Psychological Measurement*, 20(1):37, 1960.
7. Committee, Technical Standardization. *Exchangeable image file format for digital still cameras: Exif Version 2.2*. Standard JEITA CP-3451, Japan Electronics and Information Technology Industries Association, 2002. URL <http://web.archive.org/web/20131018091152/http://exif.org/Exif2-2.PDF>.
8. Cooper, Gregory F. and Edward Herskovits. “A Bayesian method for the induction of probabilistic networks from data”. *Machine Learning*, 9(4):309–347, 1992.
9. Farid, Hany. “Image forgery detection”. *Signal Processing Magazine, IEEE*, 26(2):16–25, 2009.
10. Fontani, Marco, Tiziano Bianchi, Alessia De Rosa, Alessandro Piva, and Mauro Barni. “A Forensic Tool for Investigating Image Forgeries”. *International Journal of Digital Crime and Forensics (IJDCF)*, 5(4):15–33, 2013.
11. Gallagher, Andrew. “Detection of linear and cubic interpolation in JPEG compressed images”. *Computer and Robot Vision, The 2nd Canadian Conference on, 2005. Proceedings.*, 65–72. IEEE, 2005.
12. Group, Independent JPEG. “libjpeg”, 2014.

13. Hall, Mark, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. “The WEKA Data Mining Software: An Update”. *SIGKDD Explorations*, 11(1), 2009.
14. Hamilton, Eric. *JPEG File Interchange Format*. Standard 1.02, Joint Photographic Experts Group, 1992. URL <http://www.jpeg.org/public/jfif.pdf>.
15. Keys, Robert. “Cubic convolution interpolation for digital image processing”. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 29(6):1153–1160, 1981.
16. Kornblum, Jesse. “Using JPEG quantization tables to identify imagery processed by software”. *Digital Investigation*, 5:S21–S25, 2008.
17. Levine, Brian Neil and Marc Liberatore. “Dex: Digital evidence provenance supporting reproducibility and comparison”. *Digital Investigation*, 6:S48–S56, 2009.
18. Luo, Weiqi, Zhenhua Qu, Jiwu Huang, and Guoping Qiu. “A novel method for detecting cropped and recompressed image block”. *Acoustics, Speech and Signal Processing, IEEE International Conference on, 2007. ICASSP 2007.*, volume 2, II–217–II–220. IEEE, 2007.
19. Mahdian, Babak and Stanislav Saic. “Blind authentication using periodic properties of interpolation”. *Information Forensics and Security, IEEE Transactions on*, 3(3):529–538, 2008.
20. Ouwerkerk, JD Van. “Image super-resolution survey”. *Image and Vision Computing*, 24(10):1039–1052, 2006.
21. Parker, Anthony, Robert Kenyon, and Donald Troxel. “Comparison of interpolating methods for image resampling”. *Medical Imaging, IEEE Transactions on*, 2(1):31–39, 1983.
22. Peterson, Gilbert. *Forensic analysis of digital image tampering*, 259–270. Advances in Digital Forensics. Springer, 2005.
23. Photography, Technical Committee. *Photography and graphic technology - Extended colour encodings for digital image storage, manipulation and interchange*. Technical Specification 22028, International Organization for Standardization, 2012. URL <https://www.iso.org/obp/ui/#iso:std:iso:ts:22028:-3:ed-2:v1:en>.
24. Popescu, Alin and Hany Farid. “Exposing digital forgeries by detecting traces of resampling”. *Signal Processing, IEEE Transactions on*, 53(2):758–767, 2005.
25. Popescu, Alin and Hany Farid. “Statistical tools for digital forensics”. *Information Hiding*, 128–147. Springer, 2005.

26. Prasad, S. and KR Ramakrishnan. “On resampling detection and its application to detect image tampering”. *Multimedia and Expo, IEEE International Conference on, 2006*, 1325–1328. IEEE, 2006.
27. Pudil, Pavel, FJ Ferri, J. Novovicova, and J. Kittler. “Floating search methods for feature selection with nonmonotonic criterion functions”. *Pattern Recognition, Proceedings of the Twelveth International Conference on, IAPR*. Citeseer, 1994.
28. Witten, Ian H., Eibe Frank, and Mark Hall. *Data mining: practical machine learning tools and techniques*. Morgan Kaufmann, United States, third edition, 2011.
29. Wittman, Todd. “Mathematical techniques for image interpolation”. *Report Submitted for Completion of Mathematics Department Oral Exam, Department of Mathematics, University of Minnesota, USA*, 2005.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE (DD-MM-YYYY) 26-03-2015		2. REPORT TYPE Master's Thesis		3. DATES COVERED (From — To) June 2013 — Mar 2015	
4. TITLE AND SUBTITLE Identifying Image Manipulation Software From Image Features				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Boyter, Devlin, T., CPT, USA				5d. PROJECT NUMBER N/A	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way WPAFB OH 45433-7765				8. PERFORMING ORGANIZATION REPORT NUMBER AFIT-ENG-MS-15-M-051	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Mr. Chad Heitzenreiter Air Force Research Laboratory/Multi-Sensor Exploitation Branch 525 Brooks Road Building 3 Rome, NY 13441 Phone: 315-330-2739 Email: chad.heitzenrater@rl.af.mil				10. SPONSOR/MONITOR'S ACRONYM(S) AFRL/RIGB	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT DISTRIBUTION STATEMENT A: APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.					
13. SUPPLEMENTARY NOTES This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.					
14. ABSTRACT As technology steadily increases in the field of image manipulation, determining which software was used to manipulate an image becomes increasingly complex for law enforcement and intelligence agencies. To combat this difficult problem, new techniques that examine the artifacts left behind by a specific manipulation are converted to features for classification. This research implemented four preexisting image manipulation detection techniques into a framework of modules: Two-Dimensional Second Derivative, One-Dimensional Zero Crossings, Quantization Matrices Identification, and File Metadata analysis. The intent is the creation of a framework to develop a capability to determine which specific image manipulation software program manipulated an image. The determination is based on each image manipulation software program having implemented the manipulation algorithms differently. These differences in the implementation will leave behind different artifacts in the resultant image. Experimental results demonstrate the framework's ability to determine the image manipulation software program.					
15. SUBJECT TERMS image,manipulation,feature,detection					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. G. L. Peterson, AFIT/ENG
U	U	U	U	121	19b. TELEPHONE NUMBER (include area code) (937) 255-3636, x4281; gilbert.peterson@afit.edu