

# Technological University Dublin ARROW@TU Dublin

Conference papers

School of Electrical and Electronic Engineering

2020

# Modern Techniques for Discovering Digital Steganography

**Michael Hegarty** 

Anthony J. Keane

Follow this and additional works at: https://arrow.tudublin.ie/engscheleart

Part of the Electrical and Computer Engineering Commons

This Conference Paper is brought to you for free and open access by the School of Electrical and Electronic Engineering at ARROW@TU Dublin. It has been accepted for inclusion in Conference papers by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie, gerard.connolly@tudublin.ie.



This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 4.0 License



## Modern Techniques for Discovering Digital Steganography

## Michael T. Hegarty and Anthony J. Keane TU-Dublin, Blanchardstown, Ireland Michael.hegarty@tudublin.ie Anthony.keane@tudublin.ie

DOI: 10.34190/EWS.20.504

**Abstract**: Digital steganography can be difficult to detect and as such is an ideal way of engaging in covert communications across the Internet. This research paper is a work-in-progress report on instances of steganography that were identified on websites on the Internet including some from the DarkWeb using the application of new methods of deep learning algorithms. This approach to the identification of Least Significant Bit (LSB) Steganography using Convolutional Neural Networks (CNN) has demonstrated some efficiency for image classification. The CNN algorithm was trained using datasets of images with known steganography and then applied to datasets with images to identify concealed data. The algorithm was trained using 5000 clean images and 5000 Steganography images. With the correct configurations made to the deep learning algorithms, positive results were obtained demonstrating a greater speed, accuracy and fewer false positives than the current steganalysis tools.

Keywords: Steganography, Deep Learning, JPEG, LSB, Darknet, Openpuff

### 1. Introduction

Steganography can be explained as the art and science of writing hidden messages. Using this method no one apart from the sender and intended recipient suspects the existence of the message. In the modern digital world, Steganography methods are many as information can be hidden not only within graph images but also within other digital carriers such as video and audio (Kessler, 2015).

According to (Zielińska *et al*, 2014), technology makes it easy to perform digitalized methods of concealing information with many of the applications available for free from internet portals. Currently, there are approximately 1200 tools available to create different forms of steganography (Hegarty, 2018). These tools can be easily used by criminals to communicate undetected and challenges the cyber security specialists to establish if covert communication has occurred. Digital evidence of communication may never be detected on a suspect's machine due to it being concealed within a carrier file (Hunt, 2012). Richer (2015) suggests that Steganography can be used to communicate with complete freedom under conditions that are monitored.

While there is the suspicion that terror groups are using technologies like Steganography to help conceal their communications on the Internet, there is little solid evidence to support that claim. In May 2011 a 22-year-old Austrian citizen Maqsood Lodin was stopped and questioned in Germany. He was traveling from Pakistan to Berlin via Hungary. During a search, a memory stick was found in his underpants. The memory stick contained two pornographic videos. Over 100 documents were concealed within data on the storage device. They contained secret documents about operational details and training manuals for al-Qaeda. Its source came from a German Magazine "Die Zeit". The journalist and author of the article Yassin Musharbash has published an investigative report on his personal blog that includes a summary of the findings (Musharbash, 2012).

There are examples of the use of Steganography in industrial espionage when in 2018, Zheng Xiaoqing and Zhang Zhaoxi were both indicted for their role in stealing proprietary trade secret information related to GE's turbine technology. The indictment identified highly sophisticated steganography measures deployed to hide the stolen GE trade secrets in JPeg images files and sending them to a personal Hotmail account (Lynch and Shepardson, 2019).

#### 2. Digital Steganalysis

Steganalysis involves examining several parameters of media type for Steganography to check for hidden data. Statistical Steganalysis is when alterations are applied to a carrier file, there is a change in the statistical information such as changes made to the LSB of the original image file in order to hide information (Ker *et al.*, 2013).

The detection of files with hidden data required that the servers/applications that store the file do not change the files or the secret payload can be destroyed. Ebay does this to files to fit its requirements and renders the files useless for retrieving hidden information. Web-portals on the clear and darknet were identified as places that steganography could survive as the files are stored in original formats. Some of these web-portals that were researched are:

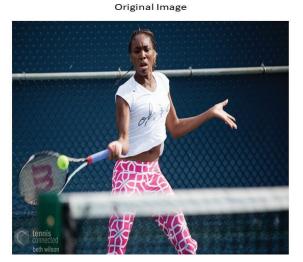
- Draugiem: The main social media platform in Latvia with 2.6 million users
- VKontakte (VK): Russia's largest social media platform claiming to have 500 million registered accounts
- ISIS Darknet: wpcxzq4ykmsxpacm.onion a propaganda outlet for Islamic state on the darknet

A dataset of 3 million images was harvested from clear-net and dark-net webportals. This data was analysed using existing steganalysis tools such as Stegdetect, Stegexpose, Outguess, and Virtual Steganographic Laboratory (VSL). Steganalysis tools will return a probability and categorise findings. Highly probable steganography images were then further analysed. Of the 3 million images analysed 0.02% (600 images) returned as suspicious of concealing data.

#### 3. Machine Learning for Steganalysis

According to (Giarimpampa, 2018), "Neural Networks have proven to be a useful tool in image Steganalysis". The use of both Support Vector Machine (SVM) and Artificial Neural Networks (ANN) algorithms have become the predominant modern Steganalysis trend.

Using Python, Anaconda, on the Spyder IDE, 5000 clean images with no concealed data was developed. 5000 steganography images were created using OpenPuff steganography tool. OpenPuff is the most downloaded steganography tool according to Softpedia with 60,352 downloads as of September 2019. Figure 1 displays an original image on the left and the carrier image (with a secret massage) on the right, this secret message was created using OpenPuff steganography software. Figure 2 presents a sample of the HEX code and the changes OpenPuff has created to hide the secret message in the LSB.



Steganography Image OpenPuff



Figure 1: Sample original image and same image with hidden message

000000555         14         88         16         38         42         0.0         0.3         14         76				
Norm         Sexion         At         F         T <tht< td=""><td>000000463174.jpg &lt;&gt; 000000463174TESTSTEGO.jpg - Hex Compare -</td><td>Beyond Compare</td><td></td><td></td></tht<>	000000463174.jpg <> 000000463174TESTSTEGO.jpg - Hex Compare -	Beyond Compare		
Norm         Sexion         At         F         T <tht< td=""><td>Session File Edit Search View Tools Help</td><td></td><td></td><td></td></tht<>	Session File Edit Search View Tools Help			
North         Sension         All         Diffs         Same         Paires         Priority         Next Diff         Year         Read           CCL_Train 1000 Clean image Dataset/train1700000046317A/grg         Colume         Clume			- A	
CX_Train 1000 Clean image DatasetUtion1V0000004631741pg C2_UserVinickHeg/Desktop/Step/0000004631741ESISEGO.pp C2_UserVinickHeg/Desktop/Step/0000004631741ESISEGO.pp C2_UserVinickHeg/Desktop/Step/000004631741ESISEGO.pp C2_UserVinickHeg/Desktop/Step/000004631741ESISEGO.pp C2_UserVinickHeg/Desktop/Step/000004631741ESISEGO.pp C2_UserVinickHeg/Desktop/Step/000004631741ESISEGO.pp C2_UserVinickHeg/Desktop/Step/000004631741ESISEGO.pp C2_UserVinickHeg/Desktop/Step/000004631741ESISEGO.pp C2_UserVinicHeg/Desktop/Step/000004631741ESISEGO.pp C2_UserVinicHeg/Desktop/Step/000004631741ESISEGO.pp C2_UserVinicHeg/Desktop/Step/000004631741ESISEGO.pp C2_UserVinicHeg/Desktop/Step/000004631741ESISEGO.pp C2_UserVinicHeg/Desktop/Step/000004631741ESISEGO.pp C2_UserVinicHeg/Desktop/Step/000004631741ESISEGO.pp C2_UserVinicHeg/Desktop/Step/000004631741ESISEGO.pp C2_UserVinicHeg/Desktop/Step/000004631741ESISEGO.pp C2_UserVinicHeg/Desktop/Step/000004631741ESISEGO.pp C2_UserVinicHeg/Desktop/Step/001631741ESISEGO.pp C2_UserVinicHeg/Desktop/Step/01620630 Step/10 PA 52 PT 82 PA 52 PA 5				
0000055FB A4 A6 14 F4 39 43 83 F5       27 78 A2 14 CA 70 62 44       F1.00c76'(4.5)         0000055FC AA A6 14 F4 39 43 83 F5       27 78 A2 14 CA 75 62 44       F1.00c76'(4.5)         0000050FC AA A6 14 F4 39 43 83 F5       27 78 A2 14 CA 75 62 44       F1.00c76'(4.5)         0000050FC AA A6 14 F4 39 43 83 F5       27 78 A2 14 CA 75 62 44       F1.00c76'(4.5)         0000050FC A5 28 0A 50 FF 80 6A 0F       78 63 B4 110 0F F6 5C A5       A*TKy.j5xC.50 Y         0000050FC A5 28 0A 50 FF 80 6A 0F       78 63 B4 110 0F F6 5C A5       A*TKy.j5xC.50 Y         0000050FC A5 28 0A 70 50 F8 155 78 155 71 0A 96 80 95 D       %000050FC A5 28 0A 90 51 CB 50 83 55 73 15 73 0A 96 80 95 D       %000050FC A5 28 0A 10 73 15 9 C9       62 BA 1E 10 FA 18 72 91       12 (-1.745)6         0000050FC A5 28 0A 70 D8 55 78 155 71 0A 96 80 95 D       %000050FC A3 30 97 61 40 28 0 P0 85       56 A4 66 E5 33 55 73 10 C       %000050FC A3 30 97 50 14 A 20 55 78 115 71 0A 96 80 95 D       %000000FC A1 A 20 57 89 51 AA A6 65 53 55 73 10 C       %000050FC A4 80 93 55 78 115 71 0A 96 80 95 D       %00000FC A1 A 20 57 89 51 AA A6 65 73 97 10 16 E1 69 9F 72       %00000FC A1 A 20 57 84 51 AA 64 55 35 5F 20 1C       %00000FC A1 A 20 57 84 51 AA 64 83 73 11 A 74 18 80 71 1A 74 18 74 80 70 1A 74 80 80 71 1A 74 74 80 70 71 1A 74 80 80 FC 75 10 86 AA A6 80 71 1A 74 78 80 70 71 1A 74 80 80 FC 74 18 74 84 81 A6 80 80 71 1A 74 74 80 70 74 80 70 74 80 78 74 80 80 76 74 80 78 74 80 78 74 80 78 74 78 80 70 74 80 78 74 80 78 74	Home Sessions All Diffs Same Rules Pormat Copy			
000005001 Ar A 20 D4 F 30 A 20 A 20 BC A5 10 D4 A5 22 L 23 BC C8 16 BD C6       ((00000000000000000000000000000000000	C:\\Train 1000 Clean Image Dataset\train1\000000463174.jpg			
Debesche Ab 26 BA 86 DC F78 68 R4 110 DF 78 63 R4 118 DF F6 5C A5         Alves, jacc. Bark         Devester BA 36 Nc FF 98 6A DF 78 63 R4 110 DF F6 5C A5         Alves, jacc. Bark           Debesche DC 52 67 DB 6A 32 E1 55 08 33 19 73 42 82 40 C2 6F         Alves, jacc. Bark         Devester BA 36 Nc FF 98 6A DF 78 51 84 18 DF F6 5C A5         Alves, jacc. Bark           Debesche DC 52 67 DB 78 110 7 119 50 C3         BC 28 AL E1 DE A1 18 72 91         La (1, 1, VES), 4, r <sup>2</sup> Devester BA 36 Nc FF 98 6A DF 78 51 97 32 42 82 40 C2 6F         Alves, jacc. Bark           Debesche DC 52 67 DB 78 51 71 95 9C 96 62 BA 1E1 DE A1 18 72 91         La (1, 1, VES), 4, r <sup>2</sup> Devester BA 36 Nc FF 98 117 91 15 57 18 49 68 95         La (1, 1, VES), 4, r <sup>2</sup> Debesche DC 52 67 DB 78 51 78 15 71 BA 96 88 95 10         CA 46 85 53 5F 291 10 6E 16 98 F2         Debesche DC 52 67 DB 88 70 AC 48         S2 64 48 65 13 6F 91 10 6E 16 98 F2         Debesche DC 52 67 DB 88 AC 46 6E 10 68 C2 78 18 84 AC 48         S2 7 91 10 6E 16 98 F2         Debesche DC 78 62 AA AF 68 EE 60 CB 5C 77 10 75 48 AA AC 88 77         L(1, 'VES), 4, r <sup>2</sup> Debesche DC 78 62 AA AF 68 EE 60 CB 5C 78 18 84 AC 44 57         L(1, 'VES), 4, r <sup>2</sup> Debesche DC 78 62 AA AF 68 EE 60 CB 5C 77 10 75 48 AA AC 88 77         L(1, 'VES), 4, r <sup>2</sup> Debesche DC 78 51 84 51 AA 86 87 77         L(1, 'VES), 4, r <sup>2</sup> Debesche DC 78 51 84 51 AA 86 87 77         L(1, 'VES), 4, r <sup>2</sup> L(1, 'VES), 4,	800065FB A4 A6 14 F4 39 43 83 F5 27 7B A2 14 CA 7D 62	14 ¤¦.ô9Cfō'{¢.Ê}bD	^	
aeeeecs11 ES BA 86 DE FF @ 6A DF 78 63 B4 1B DF F6 5C AS       a**by.jBxc.80;kr       beeeecs12 ES BA 86 NE FF @ 6A DF 78 63 B4 1B DF F6 5C AS       a**by.jBxc.80;kr         beeeecs12 ES CA 36 DF 15 F0 1A D7 19 59 C9 2E 15 S0 93 19 73 42 82 40 C2 BF       R6j 4U.3.36; M.A       beeeces12 ES BA 86 NE FF @ 6A DF 78 63 B4 1B DF F6 5C AS       a**by.jBxc.80;kr         beeeces2 AF	00006608 A6 28 D4 F8 86 A5 31 D0 A5 32 12 3E C8 16 BD	C0 : (Ô¢+¥1Đ¥2.>È.%À		
Debeoese:       3       1       3				
000000000000000000000000000000000000	000066DC 52 67 D3 6A 92 E1 55 09 33 19 73 42 82 4D C2		10000000	
000000723 BA 9D 51 CB D5 06 AD D5 76 11 55 77 10 A9 68 09 50       0, 0000 - 0,0000 - 0,0000 - 0,000 - 0,000 - 0,000 - 0,000 - 0,000 -	000066EC 31 BF 7B 11 B7 19 59 C9 62 BA 1E 1D EA 18 72	91 1¿{YÉbºê.r'		
000005/34 98 76 01 40 28 70 100 83 36 4 40 6 25 35 01 20 65 16 99 72 10       060007+40'.n0       060007	0000672A BA 9D 51 CB D5 0A 3D 55 F8 15 57 10 A9 68 09			
989865778 D4 EA D6 57 A9 51 BA AC B5 56 91 1D 6E 16 98 F2       0600002 + M <sup>2</sup> (n	0000673A 90 76 01 40 28 7D D0 85 3C 4A 06 E5 33 5F 29			
Beeberg 248         36         66         24         47         47         66         52         25         66         40         44         46         66         52         55         76         77         77         76	200006778 D4 EA D6 57 A9 51 AA AC 85 56 91 1D 6E 16 90	F2 ÔêÖW0Qª¬µV'.nò		
Debelog 1/2 gr L 46 38 0 62 11 39 89 32 / 3 80 44 21 30 67 74 34 64 74       Oeeeeerra L 75 84 48 11 46 88 70 80 77 54 48 A8 C 83 72       1.xxai Y. rekt 47         Debelog 7/2 gr L 7 58 48 11 46 88 70 80 77 54 48 A8 C 83 72       1.xxai Y. rekt 47       Oeeeesra 16 7 54 48 11 40 88 70 80 77 54 48 48 C 83 72       1.xxai Y. rekt 47         Debelog 7/2 gr L 7 58 48 11 40 88 70 80 77 78 78 89       00 77 54 74 78 88 70 80 77 85 48 11 40 88 77 88 99       00 77 54 74 78 88 70 80 77 54 74 88 70 80 77 54 74 88 70 90 75 74 78 88 70 78 78 89       00 77 54 74 78 88 70 80 77 75 74 78 88 70 90 77 54 74 78 88 70 90 77 54 74 78 88 70 90 77 54 74 78 88 70 90 77 54 74 78 88 70 90 77 54 74 78 88 70 78 78 89       00 77 54 74 78 78 89 70 80 77 54 74 78 88 70 90 77 54 74 78 88 70 90 77 54 74 78 88 70 90 77 54 74 78 88 70 90 77 54 74 78 88 70 90 77 54 74 78 78 89 70 70 10 78 74 78 74 78 78 78 90 90 77 80 78 74 78 78 78 78 89       00 77 78 74 74 74 78 78 89 70 72 78 74 78 88 70 77 75 74 78 78 89 70 72 78 74 78 78 78 78 78 78 78 78 78 78 78 78 78	00006788 58 68 DD 28 4A AF 68 EE 6D CB 5C 2F B1 08 AC			
Deedees21 (7) GE 17 SB 48 E1 A B 68 70 80 77 5 M8 48 C 37 22       Discrete (7) GE 17 SB 48 E1 A B 68 70 80 77 5 M8 48 C 37 24       Discrete (7) GE 17 SB 48 E1 A B C 18 80 70 80 77 88 89       Gebes 28 77 EF F 60 BE C 3 D C 83 BE 14 E L F 8 B A 77 B8 89       Gebes 28 77 C 7 F F 7 60 F 6 BE C 3 D C 83 BE 14 E L F 8 B A 77 B8 89       Gebes 28 77 D C 27 BE 7F 60 BE C 3 D C 83 BE 14 E L F 8 B A 77 B8 89       Gebes 28 77 C 7 F 7 B 7 87 B 89       Gebes 28 77 B 2 F 7 8 7 B 7 8 8 9 G 62 A A 8 9 39 9 E D C C 13 G 7 3 F 7 18 B 8 0 F A 4 5 D A 4 4 D C A 1 9 E C A 1 8 E C A	000067C9 8C 48 33 03 62 11 59 89 52 75 3D FA 54 8A A4			
Debelogs P C/ FE /F 60 B2 80 DE 43 80 B1 C 1 C 63 C C 7              yff: "Direct P 1 P 1 P 1 P 1 P 1 P 1 P 1 P 1 P 1 P	000067D9 CE 17 58 48 E1 A6 88 FD 8D F7 F5 48 AA 8C 83			
00000831 FD A/ AS AD 60 52 AS 02 43 83 65 60 AC 45 99 66 A2 AB 59 09 CL       00000831 FD A/ AS AD 60 52 AS 02 43 63 99 9E CL       021/21/21/21/21/21/21/21/21/21/21/21/21/2				
Deeleesby 163 CA ED 94 GA BS 46 62 BS 95 Az AK 18 25 93 Z LD Deeleesby 163 CA ED 94 GA BS 46 62 BS 95 Az AK 18 25 93 Z LD Deeleesby 163 CA ED 94 GA BS 46 62 BS 95 Az AK 18 25 93 Z LD Deeleesby 163 CA ED 94 GA BS AF 73 BS DEELE LL, L,				
000005877 CD 42 39 7A 0A 4D CA 10 40 CA 10 CA 1				
000005085 60 48 59 AA 05 08 30 68 FF 06 51 26 F7 61 74       0.6 126 EC a0 26 56       0.6 126 EC a0 26 56<				
000006805 EB 08 10 0 /9 10; /A 06 04 36 A7 AE 36 P2 40 F1 /2       000006805 C2 A3 AD D8 88 DB 90 36 07 A2 D2 F3 DF 36 C2 PF 1 *0 /0.6, 4*0/11         000006805 D2 06 0A AD B8 DB 90 36 07 A2 D2 F3 DF 36 C2 PF 1 *0 /0.6, 4*0/11       000006805 (6, 4*0/11)       000006805 (6, 4*0/11)         000006805 D2 07 A5 D6 E5 25 13 34 55 22 E6 17 A8 PF 36 C2 PF 1 *0 /0.6, 4*0/11       000006805 (6, 4*0/11)       000006805 (6, 4*0/11)       000006805 (6, 4*0/11)         000006805 D2 D7 A5 D6 E5 51 33 45 25 22 EF 17 A9 P9 67 C3 22 EF 17 A9 P9 67 C3 22 EF 17 A9 P9 67 C3 24 E7 19 P3 A7 A8 P2 D3 PF 36 A2 E8 P8 A2 PF 37 P0 67 C3 24 E7 19 P3 A7 A8 P2 D3 PF 36 A2 E8 P8 A2 PF 37 P0 67 C3 24 E7 19 P3 A7 A8 P2 D3 PF 36 A2 E8 P3 A2 PF 37 P0 10 C C A A8 P7 98 2D 3P X30.E15 A8 P* 1 - 20000691F BC A5 D6 E4 C8 A7 D0 16 C4 CA BA B7 98 2D 3P X30.E15 A8 P* 1 - 20000691F BC A5 D6 E4 C8 A7 D0 16 C4 CA BA B7 98 2D 3P X30.E15 A8 P* 1 - 20000691F BC A5 D6 E4 C8 A7 D0 16 C4 CA BA B7 98 2D 3P X30.E15 A8 P* 1 - 20000691F BC A5 D6 E4 C8 A7 D0 16 C4 CA BA B7 98 2D 3P X30.E15 A8 P* 1 - 20000691F BC A5 D6 E4 C8 A7 D0 16 C4 CA BA B7 98 2D 3P X30.E15 A8 P* 1 - 20000691F BC A5 D6 E4 C8 A7 D0 16 C4 C4 BA B7 98 2D 3P X30.E15 A8 P* 1 - 20000691F BC A5 D6 E4 C8 A7 D0 16 C4 C4 BA B7 98 2D 3P X30.E15 A8 P* 1 - 20000691F BC A5 D6 E4 C8 A7 D0 16 C4 A8 B7 98 2D 3P X30.E15 A8 P* 1 - 20000691F BC A5 D6 A4 C8 B0 C7 35 C8 B8 B4 A5 C8 D0 04 A1 C7 - 20000691F BC A5 D6 A4 A8 DF 77 D1 07 98 A8 A4 1F 11 B9 16 A3 00 XEEBMN 1 1, 100087         000069912 AF C7 B1 6F B1 38 C8 A4 C8 P3 D7 C7 35 C8 B8 B4 B1 C7 24 P* 1 1, 100087       000069912 AF C8 D1 69 A5 A4 A8 B9 D0 A1 S C4 - 200 - 200 57 C8 B1 68 D5 C1 C7 B1 F1 2, 7 1, 100087         000069912 AF B1 93 06 D4 A5 A6 A8 D0 CC A7 37 96 FF 11 B2 D4 + 11       000069912 AF B1 5				
Deepeesers 69       69       AA D8       88 D8       89 05       67       72 044-x641/1851         Deepeesers 69       F9       73 D9       44 F8       75 C5       55 C8       55 C8       75 C8       74 044-x641/1851         Deepeesers 69       F3       A5 52       2E F1       74 07 46 <td></td> <td></td> <td></td> <td></td>				
Deceessers 9f 99 71 D9 34 F8 75 87 E5 92 E6 73 04 5E 86       F 78 50 E5 51 83 45 52 2E F1 74 9F 96 7C 32 E6 73 040 FR AtV-12       Pedeessers 97 50 E5 51 83 45 52 2E F1 74 9F 96 7C 32 E6 73 040 FR AtV-12       Pedeessers 97 50 E5 51 83 45 52 2E F1 74 9F 96 7C 32 E6 73 040 FR AtV-12       Pedeessers 97 50 E5 51 83 45 52 2E F1 74 9F 96 7C 32 E6 73 040 FR AtV-12       Pedeessers 97 50 E5 51 83 45 52 2E F1 74 9F 96 7C 32 E6 73 040 FR AtV-12       Pedeessers 12 F1 74 9F 96 7C 32 E6 73 FR FT				
Beeeessie B7 A5 D6 E5 51 83 45 52 2E F1 A3 9F 98 /0 32 E       Value 15 / Å E 4 - A       Deege 51 E 3 A D E 5 1 83 45 52 2E F1 A 3 9F 98 /0 32 E       Value 15 / Å E 4 - A       Deege 51 E 3 A D E 5 1 83 45 52 2E F1 A 3 9F 98 /0 32 E       Value 15 / Å E 4 - A       Deege 51 E 3 A D E 5 1 83 45 52 Z       Deege 51 E 3 A D E 5 1 83 45 52 Z       Deege 51 E 3 A D E 5 1 83 45 52 Z       Deege 51 E 5 1 83 45 52 Z       Deege 51 E 3 A D E 5 1 83 54 52 Z       Deege 51 E 3 A D E 5 1 83 54 52 Z       Deege 51 E 3 A D E 5 1 83 54 52 Z       Deege 51 E 3 A D E 5 1 83 54 52 Z       Deege 51 E 3 A D E 5 1 83 54 52 Z       Deege 51 E 3 A D E 5 1 83 54 52 Z       Deege 51 E 3 A D E 5 1 83 54 52 Z       Deege 51 E 3 A D E 5 1 83 54 52 Z       Deege 51 E 3 A D E 5 1 83 54 52 Z       Deege 51 E 3 A D E 5 1 83 54 52 Z       Deege 51 E 3 A D E 5 1 83 54 52 Z       Deege 51 E 3 A D E 5 1 83 54 52 Z       Deege 51 E 3 A D E 5 1 83 54 52 Z       Deege 51 E 3 A D E 5 1 2 55 2 Z       Deege 51 E 3 A D E 5 1 2 55 2 Z       Deege 51 E 3 A D E 5 1 2 55 2 Z       Deege 51 E 3 A D E 5 1 2 55 2 Z       Deege 51 E 2 5 2 Z       Deege 51 E 7 Z       Deege 51 E 2 5 2 Z       Deege 51 E 7 Z       Deege 51 E				
000005051 EC 9A DA 85 CB 4C 8A 70       12 C 4 CB 0A 87 92 D 37       12 0 4 CB 0A 97 73       12 0 4 CB 0A 73 73       12 0 4 CB 0A 73 73       12 0 4 CB 0A 73 73 75				
000005932 FC 01 16 F 12 93 04 F3 36       05 7 0A 00 5 08 80 04 22       020 000 10 81 5 1       020 000 10 81 5 1       0000050 4 81 5 1       00				
000005945       25       A3       50       77       70       78       88       11       11       10       14				
000000955 14 F8 14 3F 88 8F 16 38 4E 2 0 C 03 E1 20 3F 13         0.4 F1/C51K41k5         0000059E 72 FC 0F 12 9C 0A CA 89 C7 35 48 80 E0 A1 57 24         r0. xE/C51K41k5           00000595 24 87 50 61 86 5F 82 D0         16 95 00 54 85 90 D0 A1         5 yat 2 D + T. 1         0000059E 24 87 50 61 86 5F 82 D0         16 95 00 54 85 90 D0 A1         5 yat 2 D + T. 1           00000595 48 85 20 D1 69 5 00 55 00 56 48 59 0D 0A1         5 yat 2 D + T. 1         0000059E 78 D0 69 AE AA 68 90 C CA 73 79 6F F1 18 2D         0.4 e <sup>2</sup> + 0.4 + 1. 1         0000059E 78 D0 69 AE AA 68 90 C CA 73 79 6F F1 18 2D         0.4 e <sup>2</sup> + 0.4 + 1. 1         0000059E 78 D0 69 AE AA 68 90 C CA 73 79 6F F1 18 2D         0.4 e <sup>2</sup> + 0.4 + 1. 1         0000059E 78 D0 69 AE AA 68 90 C CA 73 79 6F F1 18 2D         0.4 e <sup>2</sup> + 0.4 + 1. 1         0.4 e <sup>2</sup>				
00000699D 72 FC 0F 12 9C 0A CA B9 (7 35 40 80 E0 A1 5/24 F0.4C.T.D)       00000591 24 B3 51 61 86 5F B2 D0 16 95 00 54 85 9D D0 A1       \$-0a1_2D,-T.L.D)         00000591 24 B7 50 61 86 5F B2 D0 16 95 00 54 85 9D D0 A1       \$-0a1_2D,-T.L.D)       00000591 24 B3 51 61 86 5F B2 D0 16 95 00 54 85 9D D0 A1       \$-0a1_2D,-T.L.D)         00000591 FB E9 36 09 AE AA 68 9D       CC CA 73 79 6F F1 1B 2D       Udef, **h.IEsyoñ.       00000591 E4 B3 51 61 86 5F B2 D0 16 95 00 54 85 9D D0 A1       \$-0a1_2D,-T.L.D)         00000591 FB E9 36 09 AE AA 68 9D       CC CA 73 79 6F F1 1B 2D       Udef, **h.IEsyoñ.       00000591 E4 B3 26 09 C3 34 92 66 49 65 B4 85 C1 D2 FE BF       K%, 3000-µ µADb         00000591 FB E9 69 AE AA 68 9D       S 51 73 56 F8 38 2C 03 D4       -au3_2(FA, Pas, D)       00000591 E4 B3 26 31 19 24 58 C5 5E 17 56 F8 38 2C 03 D4       -au3_2(S[A^A, Pas, D)				
00006090 24 B7 50 61 86 5F 82 00         16 95 00 54 85 90 00 A1         3 40 0 0 A1         9 40 0 0 0 A1         9 40 0 0 0 A1         9 40 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0				
00006590D H8 E9 30 99 AE AA 68 90 CC CA 73 79 GF 71 B 20 00 A CA 74 79 GF 71 A CA 74 74 74 74 74 74 74 74 74 74 74 74 74				
00000508 48 52 50 50 34 62 64 56 55 64 56 56 64 56 52 64 56 56 10 10 10 10 10 10 10 10 10 10 10 10 10				
00006A06 7F 0A D8 F1 07 77 36 F8 81 4F CC F1 F8 7D 58 CF Ch. C60 OTAGX				00006A06 7F 0A D8 P 05 C7 36 F8 81 4F CC F1 F8 7D 58 CF 0ñ. C60. 01ñ0}XI

Figure 2: HEX code of original image and HEX code of image with hidden message

The Deep Learning algorithm Convolutional Neural Networks (CNN) was deployed to train and test the image datasets. The algorithm has been used for steganalysis purposes because of efficiency in image classification (Liu., Yang, and Kang, 2017).

The CNN algorithm Python code is designed to be able to perform image classification on the datasets being trained and tested. The convolution code will extract the features and learn the image feature matrix also known as the feature map. The Pooling code will reduce the dimensions of each feature map but keep the most important data. The Flattening code will convert the matrix into a linear array and Fully Connecting code will connect the convolutional network to the neural network and then compile the network. The Epoch code is set to 100 which means the dataset will pass through the neural network 100 times which will result in an increase in accuracy and a decrease in loss .

The following metrics were used:

- The **Training Accuracy** represents the accuracy that is applied to the CNN model on the training data function as a percentage and how well the CNN model was trained.
- The **Training Loss** represents the error on the training set of that data. A decrease in the Training Loss means the CNN model is improving.
- The **Validation Accuracy** is used after the CNN has been trained for adjusting the networks hyperparameters and comparing how the changes to them affect the predictive accuracy of the CNN model.
- The **Validation Loss** is used after the CNN has been trained and represents the errors discovered after the test was used to test the predictive accuracy of the trained CNN model.

#### 4. Initial Experiments

The CNN utilised the algorithm to train 5000 clean images and test 5000 stego-images testing different configurations to decrease the loss and increase validation

Experiments 1-3 were tested using the Final Layer Activation Function <u>Sigmoid</u> which is used for binary classification in logistical regression modelling.

Experiments 4-6 were tested using the Final Layer Activation Function <u>Softmax</u> which is used for multiple classifications in logistical regression modelling.

Results of the Current Experiments to date												
Experiment	Final Layer	1 <sup>st</sup> Layer	Epoch per	Validation	Training	Validation	Training	Validation				
Number	Activation	Filter	Steps	Steps	Loss	Loss	Accuracy	Accuracy				
	Function											
Experiment 1	Sigmoid	32, size of	80	8	0.011771	5.19850	0.99558	0.57692				
		3x3			Decrease	Increase	Increase	Increase				
Experiment 2	Sigmoid	32, size of	800	80	0.00416	6.78965	0.99908	0.56939				
		3x3			Decrease	Increase	Increase	Increase				
Experiment 3	Sigmoid	64, size of	80	8	0.00809	5.78904	0.99759	0.57692				
		7x7			Decrease	Increase	Increase	Increase				
Experiment 4	Softmax	64, size of	80	8	8.22109	8.78875	0.48432	0.44872				
		7x7			Consistent	Consistent	Consistent	Consistent				
Experiment 5	Softmax	32, size of	80	8	8.19546	9.06127	0.48593	0.43162				
		3x3			Consistent	Consistent	Consistent	Consistent				
Experiment 6	Softmax	32, size of	800	80	8.09873	8.55683	0.49200	0.46327				
		3x3			Consistent	Consistent	Consistent	Consistent				

#### Table 1: Results of Experiments

#### 4.1 Comparison of Experiment 1 & 3: Sigmoid

Results showed a decrease in the Training Loss of 0.003681 but an increase of 0.59054 in the Validation Loss. The Training Accuracy increased by 0.00201 and the Validation Accuracy was identical 0.57692. This comparison of results indicates when configuring the 1st Layer Filter of "32, size of 3x3" to "64, size of 7x7" it can improve the CNN model by decreasing the Training Loss and increasing the Training Accuracy. The experiment had a Validation prediction loss increase of 0.59054% but showed improvement when compared to the 1st Comparison as the Validation Accuracy remained balanced.

#### 4.2 Comparison of Experiment 5 & 6: Softmax

Examining and comparing experiments 5 and 6, results showed a decrease in the Training Loss of 0.9673 and a decrease in the Validation Loss of 0.50444. The Training Accuracy increased by 0.00607 and the Validation Accuracy increased by 0.03165. This comparison of results indicates when configuring settings for the Epoch per steps from 80 to 800 and the Validation steps from 8 to 80 it can improve the CNN model results as both the training and validation Losses decreased. Also both the training and validation Accuracy results were increased.

#### 4.3 Comparison of Experiment 2 & 6

Comparing experiment 2 with Sigmoid configurations and experiment 6 with Softmax configurations. Using the configurations for the Epoch per steps 800 and the Validation steps 80 with the 1st Layer Filter of "32, size of 3x3" being set. The results showed the Training Loss increased by 8.09873 from Sigmoid to Softmax and the Validation Loss increased by 1.76718. The results also showed a decrease in Training Accuracy by 0.50708 and a decrease in the Validation Accuracy by 0.10612. The results in this comparison are significantly larger than the other experiments as the Sigmoid is used for binary classification and the Softmax is used for multiple classifications.

#### 5. Conclusions

As the configuring the Epoch per steps to 800 plus Validation steps to 80, and configuring 1st Layer Filter of "32, size of 3x3" to "64, size of 7x7" have showed that they can improve the performance of the CNN model for both Accuracy and Loss.

These low results for Softmax for the Validation Accuracy, which show a decrease of 0.1% and a Validation Loss increase of 2 on average when all 6 experiments were compared may indicate detection of steganalysis when compared to the Sigmoid results.

When implementing the deep learning CNN algorithm, results showed that increasing configurations such as the 1st Layer Filter, size, the epoch steps and the validation steps in the CNN code. The CNN model can improve the training and validation accuracy, also decreasing the loss in the training and validation of the CNN model. This leads to improved accuracy in the detection of OpenPuff steganography in a large dataset

This work in progress will now examine increased secret message size within the carrier image and various settings in the CNN model to improve accuracy.

#### References

Giarimpampa, D., 2018. Blind Image Steganalytic Optimization by using Machine Learning.

- Types of Machine Learning. (2018). https://towardsdatascience.com/machine-learning-for-beginners-d247a9420dab [Accessed 14 Aug. 2019]
- Hegarty, M. (2018), "Steganography, The World of Secret Communications". CreateSpace Publishing Platform. ISBN-10: 1986125424
- Hunt, R., 2012, December. New developments in network forensics—Tools and techniques. In 2012 18th IEEE International Conference on Networks (ICON) (pp. 376-381). IEEE.
- Kessler, Gary C. "An Overview of Steganography for the Computer Forensics Examiner (Updated Version, February 2015)." (2015).
- Ker, A.D., Bas, P., Böhme, R., Cogranne, R., Craver, S., Filler, T., Fridrich, J. and Pevný, T., 2013, June. Moving steganography and steganalysis from the laboratory into the real world. In Proceedings of the first ACM workshop on Information hiding and multimedia security (pp. 45-58). ACM.
- Liu, K., Yang, J. and Kang, X., 2017, May. Ensemble of CNN and rich model for steganalysis. In 2017 International Conference on Systems, Signals and Image Processing (IWSSIP) (pp. 1-5). IEEE.
- Lynch, S. Shepardson, D. "U.S. accuses pair of stealing secrets, spying on GE to aid China" Reuters News Agency, April 23, 2019. https://www.reuters.com/article/us-usa-justice-ge/us-accuses-pair-of-stealing-secrets-spying-on-ge-to-aid-china-idUSKCN1RZ240 [Accessed: 15th August 2019]
- Morrison, B. (2004) "Ex-USA TODAY reporter faked major stories" USA Today. 3/19/2004 4:13 AM
- http://usatoday30.usatoday.com/news/2004-03-18-2004-03-18\_kelleymain\_x.htm [Accessed:10th September 2019] Provos N and Honeyman P. 2001. Detecting Steganographic Content on the Internet: Center for Information Technology Integration University of Michigan (August 2001).
- Richer, P., (Updated Version, February 2015). Steganalysis: Detecting hidden information with computer forensic analysis. SANS/GIAC Practical Assignment for GSEC Certification, SANS Institute, 6.
- Van Heerden, H. "A First Course in Ethical Hacking" Second Edition, 07 November 2014

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.