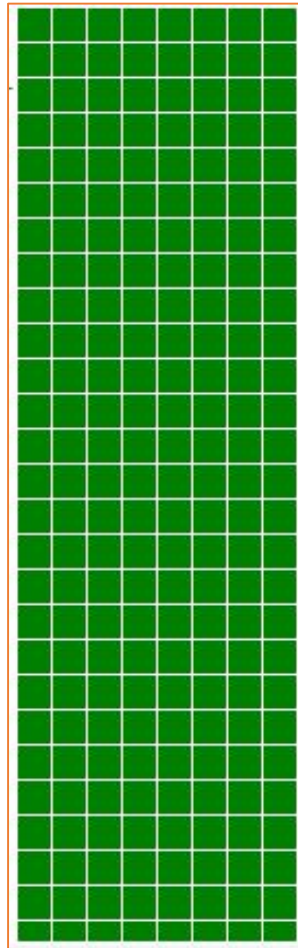# New file carver

Igor Loskutov - Rusolut

Ideal case

NAND memory device

ECC map

No lost blocks

Markers table 0

		Block markers					Page markers	
Use	Bank	LBN	Header	Address	LB	RB	LPN	Address
☑	00	1000	FF	0108114000	☐	☐		
☑	00	1001	FF	00A2F30000	☐	☐		
☑	00	1002	FF	00A206C000	☐	☐		
☑	00	1003	FF	010A4F0000	☐	☐		
☑	00	1004	FF	00A3368000	☐	☐		
☑	00	1005	FF	010C278000	☐	☐		
☑	00	1006	FF	0107688000	☐	☐		
☑	00	1007	FF	010A928000				

Reached end of the list
No errors found

OK

☑	00	1010	FF	0106E18000	☐	☐		
☑	00	1011	FF	0109A64000	☐	☐		
☑	00	1012	FF	010DBC8000	☐	☐		
☑	00	1013	FF	010B5D0000	☐	☐		
☑	00	1014	FF	0108FD8000	☐	☐		
☑	00	1015	FF	010A0B8000	☐	☐		
☑	00	1016	FF	0107CDC000	☐	☐		
☑	00	1017	FF	010C05C000	☐	☐		

Good file system metadata

Log image 0

▶ MBR ▶ Volume0 (Microsoft FAT32) 3.73 GB ▶ Root ▶ PHOTOS

Dump
- 🖥 MBR
  - ◢ 🔵 Volume0 (Microsoft FAT32) 3
    - ◢ ☐ 📁 Root
      - ▷ ☐ 📁 backup_photos
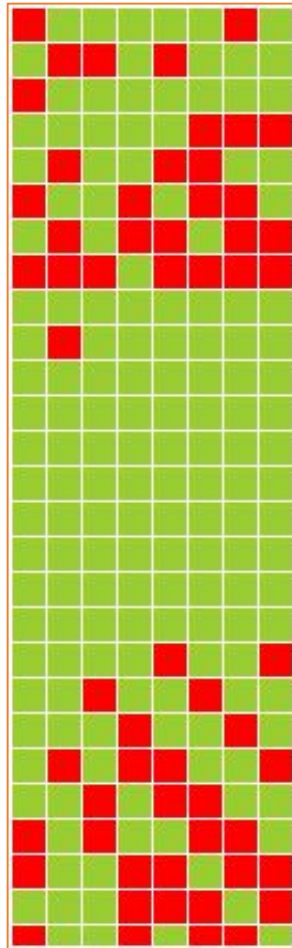   - ☐ 📁 DISTRIB
      - ☐ 📁 PHOTOS
    - 📄 FAT0
    - 📄 FAT1
    - 📄 FSInfo

	Name	Ext	Size	Last n
☐ 📁	..			
☐ 📄	P1050430	JPG	1.95 MB	02/27/
☐ 📄	P1050431	JPG	1.94 MB	02/27/
☐ 📄	P1050432	JPG	1.93 MB	02/27/
☐ 📄	P1050433	JPG	1.99 MB	02/27/
☐ 📄	P1050434	JPG	1.77 MB	02/27/
☐ 📄	P1050436	JPG	1.80 MB	02/27/
☐ 📄	P1050437	JPG	1.55 MB	02/27/
☐ 📄	P1050438	JPG	1.53 MB	02/27/
☐ 📄	P1050439	JPG	1.36 MB	02/27/
☐ 📄	P1050448	JPG	2.41 MB	03/07/
☐ 📄	P1050449	JPG	1.77 MB	03/07/
☐ 📄	P1050450	JPG	1.99 MB	03/07/
☐ 📄	P1050451	JPG	2.36 MB	03/07/
☐ 📄	P1050452	JPG	2.22 MB	03/07/
☐ 📄	P1050453	JPG	2.35 MB	03/07/
☐ 📄	P1050454	JPG	2.35 MB	03/07/
☐ 📄	P1050455	JPG	2.34 MB	03/07/
☐ 📄	P1050456	JPG	2.36 MB	03/07/
☐ 📄	P1050457	JPG	2.26 MB	03/07/
☐ 📄	P1050458	JPG	2.22 MB	03/07/
☐ 📄	P1050459	JPG	2.10 MB	03/07/
☐ 📄	P1050461	JPG	1.87 MB	03/07/
☐ 📄	P1050463	JPG	1.87 MB	03/07/
☐ 📄	P1050464	JPG	1.90 MB	03/07/
☐ 📄	P1050465	JPG	1.95 MB	03/07/
☐ 📄	P1050466	JPG	2.07 MB	03/07/

ruSolut
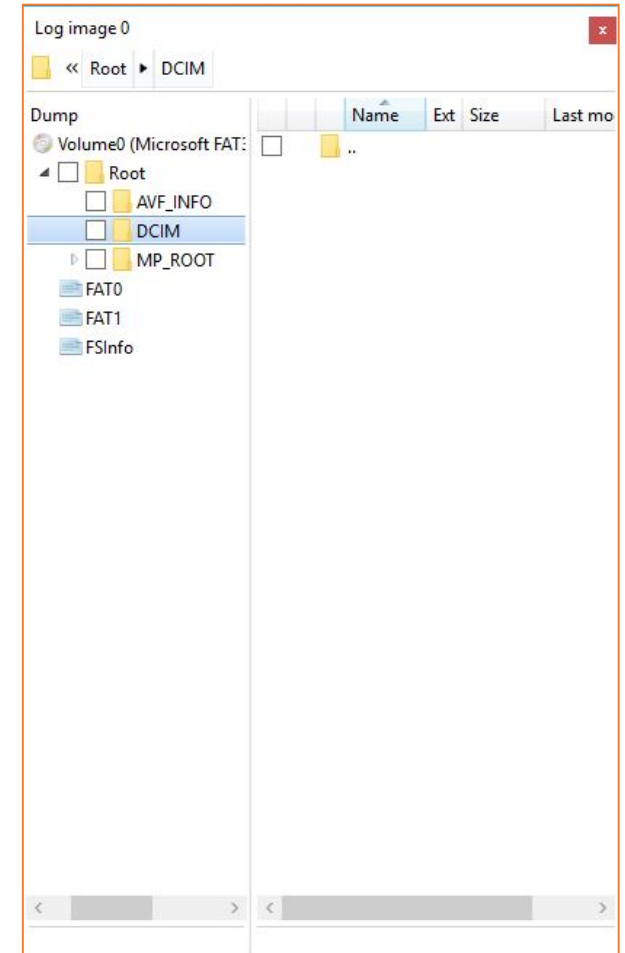
Real case

NAND memory device

ECC map

Lost blocks

Corrupted file system metadata

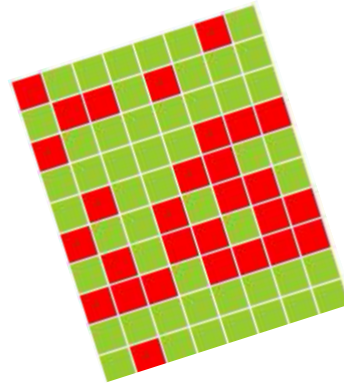| Markers table 0								
	Block markers					Page markers		
Use	Bank	LBN	Header	Address	LB	RB	LPN	Ad
☑	00	1003	FF	010A4F0000	☐	☐	00000000	00
☑	00	1004	FF	00A3368000	☐	☐	00000001	00
☑	00	1005	FF	010C278000	☐	☐	00000002	00
☑	00	1006	FF	0107688000	☐	☐	00000003	00
☑	00	1007	FF	010A928000	☐	☐	00000004	00
☑	00	1008	FF	010A2D4000	☐	☐	00000005	00
☑	00	1009	FF	010AF7C000	☐	☐	00000006	00
☑	00	100A	FF	010BC24000	☐	☐	00000007	00
☑	00	100B	FF	010854C000	☐	☐	00000008	00
☑	00	100C	FF	010962C000	☐	☐	00000009	00
☑	00	100D	FF	010C6B0000	☐	☐	0000000A	00
☑	00	100E	FF	010BA08000	☐	☐	0000000B	00
☑	00	100F	FF	01054C8000	☐	☐	0000000C	00
☑	00	1010	FF	0106E18000	☐	☐	0000000D	00
☑	00	1011	FF	0109A64000	☐	☐	0000000E	00A
☑	00	1012	FF	010DBC8000	☐	☐	0000000F	00
☑	00	1013	FF	010B5D0000	☐	☐	00000010	00
☑	00	1014	FF	0108FD8000	☐	☐	00000011	00
☑	00	1015	FF	010A0B8000	☐	☐	00000012	00
☑	00	1016	FF	0107CDC000	☐	☐	00000013	00
☑	00	1017	FF	010C05C000	☐	☐	00000014	00
☑	00	1018	FF	01056E4000	☐	☐	00000015	00
☑	00	1019	FF	010D358000	☐	☐	00000016	00
☑	00	101A	FF	010D13C000	☐	☐		

Position 0x2 from 0x7FF   Position 0x0 from

Log image 0

« Root ▶ DCIM

Dump | | Name | Ext | Size | Last mo
Volume0 (Microsoft FAT: ☐
▲ ☐ Root
☐ AVF_INFO
☐ DCIM
▷ ☐ MP_ROOT
FAT0
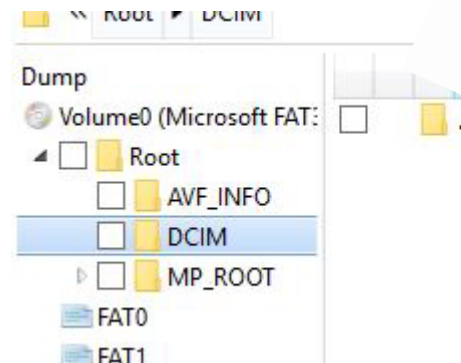FAT1
FSInfo

..

ruSolut

Typical problems when file carving should be used

- Bit errors
- Lost logical blocks
- Corrupted file system metadata
- Deleted files
- Data recovery from eMMC NAND

Markers table 0

Use	Bank	LBN	Block markers			
			Header	Address	LB	RB
☑	00	1003	FF			
☑	00	1004	FF	010A4F0000	☐	☐
☑	00	1005	FF	00A3368000	☐	☐
				010C278000	☐	☐

Root ▶ DCIM

Dump
- Volume0 (Microsoft FAT: ☐ ..
  - ☐ Root
    - ☐ AVF_INFO
    - ☐ DCIM
    - ☐ MP_ROOT
  - FAT0
  - FAT1

ruSolut

Classic carving approach

The file header is the standard combination of several bytes at a standard offset at the start of the file.

The file footer is the standard combination of bytes at a standard offset at the end of the file.
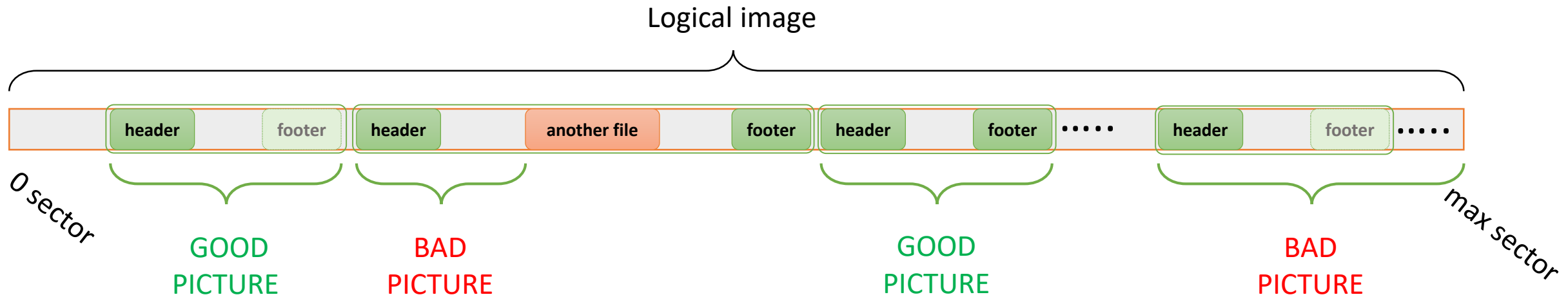
0xFF D8 – JPEG header
0xFF D9 – JPEG footer

Classic carving approach

Logical image

Searching for headers and footers is the classic carving approach. It assumes that the data between the header and footer or the header and next header is one file.

Classic carving approach

Logical image

| header | footer | header | another file | footer | header | footer | ..... | header | footer | ..... |

0 sector

GOOD
PICTURE

BAD
PICTURE

GOOD
PICTURE

BAD
PICTURE

max sector

This approach may find many false positive results in the case of file fragmentation, lost blocks, or data corruption since file integrity is not checked.

ruSolut

Results of classic carving approach
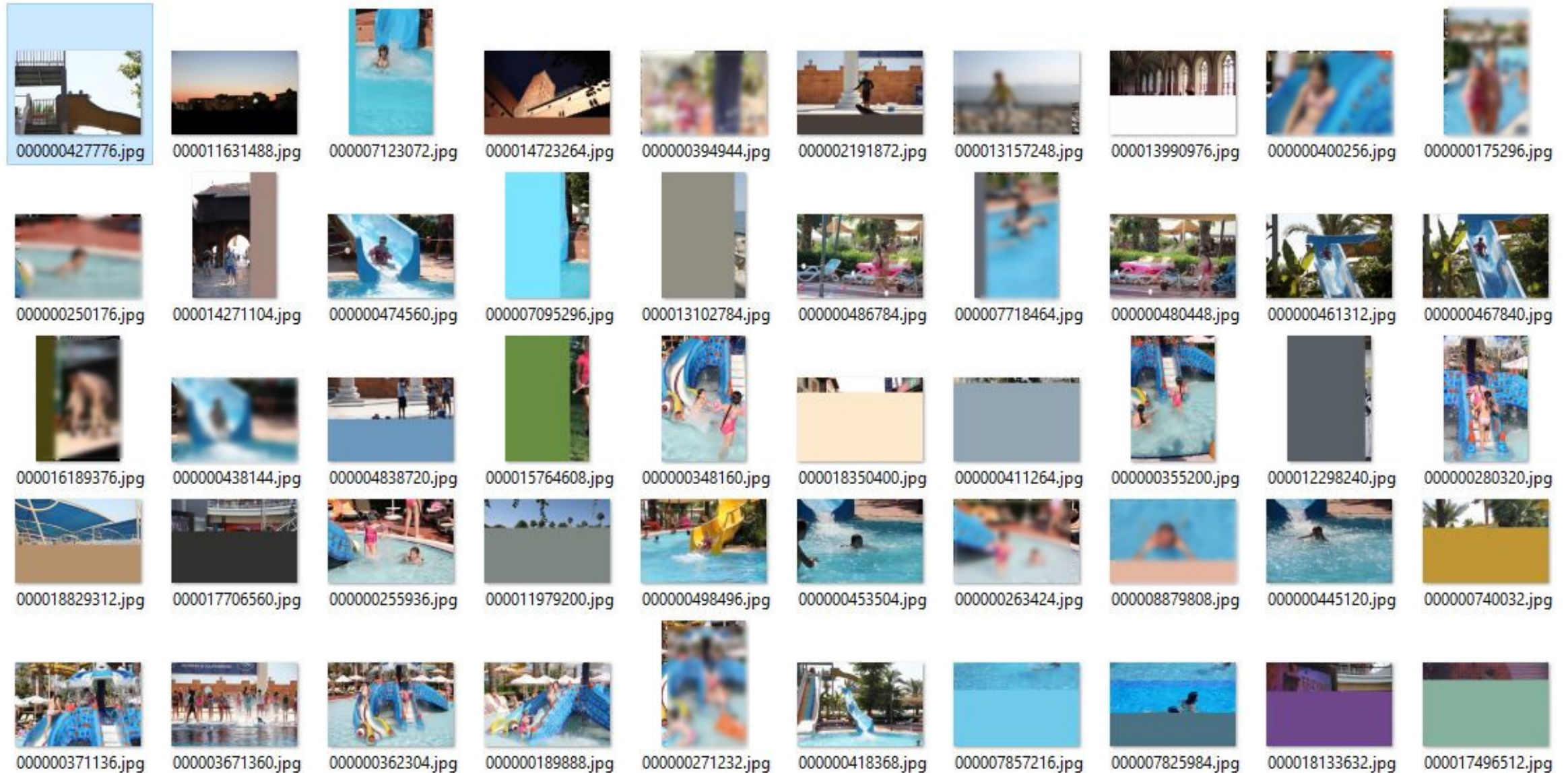
The new VNR carver checks file integrity and can separate good files from bad ones.

JPEG carving

JPEG file

First segment/ header

Last segment/ footer
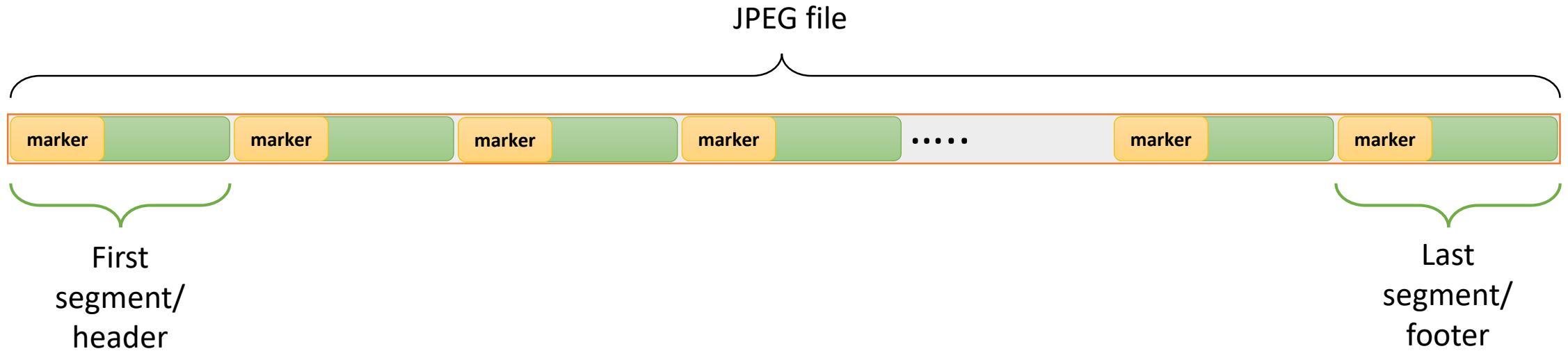
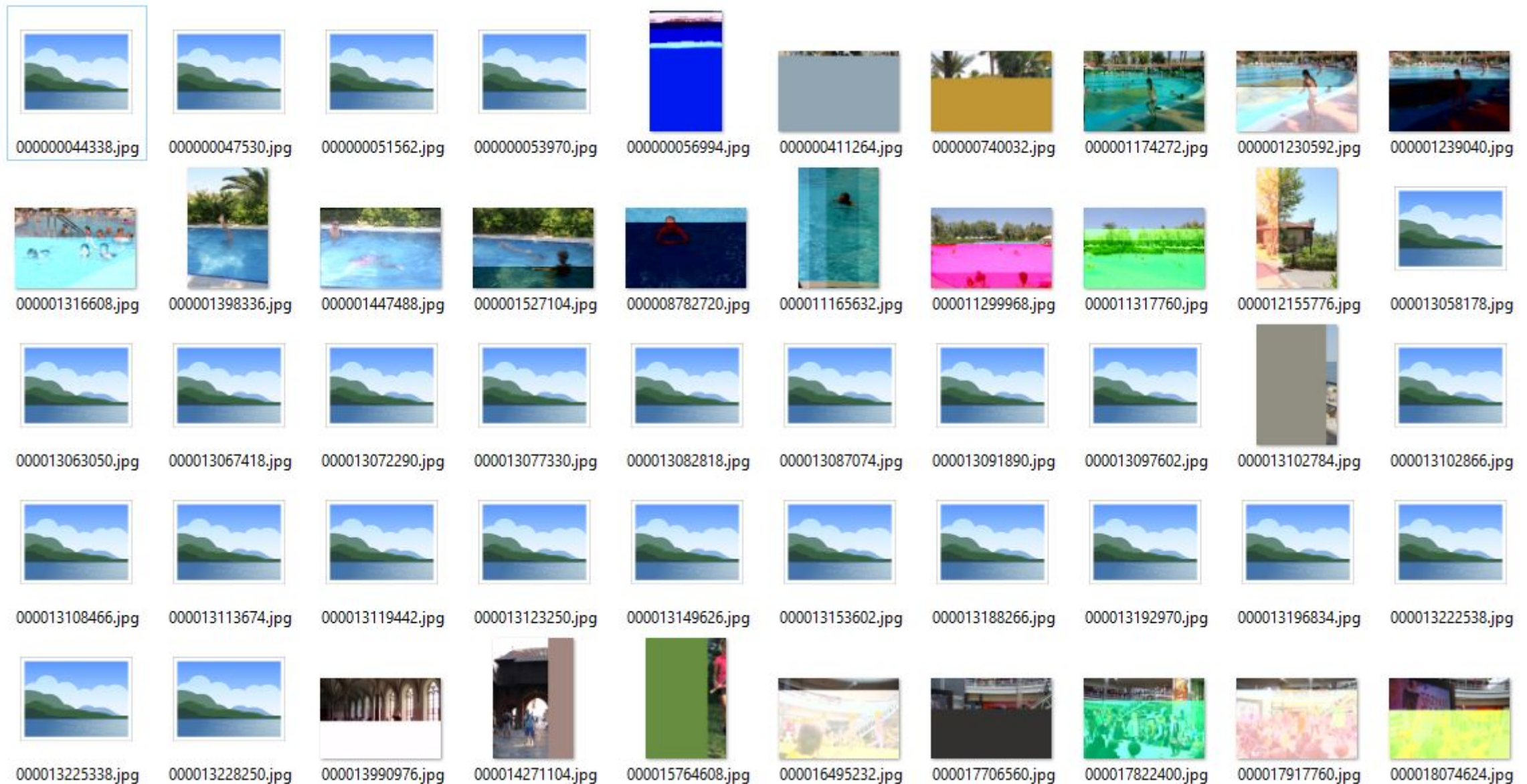A JPEG image consists of a sequence of *segments*, each beginning with a *marker*.

VNR carver checks the order and structure of the segments, and detects if the file is good or damaged.

Error types

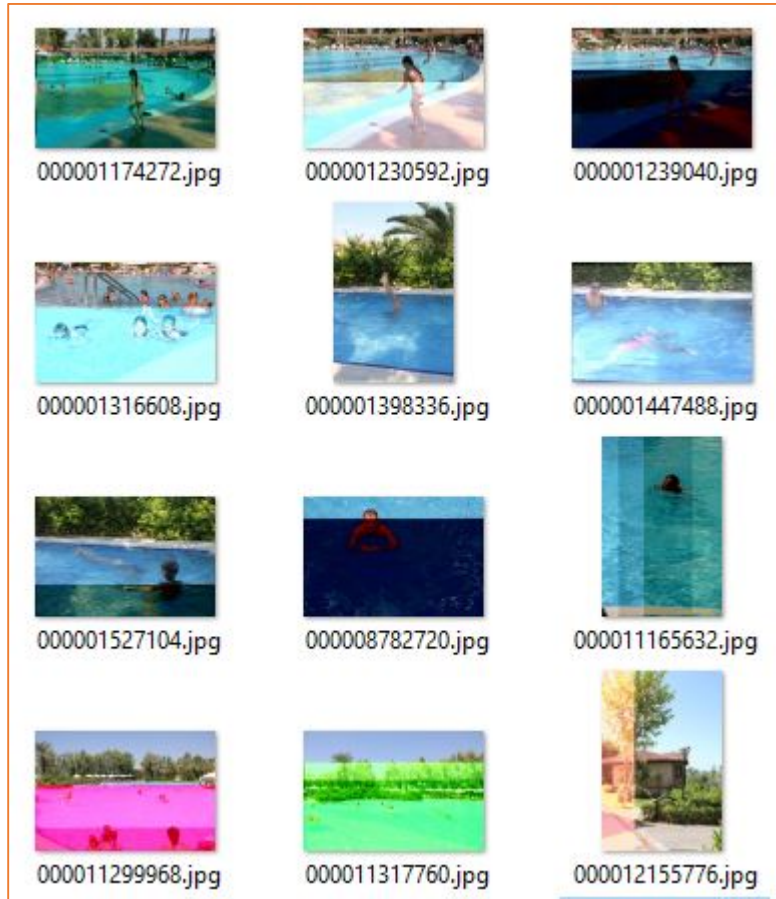Moreover, VNR carver rates the level of file damage and group bad pictures into three categories:

- Critical damage
- Medium damage
- Bit error

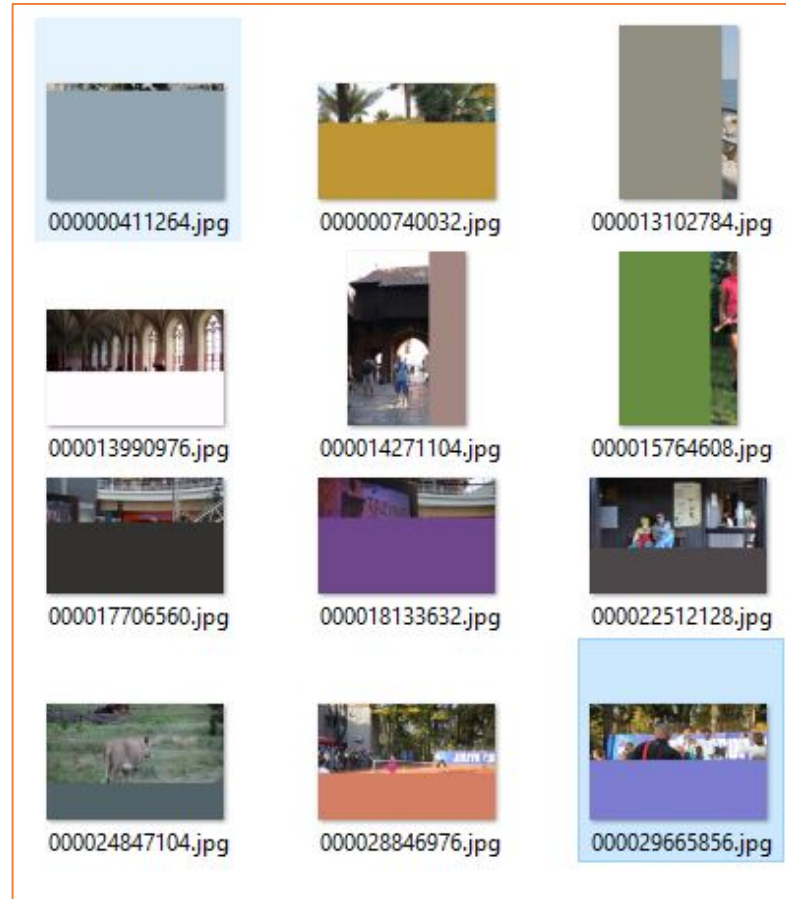# Bad pictures without grouping by level of damage

000000044338.jpg 000000047530.jpg 000000051562.jpg 000000053970.jpg 000000056994.jpg 000000411264.jpg 000000740032.jpg 000001174272.jpg 000001230592.jpg 000001239040.jpg

000001316608.jpg 000001398336.jpg 000001447488.jpg 000001527104.jpg 000008782720.jpg 000011165632.jpg 000011299968.jpg 000011317760.jpg 000012155776.jpg 000013058178.jpg

000013063050.jpg 000013067418.jpg 000013072290.jpg 000013077330.jpg 000013082818.jpg 000013087074.jpg 000013091890.jpg 000013097602.jpg 000013102784.jpg 000013102866.jpg

000013108466.jpg 000013113674.jpg 000013119442.jpg 000013123250.jpg 000013149626.jpg 000013153602.jpg 000013188266.jpg 000013192970.jpg 000013196834.jpg 000013222538.jpg

000013225338.jpg 000013228250.jpg 000013990976.jpg 000014271104.jpg 000015764608.jpg 000016495232.jpg 000017706560.jpg 000017822400.jpg 000017917760.jpg 000018074624.jpg

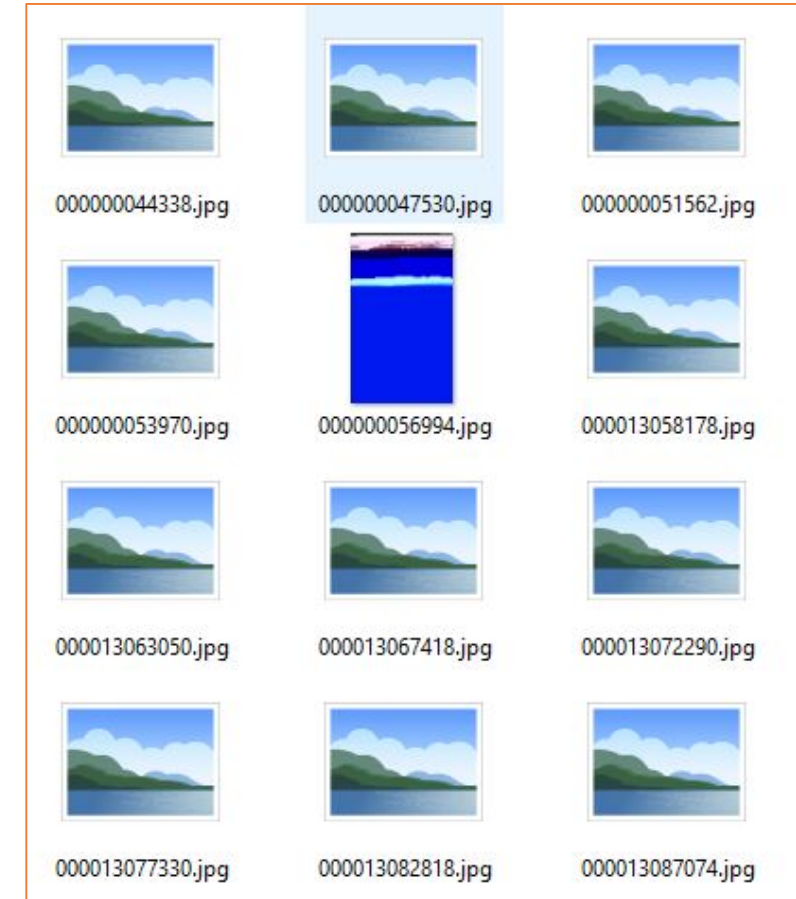# Bad pictures recovered by VNR carver
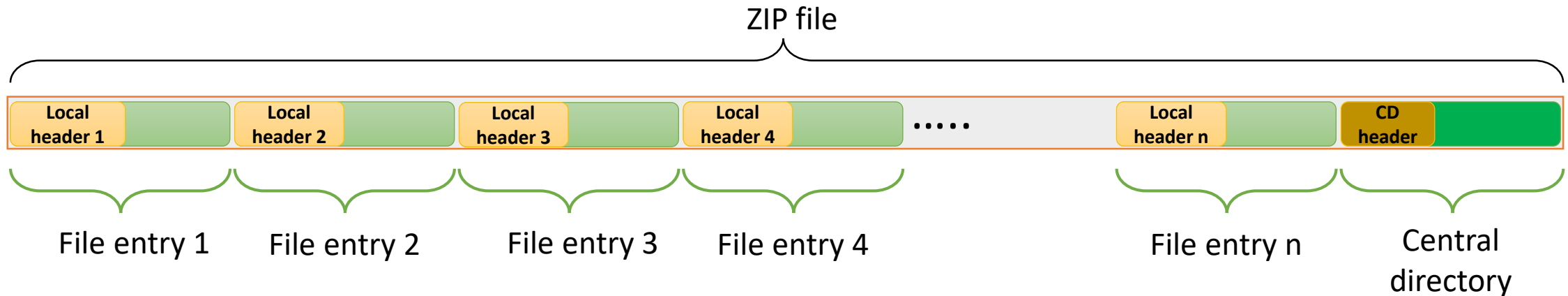
Bit error

Medium damage

Critical damage

ZIP carving and file extraction

ZIP files are archives that store multiple files. A *directory* is placed at the end of a ZIP file. This identifies what files are in the ZIP and identifies where in the ZIP that file is located.

VNR carver checks the integrity of the whole ZIP in general and the internal integrity of each file in particular. This allows it to detect AND recover partly damaged archives which contains damaged files or archives with a damaged *central directory*.

As a result VNR shows GOOD, BAD, and PARTLY DAMAGED ZIP files.

Types of recovered ZIP files

GOOD file – all files inside the ZIP are good and have the correct CRC, The central directory is found and has a good structure, the Footer of the Central directory is found.

PARTLY DAMAGED file – at least one file is good and has the correct CRC, The Central directory is found and has a good structure, the Footer of the Central directory is found.
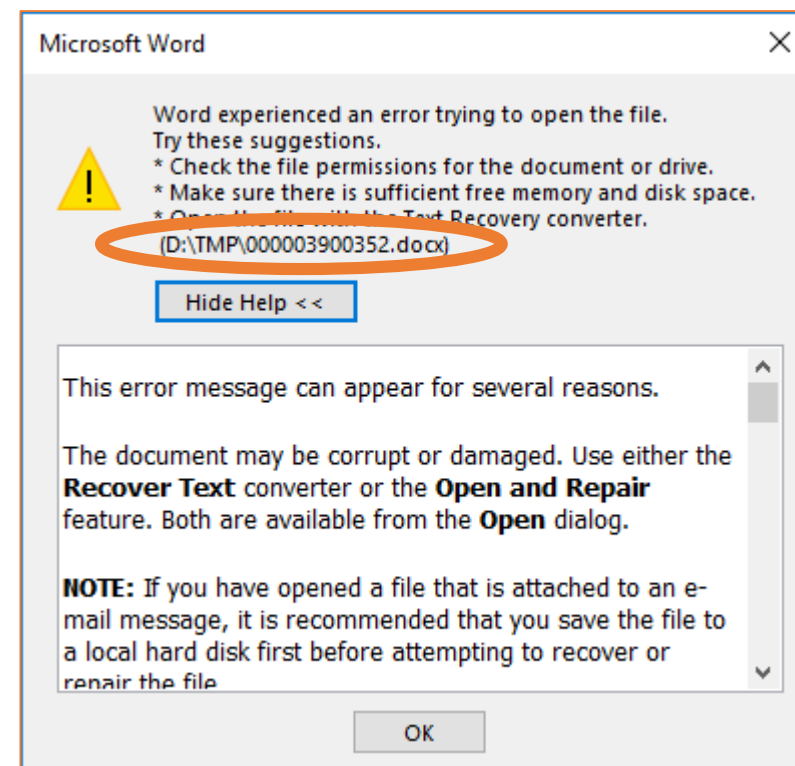
BAD file – the Local header is damaged/ ALL files inside the ZIP are damaged/ the Central directory is not found or is damaged.

ruSolut

Office files carving

VNR carver recovers WORD documents, SPREADSHEETS, and PRESENTATIONS made in modern Office applications. Supported formats are:

- *.docx*
- *.xlsx*
- *.pptx*
- *.odt*
- *.ods*
- *.odp*
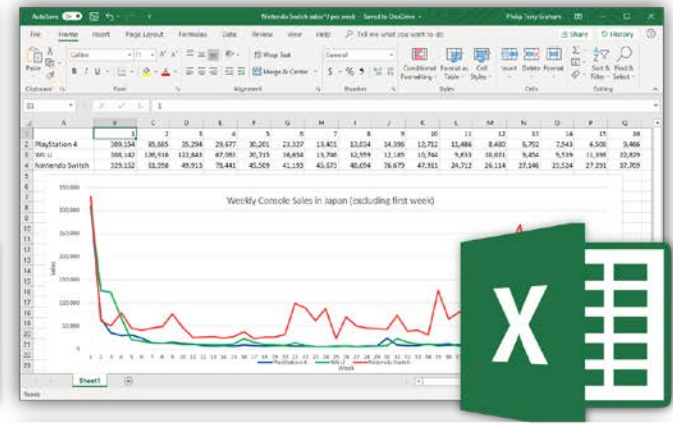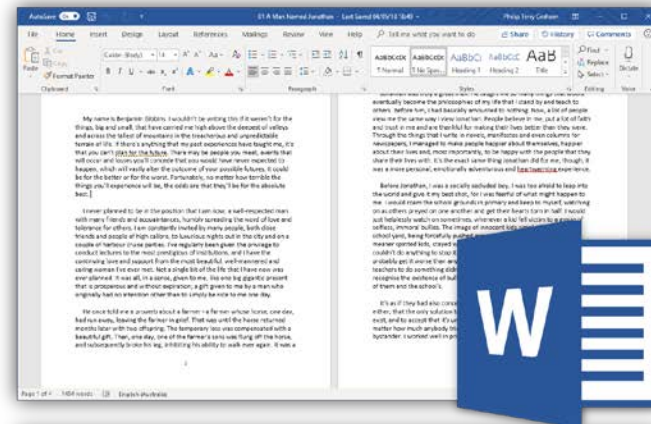
Types of recovered office files

Since modern office applications use a ZIP-compressed XML-based file format, one office file is basically a ZIP archive with XML files inside. Then, it's possible to separate recovered results into three groups, the same as for ZIP files.

GOOD file – all XML files inside the office file are good and have a correct CRC, the Central directory is found and has a good structure, the Footer of the Central directory is found.

PARTLY DAMAGED file – at least one XML file is unpackable and has a good CRC, the Central directory is found and has a good structure, the Footer of the Central directory is found.

BAD file – the Local header is damaged/ the Central directory is not found or is damaged.

ruSolut

Partly damaged .docx file

Partly damaged .docx file

Block	Name	Extension	Size	Status	Error
3582976	000003582976.docx	docx	877,001	Partly Damaged	
3022208	000003022208.docx	docx	877,414	Partly Damaged	
2982144	000002982144.docx	docx	876,759	Partly Damaged	
3049600	000003049600.docx	docx	876,605	Partly Damaged	
2580576	000002580576.docx	docx	974,399	Partly Damaged	
4887840	000004887840.docx	docx	875,911	Partly Damaged	
3055872	000003055872.docx	docx	152,008	Partly Damaged	
2606112	000002606112.docx	docx	972,572	Partly Damaged	
2614944	000002614944.docx	docx	861,371	Partly Damaged	
2970496	000002970496.docx	docx	876,728	Partly Damaged	
2968768	000002968768.docx	docx	877,725	Partly Damaged	
3047872	000003047872.docx	docx	874,446	Partly Damaged	
3868864	000003868864.docx	docx	878,057	Partly Damaged	
4869024	000004869024.docx	docx	875,162	Partly Damaged	
4855200	000004855200.docx	docx	876,759	Partly Damaged	
3900352	000003900352.docx	docx	151,972	Partly Damaged	
3023936	000003023936.docx	docx	877,632	Partly Damaged	
3019456	000003019456.docx	docx	1,400,598	Partly Damaged	
4870752	000004870752.docx	docx	875,168	Partly Damaged	
4850016	000004850016.docx	docx	876,850	Partly Damaged	
3061568	000003061568.docx	docx	875,403	Partly Damaged	
3027392	000003027392.docx	docx	877,184	Partly Damaged	
3015808	000003015808.docx	docx	875,688	Partly Damaged	
3029120	000003029120.docx	docx	875,702	Partly Damaged	
3002432	000003002432.docx	docx	878,814	Partly Damaged	
3000704	000003000704.docx	docx	877,659	Partly Damaged	
3581248	000003581248.docx	docx	877,037	Partly Damaged	

Name	Size
[Content_Types].xml	2,438
_rels/.rels	590
word/_rels/document.xml.rels	2,212
word/document.xml	41,821
word/_rels/footer1.xml.rels	290
word/footer2.xml	6,323
word/footer1.xml	6,113
word/endnotes.xml	1,606
word/footnotes.xml	1,612
word/_rels/header1.xml.rels	290
word/_rels/footer2.xml.rels	289
word/header1.xml	5,941
word/media/image3.jpg	40,167
word/theme/theme1.xml	7,021
word/media/image1.jpeg	30,409
word/media/image2.jpeg	49,562
word/settings.xml	8,814
word/styles.xml	24,047
customXml/itemProps1.xml	341
customXml/_rels/item1.xml.rels	296
customXml/item1.xml	218
docProps/app.xml	1,003
word/stylesWithEffects.xml	24,800
word/webSettings.x?l	2,074
word/fontTable.xml	2,531
docProps/core.xml	697
word/numbering.xml	22,431

Checked files: 7135, Good files: 4180, Bad files: 2846

Save all files | Save good files | Save bad files | Save partly damaged files | Search | Cancel
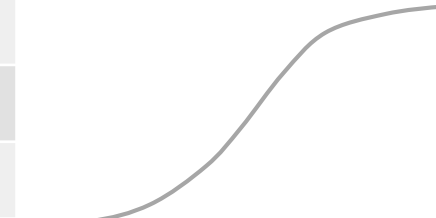
Metadata information is usually stored in the folder *docProps*. Two or more XML files are stored inside that folder: *app.xml* that stores metadata information extracted from the Word application itself and *core.xml* that stores metadata from the document itself, such as *the author name, last time it was printed, etc.*

Word folder contains the actual content of the document. An XML file called *document.xml* is the main document, containing most of the content of the document itself.
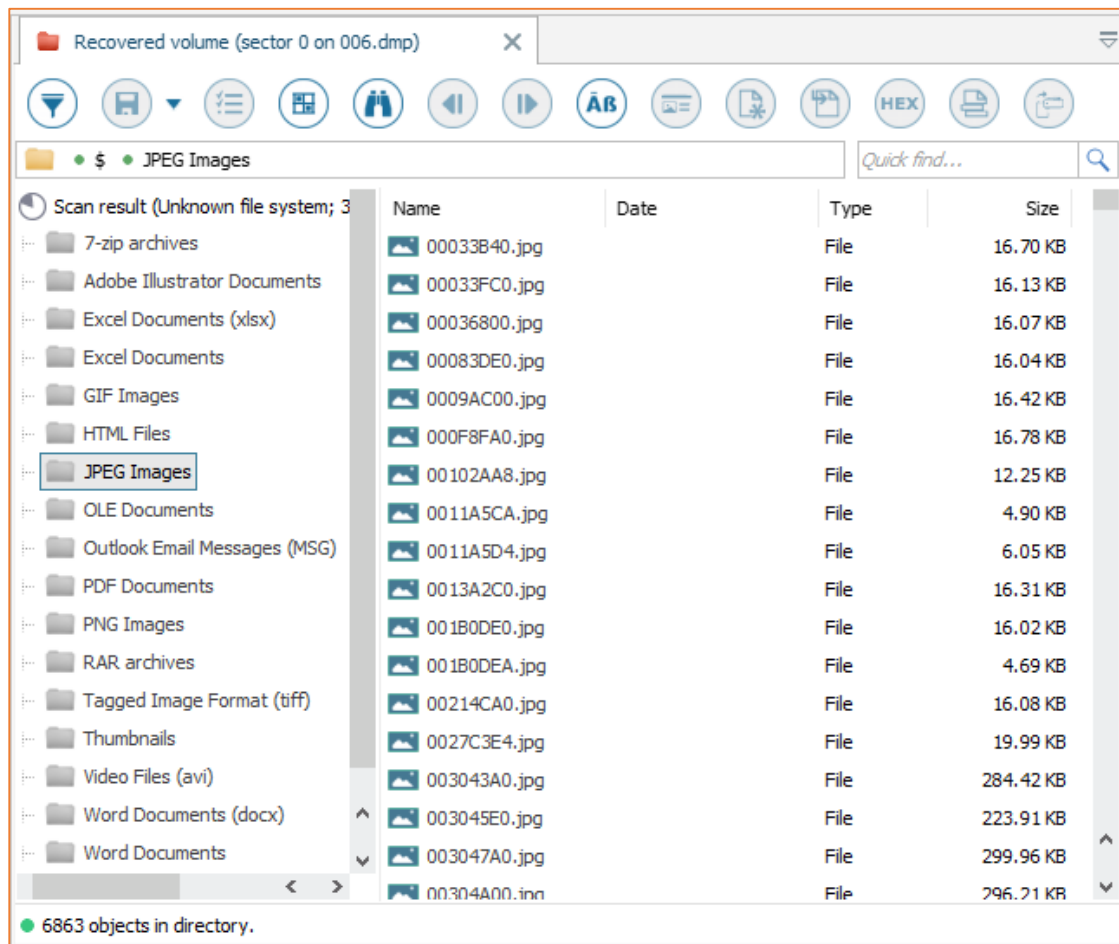
Partly damaged .docx file

CORE.XML

DOCUMENT.XML

Comparison of results

Software #1

FILE extension	Total amount
ZIP	23
DOCX	276
XLSX	18
JPEG	6863

Comparison of results

Software #2

FILE extension	Total amount	Good	Bad
ZIP	3	0	3
DOCX	280	162	118
XLSX	15	14	1
JPEG	6774	4077	2697

Comparison of results

Software #3

FILE extension	Total amount	Good	Bad
ZIP	126	15	111
DOCX	172	159	13
XLSX	17	17	0
JPEG	6775	5327	1448

Comparison of results

VNR carver

FILE extension	Total amount	Good	Bad
ZIP	117	3	114
DOCX	172	66	106
XLSX	17	15	2
JPEG	6828	4096	2732

Comparison of results

JPEG

	Total	Good	Bad
SW 1	**6863**	X	X
SW 2	**6774**	4077	2697
SW 3	**6775**	5327	1448
VNR	**6828**	4096	2732

Bad pictures

BIT error	Medium damage	Critical damage
1509	1085	138

ZIP

	Total	Good	Bad	Partly bad
SW 1	**23**	X	X	X
SW 2	**3**	0	3	X
SW 3	**126**	15	111	X
VNR	**117**	3	113	1

DOCX

	Total	Good	Bad	Partly bad
SW 1	**276**	X	X	X
SW 2	**280**	162	118	X
SW 3	**172**	159	13	X
VNR	**172**	66	0	106

XLSX

	Total	Good	Bad	Partly bad
SW 1	**18**	X	X	X
SW 2	**15**	14	1	X
SW 3	**17**	17	0	X
VNR	**17**	15	0	2

ruSolut