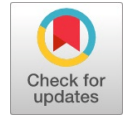# An Medical Image File Formats and Digital Image Conversion

P. Sriramakrishnan, T. Kalaiselvi, S.T. Padmapriya, N. Shanthi, S. Ramkumar, N. Kalaichelvi

Abstract: *Medical image file formats make the confusing aspect to young researchers who start work with medical images. Medical image format conversion is still a tedious task due to the different structure of files. Digital image conversion from the medical image is an important pre-processing step to the process and visualizes the data. This article presents an overview of the major medical image file formats such as Analyze, neuro-imaging informatics technology initiative (NIFTI), MINC, and digital imaging and communications in medicine (DICOM). Then the characteristics and strengths of the various formats are discussed. Also the article describes the easiest way of digital image conversion from medical image format.*

Keywords: **Medical image, file formats, NIFTI, DICOM, Analyze, file conversion**

I. INTRODUCTION

Rapid development placed in the medical industry after the arrival of medical imaging modalities such as ultrasound, computed tomography (CT), positron emission tomography (PET) and magnetic resonance imaging (MRI). These modalities are playing an vital role in disease identification and treatment procedure. Modalities produce huge volumes of medical images to visualize the internal structures of human organs [1]. Patient data obtained from modalities are differing from the digital image format. Medical image is a visual representation of internal organs. Medical imaging in the field of clinical and research areas supplies a variety of image formats. Medical image file formats contain more information about the patient's details

and equipment setting. Medical image data was aligned in the form of an array of picture element called as the pixel in two dimensional and voxel in three dimensional [2].

Image format describes the standardized procedure to store the data inside an image. Medical image dataset contains more images to represent the three-dimensional view. Medical image intimates the information about the number of pixels in the horizontal and vertical direction, number of bits per pixel and number of images per subject. Further, medical format needs to be converting into digital format for process and visualize the images. Medical image format conversion is a critical task for researchers that can be done by open source software packages such as Mango, MRIcro, 3D slicer, MicroDicom, etc., [5].

This papers gives more details about medical imaging file formats and motivate the young researchers in the field of medical image processing. This article aims to increase the awareness about format conversion between medical images to digital images. This paper discusses the strength and limits of each file format in detail manner. Remaining chapters was organized as follows. Section II discussed the limits and strength of the medical image formats. Discussion about the medical image file conversion is given in section III. Section IV concludes the paper.

II. MEDICAL IMAGE FORMAT

Representation of medical images requires two major components namely, metadata and image data and shown in Figure 1. Metadata provides the structure and information about the image. Metadata can be automatically added to an image by the capturing device [3]. Metadata stored at the beginning of an image file or separate file as a header which containing technical, descriptive and administrative metadata. Technical metadata includes image dimension, depth of pixel and spatial resolution, camera settings, and dots per inch. Descriptive metadata contains image creator, keywords related to the image, captions, titles and comments. Administrative metadata is added manually with the information such as licensing rights, restrictions on reuse, and contact information for the owner.

Image data is the raw intensity values to represent the pixels. Each pixel has its own memory space to save the intensity in term of bits called depth per pixel. In, binary image each pixel is store in a single bit either zero or one. X-ray, CT, MRI are produced grey scale images with 8 bits or 16 bits. The number of gray levels between black to white

Retrieval Number: A10931291S419/2019©BEIESP
DOI:10.35940/ijeat.A1093.1291S419
Journal Website: www.ijeat.org

74

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)
© Copyright: All rights reserved.

depends on number bits used to represent the pixel. PET and SPECT scanner has given colour images with 24 bits per pixel in a respective red-green-blue palette. Size of the image data is calculated from Eqn. (1) using metadata, number of rows, number of columns and bits per pixel or voxel.
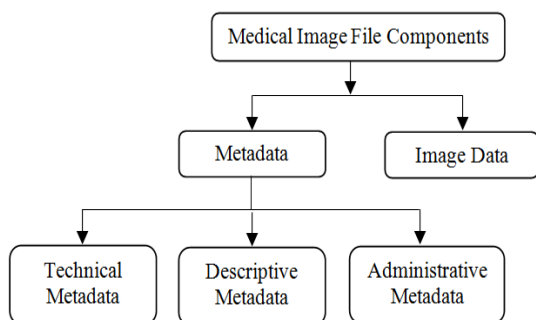

Fig. 1. Medical image file components formats

$$Image\ size = size\ of\ metadata + no.of\ rows * no.of\ columns * bits\ per\ pixel$$

Medical image files were stored in two ways. The first is metadata stored in an image itself. The main advantage of the first method is all the image information stored in a single file. On the other hand problems may arise when extracting or modifying information from the file because of its size and complex structure. The second method uses separate file to store the metadata information called a header with .hdr extension. Famous file formats used in the medical industry is shown in Figure 2. General characteristics of medical image file format are listed in Table 3. Remaining sections of this article talks about the medical image file formats in detail.
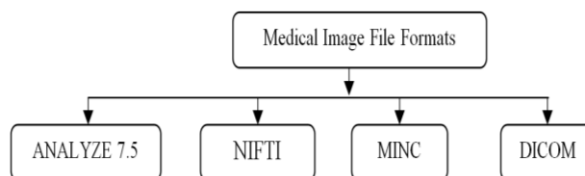

Fig. 2: Major medical image file formats

Table 1: General characteristics of famous medical image formats

Format	Year	Extension	Header Length (in Bytes)	Files per Scan	Produced by
Analyze 7.5	1980	.img/.hdr	348 bytes	Two	PET
NIFTI – I	2000	.nii	352 bytes	More number of Slices	PET
MINC 2.0	1992	.mnc	Highly standardized and extensible	One	Diffusion MRI (dMRI)
DICOM	1993	.dcm	128 bytes File Preamble + 4 bytes DICOM prefix	More number of Slices	CT, MRI, X-ray, fluoroscopy, angiography, mammography, PET, SPECT, Endoscopy, microscopy

Table 2: Analyze 7.5 Data types and Sizes

S. No	Data type Name	Name of Data types	Size of Data types
1	DT_NONE	None data type	0
2	DT_UNKNOWN	Unknown data type	0
3	DT_BINARY	Binary data type	1
4	DT_UNSIGNED_CHAR	Unsigned char	8
5	DT_SIGNED_SHORT	Signed short	16
6	DT_SIGNED_INT	Signed integer	32
7	DT_FLOAT	Floating point	32
8	DT_COMPLEX	Complex	64
9	DT_DOUBLE	Double precision	64
10	DT_RGB	-	128
11	DT_ALL	-	255

Table 3: ANALYZE 7.5 Header file important properties and Values of IBSR dataset

Parameters	Value	Type	Field Size (in Bytes)
HdrFileSize	348	int32	4
HdrDataType	'dsr'	char	10
DatabaseName	'..\work_t1\ibsr\s'	char	18
Extents	0	int32	4
SessionError	0	int16	2
Regular	1	char	1
Dimensions	[256 65 256 1]	int16	16
VoxelUnits	'mm'	char	2
ImgDataType	'DT_SIGNED_SHORT'	int16	2
BitDepth	16	int16	2
PixelDimensions	[1 3.1000 1]	float	32

75

A.Analyze 7.5

Analyze 7.5 is a file format for storing MRI data developed by the Biomedical Imaging Resource (BIR) at Mayo foundation in late 1980. Analyze 7.5 contains two files to represent the image data and metadata. The files are available in the extension of ".img" and ".hdr". The .img file contains the information in the form of voxel raw data.

The .hdr file contains information about the img file, such as a number of voxels in each dimension and voxel size. Size of Analyze 7.5 header has 348 bytes. The header file was constructed in C structure contains three substructures such as header key (40 bytes), image dimension (108 bytes), and data history (200 bytes). Header key and image dimension are essential structure and data history is an optional one. Analyze header is a flexible one and that could be adapt with new user-defined data types.

Header key contains several elements namely sizeof_header, extends, and regular. Element sizeof_header indicates the size of the header in byte representation. Element regular mention all images in the volume are the same size. Image dimension has several elements such as X,Y, Z dimension of data, spatial units of measure for a voxel, datatype, pixel dimension in millimetre. Table 2 shows the various available data types to Analyze 7.5 with their respective size. The Image Processing Toolbox of MATLAB has given analyze75info and analyze75read functions to read Analyze .hdr and .img files respectively. IBSR20 dataset header files important parameters are given in Table 3 using analyze75info function.

Analyze format was frequently used with SPM, FreeSurfer, AIR, MRIcro and Mango software. Positron emission tomography can give the output of file format in Analyze 7.5. This file format does not specifies the information about orientation in space. Therefore, every Analyze file requires the space orientation from .mat file. Analyze files do not contain time information. This drawback overcomes by the NIFTI format.

B. NIFTI

NIFTI overcomes the drawbacks of Analyze 7.5 developed at the early of 2000 by National Institute of Health. NIFITI file format is more compatibility with the Analyze 7.5, data stored in NIFTI format also uses a pair of files, ".img" and ".hdr" [4]. A process with a pair of file is inconvenient and also error-prone. To address these issues, NIFITI allows for storing the header and data as a single file with ".nii" extension. During transmit the NIFITI file through a network, the deflate algorithm helps to compress and decompress the file. The compressed version of single and pair of files have the extension as ".nii.gz" and ".hdr/.img.gz".

In NIFTI file format, the first three dimensions store the three spatial data, *x, y, z* and fourth dimension reserve for time point *t*. The header of NIFTI file format that the size would be 348 bytes in case of ".hdr/.img" and a size of 352 bytes in the case of ".nii". NIFTI header parameters are very similar like ANALYZE 7.5 which given in Table 3. Information of header structure can be read through niftiinfo() function using MATLAB. NIFTI header info of Brain Imaging of Normal Subjects (BRAINS) dataset is given in Table 4.

NIFTI was created for handling Neuro-imaging but can be used for other fields as well. NIFTI has several features such as raw data saved in the 3D image which contains two affine coordinates to relate voxel index to spatial index. NIFTI has an advantage as storing two files per 3D scans instead of handling multiple Analyze files. NIFTI file can allow storing some addition parameters such as key acquisition parameters, and experimental design.

TABLE 4: NIFTI header information of brains dataset

Parameters	Value
Filename	group00001_T1w_CSF_probability.nii
Filesize	40894816
Description	'DAD210715'
ImageSize	[256 156 256]
xPixelDimensions	[1 1.3000 1]
Datatype	'single'
BitsPerPixel	32
SpaceUnits	'Millimeter'
TimeUnits	'Second'
MultiplicativeScaling	1

C. MINC

MINC 1.0 file format was designed for a flexible way to store medical data upon the NetCDF (Network Common Data Format). It was developed by Montreal Neurological Institute (MNI) from1992. MINC file has a platform independent format based on NetCDF [6]. MINC adapts all the features of NetCDF standard with patient information, image information, acquisition information, etc.

A more recent version of MINC format (MINC 2.0) has been restructured by hierarchical data format version 5 (HDF5) for hierarchical structure, internal compression, 64-bit file sizes, and other more modern features. Header or metadata of MINC in HDF5 can construct hierarchical of group or dataset and that contains any number of metadata entries. Metadata represents in any form of data type including, integer, floating point, or character string. There are three major subgroups available under MINC 2.0 namely, info, dimension, and image.

Table 5: MINC header information of figshare dataset

Parameters	Value
Filename	subject04.mnc
Pixsize	[0.5000 0.5000 0.5000]
Space_start	[-90.2500 -126.2500 -72.2500]
Mag	255
Float	0
Min	0
Num_of_slice	362
Imfm	[362 434]
Frames	1
FileType	'mnc'

The info subgroup carries the information about data related to scan, patient identification details, modality, and experimental settings. This field is an optional one for the large datasets. Dimension subgroups talk about the property of MINC file dimension. This has several subgroups xspace, yspace, zspace, time, and vector_dimension.

Parameters xspace, yspace, zspace represents the data in the axis on patient left to right, patient posterior to patient anterior, and inferior to patient superior respectively. Time defines time samples of scanning and vector_dimension represent axis used for components of a vector field. Image subgroup contains actual image data of MINC 2.0 through subgroups such as image-min, image-max, and image. A lowest and high value of each slice is stored in image-min and image-max respectively. In MATLAB, loadminc() function helps to read the MINC file format interims of metadata and volume data. We are using Figshare dataset with MINC format for extracting header information. Table 5 reports the subject04 header information of Figshare dataset.

D. DICOM

Digital imaging and communication in medicine (DICOM) is a communication standard developed by American College of Radiology (ACR) and National Electrical Manufacturers Association (NEMA) in 1993. DICOM is a generic file format used to handling, storing, printing, transmitting and displaying medical images. DICOM has a protocol to exchange a secure medical image transfer between two or more entities [7]. DICOM integrates the image acquisition devices, servers, PACS, workstations, and printers from different manufacturers. DICOM standard is accepted by medical imaging modalities such as ultrasound, CT, MRI, and radiation therapy. DICOM standard includes the exchange protocols, compression, visualization, and result analysis. DICOM is an open source standard and free to use. DICOM is a platform independent methods and the file format easily convert into digital image file formats (.bmp, .tif, .jpg). DICOM generates a single file for each image.

DICOM file consists of both metadata (header) and image data packed in a single file and are represented as '.dcm'. DICOM header contains patient information, acquisition parameters of modality, an operator identification, image dimensions, and matrix size. This header is organized into a series of tags and tags are organized into a group of data elements. DICOM header consists of 128 bytes followed by 4 bytes of DICOM prefix. MATLAB contains dicominfo( ) function to read the header information of DICOM file. Table 6 shows the details of DICOM header of IBSR18 dataset. The remaining portion of the DICOM file contains the image data.

III. DISCUSSION

The pervious section gives an outline about general medical image file formats. This section discusses the format conversion from medical image formats to digital image formats. Medical image file conversion is a pipeline process which will be combined by one or more software. This article discusses Mango and MRIcro software to simple file format conversion. They can load the dataset from medical format, view and store to digital format. DICOM is well known generic file format, which can support to convert into the digital file format such as tiff, bmp, jpg. Therefore, any medical image format needs to convert into DICOM format for digital conversion.

Mango (Multi-image Analysis GUI) is a software for viewing, editing and analyzing volumetric medical images [11]. This is non-commercial software for educational and scientific uses. This supports to load and convert various formats like Analyze 7.5, NEMA-DES, MINC, NIFTI and DICOM [8-10]. This can be capable to load and display the 2D, 3D and 4D volumes. Mango provides tools for creation and editing of region of interest (ROI), surface rendering, filtering, image registering, and histogram analysis. This software can extend user-defined plug-ins like behaviour analysis, disease analysis, skull stripping (BET), paradigm analysis, and post matching co-registration.

This software can load and view any supported format of medical image file formats through the open option in the menu bar. Then the stack of 2D slices is opened in the new window with three different orientations such as axial, coronal and sagittal. Next save as menu available in the new window helps to save the different file format. User can choose the destination header file format available in the right side window. According to the analysis, DICOM header format is a generic file format for converting digital images. This article suggests that the any medical format should be converted into DICOM format. This DICOM header helps to convert into digital images for visualization in Windows operating systems.

A lot of open source software's are available to load and view the DICOM header. MRIcro is a standalone program to analyze MRI, fMRI and PET images. This software can display the following formats: NIfTI (.nii, .nii.gz, .hdr/.img), Bio-Rad Pic (.pic), NRRD, Philips (.par/.rec), ITK MetaImage (.mhd, .mha), AFNI (.head/.brik), Freesurfer (.mgh, .mgz), and DICOM (.dcm). Final output can be exported into BMP, JPEG, PNG or TIF file formats format. The 'open Analyze format allows you to load and view an Analyze or NIfTI format image directly and that can be export into any one of the digital file formats. Plenty of features are available in MRIcro such as, histogram smoothing, 3D volume rendering, ROI selection, rotation, contrast adjustment, overlays and selecting a coordinate position. Also several supporting features are also incorporated such as automatic skull stripping using BET method, remove air speckles and smooth image.

MRIcro can load the DICOM header given by the Mango software through import open header option. The corresponding dataset is loaded in the MRIcro viewer and in case of MRI datasets orientation (axial, coronal, sagittal) options also available in the slice viewer menu. Researchers can make necessary pre-processing (contrast and brightness adjustments, histogram adjustment, rotation, skull stripping) before export the dataset. Finally, save as picture option in the file menu helps to save the dataset into any one of the supported digital file formats (BMP, JPEG, PNG or TIF).

IV. CONCLUSION

The article discusses the medical image file format given by the modalities and conversion into the digital file format. These images are viewed in windows systems and which will facilitate the further researches on artificial intelligence and machine learning areas. Finally this article concluded that the simple procedure for converting medical formats to digital formats using open source software.

77

REFERENCES

1. J.L. Prince & J.M. Links, *Medical imaging signals and systems*. Upper Saddle River, NJ: Pearson Prentice Hall, 2006.
2. M. Larobina & L. Murino, Medical image file formats. Journal of digital imaging, vol. 27, no. 2, pp. 200-206, 2014.
3. D. Korenblum, D. Rubin, S. Napel, C. Rodriguez, & C. Beaulieu, Managing biomedical image metadata for search and retrieval of similar images, Journal of digital imaging, vol. 24, no.4, pp. 739 -748, 2011.
4. K. J. Gorgolewski, T. Auer, V.D. Calhoun, R.C. Craddock, S. Das, E.P. Duff, ... & D.A. Handwerker, The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Scientific Data, vol.3, pp.160044, 2016.
5. B. Whitcher, V.J. Schmid, & A. Thornton, Working with the DICOM and NIfTI Data Standards in R, Journal of Statistical Software, vol.6, 2011.
6. R.D. Vincent, P. Neelin, N. Khalili-Mahani, A.L. Janke, V.S. Fonov, S.M. Robbins, ... & D. MacDonald, MINC 2.0: a flexible format for multi-modal images, Frontiers in neuroinformatics, vol.10, no.35, 2016.
7. D.R. Varma, Managing DICOM images: Tips and tricks for the radiologist. The Indian journal of radiology & imaging, vol. 22, no.1, 4, 2012.
8. J.L. Lancaster, M.D. Cykowski, D.R. McKay, P.V. Kochunov, P.T. Fox, W. Rogers, ... & J. Mazziotta, Anatomical global spatial normalization. Neuroinformatics, vol.8, no.3, pp.171-182, 2010.
9. J.L. Lancaster JL, A.R. Laird, S.B. Eickhoff, M.J. Martinez, P.M. Fox, P.T. Fox, Automated regional behavioral analysis for human brain images. Frontiers in Neuroinformatics, 6, 23, 2012.
10. Mango Software, http://ric.uthscsa.edu/mango/, Last Accessed 22nd May 2019.
11. B. Liu, M. Zhu, Z. Zhang, C. Yin, Z. Liu, & J. Gu, Medical image conversion with DICOM. In 2007 Canadian Conference on Electrical and Computer Engineering. IEEE, pp. 36-39, 2007