

TOOLS FOR MANUAL 3D MEDICAL IMAGE SEGMENTATION AND DATA PREPROCESSING FOR AUTOMATIC 3D MEDICAL IMAGE SEGMENTATION

X.N.Zayniddinov

Head of the "Artificial Intelligence" Department of Tashkent University of Information Technologies named after Muhammad al-Khwarizmi (TUIT), Doctor of Technical Sciences, Professor.

E-mail: tet2001@rambler.ru

R.M.Normatov

Master student at Tashkent University of Information Technologies named after Muhammad al-Khwarizmi (TUIT).

E-mail: aou1996@gmail.com 99 851 09 96

B.R. Azimov

Associate Professor of the Department of Artificial Intelligence at Tashkent University of Information Technologies named after Muhammad al-Khwarizmi (TUIT).

E-mail: bunyodbekazimov11@gmail.com 90 268 48 01

S. A. Gafurov

Master student at Tashkent University of Information Technologies named after Muhammad al-Khwarizmi (TUIT).

E-mail: sultontatu21317@gmail.com 97 263 33 93

Abstract: This paper shows the useful tools for 3D medical image preparation and semi-automatic segmentation, moreover, some python code examples are shown for data preprocessing. ITK-Snap and 3D Slicer became useful tools for above problems. The result show that, usage of both tools decreases research time and highly effective.

Key words. NIFTI, DICOM, ITK-Snap, 3D Slicer, nibabel, Segmentation

1. Introduction. Nowadays, medical image processing and further using of data from medical images in order to detect illnesses or segmentation is becoming important in both, IT and medicine engineers. However, in this field of study, finding data, especially labelled data, to train with deep learning and machine learning tasks is difficult. While during research on prostate segmentation and prostate cancer detection task, we had a difficulty to find data for this, there is no labelled data. We got data from decathlon database Task 5, which is not labelled. In order to labelling medical images, we used ITK-Snap software which is very useful for semi-automatic segmentation. After labelling some of data for further train our neural network for automatic segmentation, we need to preprocess 3D medical image data. In decathlon database, all images are on NIfTI (Neuroimaging Informatics Technology Initiative) file format with not equal slices on each patient. In order to make all them equal slices we had to get all slices from each patient NIfTI file, for this reason, with the help of 3D Slicer software, which gets slices and makes DICOM file, and some codes on python we made equal slices on all NIfTI images. In this article, we will try to explain usage of these both software and data preprocessing for prostate image segmentation.

2. Usage of ITK-Snap and 3D Slicer software for medical image segmentation.

Firstly, we should give some information about DICOM and NifTI file formats. Digital Imaging and Communications in Medicine (DICOM) is a technical standard for the digital storage and transmission of medical images and related information. It includes a file format definition, which specifies the structure of a DICOM file, as well as a network communication protocol that uses TCP/IP to communicate between systems. The primary purpose of the standard is to facilitate communication between the software and hardware entities involved in medical imaging, especially those that are created by different manufacturers. Entities that utilize DICOM files include components of picture archiving and communication systems (PACS), such as imaging machines (modalities), radiological information systems (RIS), scanners, printers, computing servers, and networking hardware [1].

The Neuroimaging Informatics Technology Initiative (NifTI) was established to work with medical and research device users and manufacturers to address some of the problems and shortfalls of other imaging standards. The NifTI standard was specifically designed to address these challenges in the neuroimaging field, focusing on functional magnetic resonance imaging (fMRI).

One of the most significant challenges neurosurgeons faced with older image formats, specifically the Analyze 7.5 file format, was the lack of information about the orientation of image objects.

Orientation was ambiguous and unclear, forcing anyone attempting to analyze an image to add detailed notes about the orientation of objects within images. In particular, there was often confusion as to which side of the brain a doctor was looking at, a significant problem that needed a solution.

NifTI solved this problem in the following way:

“In the NifTI format, the first three dimensions are reserved to define the three spatial dimensions — x, y and z —, while the fourth dimension is reserved to define the time points — t.” Other dimensions store “voxel-specific distributional parameters [and] vector-based data.”

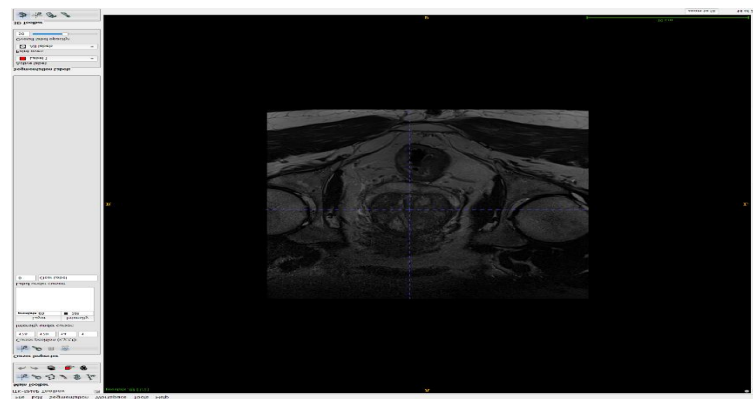
According to NifTI, “the primary goal of NifTI is to provide coordinated and targeted service, training, and research to speed the development and enhance the utility of informatics tools related to neuroimaging.” [2].

From the above information, it can be known that, it would be best for researchers and make train neural network easier if one can work with DICOM and NifTI formats.

As mentioned above, in our example we used DECATHLON Task 5 dataset, which is prostate NifTI images which is not labelled. For labelling, these images on NifTI format we use ITK-Snap.

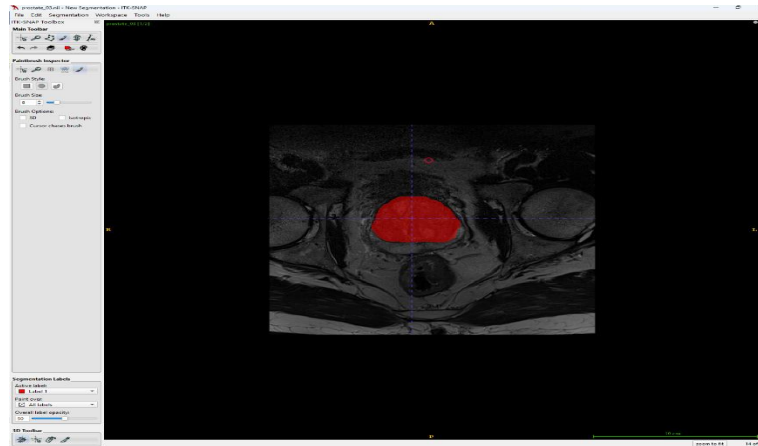
ITK-SNAP is a free, open-source, multi-platform software application used to segment structures in 3D and 4D biomedical images. It was originally developed at the University of North Carolina by student teams led by Guido Gerig (NYU Tanden School of Engineering), who envisioned a tool that would be easy to learn, with a limited feature set centered specifically on the task of image segmentation. Current ITK-SNAP development is led by Paul Yushkevich, Jilei Hao, Alison Pouch, Sadhana Ravikumar and colleagues at the Penn Image Computing and Science Laboratory (PICSL) at the University of Pennsylvania [3].

To use this software, we should install it first, then, load our image in any image format, for our case it is NifTI file format. Then it opens like below:



Picture 1. Prostate image before manual segmentation

After opening image as shown in picture 1, we should paintbrush mode and start segmenting our data. For each slices need to segment separately. The result of one slice's manual segmentation shown on picture 2.

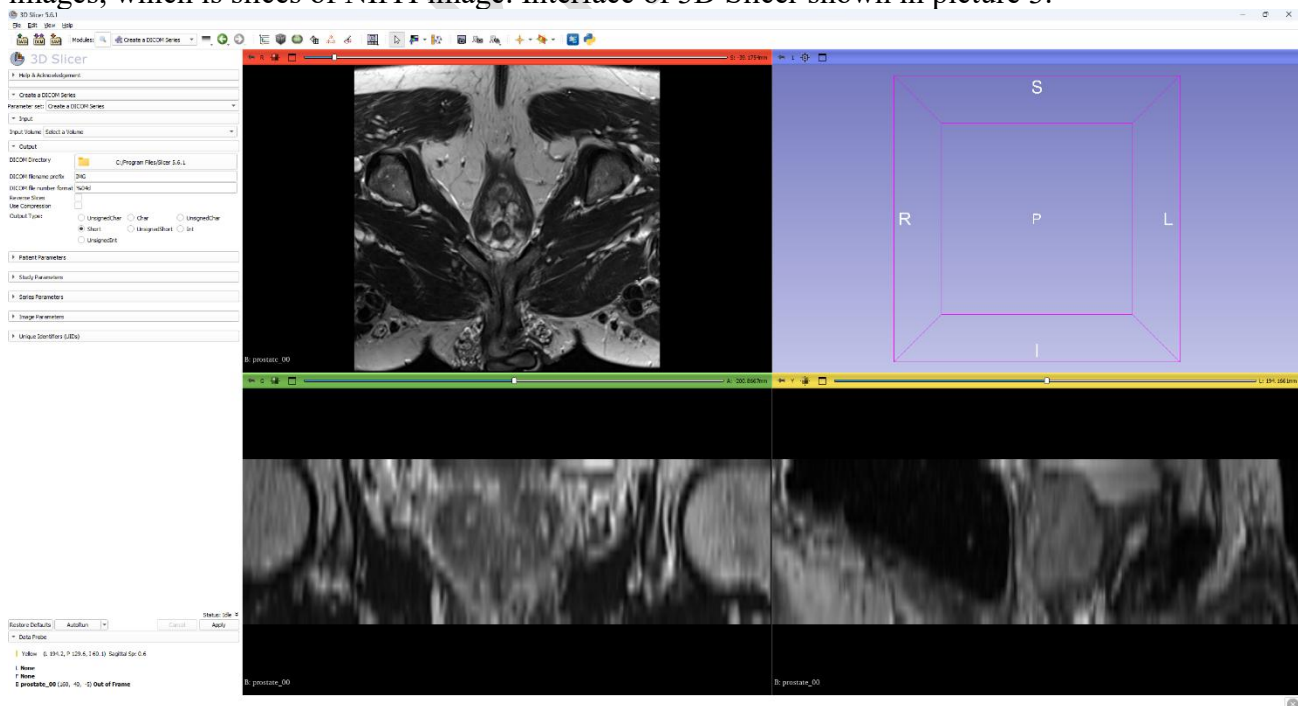


Picture 2. Prostate image after segmentation

With the use of this software we can manual segmentation, region of interest selection, resampling, select active contour parameters, see active contour results and etc. For our case, we just made manual segmentation for further training neural network.

3D Slicer is a free, open source software for visualization, processing, segmentation, registration, and analysis of medical, biomedical, and other 3D images and meshes; and planning and navigating image-guided procedures [4].

With the help of 3D Slicer we got all slices from NIFTI file as separate image in DICOM format. In order to do it, we first should upload each of our NifTI images to 3D Slicer, then, find Create DICOM Series, then choose DICOM directory, then apply. On chosen directory we can find DICOM images, which is slices of NifTI image. Interface of 3D Slicer shown in picture 3.



Picture 3. 3D Slicer Interface

After getting DICOM slices, we should check all NifTI files and choose the one with minimal number of slices, then from other NifTI files we should delete some slices in order to make slice numbers for all NifTI files same. For our case, it was 74 slices per NifTI file. This process made with python code:

```
for kasal in glob(kirish_path + '/*'):
    kasal_name = os.path.basename(os.path.normpath(kasal))
    papkalar_soni = int(len(glob(kasal+'/*'))/73)
    for i in range(papkalar_soni):
        chiqish_path_name = os.path.join(chiqish_path, kasal_name + '_' + str(i))
        os.mkdir(chiqish_path_name)
        for i, file in enumerate(glob(kasal+'/*')):
            if i==73+1:
                break;
            shutil.move(file, chiqish_path_name)
```

Picture 4. Slice of code for moving equal amount of slices to another folder for further converting them to NifTI

In the code which shown on picture 4, DICOM files location given on a kirish_path, which saves all folders for each patient. Then on kasal_name get same name as folders, then from each folder we get 74 slices and move them another folder which is chiqish_path_name, which is equal to kasal_name + numerator. Till here we got folders for each patient with same slices. After that we should convert them to NifTI.

```
kirish_path_images = 'D:/Project Segmentation/Task05_Prostate/dicom_groups/images/*'
chiqish_path_labels = 'D:/Project Segmentation/Task05_Prostate/dicom_groups/labels/*'
chiqish_path_images = 'D:/Project Segmentation/Task05_Prostate/nifti_files/images'
chiqish_path_labels = 'D:/Project Segmentation/Task05_Prostate/nifti_files/labels'
list_images = glob(kirish_path_images)
list_labels = glob(kirish_path_labels)

for kasal in list_labels:
    kasal_name = os.path.basename(os.path.normpath(kasal))
    dicom2nifti.dicom_series_to_nifti(patient, os.path.join(chiqish_path_labels, kasal_name + '.nii.gz'))
```

Picture 5. Code of converting DICOM files to NIFTI

By using dicom2nifti package, we successfully converted each patient's DICOM files to NifTI. After that we should check, if all NifTI files have necessary data. By this, we meant, sometimes some NifTI data may only have background. To check it we use nibabel package as shown in picture 6.

```
kirish_nifti_file_path = 'D:/Project Segmentation/Task05_Prostate/nifti_files/labels/*'
list_labels = glob(kirish_nifti_file_path)
for kasal in list_labels:
    nifti_file = nib.load(kasal)
    file_data = nifti_file.get_fdata()
    check_unique = np.unique(fdata)
    if len(np.unique) == 1:
        print(kirish_nifti_file_path)
```

Picture 6. Checking for having only background

Nibabel helps us to read and write access to common neuroimaging file formats, including NifTI. In our case, all NifTI files have necessary data. So, we can say our dataset is ready for further processing and training.

3. USE IN IMAGE PROCESSING.

This paper is designed to show software for manual segmentation of 3D medical data for further train, necessary medical image file formats and software to work with these files and necessary medical image data preprocessing steps.

Data preprocessing includes the following steps:

1. Not labelled dataset is manually labelled with ITK-Snap and saved as segmentation file.
2. All NIfTI files are converted to DICOM files with 3D Slicer.
3. From all NIfTI images found the least number of slices and from all NIFTIs taken that much DICOM files and it's converted again to NIFTIs in order to make train accurate and without errors. Furthermore, new converted NifTI files are checked for not having only background slices.

Python language is used to develop some process. In the beginning, the input image folder normalized with **glob** library and necessary part of DICOM files are moved to new folder with **OS** package, then, with the **dicom2nifti** package they are converted to nifty files. By using nibabel package it has checked for not having only background image. Moreover, final NifTI files are compressed as .gz format to take less place from our memory.

4. RESULT AND DISCUSSION.

In order to show labelling and data preprocessing steps we used Decathlon's not labelled 3D prostate medical image NIfTI dataset.

On the picture 1, we can see not labelled image and labelled image on the picture 2. After comparing some software we realized that these software can be very useful for beginners and even for advanced researchers too. It is easy to understand and free, open source.

5. CONCLUSIONS Software and some code implantations of 3D medical image manual segmentation and data preprocessing shown above. The task from this paper is showing the first steps for medical image processing and making them suitable for train for medical tasks.

This is the part of our main project which is creating automatic medical image segmentation and prostate cancer detection neural network. We had a difficulty to find prostate labelled data, for this reason, this can be helpful for other researchers who are having same problem. Moreover, we are glad to share this manual segmented data in the link [6]. Furthermore, for next steps we publish this dataset further processing and training neural network for automatic segmentation.

Literature:

- [1] "<https://en.wikipedia.org/wiki/DICOM> "
- [2] "<https://encord.com/blog/dicom-and-nifti-comparison/>"
- [3] "<http://www.itksnap.org/pmwiki/pmwiki.php>"
- [4] "<https://www.slicer.org/>"
- [5] "<https://drive.google.com/file/d/1Er-WmNgv3OO4ZK5QFHD0StaeWW7-EQho/view?usp=sharing>"