



+ Code + Text



## ▸ Загрузка Библиотек



↳ 2 cells hidden

## ▾ Читаю и вывожу файл DICOM

```
[ ] # проверка содержимого списка patients
print(patients)
```

```
['09-18-2008-StudyID-NA-69331']
```



```
import pydicom
```

```
# Указываем путь к файлу DICOM пациент LUNG1-001
```

```
dicom_file_path = '/content/drive/MyDrive/AI_стажировка/Текущий проект/Исходные данные/Исходный датасет/manifest-1586'
```

```
# Читаем файл DICOM
```

```
ds = pydicom.dcmread(dicom_file_path)
```

```
# Выводим содержимое файла DICOM
```

```
print(ds)
```

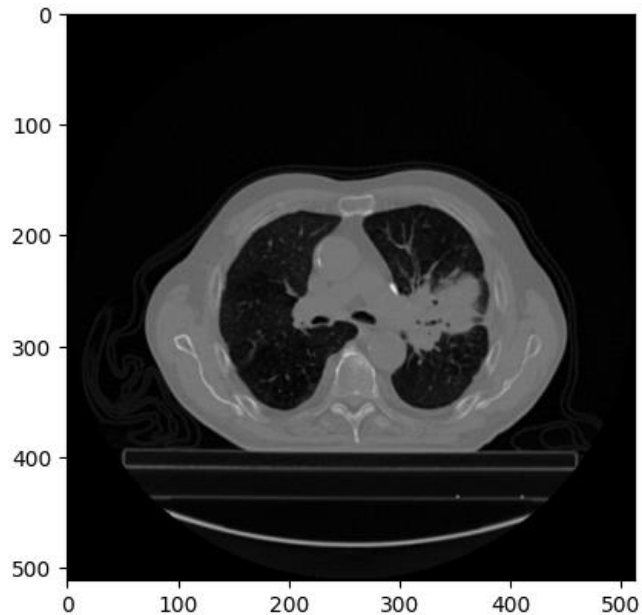
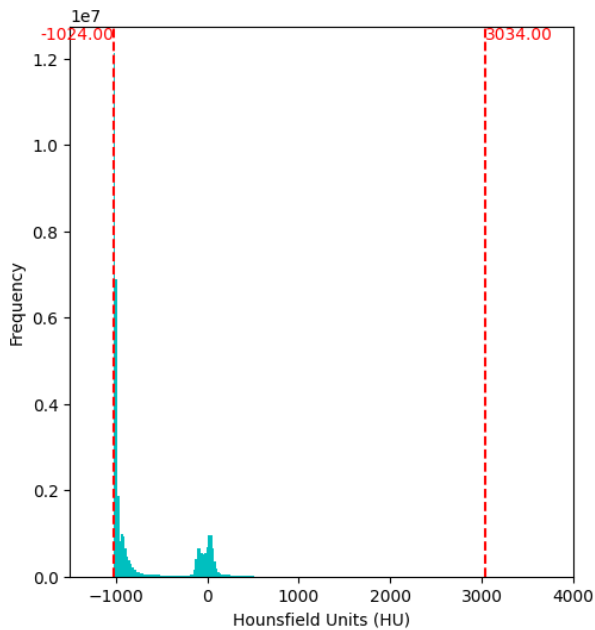
```
Dataset.file_meta -----
(0002, 0000) File Meta Information Group Length  UL: 206
(0002, 0001) File Meta Information Version       OB: b'\x00\x01'
(0002, 0002) Media Storage SOP Class UID        UI: CT Image Storage
(0002, 0003) Media Storage SOP Instance UID     UI: 1.3.6.1.4.1.32722.99.99.106464456875572292175554838571095963812
(0002, 0010) Transfer Syntax UID                UI: Explicit VR Little Endian
(0002, 0012) Implementation Class UID          UI: 1.3.6.1.4.1.22213.1.143
(0002, 0013) Implementation Version Name       SH: '0.5'
(0002, 0016) Source Application Entity Title    AE: 'POSDA'
-----
(0008, 0005) Specific Character Set             CS: 'ISO_IR 100'
(0008, 0008) Image Type                        CS: ['ORIGINAL', 'PRIMARY', 'AXIAL', 'CT_SOM5 SPI']
(0008, 0016) SOP Class UID                     UI: CT Image Storage
(0008, 0018) SOP Instance UID                   UI: 1.3.6.1.4.1.32722.99.99.106464456875572292175554838571095963812
(0008, 0020) Study Date                        DA: '20080918'
(0008, 0021) Series Date                       DA: '20080918'
(0008, 0022) Acquisition Date                  DA: '20080918'
(0008, 0023) Content Date                      DA: '20080918'
(0008, 0030) Study Time                        TM: '110915.625000'
(0008, 0031) Series Time                       TM: '121748.328000'
(0008, 0032) Acquisition Time                  TM: '113837.981607'
(0008, 0033) Content Time                     TM: '113837.981607'
(0008, 0050) Accession Number                  SH: '2819497684894126'
(0008, 0060) Modality                          CS: 'CT'
(0008, 0070) Manufacturer                     LO: 'SIEMENS'
(0008, 0090) Referring Physician's Name        PN: ''
(0008, 1090) Manufacturer's Model Name         LO: 'Biograph 40'
(0009, 0010) Private Creator                   LO: 'SIEMENS CT VA1 DUMMY'
(0010, 0010) Patient's Name                    PN: 'LUNG1-001'
(0010, 0020) Patient ID                       LO: 'LUNG1-001'
(0010, 0030) Patient's Birth Date              DA: ''
(0010, 0040) Patient's Sex                     CS: 'M'
(0010, 1010) Patient's Age                     AS: '083Y'
(0010, 1030) Patient's Weight                  DS: '72.0'
(0012, 0062) Patient Identity Removed          CS: 'YES'
```

## Загрузка слайсов легкого из папки отдельного пациента с файлами DICOM

## Преобразование пиксельных значений

Преобразовываю список срезов (slices) в трехмерный массив, затем конвертирую значения пикселей из исходных значений в единицы HU

Гистограмма распределения пикселей по HU и отображение одного из срезов (80-й срез) на экране



**Следующий этап - `resampling` для получения изоморфного разрешения – одинакового в 3х измерениях**

**Здесь хочу добиться одинаковых свойств пространства по всем направлениям.**

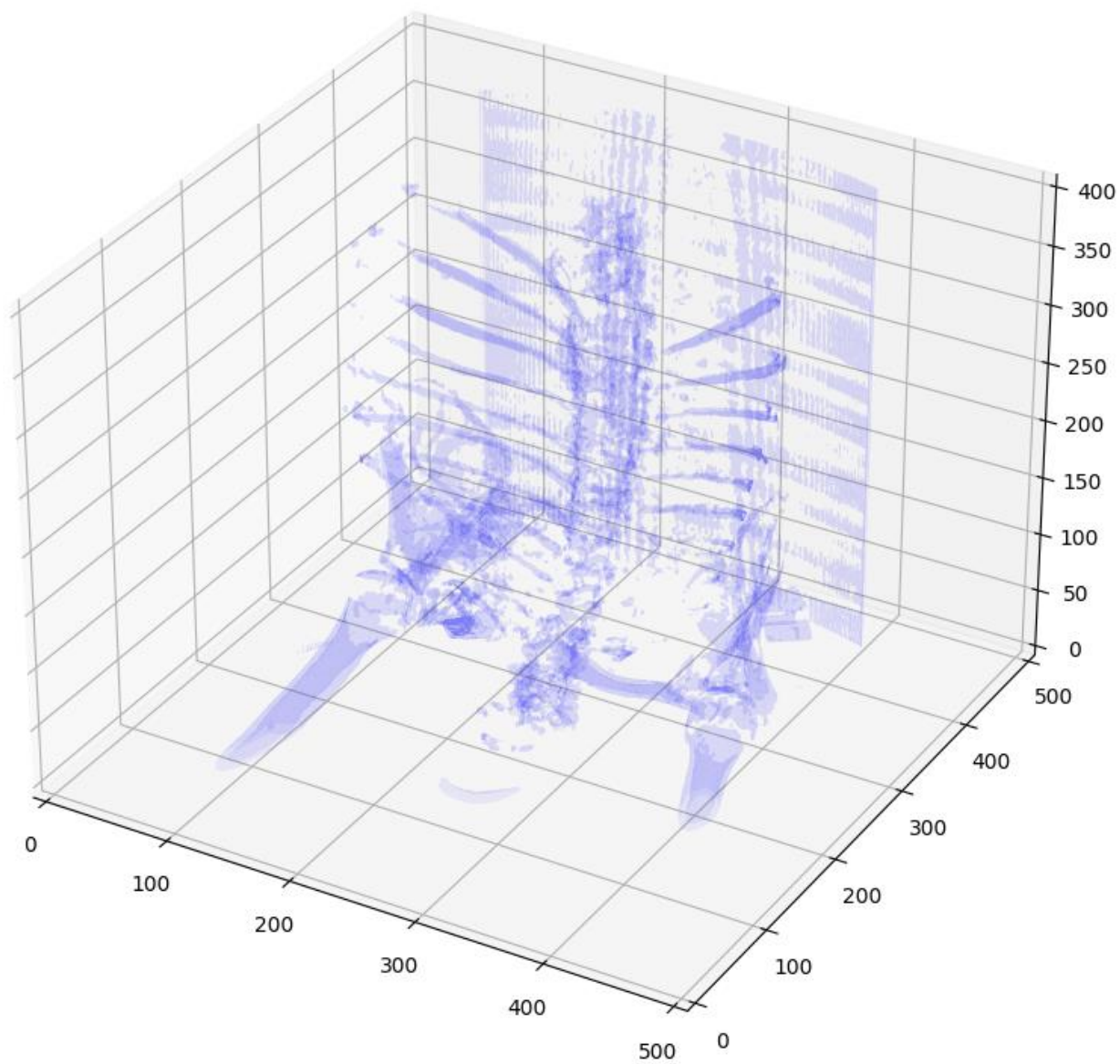
**Функция `resample` изменяет размер изображения, так чтобы расстояние между пикселями стало одинаковым по всем трем измерениям**

**Shape before resampling      (134, 512, 512)**

**Shape after resampling      (402, 500, 500)**

# Визуализация в 3D

Создаю 3D визуализацию изображения



# Сегментация лёгких

**1. Пороговая обработка изображения**

**2. Определим метку воздуха вокруг человека**

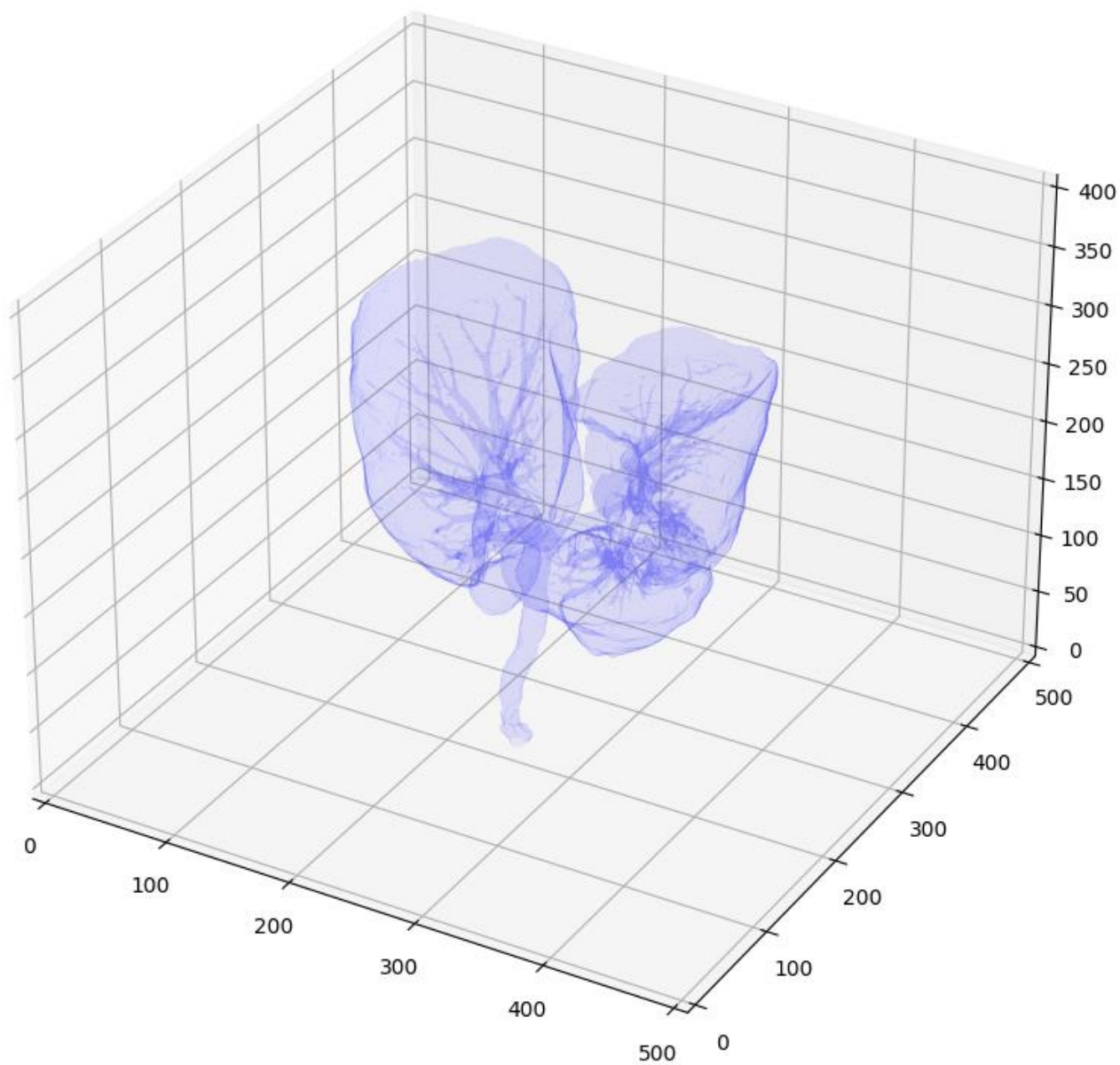
**3. Для каждого среза определяем наибольшую связную компоненту (тело + воздух вокруг человека) и устанавливаем остальные на 0. Это заполняет структуры в легких на маске.**

**Этот шаг говорит о том, что для каждого горизонтального среза ищем наибольшую область, которая связана вместе.**

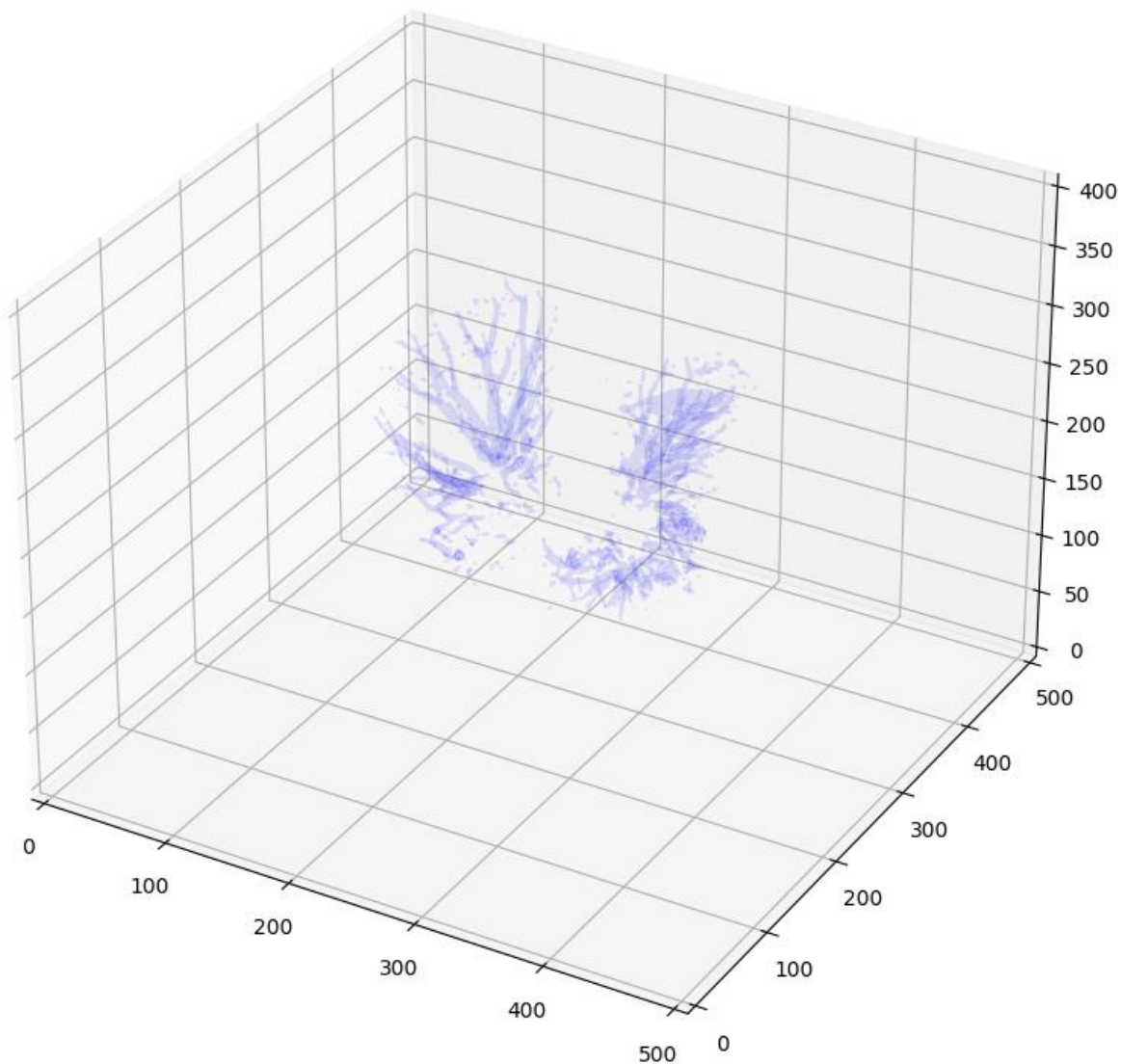
**Это будет включать в себя тело человека и воздух вокруг него.**

**Затем устанавливаем все остальные области (которые не являются частью этой наибольшей области) равными нулю.**

**4. Оставляем только наибольший воздушный карман**



**Визуализация структуры внутри лёгких**



Наши значения варьируются от -1024 до приблизительно 600. Крайнее значение 3034 нам не нужно. - Все, что выше 400, нам не интересно, так как это кости

Чтобы сэкономить место в памяти, не надо делать нормализацию и центрирование нуля заранее, сделаем во время обучения. Сейчас наше изображение - это int16, которое меньше float32 и легче сжимается.

Далее этап Нормализации значений и возможно центрирование на нуле