

Práctica 04

DOCENTE	CARRERA	CURSO
Vicente Machaca Arceda	Maestría en Ciencia de la Computación	Algoritmos y Estructura de Datos

PRÁCTICA	TEMA	DURACIÓN
04	Proyecto Final	3 horas

1. Datos de los estudiantes

- Grupo: V
- Integrantes:
 - Angel Yvan Choquehuanca Peraltilla
 - Estefany Pilar Huaman Colque
 - Eduardo Diaz Huayhuas
 - Gustavo Raul Manrique Fernandez

2. Introducción

Actualmente, el aumento y la digitalización de los hospitales, especialmente en el diagnóstico por imagen, y la necesidad de comunicaciones médicas, ha puesto de relieve la necesidad de estandarizar los protocolos de comunicación y los formatos de la información en sanidad. Uno de los estándares más exitosos hasta la fecha es DICOM (Digital Imaging and Communications in Medicine).

Para la detección de patologías como el COVID-19, se utiliza algoritmos avanzados para la detección de estas anomalías.

En este proyecto final, utilizaremos el algoritmo KDTree que nos permitirá detectar estas anomalías a partir de imágenes DICOM y con ello generar diagnósticos preliminares con el fin de acelerar la atención hospitalaria.

3. Objetivos

- Aplicar un algoritmo computacional para la detección de patologías en imágenes DICOM
- Mediante la escala Hounsfield, determinar la aparición de patologías presentes en tomografías usando un algoritmo Computacional

4. Marco Teorico

4.1. Que es DICOM



Figura 1: DICOM

DICOM es un protocolo estándar de comunicación entre sistemas de información y a la vez un formato de almacenamiento de imágenes médicas que aparece como solución a los problemas de interoperabilidad entre tipos de dispositivos.

Una imagen médica por sí misma no aporta suficiente información. Para que sea correctamente interpretada es necesario que vaya acompañada de datos del paciente y de la adquisición. Por eso formatos tradicionales como el .jpeg o el .png se quedan cortos.

El trabajo propuesto parte del interés por aplicar conocimientos del área de ingeniería para brindar soluciones a problemas del área de medicina. En particular se trabajará en el desarrollo de una solución que ayude a los profesionales de la salud a subir archivos DICOM (Digital Imaging and Communications in Medicine) o imágenes derivadas de estos, y poder previsualizarlos, elegir cuales son de interés, y almacenarlos en un formato estándar. DICOM es un estándar reconocido mundialmente para el intercambio de imágenes médicas. Está pensado para el manejo, visualización, almacenamiento, impresión y transmisión de las mismas. Además, incluye la definición de un formato de archivo y de un protocolo de comunicación de red. El protocolo de comunicación es un protocolo de aplicación, el cual hace uso del protocolo TCP/IP para comunicarse entre sistemas. Los archivos DICOM pueden ser compartidos entre dos entidades que sean capaces de recibir imágenes e información del paciente en formato DICOM. El formato de archivo DICOM tiene la particularidad de tener tanto información de la imagen como información del paciente (metadata). Fue desarrollado con el objetivo de estandarizar el intercambio de imágenes médicas, aún cuando cada fabricante de equipos de imágenes médicas trabaje con su propio enfoque, mientras que formatos como Nifti, Analyze o Minc, apuntan a facilitar y mejorar el análisis de post procesamiento.

Como se puede observar, es muy importante tener un estándar que se utilice en todos los hospitales para el mismo o similar examen, ya que esto ayuda a evitar problemas como por ejemplo si el paciente se muda de un hospital a otro. DICOM también proporciona interconectividad entre diversos sistemas médicos y soporta todas las ramas de la medicina, siendo actualmente un estándar exhaustivo. Además DICOM soporta información de imágenes comprimidas, mediante un mecanismo que permite a documentos con distinto formato ser encapsulados y transformados en un archivo DICOM. Los esquemas de compresión soportados por DICOM incluyen los siguientes: JPEG, Run-Length Encoding (RLE), JPEG-LS, JPEG-2000, MPEG2/MPEG4, y Deflated. Un archivo DICOM encapsulado incluye metadata relacionada al documento nativo y además la metadata necesaria para crear la cápsula DICOM. La innovación del formato de archivo DICOM fue establecer que la información del píxel no puede ser separada de la descripción del procedimiento médico que llevó a la formación de la imagen en sí misma. En otras palabras, una imagen que está separada de su metadata se convierte en una imagen médica sin sentido. Metadata y la información del píxel se juntan en un único archivo, y el encabezado DICOM, además de la información sobre la matriz de la imagen, contiene una descripción muy completa sobre el procedimiento utilizado para generar la imagen, además de contener información relevante

sobre el paciente. Por todo esto, el encabezado es dependiente de la modalidad de la imagen y varía en tamaño. Esto trae aparejada una desventaja importante que radica en la necesidad de un software especial para poder abrir los archivos DICOM, ya que los distintos sistemas operativos no reconocen este formato de forma nativa [4]. Los visores DICOM pueden ser software propietario provisto junto al sistema de imágenes médicas, o de terceros. Dentro de estos últimos, existen herramientas de variado desarrollo, pero con la gran desventaja de que son pagas (RadiAnt, MicroDicom, EMV Dicom Viewer, IrfanView, 3DimViewer) mientras otras requieren de conocimientos técnicos para su implementación y puesta en producción (DICOM Web Viewer, Mango). Otra desventaja que tiene el estándar DICOM está relacionada con los datos de entrada de la información. Esto se debe a la posibilidad de incorporar demasiados campos opcionales al encabezado de un archivo por parte de los fabricantes de equipos de imágenes médicas. Esta desventaja condujo a inconsistencias en el completado de estos campos. Algunos atributos dentro del encabezado de imágenes DICOM muchas veces están incompletos, se dejan en blanco o completan con información incorrecta. Todo esto es causado por la flexibilidad en la definición de la construcción de los archivos DICOM. El encabezado de un archivo DICOM consiste en un preámbulo y un prefijo, seguido de una secuencia de elementos de datos. El preámbulo contiene bytes reservados que no son utilizados, mientras que el prefijo contiene el texto ASCII "DICM". Luego del preámbulo y el prefijo viene una secuencia de elementos de datos. Cada uno de ellos consiste en una etiqueta DICOM, un código ASCII de dos caracteres indicando el valor de representación, el tamaño del elemento, y el valor. Los elementos de datos en el archivo están ordenados por el número de etiqueta DICOM. La imagen en sí misma es solamente otro elemento de datos cuyo valor de representación indica si la información corresponde a una sola imagen, varios cuadros de un estudio o un ciclo de video.

4.2. Características del Formato DICOM

El formato DICOM cuenta con objetos IOD (Information Object Definition), formados por la imagen y su información asociada (Son una representación lógica de objetos del mundo real) y DIMSE (DICOM Message Service Element), operaciones que pueden realizarse sobre un objeto. IOD y DICOM forman SOP, la unidad funcional de DICOM.

Un IOD se compone de IEs (Entidades de información) (Hay IE de paciente, de estudio, de serie, de equipo, de imagen...) que a su vez se componen de uno o varios módulos que a su vez se contienen varios atributos. Un atributo se define con nombre, etiqueta, tipo y descripción.

4.3. Composición del estandar DICOM



Figura 2: Diagrama del proceso para el análisis de imágenes DICOM

En el estándar DICOM la información se define mediante un modelo que refleja el mundo real. La imagen es el núcleo de información de un fichero DICOM. Cada fichero contiene, además de la imagen, información sobre el paciente (identificación demográfica y de identificación), el estudio en el que se encuadra la toma de la imagen, la serie a la que pertenece la imagen e información sobre la propia imagen.

4.4. Importancia del estandar DICOM

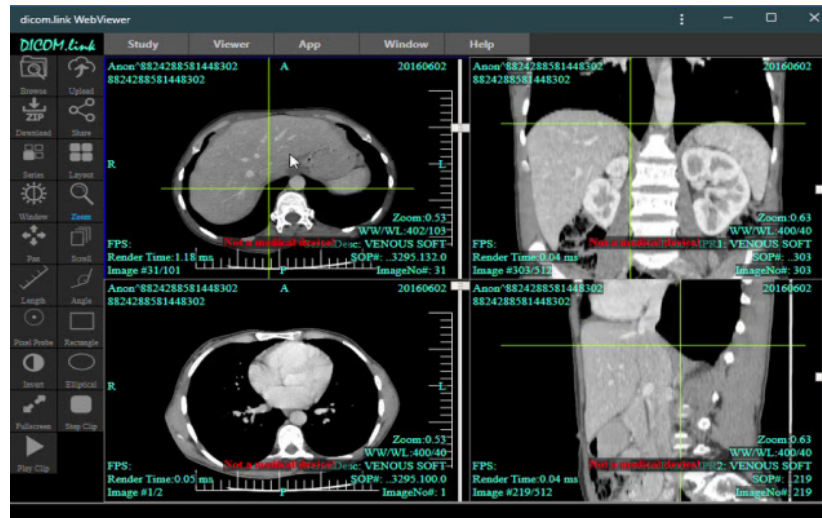


Figura 3: Uso de Software para analizar patologías o anomalías usando DICOM

DICOM permite una identificación unívoca de objetos. Cada fichero DICOM tiene un UID único compuesto por varios números.

Las comunicaciones DICOM se adaptan al estándar OSI para el intercambio de información. La AE (Entidad de Aplicación) se encarga de las comunicaciones de modo que para cada servicio existe un AE cliente y un AE aplicación.

Gracias a sus características y a su nivel de implantación, hoy día DICOM es mundialmente reconocido para el manejo, almacenamiento, impresión y transmisión de imágenes médicas.

4.5. Importancia de la Escala Hounsfield

La escala de Unidades Hounsfield (símbolo HU del inglés 'Hounsfield Units') es el resultado de la transformación de la escala de coeficientes de atenuación lineal de rayos X en una nueva escala en la cual el valor de atenuación del agua destilada en Condiciones Normales de Presión y Temperatura (CNPT) se define como 0 unidades de Hounsfield (HU).

En la tomografía computarizada se hace uso de una escala para determinar la densidad de los tejidos, esta es la escala de unidades Hounsfield (UH), que mediante la atenuación de rayos X se puede determinar rangos entre -1000 a +1000, cuyos valores dependerá del tejido en cuestión siendo -1000 un valor dado para la radiosensibilidad del aire, 0 para el agua destilada y +1000 para tejidos óseos.

TIPO DE TEJIDO	HU
Hueso compacto	¿250
Hueso esponjoso	50-300
Sangre coagulada	70-90
Tiroides	60-80
Hígado	50-70
Sangre entera	50-60
Sustancia gris	37-41
Músculo	35-50
Páncreas	30-50
Riñón	20-40
Sustancia blanca	20-34
Plasma	27 ± 2
LCR	5-10
Grasa	-80 hasta -100
Pulmón	-950 hasta -550

5. Metodología y Desarrollo

5.1. Librería Pydicom

pydicom es un paquete de Python puro para trabajar con archivos DICOM. Le permite leer, modificar y escribir datos DICOM de una manera fácil y "pitónica". Como un paquete de Python puro, pydicom puede ejecutarse en cualquier lugar donde se ejecute Python sin ningún otro requisito, aunque si está trabajando con Pixel Data, le recomendamos que también instale NumPy. Elementos centrales en pydicom modelo de objetos pydicom, descripción de clases, ejemplos

conjunto de datos Dataset el objeto principal con el que trabajará directamente. Datasetenvuelve un diccionario, donde la clave es la etiqueta DICOM (como un BaseTagobjeto, que se describe a continuación) y el valor es la DataElementinstancia correspondiente (que también se describe a continuación). Implementa la mayoría de los métodos de dict, por lo que en su mayoría se comporta como un envuelto dict. Esto permite el acceso directo a los elementos de datos a través de sus etiquetas, como se muestra a continuación. El iterador de a Datasetproduce DataElementinstancias, por ejemplo, los valores del diccionario en lugar de las claves que normalmente se obtienen al iterar a dict. Para la implementación del estándar DICOM, Python cuenta con la librería llamada pynetDicom ('Documentación de pynetdicom - documentación de pynetdicom 1.4.1', n.d.), la cual facilita la comunicación entre los usuarios (SCU) y los proveedores (SCP), adicional a esto, permite llevar a cabo la ejecución de diversos servicios, entre los que se encuentran: C-Echo, C-Storage, C-Find, C-Get, C-Move, C-Cancel, Modality Worklist. (Pianyk, 2012) Pydicom también permite la lectura de las imágenes médicas y esto es uno de sus más grandes fuertes, debido a que estas imágenes en formato DICOM estas codificadas, lo que las hace de difícil acceso o procesamiento si no se cuenta con la aplicación o programa adecuado para su decodificación.

5.2. Lenguaje Utilizado

- Javascript y librerías:
- MITK (Medical Imaging Interaction Toolkit)

5.3. Repositorio del Proyecto

Todo el documento estará alojado en el siguiente Repositorio: <https://github.com/AngelYvan/mcc-final-grup5> Para el despliegue de la aplicación se utilizará:

5.4. Preparación del entorno de desarrollo

- Software MITK:

MITK es un programa de software libre empleado para el análisis de imágenes médicas de tomografía computerizada. Este programa trabaja con archivos DICOM (Digital Imaging and Communication in Medicine) que es el formato estándar para el procesamiento de este tipo de imágenes. En ella se puede encontrar versiones actualizadas del programa así como un manual para el usuario con explicaciones de las herramientas de las que dispone y otro manual para desarrolladores.

En el diseño del código se ha trabajado contando con las librerías “Medical Imaging Interaction Toolkit (MITK) debido a la facilidad para el manejo de las imágenes en formato DICOM que se están utilizando.

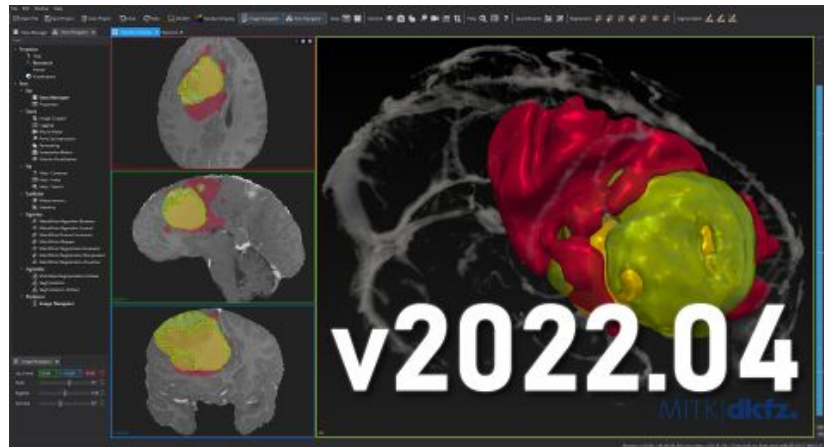


Figura 4: Software MITK

5.5. Procesamiento de la Imagen

Se tendrá una base de datos con imágenes de tomografía con 3 vistas (Superior, Frontal y Lateral). Donde el algoritmo KDTree hará un análisis de puntos para determinar las patologías preexistentes en el paciente.

5.6. Utilizando el Algoritmo KDTree

5.7. Output: Detectando las patologías

6. Resultados

- Se logró aplicar un algoritmo computacional para la detección de patologías en imágenes DICOM
- Se determinó la presencia de patologías en imágenes DICOM mediante el uso de algoritmos computacionales y basados en la escala Hounsfield

7. Conclusiones