

# Tècniques de codificació de vídeo basades en la predicció dins l'estàndard JPEG2000

Ferran Ruiz Lopez

**Resum**—La compressió d'imatges mèdiques en escales de grisos es un tema estudiat i discutit en quan al gran volum de dades que generen i la seva necessitat de reduir-lo degut a les nombroses proves que es realitzen en l'àmbit hospitalari i clínic. Aquestes, pertanyen a un context peculiar respecte a imatges naturals on no existeix tanta similitud entre píxels veïns de tal manera que es poden extreure estimacions o prediccions i en conseqüència una explotació eficient de la redundància que donarà lloc a una compressió amb un rati més elevat. La idea principal d'aquest treball es fer recerca en la compressió d'imatges enfocant exclusivament en l'àmbit mèdic dels arxius generats per proves com **MRI** (Imatge de ressonància magnètica), **CT** (Tomografia computacional), i seqüències angiogràfiques de **rajos-X**. L'esquema de compressió utilitzat és l'estàndard JPEG2000, on s'han fet modificacions amb idees dels mètodes de predicció pròpies de la codificació de vídeo HEVC però adaptant i aplicant dintre del context en qüestió.

**Paraules clau**— DPCM, HEVC, JPEG2000, codificació intra.

**Abstract**— Compression of grayscale medical images is a studied and discussed subject about the large volume of data generated and the need to reduce it due to the numerous tests performed in the hospital and clinic. These belong to a particular context with respect to natural images where there is so much similarity between neighboring pixels so they can extract estimates or predictions and therefore an efficient exploitation of the redundancy that will lead to a higher compression ratio. The main idea of this work is to research the image compression focusing exclusively on the medical files generated by tests such as MRI (magnetic resonance imaging), CT (computer tomography), and X-ray angiographic sequences. The compression scheme used is JPEG2000 standard, where modifications have been made with the own ideas of prediction methods for video encoding HEVC but adapted and applied in this context.

**Index Terms**— DPCM, HEVC, JPEG2000, intra coding.

## 1 INTRODUCCIÓ

Avui en dia les imatges mèdiques digitals, com les Resonàncies magnètiques, tomografies computacionals o seqüències angiogràfiques, són elements clau per la pràctica assistencial tant en hospitals com en clíniques. La seva utilitat fa que actualment s'estigui generant un gran volum d'imatges mèdiques. Les quals s'han d'emmagatzemar i/o transmetre per la xarxa.

En l'actualitat el sistema d'arxius i comunicació utilitzat tant per a l'emmagatzematge, distribució i visualització d'imatges mèdiques s'anomena Picture Archiving and Communications Systems (PACS). Aquest sistema facilita l'accés i la transmissió d'aquestes dades a través de xarxes amb diverses capacitats d'ample de banda. La majoria de PACS utilitzen l'estàndard anomenat Digital Imaging and Communication in Medicine (DICOM) [1] per al intercanvi i emmagatzamament d'imatges mèdiques.

Els estàndards de compressió sense pèrdua que s'utilitzen avui dia reconeguts i compatibles amb DICOM són H.264/AVC, JPEG2000 i JPEG-LS.

Per tant, el disseny de tècniques de compressió basades en mètodes compatible amb DICOM són una bona solució per reduir l'emmagatzematge i la càrrega de transmissió d'imatges mèdiques, mentre que manté la compatibilitat amb PACS.

Per aquest motiu l'esquema de compressió utilitzat per aquest treball es JPEG2000 juntament amb un dels dos mètodes que proporciona el successor de l'estàndard H.264/AVC que és High Efficiency Video Coding (HEVC) amb el mètode intra-picture coding.

S'ha escollit aquest mètode pel rendiment que proporciona en quan a prediccions de píxels veïns basats en blocs separats independentment de la imatge.

La motivació d'aquest treball consisteix en fer recerca en la compressió d'imatges mèdiques ja que tenen una relació de compressió baixa segons imatges naturals com per exemple paisatges, cares humanes, objectes, etc on aquestes presenten unes propietats estadístiques associades a la imatge amb certes semblances dintre d'elles. Podent ser colors similars en varies regions, i certs similituds entre

• E-mail de contacte: fate.rl626@gmail.com

• Menció realitzada: *Tecnologies de la Informació*.

• Treball tutoritzat per: Joan Bartrina-Rapesta (dEIC)

• Curs 2013/14

píxels veïns de tal manera que d'elles es poden extreure estimacions o prediccions i en conseqüència una explotació eficient de la redundància que donarà lloc a una compressió amb un rati més elevat. En canvi en les imatges anatòmiques no es presenta tanta informació útil degut a les condicions en que són preses mitjançant aparells tècnics que fan ús de rajos X que incideixen en el cos per estudiar-lo. Alhora aquestes proves produeixen que en la imatge resultant s'insereixi soroll produït per les radiacions dels dispositius electrònics encarregats de prendre aquestes fotografies com podria ser un colímetre (en una tomografia computacional). A demés les proves resultants presenten una gran quantia de tonalitats en escales de grisos de fins a 2367 valors en funció de l'atenuació que es produeixi d'aquests rajos junt amb les nombroses cantonades canviant, contorns variants, i canvis bruscs presentades per la naturalesa interna del cos humà com poden ser ramificacions arterials.

Tenint en compte el context de les característiques associades a les imatges mèdiques i que els estàndards mencionats anteriorment no estan enfocats i ideats per aquestes, es busca doncs aplicar tècniques per intentar explotar correlacions de píxels que puguin haver-hi en aquestes imatges basant-se en les funcionalitats de codificador de forma d'onda anomenat Difference pulse code modulation (DPCM) de HEVC.

En aquest TFG es pretén fusionar l'esquema de compressió de l'estàndard JPEG2000 i la part intra-coding de HEVC per donar lloc a un esquema de compressió amb una etapa addicional que incorpora tècniques de predicció amb diferents modes angulars amb la finalitat de reduir el volum ocupat per les imatges comprimides generades, sense introduir cap mena de pèrdua en la imatge un cop descomprimida.

JPEG2000 és un sistema basat en transformada wavelet, que organitza els coeficients resultants de la transformada en estructures rectangular anomenades codeblocs. En el nostra cas la predicció s'aplicarà després de la transformada wavelet a nivell de codeblocs, just després de l'etapa de quantització i abans de la codificació. Per a cada codebloc s'escollirà quin es el millor mode de predicció segons diferents criteris. Els criteris utilitzats permeten la minimització de l'energia, la variància i l'entropia de cadascun dels codeblocs.

Per poder comparar la imatge original amb la imatge comprimida i descomprimida amb el TFG, s'utilitza l'aplicació Infima 1.0 [3]. Aquesta aplicació compara les imatges i mostra informació sobre les diferents mesures de distorsió. Mentre que l'aplicació que s'utilitza per visualitzar-les s'anomena ImageJ [4].

L'organització del document està distribuïda de la següent manera. En la secció 2 s'esmenaran els objectius del treball. Seguidament en la secció 3 l'estat del art on es definiran els conceptes i les idees més rellevants que formen part del context del treball. En la secció 4 el moment d'actuació del TFG junt amb les eines utilitzades i els

criteris de selecció. La metodologia emprada en la secció 5 argumentant i descrivint els procediments utilitzats i la seva adequació per obtenir els objectius proposats. En la secció 6 la discussió dels resultats obtinguts explicant i interpretant que aporten en el context. I per últim en la secció 7 l'exposició de les conclusions i les aportacions realitzades d'aquest TFG.

## 2 OBJECTIUS

Els objectius que s'exposaran a continuació contenen les tasques que s'han dut a terme pel seu desenvolupament satisfactori.

1. L'objectiu principal és la realització i funcionament de la fusió entre l'estàndard JPEG2000 i els models de predicció angulars de HEVC.
2. Aprenentatge teòric.
  - 2.1. Comprensió i reconeixement del flux d'execució que segueix l'estàndard JPEG2000 per saber el moment precís d'aplicació.
  - 2.2. Comprensió i reconeixement de les estructures de dades amb les que es treballen.
3. Desenvolupament de la programació.
  - 3.1. Preparació de l'entorn compilant les aplicacions i instal·lant el conjunt d'eines necessàries pel desenvolupament.
  - 3.2. Realització del codi en Java dels modes angulars de predicció de HEVC.
  - 3.3. Realització de diferents criteris de selecció per escollir el millor mode per cada bloc.
  - 3.4. Realització del codi en Java encarregat de construir i reconstruir els blocs segons els modes.
4. Fer proves i validacions de compressió/descompressió provant amb diferents paràmetres del TFG comprovant i anotant els resultats.
  - 4.1. Familiarització amb el funcionament i paràmetres amb els que opera BOI 3.0.
  - 4.2. Familiarització amb el funcionament i paràmetres amb els que opera l'aplicació Infima 1.0.

En les subseccions següents es plantegen les etapes i estructures de dades aportades per JPEG2000 juntament amb els modes de predicció utilitzats.

### 3 ESTAT DE L'ART

#### 3.1 JPEG2000

JPEG2000 és un estàndard de compressió que conté 3 etapes principals: transformada wavelet, quantització i codificació.

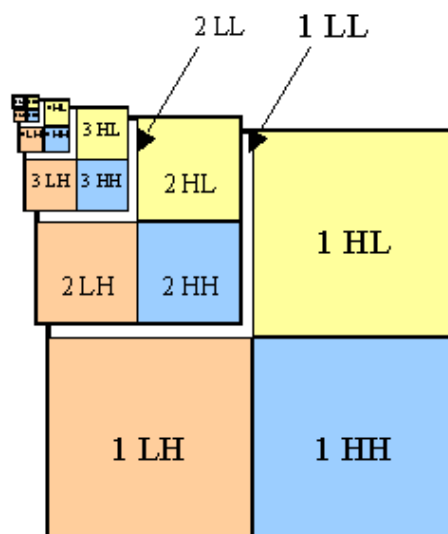
Abans de tot el procés de compressió en cas de que sigui necessari es divideix la imatge en rajoles (tiles) si les resolucions són altes i ocupen molt espai en memòria. A aquestes particions de la imatge se'ls aplicarà independentment tot l'esquema de compressió on l'algorisme de reorganització final les ajuntarà.

Per dur a terme una compressió el millor factible possible es necessita conèixer a priori el tipus d'informació que conté l'arxiu. Aquest coneixement permetrà realitzar una sèrie d'assumpcions com el tipus de símbols utilitzats, la naturalesa del document, els tipus de repeticions que ens podem trobar, etc. essencials per al compressor, ja que li permetran adaptar més acuradament els seus paràmetres per comprimir més eficientment la informació. D'aquesta manera, abans d'aplicar un tipus de codificació, es pot transformar i adaptar la informació per aprofitar al màxim les possibles redundàncies i donar al codificador un missatge més idoni per ser comprimit.

Aquest punt és on entra la transformada Wavelet. Existeixen nombroses transformades de filtres que detecten els contorns de les imatges però JPEG2000 utilitza un filtre CDF (Cohen, Daubechies i Feauveau) 9/7 per la transformada no reversible i LeGall 5/3 per la reversible [6]. Aquesta transformada actua com a filtre de freqüències de les mostres originals de la imatge de forma que separa les altes i baixes freqüències donant com a resultat una altra imatge on les energies queden concentrades en el que s'anomenen subbandes. El resultat del filtre passa-baixes s'emmagatzema a la part esquerra de la imatge on el filtre es passa en direcció vertical per extreure'n els detalls verticals, mentre que el resultat del passa-altes es troba a la dreta on els filtres es passen en horitzontal i diagonal per extreure'n els respectius detalls. D'aquesta forma, s'obté una imatge on a la seva part esquerra hi tenim tots els coeficients de freqüències baixes i a la dreta els de freqüències altes.

Sempre que es torna a fer una altre nivell de transformada, aquesta s'aplica sobre la subbanda LL on es manté la imatge original a un nivell de resolució  $\frac{1}{4}$  respecte l'original cada vegada.

En la Figura 1 es mostra com queda la organització d'aquestes subbandes.



**Fig. 1.** Subbandes HL, LH i HH on s'hi troben els detalls horitzontals, verticals i diagonals respectivament.

La quantització és el procés en el qual es preparen els coeficients provinents de l'etapa de transformada per donar-los un rang dins els enters per poder ser correctament interpretats a l'etapa de codificació. En el cas d'utilitzar el camí amb pèrdua es realitza una alineació (es pondera el coeficient segons la subbanda a la que pertany) que introdueix certa pèrdua. Per altra banda pel camí reversible aquesta etapa no es realitza cap tipus d'operació de quantització degut a que no es vol introduir pèrdua i s'utilitzen els valors enters originals sense ponderar.

Una vegada realitzades les transformacions a diferents nivells, obtenint les subbandes i realitzada la quantització, aquestes subbandes es divideixen en una estructura de blocs de mides diferents anomenats codeblocs que contenen els coeficients i es passen a la següent etapa de predicció que s'explica més endavant.

Seguidament la informació tractada per la predicció s'envia cap al codificador de plans de bits (BPE) i al codificador aritmètic MQ Coder. Aquest codificador és binari, adaptatiu i contextual. Això indica que els símbols que rep són binaris, adapta les probabilitats dels símbols segons els que ja li han anat arribant i té en compte el context dels símbols que codifica per encertar més acuradament les probabilitats i aconseguir major eficiència en la compressió.

#### 3.2 HEVC: Intra prediction modes

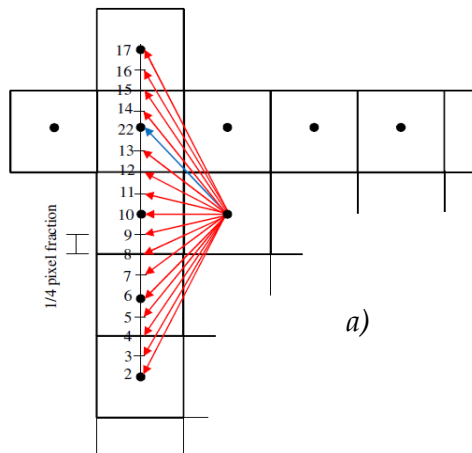
La recent introducció de l'estàndard HEVC obra noves possibilitats per la compressió sense pèrdua. S'ha demostrat que HEVC proporciona millores significatives en la eficiència de codificació de seqüències de vídeo comparat al seu predecessor H.264/AVC aconseguint reduccions de bit-rate en un rang del 50% amb una percepció de qualitat igualada [7].

HEVC presenta dos modes de codificació, inter-picture i intra-picture. Per aquest TFG s'utilitza el concepte de intra-picture coding basat en la predicció d'un píxel agafant com a referència píxels veïns del seu voltant en la mateixa imatge.

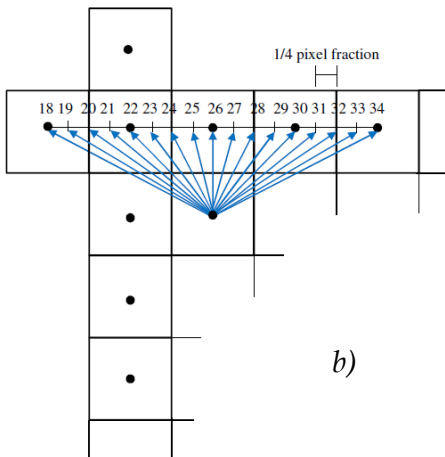
El concepte més rellevant del context d'aquest treball es que HEVC es basa en les funcionalitats de predicció DPCM dintre de intra-prediction, on la idea principal es fer una predicció d'una mostra, calcular la diferència de la mostra original respecte la predicció i codificar l'error de predicció resultant. Es a dir, al tindre un *Sample*, a aquest se li fa una predicció amb certs paràmetres de referència i s'obté una *Prediction*. Després es fa la diferència  $Sample - Prediction = Residual$  i doncs es codifica la *Residual* obtinguda. Si s'ha fet una bona predicció i aquesta s'apropa significativament al *Sample* amb valors semblants, s'obindrà una *Residual* amb valors inferiors als originals i en conseqüència s'assolirà una millor codificació del conjunt de valors predits.

Per dur a terme aquestes prediccions s'utilitzen 34 modes angulars definits en l'article [8].

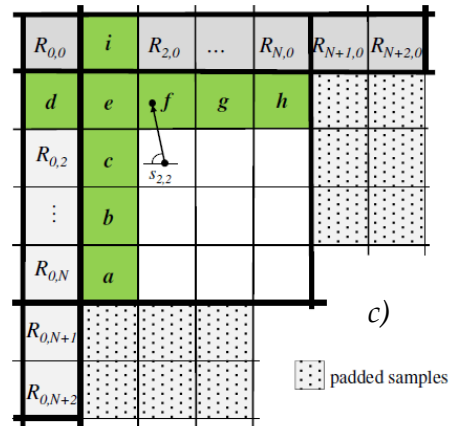
Aquests es mostren a la Taula 1 i en la següent Figura 2.



a)



b)



c)

padded samples

**Fig. 2.** Direccions de predicció associades als modes angulars DPCM, classificats com (a) verticals i (b) horitzontals. (C) Principi de Predicció dels modes angulars de DPCM proposats.

Els píxels veïns de referència (quadrats verds) que no estan disponibles per a la predicció, s'omplen amb mostres del contorn del bloc actual (padded samples).

Vertical angular modes			
Mode	Prediction operation	Mode	Prediction operation
2	$P_{x,y} = a$	10	$P_{x,y} = c$
3	$P_{x,y} = (3a + b) \gg 2$	11	$P_{x,y} = (3c + e) \gg 2$
4	$P_{x,y} = (a + b) \gg 1$	12	$P_{x,y} = (c + e) \gg 1$
5	$P_{x,y} = (a + 3b) \gg 2$	13	$P_{x,y} = (c + 3e) \gg 2$
6	$P_{x,y} = b$	14	$P_{x,y} = (3e + i) \gg 2$
7	$P_{x,y} = (3b + c) \gg 2$	15	$P_{x,y} = (e + i) \gg 1$
8	$P_{x,y} = (b + c) \gg 1$	16	$P_{x,y} = (e + 3i) \gg 1$
9	$P_{x,y} = (b + 3c) \gg 2$	17	$P_{x,y} = i$
Horizontal angular modes			
18	$P_{x,y} = d$	26	$P_{x,y} = f$
19	$P_{x,y} = (3d + e) \gg 2$	27	$P_{x,y} = (3f + g) \gg 2$
20	$P_{x,y} = (d + e) \gg 1$	28	$P_{x,y} = (f + g) \gg 1$
21	$P_{x,y} = (d + 3e) \gg 2$	29	$P_{x,y} = (f + 3g) \gg 2$
22	$P_{x,y} = e$	30	$P_{x,y} = g$
23	$P_{x,y} = (3e + f) \gg 2$	31	$P_{x,y} = (3g + h) \gg 2$
24	$P_{x,y} = (e + f) \gg 1$	32	$P_{x,y} = (g + h) \gg 1$
25	$P_{x,y} = (e + 3f) \gg 2$	33	$P_{x,y} = (g + 3h) \gg 2$
		34	$P_{x,y} = h$

**Taula 1.** Modes angulars proposats amb les seves corresponents operacions respecte els valors referència del bloc.

Aquests modes són els implementats en la nova etapa afegida a JPEG2000 i cobreixen un ampli rang de direccions amb precisió de píxel d'  $\frac{1}{4}$ . Sempre prenen com a referència la primera columna i fila del bloc a predir en qüestió, sigui de la mida que sigui, es a dir, aquets valors sempre es mantindran respecte als originals, ja que són necessaris per la predicció i la recuperació de la totalitat del bloc mentre que els demes valors canviaran segons el mode de predicció candidat a ser escollit i codificat.

Com a exemple pel cas del mode 1, aquest pren la mitjana dels seus veïns com es mostra a la següent Figura 3.

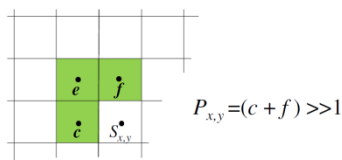


Fig. 3. Mode DC proposat i la corresponent operació de predicció. El valor S a la posició x,y és el píxel candidat a ser predit mentre que els demés valors són referències.

### 4 TÈCNiques DE PREDICCIÓ DINS JPEG2000

En aquest TFG s'afegeix la nova etapa de predicció i predicció<sup>-1</sup> que correspon al procés invers.

En aquesta etapa recau tota la importància d'aquest TFG on es produeix la predicció a nivell dels codeblocs tractats provinents de l'etapa de quantització. Com a exemple en una imatge de 64 components amb una resolució de 512x512 cadascuna, un cop aplicada la transformada wavelet s'obtenen 4480 blocs de mides variables com poden ser de 16x16, 16x32, 16x64, 32x64, 64x64. En cadascun dels blocs es corren els diferents modes i es guardarà en un fitxer auxiliar el mode que ha assolit la millor puntuació en funció del criteri de selecció escollit (mínima variància, energia o entropia). En la següent Figura 4 es mostra el moment d'aplicació.

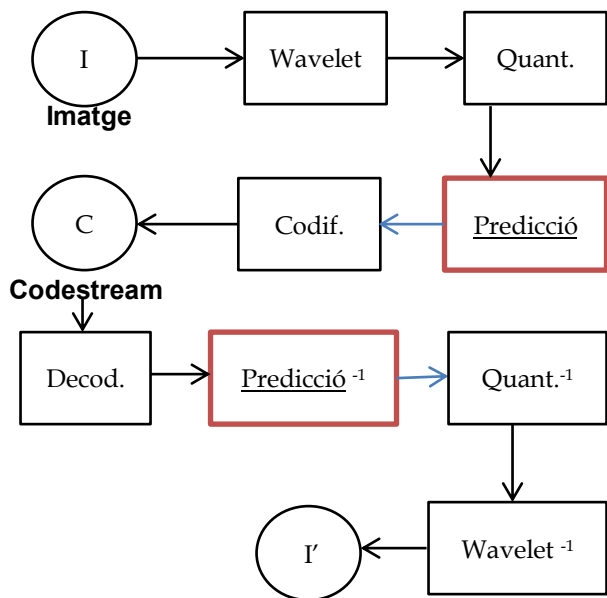


Fig. 4. Esquema general del TFG en el tractament de compressió i descompressió d'imatges. Pel cas de compressió lossless I=I' i lossy I≠I'.

#### 4.1 Criteris d'elecció

Al córrer els diferents modes per cadascun dels blocs s'ha buscat el que generava una predicció amb millor puntuació respecte els demés i respecte el bloc origina. La decisió

d'escollir-ne el més adient ha estat en funció de diversos criteris. Aquest criteris són els següents:

- Variància: Aquest criteri ha estat escollit perquè calcula les variacions que hi ha en un seguit de valors. Com més elevat és el resultat més variacions existeixen dins del conjunt de valors. Per tant es busca la mínima variància ja que així el codificador obtindrà valors més semblants. Aquest criteri s'ha integrat i aplicat a partir de (1).

$$V = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 \tag{1}$$

$x_i$  és el valor de cada píxel.  
 $\bar{X}$  és la mitja dels valors del bloc en qüestió.

- Energia: L'energia és el càlcul que aporta informació sobre la grandària d'un conjunt de valors, es a dir, com més energia tingui un bloc significa que els seus valors dels píxels són alts. Aleshores es busca la mínima energia per obtenir valors petits a ser codificats. Aquest criteri s'ha integrat i aplicat a partir de (2).

$$E = \sum_{i=1}^n (x_i)^2 \tag{2}$$

$x_i$  és el valor de cada píxel.

- Entropia: L'entropia és el càlcul de la probabilitat d'aparició d'un cert valor. Es busca doncs obtenir aquell mode de predicció que ha generat un bloc amb l'entropia mínima i que per tant conté valors més semblants i més probables de que es repeteixin. Com més gran és, més desordre existeix dins del bloc en qüestió. Aquest criteri s'ha integrat i aplicat a partir de (3).

$$H = \sum_{i=1}^n -P_i \log_2 P_i \tag{3}$$

$P_i$  és la probabilitat de cada píxel.

A continuació es presenten les eines emprades pel desenvolupament del projecte.

#### 4.2 Eines utilitzades

Al llarg d'aquest TFG els arxius, eines, i aplicacions utilitzades pel desenvolupament han estat els següents:

- Entorn de desenvolupament Eclipse.

Tot el TFG ha estat desenvolupat amb JAVA mitjançant l'IDE Eclipse Kepler ja que la implementació també està programada amb aquest llenguatge.

- El codi font de l'estàndard JPEG2000: **BOI\_3.0**.

La implementació de l'estàndard ha estat obtinguda del departament del DEIC [2] i s'ha importat al IDE per fer les modificacions pertinents. Conté l'arxiu `build.xml` per la seva compilació.

Per fer el seguit de proves de compressió/descompressió un cop compilada l'aplicació s'ha fet ús de la documentació proporcionada per la comanda `java -classpath dist\boi.jar boi.Encoder -h` on s'especifiquen un conjunt de paràmetres necessaris i d'altres opcionals per utilitzar l'aplicació.

Per exemple el paràmetre `(-c)` i seguit d'un número 0 o 1 especifica quin tipus de compressió es durà a terme, `lossy` o `lossless` respectivament. Es poden especificar els nivells de transformada wavelet `(-w)` per defecte 5 nivells. També els fils d'execució `(-tn)`, etc.

- El codi de l'aplicació per fer la comparació entre dues imatges i obtenir informació sobre diferents mesures de distorsió: **Infima 1.0**.

Aquesta aplicació obtinguda a partir de la referència [3] utilitzada al llarg del TFG ha estat necessària per observar les semblances o diferències de les imatges comprimides ja sigui utilitzant `lossy` o `lossless` respecte les seves imatges originals.

Aporta informació oferint varies mesures de distorsió però la que s'ha tingut principalment en compte ha estat el PSNR la qual indica el soroll de l'error introduït per la compressió. Aquesta mesura es expressada en dB i com més baixa és pitjor, significa que atenua la semblança entre les imatges.

- Aplicació per compilar BOI 3.0 i Infima 1.0: **Apache ant 1.9.3** [4].

Apache Ant és l'eina utilitzada per la fase de compilació i construcció de l'estàndard JPEG2000 i l'aplicació Infima que compara dues imatges. Aquest programari serveix per a processos d'automatització de compilació, similar a Make però desenvolupat en llenguatge Java i requereix la plataforma Java, així que és més apropiat per a la construcció de projectes d'aquest tipus.

Aquesta eina es basa en arxius de configuració XML i classes Java per a la realització de les diferents tasques, sent idònia com solució multi-plataforma.

- Conjunt d'eines de desenvolupament per executar i programar: **JDK 1.7**.

Aquestes eines han proporcionat l'entorn per executar i utilitzar les diferents estructures de dades pel desenvolupament.

- Aplicació per visualitzar imatges en format RAW: **ImageJ 1.47** [5].

ImageJ és un programa de processament d'imatge digital

de domini públic programat en Java desenvolupat al National Institutes of Health. Permet la visualització i processat d'imatges radiològiques entre d'altres i suporta varis formats inclòs RAW, format de les imatges amb les que es treballa en aquest TFG.

- Varies imatges en format RAW per fer proves.

Algunes d'elles són imatges no mèdiques i d'altres si que ho són, com tomografies computacionals, angiografies...

Totes elles en escales de grisos a diferents resolucions paràmetres i components.

Seguidament es descriu la metodologia utilitzada per assolir els objectius plantejats.

## 5 METODOLOGIA

Les tasques i procediments a realitzar per obtenir l'objectiu principal ha recaigut en el desenvolupament de la programació en JAVA junt amb les eines proporcionades. S'ha dut a terme a través del IDE Eclipse amb el que s'ha creat un *Package* anomenat **TFG Ferran**. Aquest s'ha afegit juntament als 9 *Packages* que ja conté el codi font de **BOI 3.0** de l'estàndard JPEG2000.

El *Package* creat consta de dues classes anomenades **Prediction** i **PredictionModes**.

Primerament s'explica la classe **Prediction** anomenada la classe controladora encarregada de comunicar i gestionar tota la informació entre el codi de BOI 3.0 i la classe **PredictionModes** a la que se l'anomena la classe treballadora, ja que aquesta realitza els càlculs i la feina tediosa amb els seus respectius mètodes.

La tasca de la classe **Prediction** és de rebre un per un els milers de codeblocs generats per la divisió de la transformada wavelet junt amb la seva posició dintre de la imatge. Per exemple una imatge de 224 components amb una resolució de 512x512 cadascuna, un cop aplicada la transformada wavelet a 5 nivells s'obtenen més de 14.000 codeblocs de mides variables com poden ser de 16x16, 16x32, 16x64, 32x64, 64x64.

La posició del bloc dintre de la imatge és una dada de vital importància ja que es guarda en un fitxer auxiliar anomenat `modes.txt` el mode que hagi estat escollit per aquella posició de bloc en concret i serà necessari per la descompressió. Per exemple si el bloc número 626 ha obtingut el mode 8 com a millor candidat de predicció segons el criteri escollit per aquella execució, el fitxer auxiliar contindrà en la posició 626 el valor número 8 i per tant la reconstrucció serà la correcta.

Arribats a aquest punt amb la informació anterior del bloc amb els seus respectius coeficients i el índex del mateix, la part més important encarregada de gestionar el comportament i execució dels passos a seguir per realitzar el treball en la part de la compressió, recau en el mètode **Predict()** de la classe **Prediction**.

És l'encarregat de comunicar a la classe **PredictionModes** quin dels 34 modes s'avalua i del qual es volen els càlculs. **PredictionModes** conté el mètode *runMode* amb els 34 intra modes *hardCoded* de HEVC proposats.

Aquest s'encarrega de fer el càlcul del mode que s'estigui executant en qüestió i retorna el resultat. Aquest mètode es basa en la matriu original ja que la predicció es sobre aquesta i té en compte els *Padded samples* (en certs modes hi ha posicions de la matriu de predicció que intenten agafar valors fora de rang, doncs fa un *padding* de valors propers).

Un cop obtingut el bloc o matriu predita es passa a calcular la matriu residual on aquesta pot ser candidata a ser codificada finalment. S'obté amb la resta de la matriu original respecte la predita amb el mètode *getResidual* de la classe **PredictionModes**.

La residual és el que es vol obtenir enlloc de la original per ser codificada i és on s'aplica el criteri de selecció (mínima variància, mínima energia o mínima entropia).

Un cop predit el bloc amb tots els modes s'avalua quin ha estat el millor i es guarda la informació necessària per la recuperació i posterior descompressió en el fitxer *modes.txt*. Si en cap cas s'ha complert el criteri de selecció respecte el bloc original podent ser perquè cap mode l'ha satisfet, aquest bloc no s'altera i retorna l'original.

En la fase de la descompressió el mètode encarregat de reconstruir els codeblocs s'anomena **Reconstruct()**. Rep per paràmetres el bloc en qüestió i el mode amb el que va ser predit en cas de que ho fos. Aquest li encarrega fer la feina a la classe **PredictionModes** (mètode *getRebuild*) i genera el bloc reconstruït segons si era un mode vertical o horitzontal (2-17 verticals, 18-34 horitzontals). Aquest punt es important a tindre en compte ja que depenent de quin mode sigui s'ha de basar en posicions del bloc ja reconstruïdes com si fos la original, sinó utilitzaria posicions que encara no han estat generades i utilitzaria valors incorrectes donant lloc a un bloc totalment diferent amb l'original. Per tant si el mode utilitzat per predir era vertical, també s'ha de reconstruir de manera vertical.

La següent tasca a realitzar té com a objectiu la fusió de BOI 3.0 junt amb el treball realitzat en aquest TFG. Aquesta unió té lloc en el Package *Operations* dins de la classe **JP2QUANTIZATION.java** on es produeix l'etapa de quantització dins del esquema que segueix l'estàndard JPEG2000. En aquesta s'ha declarat la classe principal aportada per aquest TFG (**Prediction**).

**JP2QUANTIZATION** conté dos constructors diferents, un associat a la compressió i l'altre a la descompressió. Per la fase de compressió s'ha iniciat l'estructura de dades on es guardaran els modes *Prediction.setAmountOfCodeblocks()*.

D'altra banda per la fase de descompressió s'han carregat els modes guardats del fitxer extern en la primera fase *Prediction.loadModes()*.

Els moments d'actuació de la primera fase ha estat tot just després de l'etapa de quantització en els mètodes

**Forward** i **ForwardTile**. Aquests dos mètodes realitzen la mateixa funció amb la diferència de que el segon mètode és per imatges amb resolucions grans les quals ocupen gran quantitat d'espai en memòria i es divideixen en rajoles (tiles).

Pel cas de la fase de descompressió el moment d'actuació és abans de la quantització<sup>-1</sup> en els mètodes **Reverse** i **ReverseTile** (segueix la mateixa lògica de divisió en rajoles).

El motiu pel qual l'aplicació del TFG ha estat després de l'etapa de quantització és degut a que aquest és el darrer procés el qual es preparen els coeficients provinents de l'etapa de transformada per donar-los a l'etapa de codificació on actua el BPE i el codificador aritmètic binari MQ. De la mateixa manera pel cas de la descompressió es produeix després de la descodificació i abans de la quantització<sup>-1</sup>.

La lògica d'aquest flux segueix la funcionalitat del DPCM de HEVC *Sample - Prediction = Residual* en la compressió i el procés invers en la descompressió *Sample = Residual + Prediction*.

A continuació s'explica un exemple:

Al tenir un determinat valor, per exemple el 10 i aplicada una divisió (pel cas de quantitzar)  $10/2=5$  i una altra operació aritmètica per exemple una resta  $5-4=1$  (pel cas de l'actuació del TFG al predir) es codificaria el número 1. Doncs l'aplicació és capaç de tornar al valor original 10. Per tant en la fase de descompressió primer faria una suma  $1+4=5$  (fer la reconstrucció) i quantització<sup>-1</sup>  $5*2=10$ .

Finalment l'objectiu de fusió d'integrar els modes proposats basat en DPCM de HEVC dins de JPEG2000 ha estat satisfactori ja que en la descompressió lossless s'obté el mateix resultat d'imatges aplicant el TFG que sense fer-ho on per verificar-ho s'ha fet ús de l'aplicació Infima obtenint resultats de màxima semblança en imatges. Per tant és capaç de comprimir i descomprimir amb les funcionalitats addicionals afegides dins de l'estàndard. El mateix succeeix pel cas de la compressió lossy on el PSNR és semblant entre la pèrdua que introdueix l'estàndard i la que introdueix el projecte.

Les últimes tasques realitzades i explicades a continuació han estat la realització de proves amb el compressor/descompressor i les verificacions per observar el funcionament efectiu del projecte.

La metodologia seguida per fer proves amb diverses imatges han estat en la línia de comandes de Windows 8 on s'han creat varis *scripts* per les crides de compressió/descompressió i comparació corresponent a cada imatge amb Infima.

Cadascuna conté en el nom d'arxiu certs paràmetres. En la Figura 5 es mostra un exemple d'una prova mèdica amb la que s'ha treballat.



**Fig. 5.** Prova mèdica anomenada **angiografia** on s'insereix un catèter al cor per fer un examen diagnòstic dels vasos circulatoris.

L'arxiu anterior s'anomena `xray.64_512_512_3_1_12_1_0_0.raw`.

Cada paràmetre s'explica en la documentació de BOI 3.0 inserint la comanda `java -classpath dist\boi.jar boi.Encoder -h`.

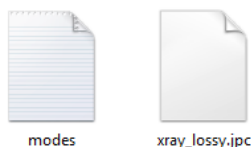
On cadascun té el següent significat:

- **64:** Nombre d'imatges (components).
- **512:** Alçada de la imatge.
- **512:** Amplada.
- **3:** Sample type.
- **1:** Signed type.
- **12:** Bit depth.
- **1:** Byte order.
- **0:** Data order.
- **0:** Component RGB.

Per realitzar les proves de compressió les comandes segueixen l'estil següent:

```
java -classpath dist\boi.jar boi.Encoder -i
..\..\IMATGES\xray.64_512_512_3_1_12_1_0_0.raw -o
xray_lossy.jpc (-c 0) (-w 0).
```

On el paràmetre `-i` correspon a la ruta de la imatge a comprimir, `-o` el nom de l'arxiu un cop comprimit, `-c 0` compressió lossy (per lossless és `-c 1`) i per últim `-w 0` és un paràmetre opcional utilitzat en algunes proves que especifica els nivells de transformada wavelet a aplicar. Una vegada executada l'anterior comanda els arxius resultants es mostren en la següent Figura 6.



**Fig. 6.** (D'esquerra a dreta). L'arxiu de text necessari per la recuperació dels blocs que han estat predits i l'arxiu resultant de la compressió en format JPC.

Es important comentar que totes les compressions generen el document `.txt` on ha de ser contemplat en la mida final del arxiu comprimit.

Aquest document de text s'anomena *side information* podent ser com a molt 1002 bytes una vegada comprimit en ZIP.

Seguidament per realitzar les proves de descompressió les comandes segueixen l'estil següent:

```
java -classpath dist\boi.jar boi.Decoder -i xray_lossy.jpc -o
xraydecodedlossy.64_512_512_3_1_12_1_0_0.raw
```

El paràmetre `-i` correspon al nom de l'arxiu d'entrada i el `-o` al nom de l'arxiu de sortida.

En la part de la recuperació dels blocs l'aplicació carrega el fitxer que conté els modes. En la següent Figura 7 es mostra una petita part de blocs que fan la recuperació de la predicció.

```
CB: 2732 28
CB: 2730 8
CB: 2803 7
CB: 2800 33
CB: 2872 24
CB: 2873 8
CB: 2870 27
CB: 2942 4
CB: 2943 29
CB: 2940 4
CB: 3010 28
CB: 3002 8
CB: 3003 8
CB: 3000 7
```

**Fig. 7.** (D'esquerra a dreta) Exemple d'alguns codeblocs recuperats junt amb el seu índex dins de la imatge i el mode de predicció que li correspon.

A continuació la metodologia seguida per comprovar la validesa d'aquest projecte ha estat realitzant proves generades també amb *scripts* en la consola de Windows amb l'aplicació Infima 1.0.

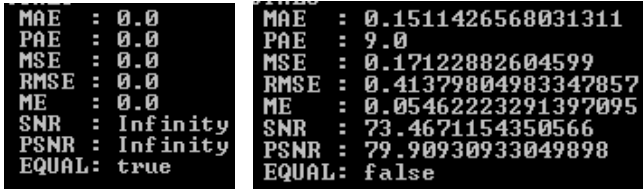
Les proves de validació segueixen l'estil de la següent comanda on es comparen les semblances i diferències amb varies mesures de distorsió entre la imatge original sense comprimir i la imatge generada per la compressió i posterior descompressió aplicant el projecte.

```
java -classpath dist\infima.jar infima.Infima -i
..\..\IMATGES\xray.64_512_512_3_1_12_1_0_0.raw ..\boi-
3_0\boi\xraydecodedlossless.64_512_512_3_1_12_1_0_0.raw
```

El paràmetre `-i` correspon a la ruta de la primera imatge a ser comparada amb la segona que no té paràmetre, només un espai respecte la ruta del primer arxiu.

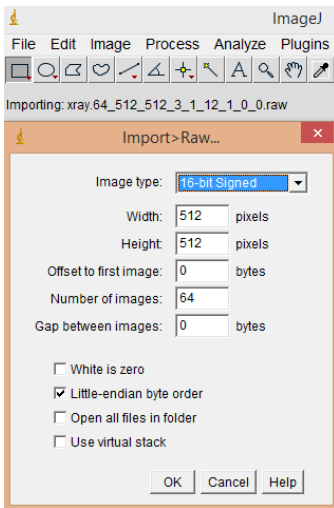
En la següent Figura 8 s'observen els resultats d'una comparació.





**Fig. 8.** (D'esquerra a dreta) Infima indica que no s'ha introduït cap error en imatges comprimides sense pèrdua. Infima indica que existeix error ja que la compressió s'ha fet lossy i per tant les imatges no son iguals.

Per últim la metodologia seguida per la visualitzar les imatges ha estat amb l'aplicació ImageJ. En la següent Figura 9 es mostra el programa.



**Fig. 9.** Interfície de l'aplicació ImageJ per a la visualització de les imatges en format RAW especificant els paràmetres que es mostren.

Per finalitzar en el següent apartat es mostra l'anàlisi dels resultats.

## 6 DISCUSSIÓ DELS RESULTATS

Les primeres proves s'han realitzat amb diferents tipus d'imatges mèdiques i no mèdiques en escales de grisos i resolucions ja que al ser un treball de recerca no es sap a priori a quines pot afavorir i quines no i per tant no s'han fet previsions de millores fins veure i contrastar els resultats.

La següent Taula 2 mostra els resultats.

	LOSSLESS	JPEG 2000	JPEG 2000 + TFG 5 WV	JPEG 2000 + TFG 0 WV
		Mida (bytes)	Mida (bytes)	Mida (bytes)
xray_64_512_512		14.523.218	14.522.286	14.898.832
n1_GRAY.1_2560_2048		2.871.928	2.872.920	3.051.719
n2_GRAY.1_2560_2048		3.459.389	3.463.661	3.460.463
bcn.1_4800_7200		24.196.784	24.202.500	24.913.956
cuprite_radiance.224_512_512		53.711.170	53.723.796	51.212.745
angio_1.151_512_512		31.535.658	31.535.501	32.648.448
ct_1.596_512_512		121.211.089	121.326.726	122.720.220

**Taula 2.** Compressió sense pèrdua realitzades amb JPEG2000 sense modificar i 5 nivells de WV i JPEG2000 junt amb el TFG amb 5 nivells i 0 nivells.

Els resultats obtinguts de compressió sense pèrdua han estat variables i la majoria negligibles pel que no es pot concretar que es satisfaci la compressió de tots els casos. En canvi s'observa un cas amb una millora significativa en la imatge cuprite radiance (224 components, resolució 512x512) on el rati de compressió ha estat del 57% amb una reducció de més de 2MB.

El criteri de selecció del mode de predicció a escollir ha estat en tots els casos la mínima entropia ja que d'aquesta manera afavoreix al codificador perquè obtindrà blocs en els que els valors seran més semblants i es podran codificar més eficientment.

Altres criteris com l'energia teòricament funcionen ja que per nombrosos blocs l'energia baixa però alhora de comprimir ocupen més que sense aplicar res. Això pot ser degut a que els blocs originals contenen valors més grans però més semblants i per contra els blocs predits contenen valors més petits però diferents. I així l'energia baixa però pel codificador es contraproductiu.

El criteri de la variància afecta més negativament que els demés podent ser per la variació que existeix de valors junt amb que tenen un valor alt i per tant s'ha de representar amb més bits per mostra.

Les proves realitzades són amb 5 i 0 nivells de transformada wavelet per comprovar si la transformada afavoreix o repercuteix en la predicció. S'observa que en gairebé tots els casos no s'afavoreix a la predicció que realitzen els modes. L'explicació d'aquest motiu té relació amb que aquesta transformació de les dades de la imatge s'adequa en cada subbanda en detalls horitzontals (HL), verticals (LH) i diagonals (HH) de manera eficient al llarg dels diferents nivells que s'apliquen. Es a dir, aquesta informació ja ha estat tractada i processada per ser codificada de manera eficient, en canvi al afegir la nova etapa de predicció, aquestes dades es maltracten i s'espalla la feina realitzada per la transformada fent que la compressió sigui contraproductiu.

Per aquest motiu i per la naturalesa de la imatge que ha obtingut bons resultats comentada anteriorment, la compressió millora amb 0 nivells de transformada.

Aquests resultats es contrasten en la següent Taula 3 amb imatges únicament de tipus radiance i 0 nivells de Wavelet on s'obtenen reduccions en la mida de l'arxiu comprimit de fins a 8MB.

LOSSLESS	JPEG 2000 0 WV	JPEG 2000 5 WV	JPEG 2000 + TFG 0 WV	Data saving
	bps	bps	bps	Mida(MB)
jasperidge_radiance.224_512_614_2_0_16_0_0_0	8.46	8.22	7.70	<u>4.35</u>
lowaltitude_radiance.224_512_512_2_0_16_0_0_0	8.43	8.90	7.72	<u>8.24</u>
lunarlake_radiance.224_512_512_2_0_16_0_0_0	8.26	7.03	6.80	<u>1.59</u>
moffetfield_radiance.224_512_512_2_0_16_0_0_0	8.67	8.14	7.71	<u>3.06</u>
yellowstone_radiance.224_512_512_2_0_16_0_0_0	7.87	7.60	7.30	<u>2.16</u>

**Taula 3.** Compressió sense pèrdua realitzades amb JPEG2000 sense modificar amb 0 i 5 nivells de wavelet i JPEG2000 junt amb el TFG amb 0 nivells de WV en imatges de tipus radiance. Mostrat en bits per sample=mida del fitxer comprimit (en bits) / n° de píxels totals.

En tots els casos s'ha comprovat i verificat a través de l'aplicació Infima, la correctesa i validesa de l'objectiu principal d'aquest treball, es a dir obtenir un retorn sense pèrdua afegida per la nova etapa en respecte la imatge original.

Finalment en la compressió amb pèrdua s'ha optat per realitzar proves de compressió en diverses imatges únicament a la subbanda on es troba la imatge amb una resolució reduïda (LL). D'aquesta manera s'introdueix menys error i el PSNR baixa infimament.

A la Taula 4 es mostren els resultats.

LOSSY	JPEG 2000		JPEG 2000 + TFG	
	Mida (bytes)	PSNR (dB)	Mida (bytes)	PSNR (dB)
xray_64_512_512	14.466.370	82.149	14.464.252	80.596
n1_GRAY.1_2560_2048	2.665.307	55.337	2.664.810	55.105
n2_GRAY.1_2560_2048	3.320.202	56.053	3.320.126	54.304
bcn.1_4800_7200	23.464.071	56.848	23.463.990	56.206
cuprite_radiance.224_512_512	54.005.733	105.771	54.000.242	105.411
angio_1.151_512_512	31.278.198	81.821	31.276.356	81.123
ct_1.596_512_512	118.934.471	105.863	118.930.746	105.260

**Taula 4.** Compressió amb pèrdua realitzades amb JPEG2000 sense modificar i JPEG2000 junt amb el TFG en diversos tipus d'imatges amb el respectiu PSNR.

En tots els casos la mida final del arxíu s'ha reduït escassament fins un màxim de 5KB en algunes imatges.

El PSNR s'ha mantingut aproximadament al voltant de la pròpia pèrdua introduïda per la quantització en la compressió lossy amb diferències inapreciables per l'ull humà.

Per finalitzar aquest article es dona peu a la següent secció final on es presenten les conclusions extretes d'aquest treball.

## 7 CONCLUSIONS

La finalitat en el context d'aquest treball es principalment la compressió sense pèrdua. On s'ha perseguit al llarg de tot el treball que es realitzessin els processos correctament de compressió i posterior descompressió amb la recuperació original de la imatge sense introduir pèrdua per part

del projecte. Per tant l'objectiu principal ha estat assolit satisfactòriament amb la inclusió de modes de predicció basats en intra coding de HEVC dins de l'estàndard JPEG2000.

Els resultats aportats no han estat beneficiosos per a tots els casos, excepte per un tipus d'imatges on s'ha aconseguit reduir notablement l'arxíu resultant de la compressió amb l'ordre d'entre 2 i 8 MB.

La naturalesa d'aquestes imatges són en aplicacions de teledetecció anomenades airborne visible/infrared imaging spectrometer (AVIRIS).

Pel cas de la compressió lossy també s'ha aconseguit reduir l'espai de l'arxíu resultant però escassament, degut a que en aquesta part, el projecte si que introdueix una pèrdua en conseqüència de les divisions dels coeficients fetes per la quantització i doncs el treball aplicat es troba únicament en les subbandes LL.

La conclusió més important que es pot extreure d'aquest treball es que els modes implementats s'han de modificar i d'adaptar més acuradament al context d'imatges mèdiques i del esquema de transformació de la informació que segueix JPEG2000.

Aquest estudi de recerca ha servit per donar peu a conclusions i reflexions posteriors que es poden fer en quan a la millora o modificacions al respecte i arribar a aconseguir resultats més significatius.

Per exemple s'ha extret la conclusió de que la transformada wavelet no afavoreix a la predicció de la redundància espacial ja que aquesta ja l'explota en les diferents subbandes separant els detalls verticals, horitzontals i diagonals. Per aquest motiu una possible extensió futura en la recerca de compressió d'aquest tipus d'imatges es adaptar nous modes pensats únicament per les diferents subbandes on es troba realment la informació tractada i extreta per la transformada.

## AGRAÏMENTS

Agraeixo el suport i dedicació empleat per part del tutor junt amb el coneixement que m'ha aportat i el bon tracte. També a persones properes i als meus pares i germà.

## BIBLIOGRAFIA

- [1] Article. V. Sanchez, F. Aulí-Llinàs, J. Bartrina-Rapesta, J. Serra-Sagrístà. "HEVC intra coding based on DPCM and edge prediction for lossless Medical image compression", 2012.
- [2] Aplicació BOI\_3.0, <http://www.deic.uab.cat/~francesc/>.
- [3] Aplicació Infima\_1.0, <http://www.deic.uab.cat/~francesc/>.
- [4] Aplicació Apache Ant 1.9.3, <http://ant.apache.org/>.
- [5] Aplicació ImageJ\_1.47, <http://imagej.nih.gov/ij/>.
- [6] Tesina. Francesc Aulí-Llinàs, "L'ESTÀNDARD DE COMPRESSIÓ D'IMATGES JPEG2000", 2004.
- [7] Article. G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," December 2012.
- [8] Article. V. Sanchez, J. Bartrina-Rapesta, "Lossless compression of medical images based on HEVC intra coding", 2013.