MedZip

3D Medical Images Lossless Compressor Using Recurrent Neural Network (LSTM)

Omniah H. Nagoor, 885020@swansea.ac.uk
Joss Whittle, joss.whittle@rfi.ac.uk
Jingjing Deng, j.deng@swansea.ac.uk
Benjamin Mora, b.mora@swansea.ac.uk
Mark W. Jones, m.w.jones@swansea.ac.uk

Department of Computer Science, Swansea University, UK

Abstract

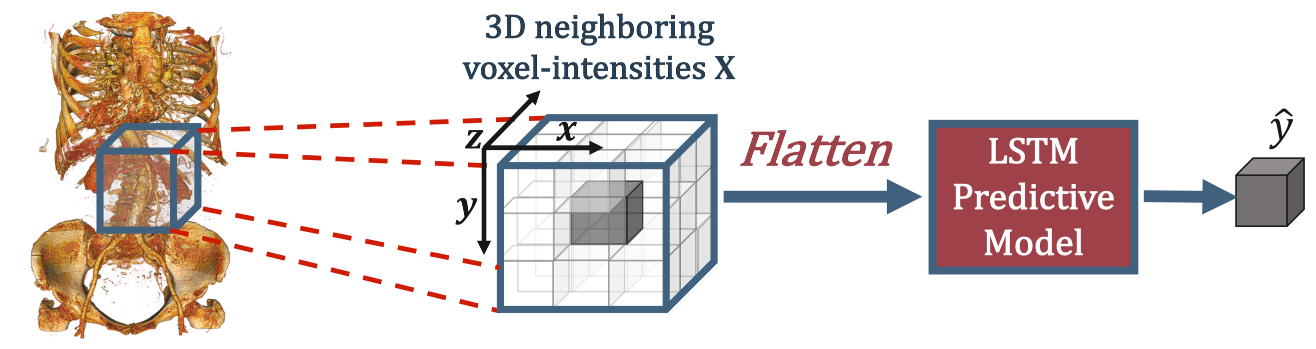
As scanners produce higher-resolution and more densely sampled images, this raises the challenge of data storage, transmission and communication within healthcare systems. Since the quality of medical images plays a crucial role in diagnosis accuracy, medical imaging compression techniques are desired to reduce scan bitrate while guaranteeing lossless reconstruction. This paper presents a lossless compression method that integrates a Recurrent Neural Network (RNN) as a 3D sequence prediction model. The aim is to learn the long dependencies of the voxel's neighbourhood in 3D using Long Short-Term Memory (LSTM) network then compress the residual error using arithmetic coding. Experiential results reveal that our method obtains a higher compression ratio achieving 15% saving compared to the state-of-the-art lossless compression standards, including JPEG-LS, JPEG2000, JP3D, HEVC, and PPMd. Our evaluation demonstrates that the proposed method generalizes well to unseen modalities CT and MRI for the lossless compression scheme. To the best of our knowledge, this is the first lossless compression method that uses LSTM neural network for 16-bit volumetric medical image compression.

A 3D volume visualization of CT scans for a patient's entire trunk (Dataset1).

Motivation

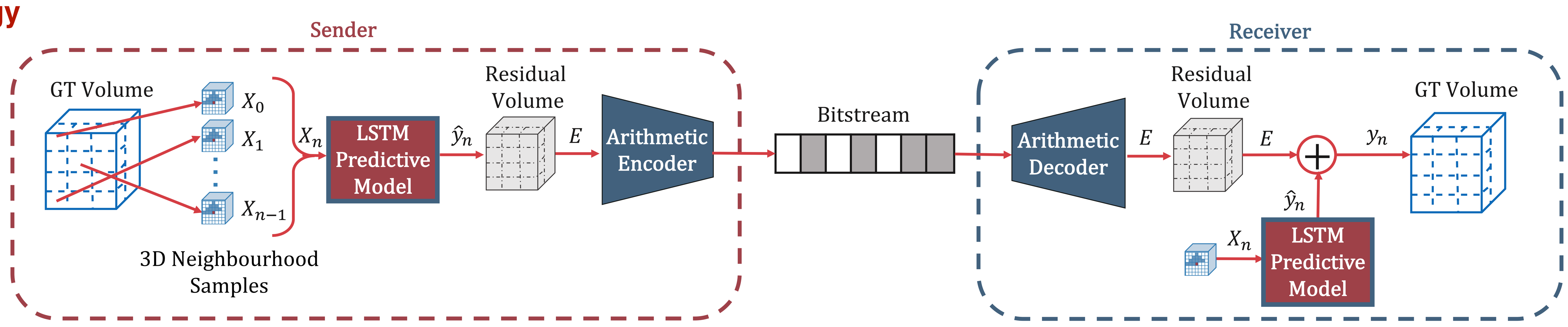
- **Medical images** contain a large amount of valuable data, which also consumes a **vast** amount of storage.
- Radiologists use these high quality and high resolution scans for **clinical** purposes, including **diagnosis** or **precise** pre-surgery planning. Therefore, keeping these scans' **quality** and **accuracy** for **accurate diagnosis** while **reducing** storage size form a significant challenge.
- The classical (non-learned) codecs may have **limited** ability in representing **non-linear** correlations or **high-dimensional** data distribution. This critical limitation rises the demand for new compression approaches with higher **flexibility** and **generalizability** in representing nonlinearity.
- Recently, the **state-of-the-art** deep neural networks models demonstrate great **potential** in representing high-dimensional data distribution for both lossy and **lossless** compression performance. Moreover, a **higher** compression ratio can be **achieved** using deep learning methods compared to traditional linear methods.

Overview



- As the LSTM model is one of the state-of-the-art sequence models, we formulated our proposed lossless compression approach as a supervised **many-to-one sequence prediction problem** and integrates the LSTM model as **3D sequence predictor** model.
- Our LSTM model takes a sequence of **3D neighbouring voxels** $\mathbf{X}$ as **input** and predicts the **next intensity value** $\hat{y}$.

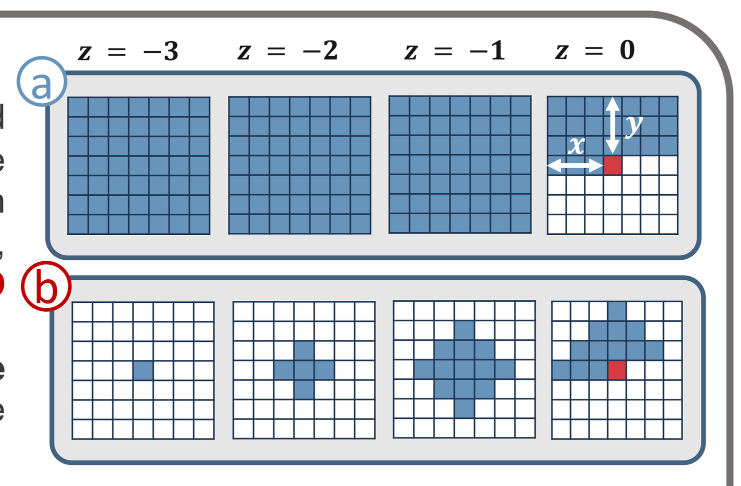
Methodology



An overview of our proposed lossless compression framework using LSTM

Local Sampling

- **Two different** 3D neighbourhood shapes **were applied to** find the input sequence that can **lead to** an **optimal compression,** namely, the (a) **3D cube** and (b) **3D pyramid** neighbourhood sequence.
- Each type introduces a **diverse coverage** of the **block around** the **target voxel**.
- The 3D pyramid sequence with (13x13, 9x9, 5x5, 1) sequence size was used as input for all the proposed models.



In both types, z=0 represents the **current slice**, the **target** voxel to be predicted is **red**, **blue** voxels are **used as input sequences** while the **white** voxels are **masked**.

Model and Training Hyper-Parameters

- The proposed **predictive** models are **Vanilla LSTM** models, which are composed of the input layer, LSTM layer with 128 cells, and a linear output layer.

Model ID	Sampling Space	Slice Thickness	Hyper Parameters
MedZip1	Random samples from volumes with pixel spacing .488	.625	Batch size=128, & learning rate=5e-5
MedZip2	Random samples from volumes with pixel spacing .625	.625	Batch size=128, & learning rate=5e-5
MedZip3	Random samples from volumes with pixel spacing .488, .578, .625	.625	Batch size=128, & learning rate=1e-4

Loss Function

- We minimize a **joint loss** function which is the sum of **Mean Absolute Error (MAE)** and the Pearson **Correlation Coefficient (PCC).**

$$L_{Joint} = MAE + \lambda(1 - |PCC|)$$

$$\frac{\sum_{i=1}^{n}|y-\hat{y}|}{n} \qquad \frac{cov(y,\hat{y})}{\sigma_y \sigma_{\hat{y}}}$$

Experimental Results

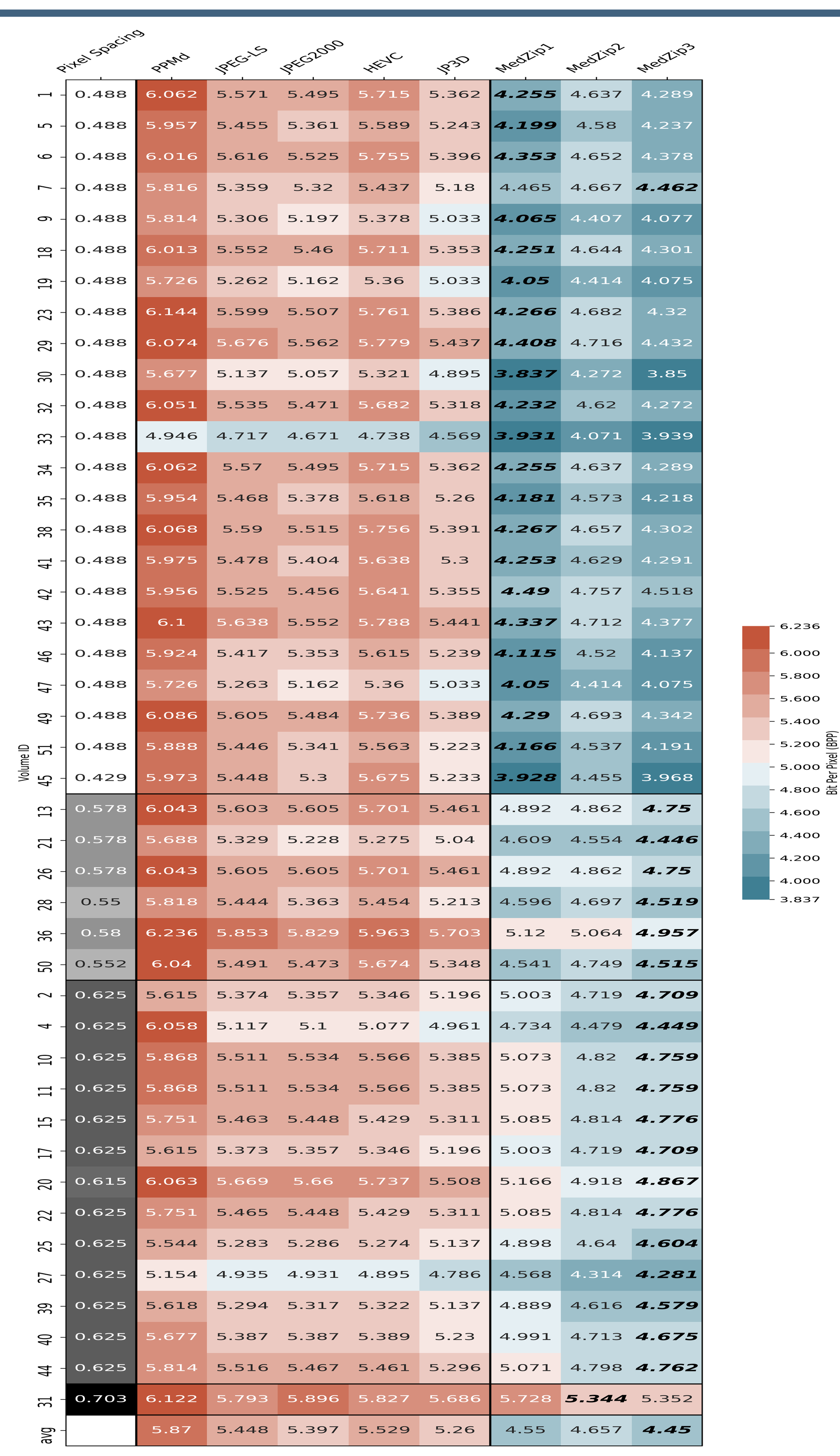
- We evaluated the compression performance in **bits-per-pixel (bpp)** of the three proposed models in comparison to the **state-of-the-art lossless** compression methods including, some well know **image** and **volumetric** codecs.
- The evaluation was conducted on two test sets:
  - Testset1 (42 volumes) – CT scans.
  - Testset2 (12 volumes) – MRI scans.

Local Sampling

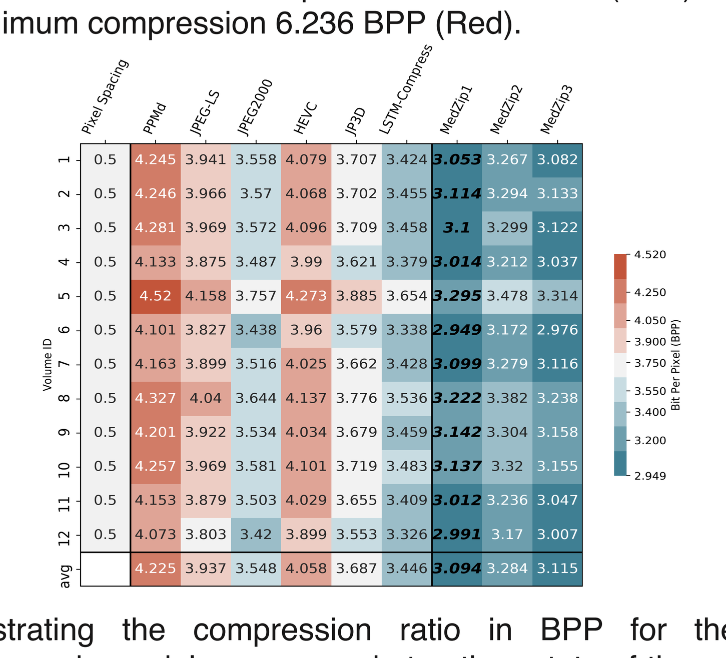
- Experimentally, **different lengths** to the **target** voxel were applied to select the 3D neighbouring size with the **best** compression performance.
- As expected, with the **increase** in the **3D cube** block size, the compression **rate** also **increases** as well as the compression **time** due to the longer sequence length.
- However, the **3D pyramid neighbourhood** demonstrates a great **balance** between the compression **time** and overall **compression achievement**. Compared to using a **full cube block**, there was **no performance loss** in terms of the size of compressed file and the training time was substantially reduced because **fewer samples** were used.

	3D Pyramid Neighbouring Sequence	3D Cube Neighbouring Sequence		
Neighbourhood Block Size	(13x13,9x9, 5x5,1x1)	(5x5x5)	(7x7x7)	(9x9x9)
Bits-Per-Pixel (BPP)	4.267	4.702	4.478	4.36
Compression Time (hh:mm:ss)	1:23:58	0:44:51	1:17:13	2:27:47

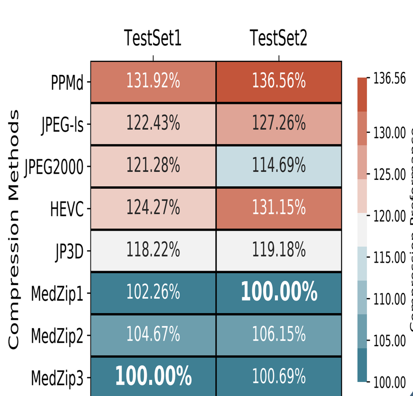
Comparing the compression **performance** (compression **ratio** (BPP) and compression **time**) of different neighboring sequence (3D pyramid & 3D cube) with different block sizes.



Illustrating the compression ratio in bits-per-pixel (BPP) for each lossless compression method on **TestSet1**. The first column is colour mapped by the pixel spacing value of each volume. The other cells are highlighted from the maximum compression 3.837 BPP (Blue) to minimum compression 6.236 BPP (Red).



Illustrating the compression ratio in BPP for the proposed models compared to the state-of-the-art lossless compression methods on **TestSet2** (16-bits volumes). The first column is colour mapped by the pixel spacing value of each volume. The other cells are highlighted from the maximum compression 2.949 BPP (Blue) to minimum compression 4.52 BPP (Red).



A summary overview of the compression performance over the **two test sets** for all the lossless methods. Cells are coloured from the best compression performance 100.00% (Blue) to the worst performance 136.56% (Red). (Less value indicates better performance).

Conclusion

- MedZip is a **novel** lossless compression approach using LSTM, specifically for compressing **3D medical images** (16 bit-depths).
- MedZip empirically demonstrates a **higher compression ratio** achieving **15% saving** compared to the state-of-the-art lossless compression standards, including JPEG-LS, JPEG2000, JP3D, HEVC, & PPMd.
- Our pre-trained LSTM models **generalized well** to unseen modality (MRI) and achieves a **higher** compression ratio compared to the other methods.

Future Work

- We believe that the proposed models would achieve more improvement by **integrating** it with **attention-based** mechanisms.

Swansea University Prifysgol Abertawe
ICPR 2020 25th International Conference on Pattern Recognition Milan, Italy 10 | 15 January 2021