MedZip

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

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



- ♦ 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** X as **input** and predicts the **next** intensity value \hat{y} .



е															ļ
-															

L	meazipi	pixel spacing .488	.020	learning rate
	MedZip2	Random samples from volumes with pixel spacing .625	.625	Batch size= ⁻ learning rate
	MedZip3	Random samples from volumes with pixel spacing .488, .578, .625	.625	Batch size= ⁻ learning rate

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			

		spacing	>	5	-00	-					
	Q	ter st	88Mg	PEGIL	PEGLE	HENC	830	Wegtri	Wedlin	Medili	
		0.488	6.062	5.571	5.495	5.715	5.362	4.255	4.637	4.289	
	- <u>م</u>	0.488	5.957	5.455	5.361	5.589	5.243	4.199	4.58	4.237	
	9 -	0.488	6.016	5.616	5.525	5.755	5.396	4.353	4.652	4.378	
		0.488	5.816	5.359	5.32	5.437	5.18	4.465	4.667	4.462	
	ი -	0.488	5.814	5.306	5.197	5.378	5.033	4.065	4.407	4.077	
	- 18	0.488	6.013	5.552	5.46	5.711	5.353	4.251	4.644	4.301	
	19	0.488	5.726	5.262	5.162	5.36	5.033	4.05	4.414	4.075	
	- 23	0.488	6.144	5.599	5.507	5.761	5.386	4.266	4.682	4.32	
	29	0.488	6.074	5.676	5.562	5.779	5.437	4.408	4.716	4.432	
	90	0.488	5.677	5.137	5.057	5.321	4.895	3.837	4.272	3.85	
	32	0.488	6.051	5.535	5.471	5.682	5.318	4.232	4.62	4.272	
	- 33	0.488	4.946	4.717	4.671	4.738	4.569	3.931	4.071	3.939	
	34	0.488	6.062	5.57	5.495	5.715	5.362	4.255	4.637	4.289	
	35	0.488	5.954	5.468	5.378	5.618	5.26	4.181	4.573	4.218	
	- 38	0.488	6.068	5.59	5.515	5.756	5.391	4.267	4.657	4.302	
	41	0.488	5.975	5.478	5.404	5.638	5.3	4.253	4.629	4.291	
	42	0.488	5.956	5.525	5.456	5.641	5.355	4.49	4.757	4.518	
	43	0.488	6.1	5.638	5.552	5.788	5.441	4.337	4.712	4.377	
	46	0.488	5.924	5.417	5.353	5.615	5.239	4.115	4.52	4.137	
	47	0.488	5.726	5.263	5.162	5.36	5.033	4.05	4.414	4.075	
	64 -	0.488	6.086	5.605	5.484	5.736	5.389	4.29	4.693	4.342	
≏	- 21	0.488	5.888	5.446	5.341	5.563	5.223	4.166	4.537	4.191	
Volume	45 -	0.429	5.973	5.448	5.3	5.675	5.233	3.928	4.455	3.968	
	- 13	0.578	6.043	5.603	5.605	5.701	5.461	4.892	4.862	4.75	
	51	0.578	5.688	5.329	5.228	5.275	5.04	4.609	4.554	4.446	
	26	0.578	6.043	5.605	5.605	5.701	5.461	4.892	4.862	4.75	
	- 28	0.55	5.818	5.444	5.363	5.454	5.213	4.596	4.697	4.519	
	9g -	0.58	6.236	5.853	5.829	5.963	5.703	5.12	5.064	4.957	
	 G	0.552	6.04	5.491	5.473	5.674	5.348	4.541	4.749	4.515	
	~ -	0.625	5.615	5.374	5.357	5.346	5.196	5.003	4.719	4.709	
	4 -	0.625	6.058	5.117	5.1	5.077	4.961	4.734	4.479	4.449	
	10	0.625	5.868	5.511	5.534	5.566	5.385	5.073	4.82	4.759	
	11 -	0.625	5.868	5.511	5.534	5.566	5.385	5.073	4.82	4.759	
	15	0.625	5.751	5.463	5.448	5.429	5.311	5.085	4.814	4.776	
	17	0.625	5.615	5.373	5.357	5.346	5.196	5.003	4.719	4.709	
	50	0.615	6.063	5.669	5.66	5.737	5.508	5.166	4.918	4.867	
	52	0.625	5.751	5.465	5.448	5.429	5.311	5.085	4.814	4.776	
	25 2	0.625	5.544	5.283	5.286	5.274	5.137	4.898	4.64	4.604	
		0.625	5.154	4.935	4.931	4.895	4.786	4.568	4.314	4.281	
	.7 6	0.625	5.618	5.294	5.317	5.322	5.137	4.889	4.616	4.579	
	 0 -	0.625	5.677	5.387	5.387	5.389	5.23	4.991	4.713	4.675	
	4 -	0.625	5.814	5.516	5.467	5.461	5.296	5.071	4.798	4.762	
	1 4	0.703	6.122	5.793	5.896	5.827	5.686	5.728	5.344	5.352	
	с Б		5.87	5.448	5.397	5.529	5.26	4.55	4.657	4.45	
	УË,	I	5.07	2.740	2.337	5.525	5.20				1

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).

	cing			~			hbre			
	Pixel Spa	Phud-	IPEG.LS	IPEG2000	- HEVC	afd.	LSTM. COL	Medzip1	Medzipz	Medzib3
- 1	0.5	4.245	3.941	3.558	4.079	3.707	3.424	3.053	3.267	3.082
- 7	0.5	4.246	3.966	3.57	4.068	3.702	3.455	3.114	3.294	3.133
m -	0.5	4.281	3.969	3.572	4.096	3.709	3.458	3.1	3.299	3.122
4 -	0.5	4.133	3.875	3.487	3.99	3.621	3.379	3.014	3.212	3.037
- <u>۱</u>	0.5	4.52	4.158	3.757	4.273	3.885	3.654	3.295	3.478	3.314
9 - 0	0.5	4.101	3.827	3.438	3.96	3.579	3.338	2.949	3.172	2.976
olume I 7	0.5	4.163	3.899	3.516	4.025	3.662	3.428	3.099	3.279	3.116
× ∞	0.5	4.327	4.04	3.644	4.137	3.776	3.536	3.222	3.382	3.238
6 -	0.5	4.201	3.922	3.534	4.034	3.679	3.459	3.142	3.304	3.158
10	0.5	4.257	3.969	3.581	4.101	3.719	3.483	3.137	3.32	3.155
11	0.5	4.153	3.879	3.503	4.029	3.655	3.409	3.012	3.236	3.047
12	0.5	4.073	3.803	3.42	3.899	3.553	3.326	2.991	3.17	3.007
bve -		4.225	3.937	3.548	4.058	3.687	3.446	3.094	3.284	3.115

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

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

(Blue) to minimum compression 4.52 BPP (Red).

		TestSet1	TestSet2	
summary overview of the	PPMd -	131.92%	136.56%	- 136.5
ompression performance	တို JPEG-ls -	122.43%	127.26%	- 130.0
ver the two test sets for all ne lossless methods. Cells	JPEG2000	121.28%	114.69%	- 125.0
re coloured from the best	Σ HEVC	124.27%	131.15%	- 120.0
ompression performance	JP3D -	118.22%	119.18%	- 115.0
00.00% (Blue) to the worst	d MedZip1 -	102.26%	100.00%	- 110.0
erformance 136.56% (Red).	O MedZip2 -	104.67%	106.15%	- 105.0
Less value indicates better	MedZip3 -	100.00%	100.69%	- 100.0
ertormance).				

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.

- 6.236

6.000 5.800

5.600

5.400 - 5.200

5.000 - 4.800 耑

4.600

- 4.200

