SSN MLRG at ImageCLEF 2022 Tuberculosis: Caverns Report using 3D CNN and Uniformizing Techniques

Dheepak S[1], Kavitha Srinivasan[1] and Raghuraman G[1]

[1]Department of CSE, Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam – 603110, India.

Abstract

Tuberculosis (TB) is a bacterial infection that mainly affects the lungs. It is a widespread chronic infectious disease, hence the analysis of Tuberculosis Computed Tomography (CT) reports has a significant impact on clinical treatment. To emphasize the importance of medical report writing, the ImageCLEF forum has introduced the Caverns Report generation task from 3D CT images this year and we participated in the same task. Due to the depth variability of the 3D CT images, we explored a pre-processing technique called Uniformizing Techniques. This pre-processing technique samples a subset of the slices using a spacing factor to equal samples from the sequence of slices to generate the desired volumetric output. The pre-processed image is fed as input to three separate binary classification networks. The results of the networks are combined to generate the report. Our team ranks the fourth position in this task and achieved a mean AUC score and min AUC score of 0.461 and 0.256, respectively.

Keywords

Tuberculosis, Computed Tomography, 3D CNN, Uniformizing Techniques, Pre-processing

1. Introduction

Tuberculosis (TB) affects 10 million people and kills 1.5 million people per year around the world, despite being a preventable and curable disease. TB is a kind of bacteria called Mycobacterium tuberculosis and it most often affects the lungs. TB is spread through the air from the TB-affected people by coughing, sneezing, or spitting. If the person with AIDS/HIV got affected by TB then it has a leading cause of death and also a major contributor to antimicrobial resistance. TB bacteria are thought to infect about a quarter of the world's population. Only 5 to 15 percent of these persons will develop active tuberculosis illness [1]. The rest have tuberculosis but aren't sick, so they can't spread the disease. Currently, tuberculosis is diagnosed mostly through an extensive assessment of the patient's clinical signs, imaging data, and laboratory examination results. A chest X-ray and a CT scan are two imaging diagnostic methods. The lymphadenopathy or miliary alteration of the lung and mediastinum can be seen on a chest X-ray. It is identified that comparatively, CT scan images have high resolution to make a detailed evaluation for the detection of tuberculosis. Therefore the analysis of Tuberculosis CT reports has a significant impact on clinical treatment.

CLEF 2022: Conference and Labs of the Evaluation Forum, September 5–8, 2022, Bologna, Italy
dheepak2020027@ssn.edu.in (D. S); kavithas@ssn.edu.in (K. Srinivasan); raghuramang@ssn.edu.in (R. G)
0000-0001-6235-5569 (D. S); 0000-0003-3439-2383 (K. Srinivasan); 0000-0001-7813-4883 (R. G)

The CLEF initiative labs are organising ImageCLEF 2022 [2], which is an evaluation campaign. This campaign includes several research projects that are open to teams from all across the world. We focus on the Tuberculosis task from the ImageCLEFmedical competition this year. Caverns Detection and Caverns Report are two sub-tasks of the ImageCLEF 2022 Tuberculosis task. We participated in the Caverns Report task, which required us to predict three binary cavern features as proposed by professional radiologists [3]. Moreover CT based Tuberculosis tasks are part of ImageCLEF from 2017 onwards for classification and prediction, where machine learning and 2D CNN approaches are adopted in the majority of papers [4, 5].

The remaining part of the paper spans the following subsections. In Section 2, the Caverns Report dataset is described. The design of the proposed system is explained in Section 3. A summary of the implementation, result, and the respective evaluation of all runs is given in Section 4 and, the conclusion and future work are summarized at the end.

2. Dataset

The Caverns Report task dataset consists of 60 training and 16 test instances [3]. In each CT image, two versions of automatically extracted lung masks, information on cavern area location, and a cavern report are all included in the dataset. A single 3D CT image is provided for all patients, with an image size of 512×512 pixels and a total number of slices of roughly 100. The CT images are all saved in the NIFTI file format with the .nii.gz extension (g-zipped .nii files). Two versions of automatically extracted lung masks were provided for all CT images. The first version of segmentation produces more accurate masks, but in the most severe cases of tuberculosis, it tends to miss big abnormal lung regions [6]. On the other hand, the second segmentation provides more rough limits but is more consistent in terms of including lesion areas [7]. The cavern report has three manually labelled binary features which characterizes the cavern (s). The existence of thick walls, calcifications, and foci around the cavern are the distinguishing features. The report comes as a simple .csv file with the following columns (including the header): ID (train case id), thick_walls(binary label for the presence of thick walls around the cavern), has_calcification (binary label for the presence of calcifications around the cavern), foci_around (binary label for the presence of foci around the cavern).

3. System Design

The architecture of our proposed system is shown in Fig. 1 and it consists of Uniformizing Technique module and 3D Convolutional Neural Network (CNN), and we will discuss each part in detail in the following subsections.

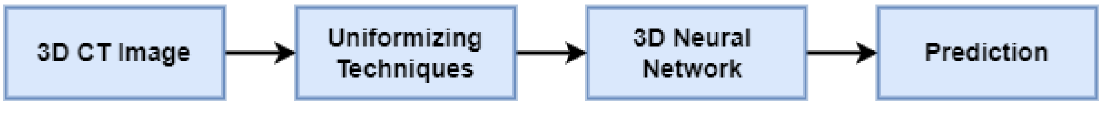

Figure 1: System Design

3.1. Uniformizing Techniques

Due to the depth variability of the 3D CT images, we explore a pre-processing technique called Uniformizing Techniques [8]. To generate the desired volumetric output, this pre-processing technique samples a subset of the slices using a spacing factor to equal sample from the sequence of slices.

3.1.1. Subset Slice Selection (SSS)

In this technique, the slices are sampled from the first, middle, and last positions of the entire volume. To achieve consistency due to depth variations, the middle slices are sampled by indexing from half of the input volume depth. The subsets are then stacked depthwise to get the desired input volume.

3.1.2. Even Slice Selection (ESS)

In this technique, a target depth N and a scan depth of size D are computed. The equation $F = \frac{D}{N}$ is then used to calculate a spacing factor. By maintaining the spacing factor F between the sequence of slices in the volumetric data, sampling is done at the slice level.

3.1.3. Spline Interpolated Zoom (SIZ)

In this technique, instead of manually selecting a subset of slices, a constant target depth size of N is pre-determined. Then, for each volume, compute its depth D and use spline interpolation [9] to zoom it along the z-axis by a factor of $\frac{1}{D/N}$, where the interpolant is an order of three. By reproducing the nearest pixel along with the depth or z-axis, the input volume is zoomed or squished. In the other experiments, similar procedures were employed [10, 11].

Figure 2: 3D Neural Network

3.2. 3D Neural Network

The proposed 3D architecture shown in Fig 2 has 17 layers, including four 3D convolutional (CONV) layers with two layers consisting of 64 filters, followed by 128 and 256 filters, all with a kernel size of 3×3×3. Following each CONV layer is a max-pooling (MAXPOOL) layer with a stride of 2 and ReLU activation, followed by batch normalization (BN) [12]. The feature extraction block is made up of four CONV-MAXPOOL-BN modules. The feature extraction block's final output is flattened and delivered to a fully connected layer with 512 neurons. We use a 60 percent effective dropout rate [13]. For the binary classification problem, the output is sent to a dense layer of two neurons with softmax activation.

4. Implementation

In this section, the proposed system implementation is explained along with the minimum software and hardware requirements.

4.1. System Specification

The hardware and software required for the development of the proposed system include, (i). Intel i7 processor with NVIDIA graphics card, 4800M at 4.3GHZ clock speed, 12GB RAM, Graphical Processing Unit, and 1TB Solid State drive, (ii). Windows 10 operating system with VSCode editor, Python 3.9 package with required libraries like TensorFlow, NumPy, scipy, nibabel, pandas, etc.

4.2. Experimental Setup

In this section, the experiment setting along with the network parameters are explained. The input image is converted into 2D slices. The slices are resized to 128×128. The resized slices are taken as input by the uniformizing techniques module. The uniformizing techniques samples at the sampling from slice level and stack them depth-wise to produce the desired 3D volume. Fig. 3 illustrates the slices of the 3D image with ID TRN_04. This image had a total of 104 slices. Fig. 4 shows the slices that were sampled by applying the Subset Slice Selection. Fig. 5 visualizes the slices that were sampled by applying the Even Slice Selection and Fig. 6 represents the slices that were sampled by applying Spline Interpolated Zoom.

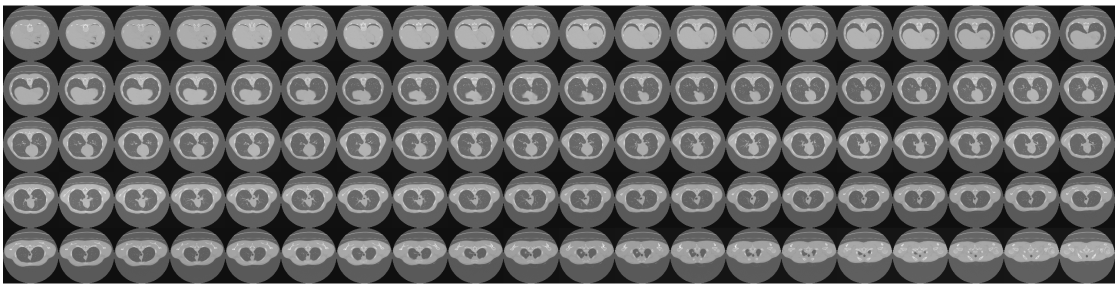

Figure 3: Slices of 3D CT Image

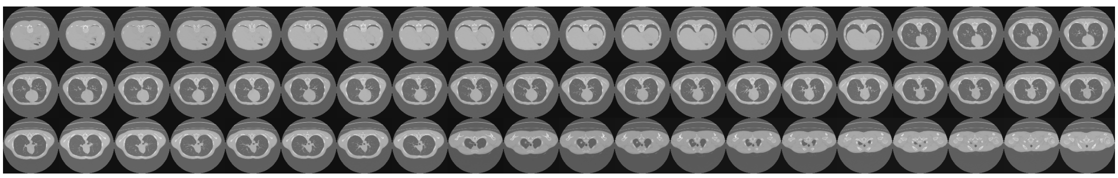

Figure 4: Slices sampled using Subset Slice Selection

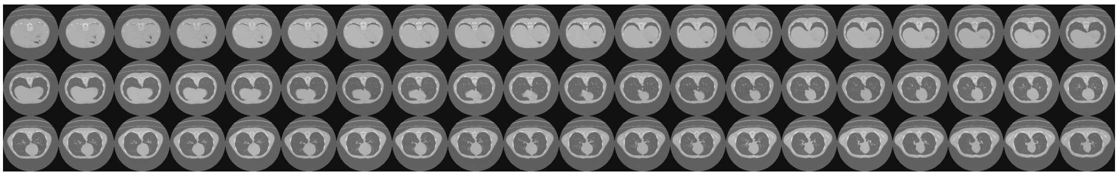

Figure 5: Slices sampled using Even Slice Selection

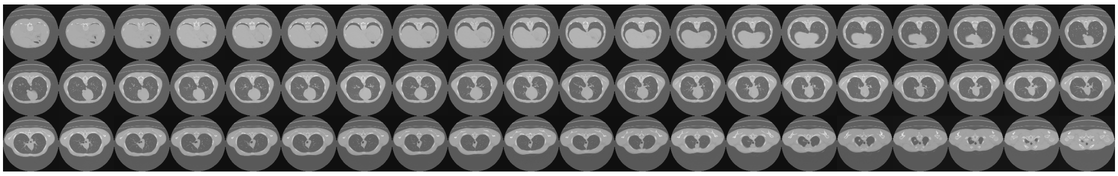

Figure 6: Slices sampled using Spline Interpolated Zoom

The pixel values were normalized by subtracting the minimum pixel value and dividing it by the difference between the maximum and minimum pixel values. The normalized image is then given as an input to the proposed 3D CNN model.

We used Stochastic Gradient Descent(SGD) optimizer with a learning rate of 10^6 and a momentum of 0.99. Weight is initialized using the Glorot initialization method [14] and minimizes the Mean average error [15] during training. The proposed network was trained for 100 epochs with a batch size of 2.

We divided the Caverns Report task into three separate binary prediction tasks based on the features and built a separate model for each one of them. The results from the models are combined to generate the report.

To ensure a fair comparison between the uniformizing methods, we set the desired input size of $128\times128\times64$ for all our experiments.

4.3. Results

The performace of the proposed system is evaluated using Area Under the ROC Curve (AUC) metric and the results are tabulated in Table 1. There were a total of 3 runs submitted for the task. One run for each one of the uniformizing techniques. When compared to SSS and ESS, we inferred that SIZ better depicts the 3D CT when downsampled. In addition to this, ESS produces slightly better results than SSS because in ESS the sampling is done consecutively. Selecting specific slices, on the other hand, does not preserve the semantic meaning of volumetric data because it is not a proper representation of the 3D CT scan, which is also intuitive. Even though ESS downsamples the volume from a subset, the sampling is done throughout the entire volume, resulting in greater performance. In comparison to SSS, ESS enhances the likelihood of sampling the TB affected segments. Because tuberculosis can affect any portion of the lung, it's impossible to know which slices should be rejected without looking at each scan individually because the annotations are provided at the volume level rather than the slice level, retrieving data from the complete volume is critical now-a-days.

Table 1
Test set Results

Uniformizing Techniques	MEAN_AUC	MIN_AUC
Subset Slice Selection	0.407	0.205
Even Slice Selection	0.400	0.231
Spline Interpolated Zoom	0.461	0.256

Table 2
Ranking of ImageCLEF 2022 Tuberculosis Caverns Report task

Participants	MEAN_AUC	MIN_AUC	Successful submissions
SDVA-UCSD	0.687	0.513	10
KDE-lab	0.658	0.317	11
KL_BP_SSN	0.536	0.413	5
SSN_Dheepak_Kavitha	**0.461**	**0.256**	**3**

In ImageCLEF 2022 Tuberculosis Caverns Report task, 4 teams participated out of 37 teams with 29 successful submissions. Among these, we have made 3 successful submissions and achieved the fourth rank. The overall ranking achieved by the teams is tabulated in Table 2.

5. Conclusion and Future Work

In this paper, we proposed a framework consisting of 3D CNN and Uniformizing Techniques to generate the Tuberculosis Caverns Report. The proposed system pre-processes the input CT image using the uniformizing techniques to sample the slices of the image to generate a volume of the desired output. The generated volume is fed as input to 3 binary classifier networks. The output of the networks is combined to generate the report. The proposed framework achieved a mean Area Under the ROC curve of 0.461. In future, different 3D network architectures will be tested, and the slice selection techniques can be improved with an attempt to construct a robust deep learning model which will generate an accurate Tuberculosis CT report.

Acknowledgments

Our profound gratitude to Sri Sivasubramaniya Nadar College of Engineering, Department of CSE, for allowing us to utilize the High Performance Computing Laboratory and GPU Server for the execution of this challenge successfully.

References

[1] W. H. Organization, Tuberculosis, 2022. URL: https://www.who.int/health-topics/tuberculosis#tab=tab_1.

[2] B. Ionescu, H. Müller, R. Peteri, J. Rückert, A. Ben Abacha, A. G. S. de Herrera, C. M. Friedrich, L. Bloch, R. Brüngel, A. Idrissi-Yaghir, H. Schäfer, S. Kozlovski, Y. D. Cid, V. Kovalev, L.-D. Ştefan, M. G. Constantin, M. Dogariu, A. Popescu, J. Deshayes-Chossart, H. Schindler, J. Chamberlain, A. Campello, A. Clark, Overview of the ImageCLEF 2022: Multimedia retrieval in medical, social media and nature applications, in: Experimental IR Meets Multilinguality, Multimodality, and Interaction, Proceedings of the 13th International Conference of the CLEF Association (CLEF 2022), LNCS Lecture Notes in Computer Science, Springer, Bologna, Italy, 2022.

[3] S. Kozlovski, Y. Dicente Cid, V. Kovalev, H. Müller, Overview of ImageCLEFtuberculosis 2022 - CT-based Caverns Detection and Report, in: CLEF2022 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Bologna, Italy, 2022.

[4] S. Kavitha, S. Poornima, N. S. Sitara, A. Sarada Devi, Classification of lung tuberculosis using non parametric and deep neural network techniques, in: 4th International Conference on Computer, Communication and Signal Processing (ICCCSP), 2020, pp. 1–5. doi:10.1109/ICCCSP49186.2020.9315211.

[5] S. Kavitha, P. Nandhinee, S. Harshana, J. S. S, K. Harrinei, ImageCLEF 2019: A 2D Convolutional Neural Network approach for severity scoring of lung Tuberculosis using CT Images, in: CLEF (Working Notes), 2019.

[6] Y. Dicente Cid, O. A. Jiménez del Toro, A. Depeursinge, H. Müller, Efficient and fully automatic segmentation of the lungs in CT volumes, in: Proceedings of the VISCERAL Anatomy Grand Challenge at the 2015 IEEE ISBI, CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, 2015, pp. 31–35.

[7] V. Liauchuk, V. Kovalev, ImageCLEF 2017: Supervoxels and co-occurrence for tuberculosis CT image classification, in: CLEF2017 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Dublin, Ireland, 2017.

[8] H. Zunair, A. Rahman, N. Mohammed, J. P. Cohen, Uniformizing techniques to process CT scans with 3D CNNs for tuberculosis prediction, in: International Workshop on Predictive Intelligence In Medicine, Springer, 2020, pp. 156–168.

[9] C. De Boor, C. De Boor, A practical guide to splines, volume 27, springer-verlag New York, 1978.

[10] M. Grewal, M. M. Srivastava, P. Kumar, S. Varadarajan, Radnet: Radiologist level accuracy using deep learning for hemorrhage detection in CT scans, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE, 2018, pp. 281–284.

[11] S. Kazlouski, ImageCLEF 2019: CT Image Analysis for TB Severity Scoring and CT Report Generation using Autoencoded Image Features, CLEF (Working Notes) 2 (2019).

[12] S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, in: International conference on machine learning, PMLR, 2015, pp. 448–456.

[13] A. Pattnaik, S. Kanodia, R. Chowdhury, S. Mohanty, Predicting tuberculosis related lung deformities from CT scan images using 3D CNN, in: CLEF (Working Notes), 2019.

[14] X. Glorot, Y. Bengio, Understanding the difficulty of training deep feedforward neural networks, in: Proceedings of the thirteenth international conference on artificial intelligence and statistics, JMLR Workshop and Conference Proceedings, 2010, pp. 249–256.
[15] C. J. Willmott, K. Matsuura, Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Climate research 30 (2005) 79–82.