Middlesex University London

UNIVERSITY OF MAURITIUS

AbuDhabi ICIP 2020

IEEE International Conference on Image Processing
25-28 October 2020, United Arab Emirates

IEEE Signal Processing Society

An enhanced deep learning architecture for classification of tuberculosis types from CT lung images

Xiaohong Gao[1], Richard Comley[1], Maleika Heenaye-Mamode Khan [2]

[1]Department of Computer Science, Faculty of Science and Technology,
Middlesex University, London NW4 4BT, UK
{x.gao@mdx.ac.uk , r.comley@mdx.ac.uk }

[2]Department of Software and Information Systems, Faculty of Information Communication and Digital Technologies, University of Mauritius, Mauritius
{m.mamodekhan@uom.ac.mu}

Introduction

- Tuberculosis (TB) is **a bacterial** infectious disease caused by Mycobacterium (M.)

- Tuberculosis is contracted through inhaling tiny droplets from the coughs or sneezes of an infected person and remains one of the top 10 causes of death worldwide.

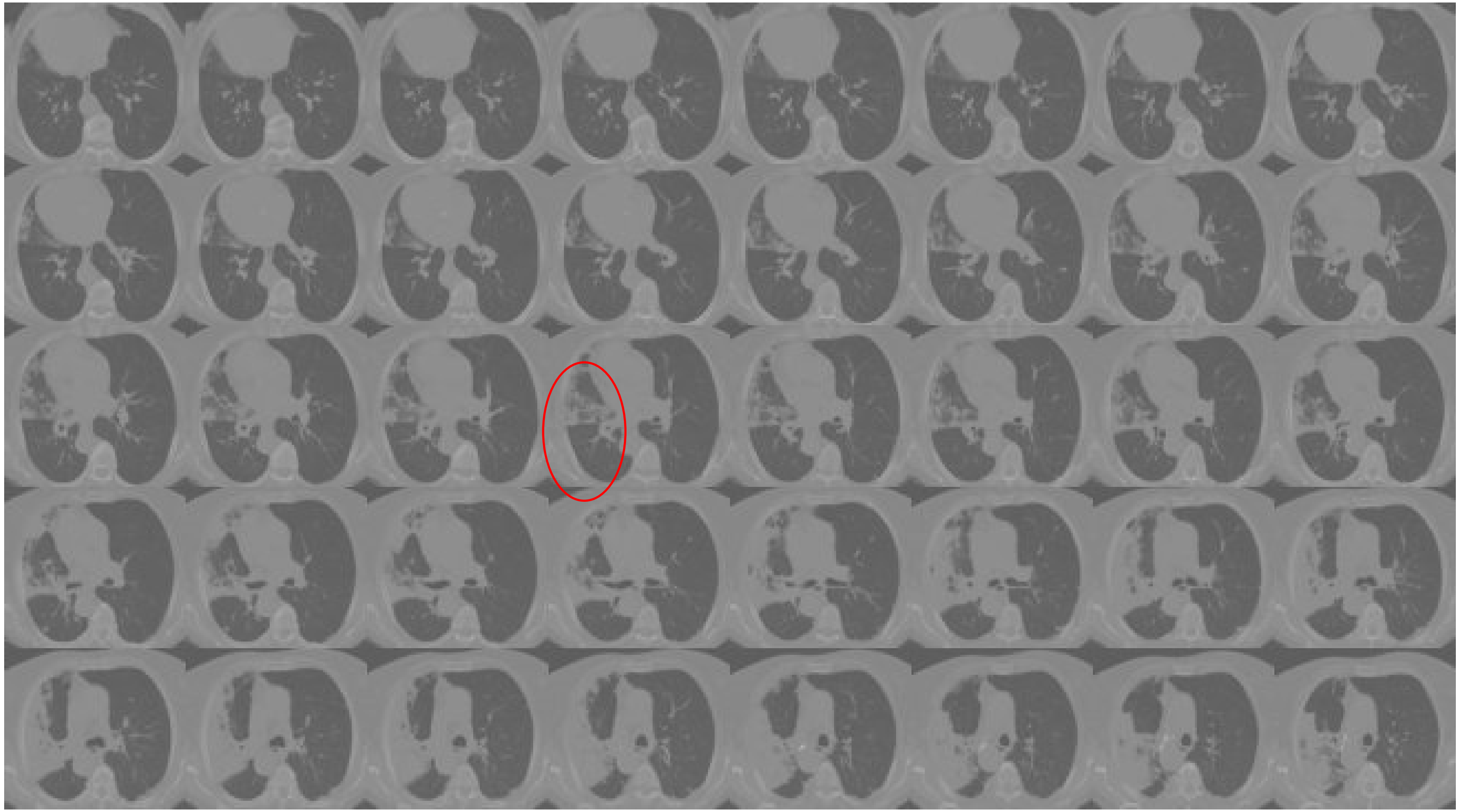
- In 2015, **10.4 million** people fell ill with TB, among them 1.8 million died of the disease**.**

- Although TB remains a serious contagious condition**, it can be cured** if treated in a timely manner with the **right antibiotics**.

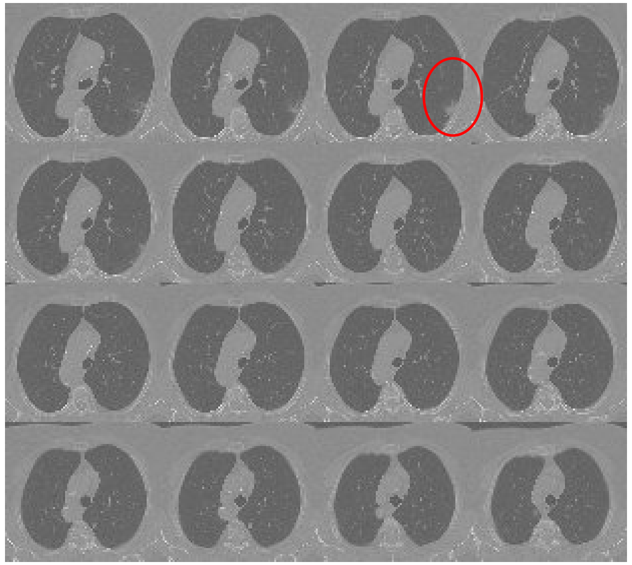
- Hence knowing the types of TB plays an important first step.

- To assist clinicians to analyze, diagnose and deliver optimal treatment for TB patients, high resolution Computed Tomography (CT) imaging is one of the tools.

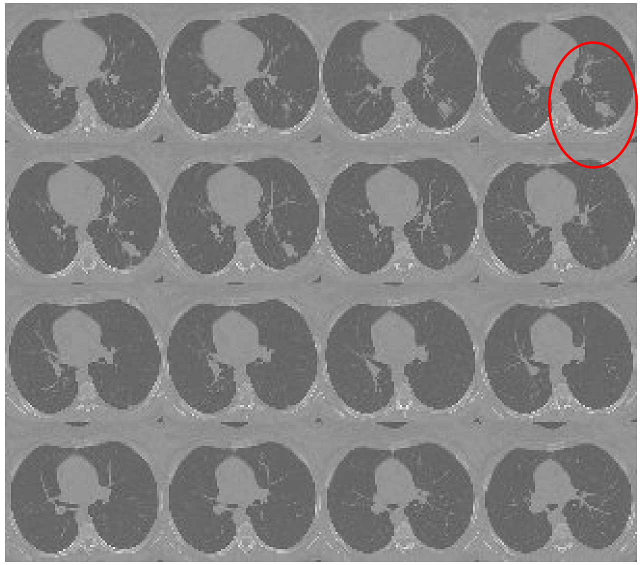
Type-1: Infiltrative TB – montage of 3D CT images



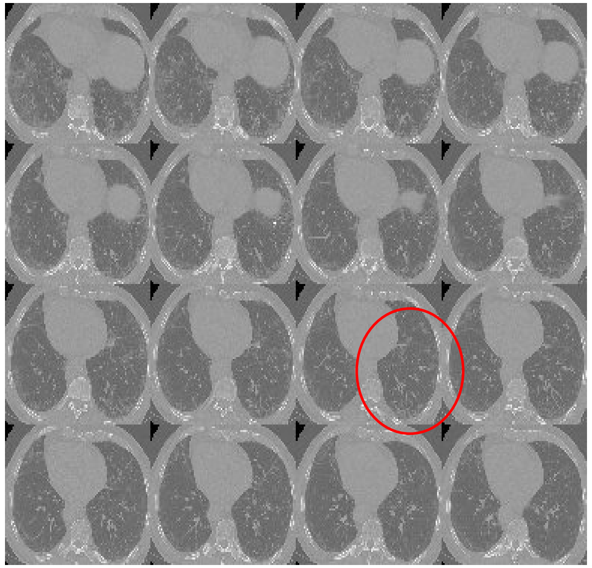
Other four types of TB



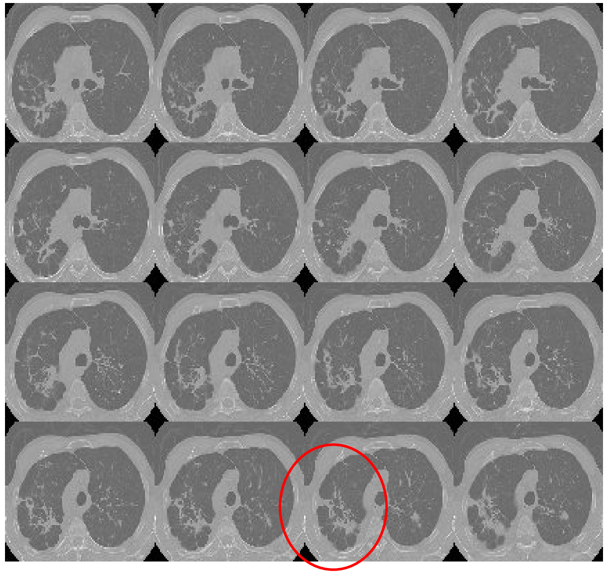
Type-2:
Focal



Type-3:
Tuberculoma



Type-4:
Miliary

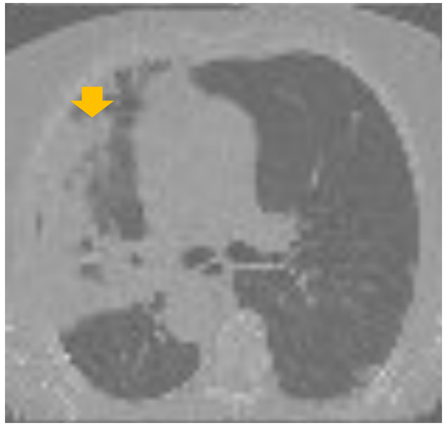


Type-5:
Fibro-cavernous

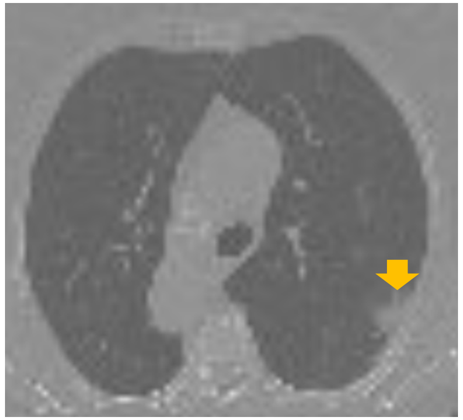
Challenges facing classification of TB types

- For Types 4 (Miliary) and 5 (Cavity), the visual features are apparent with wide-spread dissemination of small spots (i.e. Mycobacterium tuberculosis**)** and dominating holes (**cavities**) respectively (arrows).

- However, for Types 1 to 3, visual features are not easily

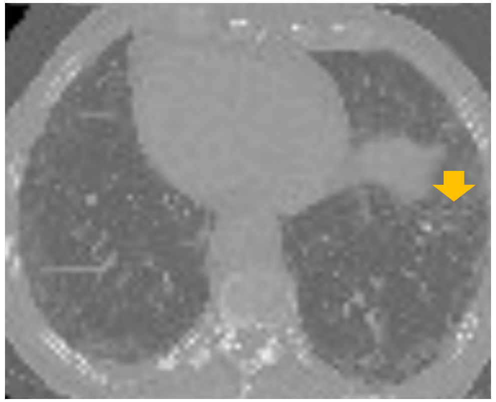


T1: Infiltrative

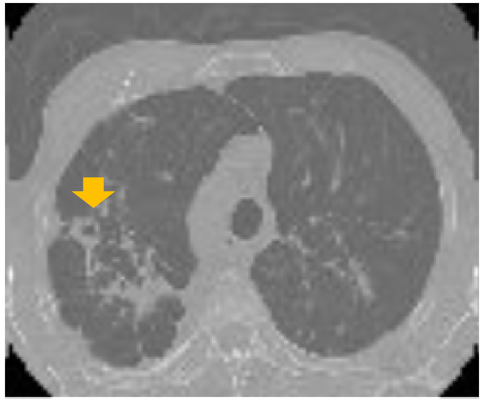


T2: Focal

T3: Tuberculoma



T4: Miliary

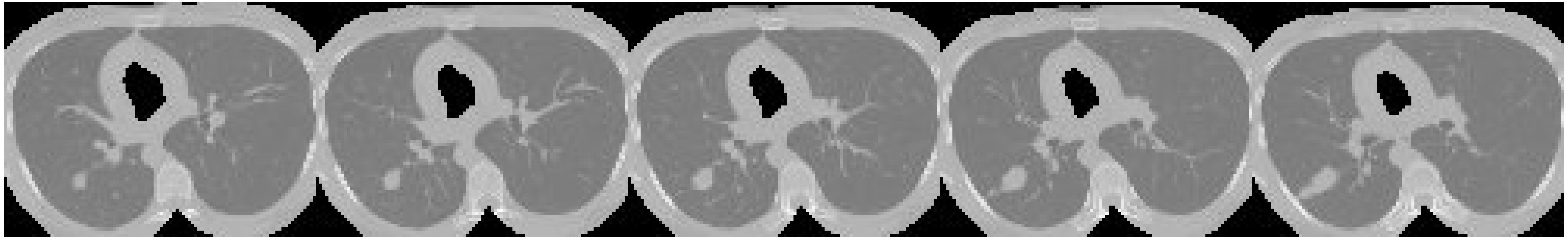


T5: Cavity

Our approach

- In this study, an **enhanced deep learning** architecture is developed for classification of TB types

- Specifically, ***depth (3rd dimension) information*** from 3D CT images was taken into account through the transfer learning that is applied built on a residual network, ResNet.



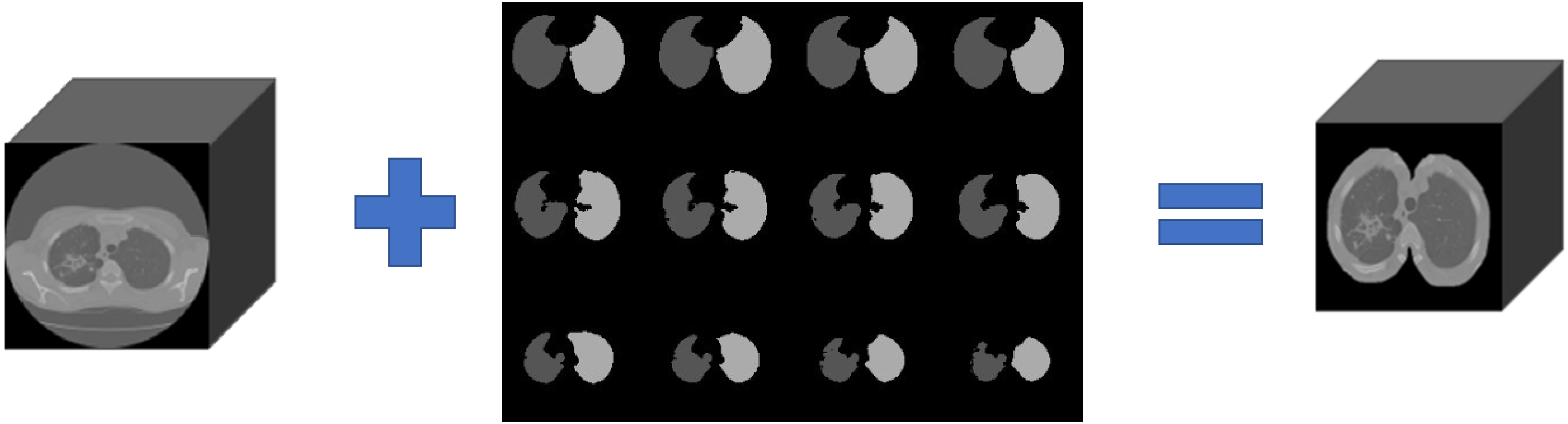
Datasets

- Data are collected from the competition organised by ImageCLEF2018 on Tuberculosis classification task (task#2) with 1008 training datasets from 677 subjects.

Type	**1**	**2**	**3**	**4**	**5**	**Total**
train	200	170	100	70	70	610
Test	176	103	54	36	29	398
Total	**376**	**273**	**154**	**106**	**99**	**1008**

Data Pre-Processing

- Each 3D dataset firstly undergoes a pre-processing stage to remove surrounding boundaries based on the provided lung masks. As a result, each volume has a dimension of $256 \times 256 \times$ depth z, with depth varying between 20 to 250 slices.



Mask montage

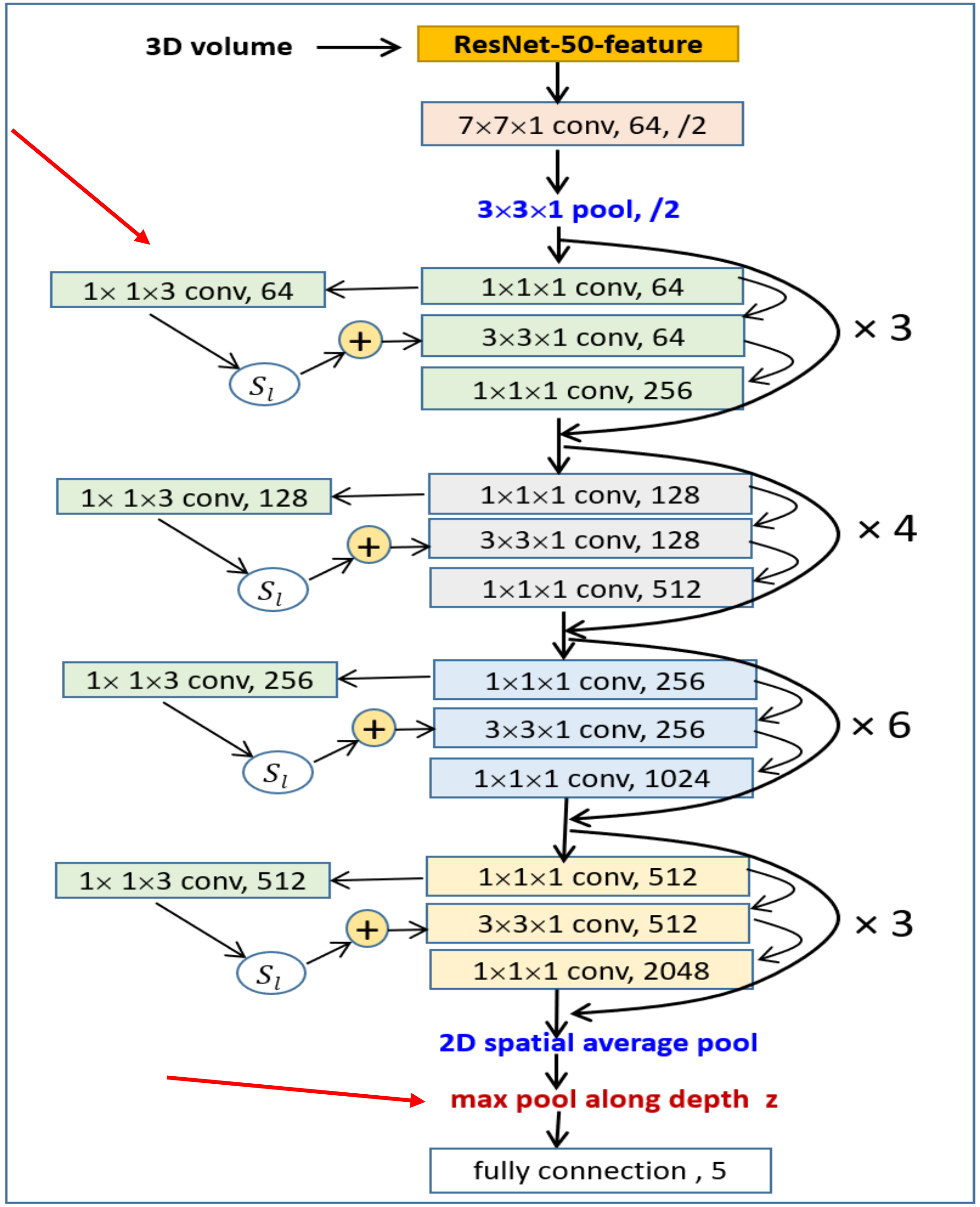
Enhanced deep residual learning – depth-resnet

- In this work, inspired by [1], an enhanced ResNet, depth-ResNet is introduced to be applied for classification of the five types of tuberculosis from CT lung images, which is built on the ResNet-50 model

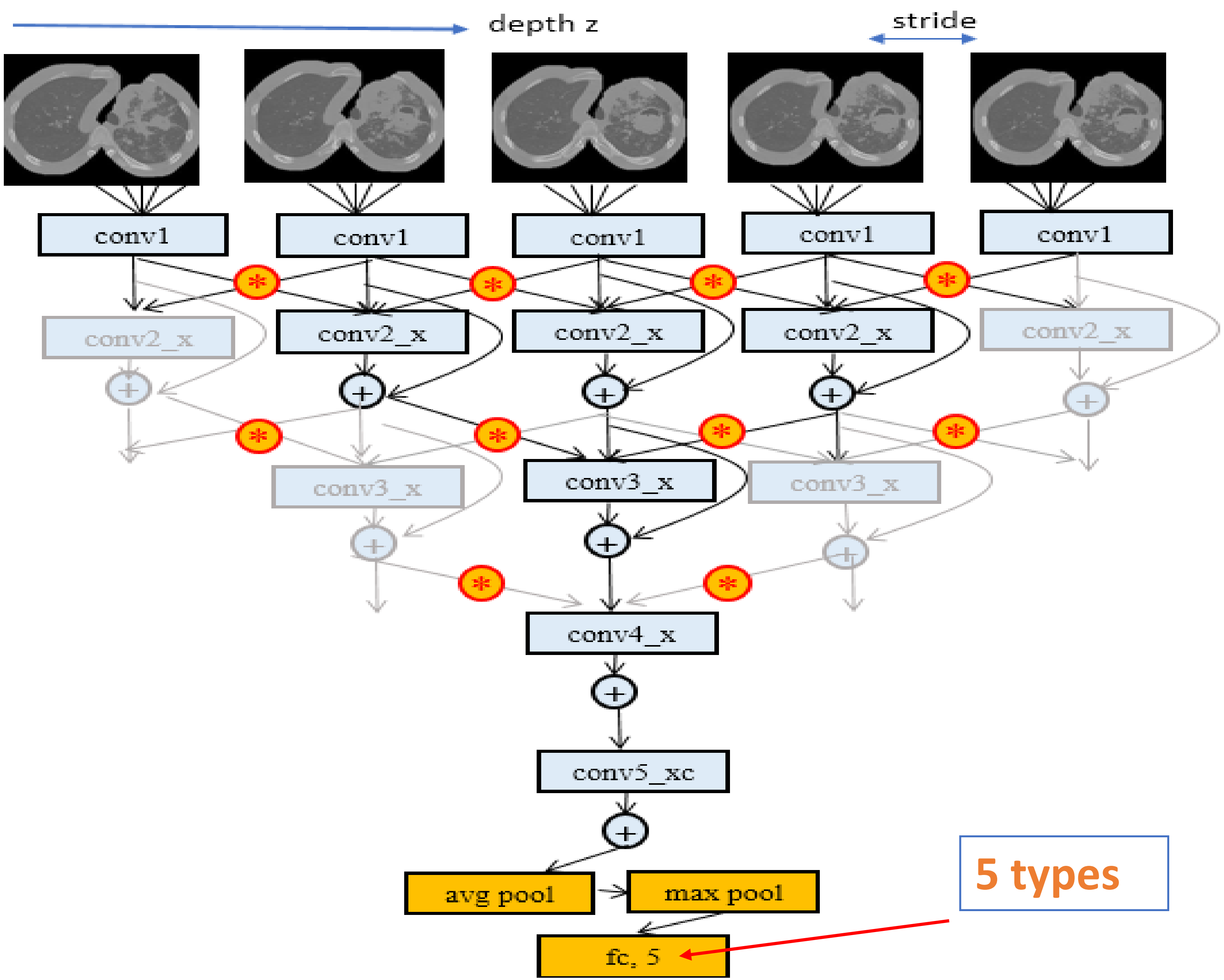
1. Feichtenhofer C, Pinz A, Wildes R, Temporal Residual Networks for Dynamic Scene Recognition, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

Architecture of depth-resnet

- To take advantage of ResNet-50 using 3x3 filters to perform spatial convolution, the depth convolution also adopts 3 pixels, i.e. 1x1x3 between the current, front and back frames.
- Considering some TB types (e.g., Type 3) have lesions only covering a few slices along the depth, a **4-frame block** is chosen in this study.
- As illustrated right, to minimize the classification errors, a global pooling layer followed by a 5-way fully connected layer, optimised using a Softmax approach is conducted.



A block in the depth-ResNet depth-resnet



5 types

Results confusion matrix for 4-slice block of depth-ResNet

Type	1	2	3	4	5	Avg
1	144		5	3	5	
2	25	83	31	12	1	
3		5	15			
4	5	1		21	1	
5	2	14	3		22	
Accuracy (%)	81.81	80.58	27.77	58.33	66.66	71.60

Results confusion matrix for 2D ResNet

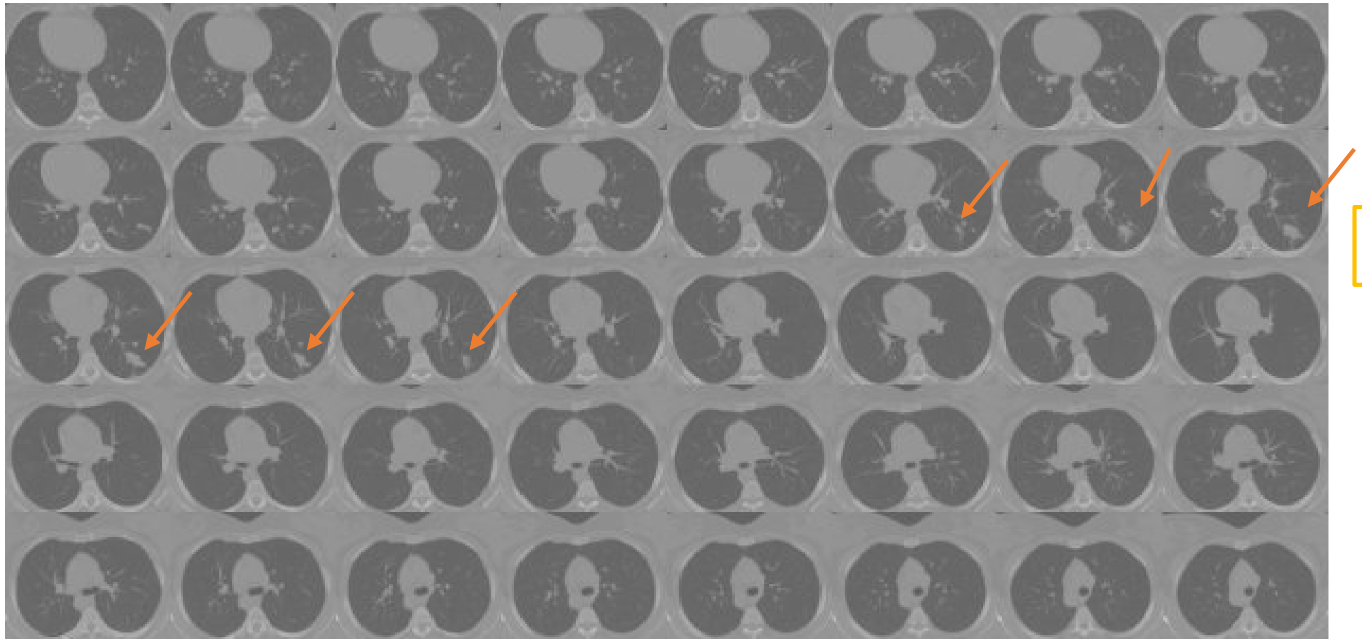
Type	1	2	3	4	5	Avg
1	135	12	7	6	4	
2	25	85	32	10		
3	3	4	11			
4	11	2	1	19	2	
5	2		3	1	23	
Accuracy (%)	76.70	82.52	20.37	52.77	84.61	68.59

Comparison of specificity (Sp) and sensitivity (Se)

Method	T	1	2	3	4	5	Avg (%)
Resnet	Se	76.76	82.52	20.37	52.77	79.31	62.33
	Sp	88.44	81.49	98.00	95.76	98.40	92.41
Depth-Resnet (8-slice)	Se	81.81	74.75	27.77	50.0	75.86	62.03
	Sp	84.41	85.01	98.00	96.79	97.36	92.31
Depth-Resnet (4-slice)	Se	81.81	80.58	27.77	58.33	75.86	64.87
	Sp	94.46	81.04	98.56	98.10	95.10	93.45

Discussion and Conclusion

- In the 2017 TB competition, the best result was achieved using the ResNet approach with averaged accuracy rate of 40.33% , which was lower than this study (71.60%).
- The biggest challenge is to detect Type 3 Tuberculoma, which has 27.77% accuracy.
- In the future, clinicians' knowledge will be incorporated to improve the classification accuracy, especially for Type 3 Tuberculoma.



Type 3 Tuberculoma (arrows)


Thank You!