

# 3D-CNN and SVM for Multi-Drug Resistance Detection

Imane Allaouzi, Badr Benamrou, Mohamed Benamrou and Mohamed Ben Ahmed

Abdelmalek Essaâdi University  
Faculty of Sciences and Techniques,  
Tangier, Morocco

**Abstract.** Recently, deep Convolutional neural networks (CNN) have obtained significant results in volumetric medical image classification task. However, a number of studies have claimed that replacing the trainable classifier (the conventional softmax function) of a CNN model with an SVM classifier can improve the classification performance. In this paper, we investigate the use of a hybrid approach which combines 3D CNN and a linear SVM to tackle the task of Multi-Drug Resistant detection.

**Keywords:** 3D-CNN, Deep Neural Networks, SVM classifier, 3D CT medical imaging, binary classification

## 1 Introduction

Tuberculosis (TB) is a serious infectious disease usually caused by the *Mycobacterium tuberculosis* (MTB) [1]. It is one of the top 10 causes of morbidity and mortality worldwide, killing 1.7 million people in 2016. The greatest problem that can happen to a patient with tuberculosis (TB) is that the organisms become resistant to two or more of the standard drugs, which is a formidable obstacle to effective TB care and prevention globally.

The motivation of ImageCLEFtuberculosis 2018 task [2,3] is to explore the use of tuberculosis CT (Computed Tomography) scans in order to help in early detection of the drug resistance (DR) status which in turn has a great importance for effective treatment. In this paper, we describe our contribution to the task of Multi-Drug Resistant (MDR) detection, which aims to assess the probability of a TB patient having resistant form of tuberculosis.

Recently, deep Convolutional neural networks (CNNs) have obtained significant results in volumetric medical image classification task [4,5,6]. However, a number of studies have claimed that replacing the trainable classifier (the conventional softmax function) of a CNN model with the Support Vector Machine (SVM) classifier can improve the classification performance [7]. In this regard, we investigate the use of a hybrid approach that combines 3D CNN and a linear SVM classifier to tackle the task of MDR detection.

The remainder of this paper is organized as follows: the dataset is described in Section 2, the proposed methodology is described in Section 3, results are presented and discussed in Section 4, and finally Section 5 draws some conclusions and future work.

## 2 Dataset:

- **MDR detection dataset:**

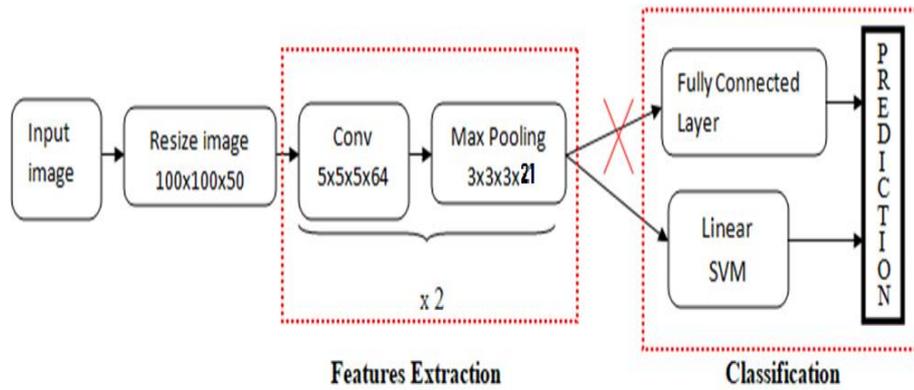
MDR detection is a Binary classification task, where the goal is to predict the probability for each of the two classes: Drug Sensitive (DS) and Multi-Drug Resistant (MDR). As shown in the table below a dataset of 3D CT images is used with 259 of training examples, and 236 of test cases.

**Table 1.** The distribution of MDR detection dataset.

Dataset	Train	Test
DS	134	99
MDR	125	137

## 3 Methodology :

This section elaborates on the details our proposed approach to address the MDR detection task from ImageCLEFtuberculosis 2018. As shown in the fig.1, our system is designed to (1) automatically extract prominent features from 3D CT scans using a 3D CNN, and then (2) classify them into predefined categories with the Linear SVM classifier.



**Fig.1.** The architecture of our model.

- **Data pre-processing:**

The input 3D CT scans are stored in Niftii file format, where the number of slices is around 50- 400, and the image size per slice of 512x512 pixels. However, slices variations can be problematic for automatic analysis using CNN. For that we have standardized the size of input images to 100x100x50.

• **3D CNN as features Extractor:**

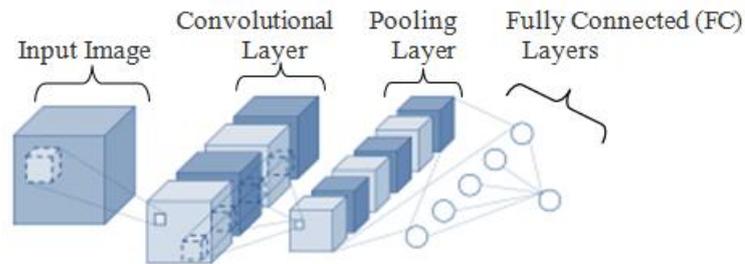
Generally, CNN architecture consist of three main types of layers (convolutional layers, pooling layers and fully connected layers) that are stacked on top of each other to form a full CNN model.

1) Convolutional layer: the convolutional layer is always the first layer of CNN, it serves as features extractor, and it consists of several features maps. Each neuron in a feature map is connected to a small region, called the local receptive field, through a set of shared weights and a single shared bias.

Two main advantages of using Convolutional layer instead of fully connected layer. The first one is parameters sharing; where sharing weights and bias are equal for all neurons in a feature map, which greatly reduces the number of parameters that in turn will result in faster training, and, ultimately, will help to build deep networks [8].The second advantage is sparsity of connections since in each layer, each output value depends only on a small number of inputs.

2) Pooling layer: The pooling layer is usually applied after convolutional layer with the aim of reducing the spatial resolution of the feature maps, speeding the computation and extracting prominent features. The most used technique for pooling is Max-pooling, where each pooling unit is equal to the largest element in a receptive field.

3) Fully connected (FC) layer: The fully connected layer input must be a vector, so we must first flatten the output features from the convolutional and pooling layers. Then, we may pass them to the output layer, where we usually use a softmax classifier or sigmoid to predict the input class label.



**Fig.1.** An example of CNN architecture.

CNNs are mainly developed for 2D images, however in the medical domain most images are 3D tensors. The traditional way to solve this problem is to convert 3D image into 2D representation, but this method leads to a loss of information, which in turn limits the performance of the overall system [4]. For that, we propose a 3D CNN model with the following architecture.

**Table 2.** Our 3D CNN architecture.

Name	Type	Kernel dims	Activation	Output dims
Data	Input	-	-	100x100x50
Conv3D_1	Convolution	5x5x5	ReLu	100x100x50x64
Max pool3D_1	Max pooling	3x3x3	-	33x33x50x21
Conv3D_2	Convolution	5x5x5	ReLu	33x33x50x64
Max Pool3D_1	Max pooling	3x3x3	-	11x11x50x21
FC1	Fully connected	-	ReLu	1024
FC2	Fully connected	-	Softmax	2

As shown in the table above, our 3D CNN has about 130 million parameters. It consists of two convolutional layers with the same padding and 64 filters with a kernel size of 5x5x5; each Convolutional layer is followed by a max-pooling layer with overlapping  $3 \times 3 \times 3$  windows, and two fully connected layers with 1024, and 2 neurons respectively.

At the first fully connected layer and in each convolutional layer, we have chosen to use a Rectified Linear Unit (ReLU)[9] as an activation function since some recent work on image recognition, has found considerable benefit in using rectified linear units through much of the network. In addition, we have adopted dropout technique with a rate of 0.5 to the first fully connected layer; this prevents hidden units from complex co-adaptations that might ultimately lead to overfitting.

Our 3D CNN was trained on the training set, with “Categorical Cross-entropy” as loss function and Adam [10] as optimizer. We have used a mini-batch size of 5 samples, and a number of epochs up to 10, where the best model was selected based on the validation loss.

However, this trained model is used only as features extractor. Therefore, we have replaced the fully connected layers (FC1 and FC2) with the SVM classifier, which takes the flattened output from the rest layers as a new feature vector.

- **Support Vector Machine (SVM) classifier:**

The SVM [11] is one of the most powerful and adopted algorithm for the classification task. Its main idea is to find the optimal hyper-plane that maximizes the margin between two classes in a given dataset.

To predict the probability of a TB patient having resistant form of tuberculosis, we have used the SVM classifier with a linear kernel. This classifier takes as input the extracted features from the pre-trained 3D CNN (with a size of 127050), and trained on the whole training set.

## 4 Results :

In this section, we present the performance results of our model on the test set:

- **Multi-Drug Resistant (MDR) Detection:**

MDR detection task consists on assigning the probability of each patient having MDR. For that, two metrics was used to evaluate model performance: Accuracy and Area Under the Curve (AUC) obtained from the ROC-curves produced with the submitted probabilities. The following table provides the results obtained by our model and the best run for the task of MDR detection.

**Table 3.** Test results obtained by our model, and best submissions for MDR detection.

Metrics	MDR Detection	
	LIST model	VISTA@UEvora model
Accuracy	0.4576	0.6144
AUC	0.5029	0.6178

Our model gives good results on MDR detection task with an AUC of 0.5029 and an accuracy of 0.4576. However, the best model has achieved an accuracy of 0.6144 and AUC of 0.6178. For that we aim to improve the performance of our model by using the provided masks[12] to reduce the size of 3D volumes before resampling them. With this we can focus on the lung tissue and we can reduce the noise from the rest of the body.

## 5 Conclusion:

In this paper, we investigate the use of a combination of 3D-CNN and SVM classifier to tackle the problem of MDR detection. Where, the 3D-CNN is used to extract relevant features from patients CT scans, while the SVM is used for the classification step. However, the test results of our submitted run are not satisfactory comparatively with the results obtained by the best model. Therefore, we aim to improve the performance of our model by adding more training data using the provided masks to reduce the size of 3D volumes before resampling them. With this we can focus on the lung tissue and we can reduce the noise from the rest of the body.

## References

1. Tuberculosis Fact sheet N°104". WHO. October 2015. Archived from the original on 23 August 2012. Retrieved 11 February 2016.
2. Dicente Cid, Y., Liauchuk, V., Kovalev, V., Müller, H. :Overview of ImageCLEFtuberculosis 2018 - Detecting multi-drug resistance, classifying tuberculosis type, and assessing severity score, CLEF working notes, CEUR, 2018.
3. Ionescu, B., Müller, H., Villegas, M., García Seco de Herrera, A., Eickhoff, C., Andrearczyk, V., Dicente Cid, Y., Liauchuk, V., Kovalev, V., Hasan, SA., Ling, Y., Farri,

- O., Liu, J., Lungren, M., Dang-Nguyen, D., Piras, L., Riegler, M., Zhou, L., Lux, M., Gurrin, M. : Overview of ImageCLEF 2018: Challenges, Datasets and Evaluation, Proceedings of the Ninth International Conference of the CLEF Association (CLEF 2018).
4. Huang, X., Shan, J., Vaidya, V. :Lung nodule detection in CT using 3D convolutional neural networks, 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pp. 379-383, VIC, Melbourne, (2017)
  5. Figueira Silva, J., Miguel Silva, J., Pinho, E., Costa, C. :3D-CNN in Drug Resistance Detection and Tuberculosis Classification, , CLEF working notes, CEUR, 2018.
  6. Golan, R., Jacob, C., Denzinger, J., :Lung nodule detection in ct images using deep convolutional neural networks,” in 2016 International Joint Conference on Neural Networks (IJCNN), pp. 243–250, Vancouver, BC, (2016).
  7. Yichuan Tang. 2013. Deep learning using linear support vector machines. arXiv preprint arXiv:1306.0239 (2013)
  8. Michael, A. Nielsen : Neural Networks and Deep Learning. Chapter 6. Determination Press (2015)
  9. Nair, V., Hinton, G. E. :Rectified linear units improve restricted Boltzmann machines. In ICML’10 (2010)
  10. Kingma, D.P., Ba, J.: Adam: A Method for Stochastic Optimization. CoRR abs/1412.6980 (2014)
  11. Cortes, C., Vapnik, V.: Support-vector Networks. Machine Learning, vol. 20, pp. 273-297, (1995).
  12. Yashin Dicente Cid, Oscar A. Jiménez-del-Toro, Adrien Depeursinge, and Henning Müller, Efficient and fully automatic segmentation of the lungs in CT volumes. In: Goksel, O., et al. (eds.) Proceedings of the VISCERAL Challenge at ISBI. No. 1390 in CEUR Workshop Proceedings (Apr 2015)