# AI-MIA: COVID-19 Detection & Severity Analysis through Medical Imaging

Dimitrios Kollias[1], Anastasios Arsenos[2], and Stefanos Kollias[2]

[1] Queen Mary University of London, UK,
d.kollias@qmul.ac.uk
[2] National Technical University of Athens, Greece

**Abstract.** This paper presents the baseline approach for the organized 2nd Covid-19 Competition, occurring in the framework of the AIMIA Workshop in the European Conference on Computer Vision (ECCV 2022). It presents the COV19-CT-DB database which is annotated for COVID-19 detection, consisting of about 7,700 3-D CT scans. Part of the database consisting of Covid-19 cases is further annotated in terms of four Covid-19 severity conditions. We have split the database and the latter part of it in training, validation and test datasets. The former two datasets are used for training and validation of machine learning models, while the latter is used for evaluation of the developed models. The baseline approach consists of a deep learning approach, based on a CNN-RNN network and report its performance on the COVID19-CT-DB database. The paper presents the results of both Challenges organised in the framework of the Competition, also compared to the performance of the baseline scheme.

**Keywords:** AIMIA, COVID19-CT-DB, COVID19 detection, COVID19 severity detection, deep learning, CT scans, medical imaging

## 1 Introduction

Medical imaging seeks to reveal internal structures of human body, as well as to diagnose and treat diseases. Radiological quantitative image analysis constitutes a significant prognostic factor for diagnosis of diseases, such as lung cancer and Covid-19. For example, CT scan is an essential test for the diagnosis, staging and follow-up of patients suffering from lung cancer in its various histologies. Lung cancer constitutes a top public health issue, with the tendency being towards patient-centered control strategies. Especially in the chest imaging field, there has been a large effort to develop and apply computer-aided diagnosis systems for the detection of lung lesions patterns.

Deep learning (DL) approaches applied to the quantitative analysis of multimodal medical data offer very promising results. Current targets include building robust and interpretable AI and DL models for analyzing CT scans for the discrimination of immune-inflammatory patterns and their differential diagnosis with SARS-CoV-2-related or other findings. Moreover, detection of the new coronavirus (Covid-19) is of extreme significance for our Society. Currently, the gold standard reverse transcription polymerase chain reaction (RT-PCR) test is identified as the primary diagnostic tool for the Covid-19; however, it is time consuming with a certain number of false positive results. The

chest CT has been identified as an important diagnostic tool to complement the RT-PCR test, assisting the diagnosis accuracy and helping people take the right action so as to quickly get appropriate treatment. That is why there is on-going research worldwide to create automated systems that can predict Covid-19 from chest CT scans and x-rays.

Recent approaches use Deep Neural Networks for medical imaging problems, including segmentation and automatic detection of the pneumonia region in lungs. The public datasets of RSNA pneumonia detection and LUNA have been used as training data, as well as other Covid-19 datasets. Then a classification network is generally used to discriminate different viral pneumonia & Covid-19 cases, based on analysis of chest x-rays, or CT scans.

Various projects are implemented in this framework in Europe, US, China and worldwide. However, a rather fragmented approach is followed: projects are based on specific datasets, selected from smaller, or larger numbers of hospitals, with no proof of good performance generalization over different datasets and clinical environments. A DL methodology which we have developed using latent information extraction from trained DNNs has been very successful for diagnosis of Covid-19 through unified chest CT scans and x-rays analysis provided by different medical centers and medical environments.

At the time of CT scan recording, several slices are captured from each person suspected of COVID-19. The large volume of CT scan images calls for a high workload on physicians and radiologists to diagnose COVID-19. Taking this into account and also the rapid increase in number of new and suspected COVID-19 cases, it is evident that there is a need for using machine and deep learning for detecting COVID-19 in CT scans. Such approaches require data to be trained on. Therefore, a few databases have been developed consisting of CT scans. However, new data sets with large numbers of 3-D CT scans are needed, so that researchers can train and develop COVID-19 diagnosis systems and trustfully evaluate their performance.

The current paper presents a baseline approach for the Competition which is a part of the Workshop "AI-enabled Medical Image Analysis – Digital Pathology & Radiology/COVID19 (AIMIA)" which occurs in conjunction with the European Conference on Computer Vision (ECCV) 2022 in Tel-Aviv, Israel, October 23- 24, 2022.

The AIMIA-Workshop context is as follows. Deep Learning has made rapid advances in the performance of medical image analysis challenging physicians in their traditional fields. In the pathology and radiology fields, in particular, automated procedures can help to reduce the workload of pathologists and radiologists and increase the accuracy and precision of medical image assessment, which is often considered subjective and not optimally reproducible. In addition, Deep Learning and Computer Vision demonstrate the ability/potential to extract more clinically relevant information from medical images than what is possible in current routine clinical practice by human assessors. Nevertheless, considerable development and validation work lie ahead before AI-based methods can be fully ready for integrated into medical departments.

The workshop on AI-enabled medical image analysis (AIMIA) at ECCV 2022 aims to foster discussion and presentation of ideas to tackle the challenges of whole slide image and CT/MRI/X-ray analysis/processing and identify research opportunities in the context of Digital Pathology and Radiology/COVID19.

The COV19D Competition is based on a new large database of chest CT scan series that is manually annotated for Covid-19/non-Covid-19, as well as Covid-19 severity ndiagnosis. The training and validation partitions along with their annotations are provided to the participating teams to develop AI/ML/DL models for Covid-19/non-Covid-19 and Covid-19 severity prediction. Performance of approaches will be evaluated on the test sets. The COV19-CT-DB is a new large database with about 7,700 3-D CT scans, annotated for COVID-19 infection.

The rest of the paper is organized as follows. Section 2 presents former work on which the presented baseline has been based. Section 3 presents the database created and used in the Competition. The ML approach and the pre-processing steps are described in Section 4. The obtained results, as well as the performance of the winning approaches in both Challenges are presented in Section 5. Conclusions and future work are described in Section 6.

## 2   Related Work

In [9] a CNN plus RNN network was used, taking as input CT scan images and discriminating between COVID-19 and non-COVID-19 cases. In [16], the authors employed a variety of 3-D ResNet models for detecting COVID-19 and distinguishing it from other common pneumonia (CP) and normal cases, using volumetric 3-D CT scans. In [22], a weakly supervised deep learning framework was suggested using 3-D CT volumes for COVID-19 classification and lesion localization. A pre-trained UNet was utilized for segmenting the lung region of each CT scan slice; the latter was fed into a 3-D DNN that provided the classification outputs. The presented approach is based on a CNN-RNN architecture that performs 3-D CT scan analysis. The method follows our previous work [12,14,13,19] on developing deep neural architectures for predicting COVID-19, as well as neurodegenerative and other [15,13] diseases and abnormal situations [4].

These architectures have been applied for: a) prediction of Parkinson's, based on datasets of MRI and DaTScans, either created in collaboration with the Georgios Gennimatas Hospital (GGH) in Athens [13,19], or provided by the PPMI study sponsored by M. J. Fox for Parkinson's Research [15], b) prediction of COVID-19, based on CT chest scans, scan series, or x-rays, either collected from the public domain, or aggregated in collaboration with the Hellenic Ministry of Health and The Greek National Infrastructures for Research and Technology [12]. Other architectures, based on capsule nets [18,5] can also be used in this framework.

The current Competition follows and extends the MIA-COV19D Workshop we successfully organized virtually in ICCV 2021, on October 11, 2021. 35 Teams participated in the COV19D Competition; 18 Teams submitted their results and 12 Teams scored higher than the baseline. 21 presentations were made in the Workshop, including 3 invited talks [11].

In general, some workshops on medical image analysis have taken place in the framework of CVPR/ICCV/ECCV Conferences in the last three years. In particular: in ICCV 2019 the 'Visual Recognition for Medical Images' Workshop took place, including papers on the use of deep learning in medical images for diagnosis of diseases. Similarly, in CVPR 2019 and 2020 the following series of Workshops took place: 'Med-

ical Computer Vision' and 'Computer Vision for Microscopy Image Analysis', dealing with general medical analysis topics. In addition, the 'Bioimage Computing' took place in ECCV 2020. Moreover, the issue of explainable DL and AI has been a topic of other Workshops (e.g., CVPR 2019; ICPR 2020 and 2022; ECAI Tailor 2020)

## 3    The COV19-CT-DB Database

The COVID19-CT-Database (COV19-CT-DB) consists of chest CT scans that are annotated for the existence of COVID-19. COV19-CT-DB includes 3-D chest CT scans annotated for existence of COVID-19. Data collection was conducted in the period from September 1 2020 to November 30 2021. It consists of 1,650 COVID and 6,100 non-COVID chest CT scan series, which correspond to a high number of patients (more than 1150) and subjects (more than 2600). In total, 724,273 slices correspond to the CT scans of the COVID-19 category and 1,775,727 slices correspond to the non COVID-19 category [1].

Annotation of each CT slice has been performed by 4 very experienced (each with over 20 years of experience) medical experts; two radiologists and two pulmonologists. Labels provided by the 4 experts showed a high degree of agreement (around 98%). Each of the 3-D scans includes different number of slices, ranging from 50 to 700. This variation in number of slices is due to context of CT scanning. The context is defined in terms of various factors, such as the accuracy asked by the doctor who ordered the scan, the characteristics of the CT scanner that is used, or specific subject's features, e.g., weight and age.
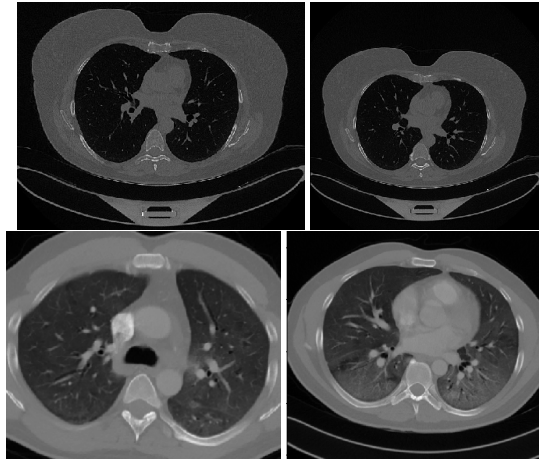


**Fig. 1.** Four CT scan slices, two from a non-COVID-19 CT scan (on the top row) and two from a COVID-19 scan (on the bottom row) including bilateral ground glass regions in lower lung lobes.

Figure 1 shows four CT scan slices, two from a non-COVID-19 CT scan (on the top row) and two from a COVID-19 scan (on the bottom row). Bilateral ground glass

regions are seen especially in lower lung lobes in the COVID-19 slices. The database has been split in training, validation and testing sets. The training set contains, in total, 1992 3-D CT scans. The validation set consists of 494 3-D CT scans. The number of COVID-19 and of Non-COVID-19 cases in each set are shown in Table 1. There are different numbers of CT slices per CT scan, ranging from 50 to 700.

**Table 1.** Data samples in each Set

| Set | Training | Validation |
|---|---|---|
| COVID-19 | 882 | 215 |
| Non-COVID-19 | 1110 | 289 |

A further split of the COVID-19 cases has been implemented, based on the severity of COVID-19, as annotated by four medical experts, in the range from 1 to 4, with 4 denoting the critical status. Table 2 describes each of these categories [17].

**Table 2.** Description of the Severity Categories

| Category | Severity | Description |
|---|---|---|
| 1 | Mild | Few or no Ground glass opacities. Pulmonary parenchymal involvement $\leq 25\%$ or absence |
| 2 | Moderate | Ground glass opacities. Pulmonary parenchymal involvement $25 - 50\%$ |
| 3 | Severe | Ground glass opacities. Pulmonary parenchymal involvement $50 - 75\%$ |
| 4 | Critical | Ground glass opacities. Pulmonary parenchymal involvement $\geq 75\%$ |

In particular, parts of the COV19-CT-DB COVID-19 training and validation datasets have been accordingly split for severity classification in training, validation and testing sets.

The training set contains, in total, 258 3-D CT scans. The validation set consists of 61 3-D CT scans. The numbers of scans in each severity COVID-19 class in these sets are shown in Table 3.

**Table 3.** Data samples in each Severity Class

| Severity Class | Training | Validation |
|---|---|---|
| 1 | 85 | 22 |
| 2 | 62 | 10 |
| 3 | 85 | 22 |
| 3 | 26 | 5 |

## 4   The Deep Learning Approach

### 4.1   3-D Analysis and COVID-19 Diagnosis

The input sequence is a 3-D signal, consisting of a series of chest CT slices, i.e., 2-D images, the number of which is varying, depending on the context of CT scanning. The context is defined in terms of various requirements, such as the accuracy asked by the doctor who ordered the scan, the characteristics of the CT scanner that is used, or the specific subject's features, e.g., weight and age.

The baseline approach is a CNN-RNN architecture, as shown in Figure 2. At first all input CT scans are padded to have length $t$ (i.e., consist of $t$ slices). The whole (unsegmented) sequence of 2-D slices of a CT-scan are fed as input to the CNN part. Thus the CNN part performs local, per 2-D slice, analysis, extracting features mainly from the lung regions. The target is to make diagnosis using the whole 3-D CT scan series, similarly to the annotations provided by the medical experts. The RNN part provides this decision, analyzing the CNN features of the whole 3-D CT scan, sequentially moving from slice 0 to slice $t - 1$.
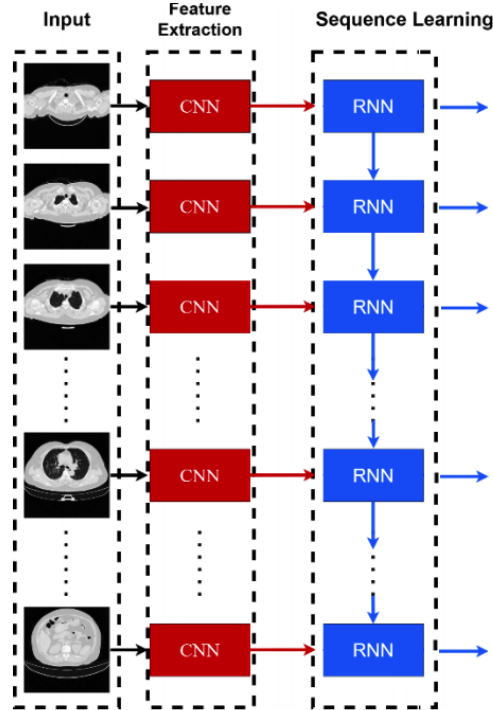


**Fig. 2.** The CNN-RNN baseline model

In the case of the COVID19 Detection Challenge, the outputs of the RNN part feed the output layer -which consists of 2 units- that uses a softmax activation function and

provides the final COVID-19 diagnosis. In the case of the COVID19 Severity Detection Challenge, the outputs of the RNN part feed the output layer -which consists of 4 units- that uses a softmax activation function and provides the final COVID-19 severity diagnosis.

In this way, the CNN-RNN network outputs a probability for each CT scan slice; the CNN-RNN is followed by a voting scheme that makes the final decision. In the case of the COVID19 Detection Challenge the voting scheme can be either a majority voting or an at-least one voting (i.e., if at least one slice in the scan is predicted as COVID-19, then the whole CT scan is diagnosed as COVID-19, and if all slices in the scan are predicted as non-COVID-19, then the whole CT scan is diagnosed as non-COVID-19). In the case of the COVID19 Severity Detection Challenge, the voting scheme is a majority voting.

The main downside of this model is that there exists only one label for the whole CT scan and there are no labels for each CT scan slice. Thus, the presented model analyzes the whole CT scan, based on information extracted from each slice.

### 4.2    Pre-Processing, Implementation Details & Evaluation Metrics

At first, CT images were extracted from DICOM files. Then, the voxel intensity values were clipped using a window/level of 350 Hounsfield units (HU)/$-1150$ HU and normalized to the range of $[0, 1]$.

Regarding implementation of the proposed methodology: i) we utilized ResNet50 as CNN model, stacking on top of it a global average pooling layer, a batch normalization layer and dropout (with keep probability 0.8); ii) we used a single one-directional GRU layer consisting of 128 units as RNN model. The model was fed with 3-D CT scans composed of the CT slices; each slice was resized from its original size of $512 \times 512 \times 3$ to $224 \times 224 \times 3$. As a voting scheme for the COVID19 Detection Challenge, we used the at-least one.

Batch size was equal to 5 (i.e, at each iteration our model processed 5 CT scans) and the input length 't' was 700 (the maximum number of slices found across all CT scans). Softmax cross entropy was the utilized loss function for training the model. Adam optimizer was used with learning rate $10^{-4}$. Training was performed on a Tesla V100 32GB GPU.

Evaluation of the achieved performance in both Challenges was made in terms of macro F1 score. The macro F1 score is defined as the unweighted average of the class-wise/label-wise F1-scores. In the first Challenge this is the unweighted average of the COVID-19 class F1 score and of the non-COVID-19 class F1 score. In the second Challenge this was similarly defined as the unweighted average of the F1 scores of the severity categories.

## 5    The COV19D Competition Results

COV19D Competition consisted of two sub-challenges, namely, COVID19 Detection Challenge and COVID19 Severity Challenge. Training and validation data from the COVID-19 category (as indicatively shown in Figures 3 and 4) and the non COVID-19
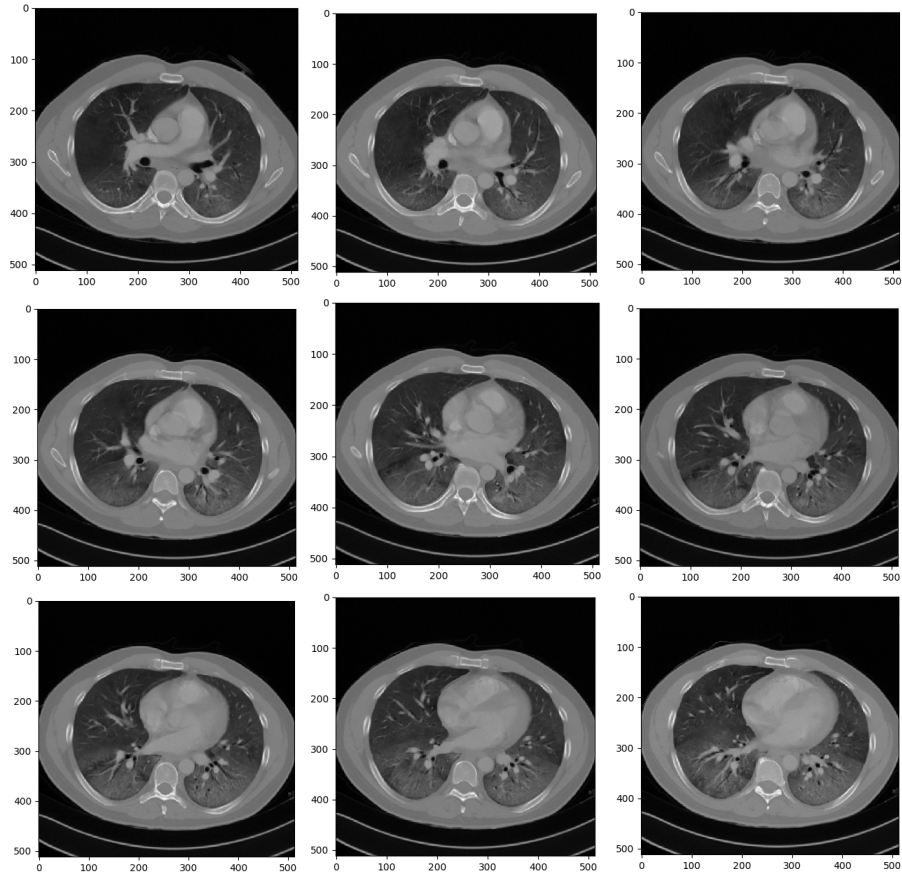
**Fig. 3.** 9 slices belonging to a CT-scan of the COVID-19 category, showing bilateral ground glass areas in lower lung lobes.

category (as indicatively shown in Figure 5) were provided and used by the participating teams for training their models.

Nine teams that submitted results scored higher than the baseline and made valid submissions regarding the COVID19 Detection Challenge. The achieved performance by each team is shown in Table 4. Six teams made valid submissions and surpassed the baseline for COVID19 Severity Challenge. The corresponding performance of each team is shown in Table 5.

### 5.1    COVID19 Detection Challenge Results

The winners of this Competition were **ACVLab** from National Cheng Kung University, Taiwan and **FDVTS** from Fudan University, China. The runner-up was **MDAP** from the University of Melbourne. All results provided by the 9 teams are shown in Table 4; links of the related Github codes and arXiv papers are also provided. In the following we give
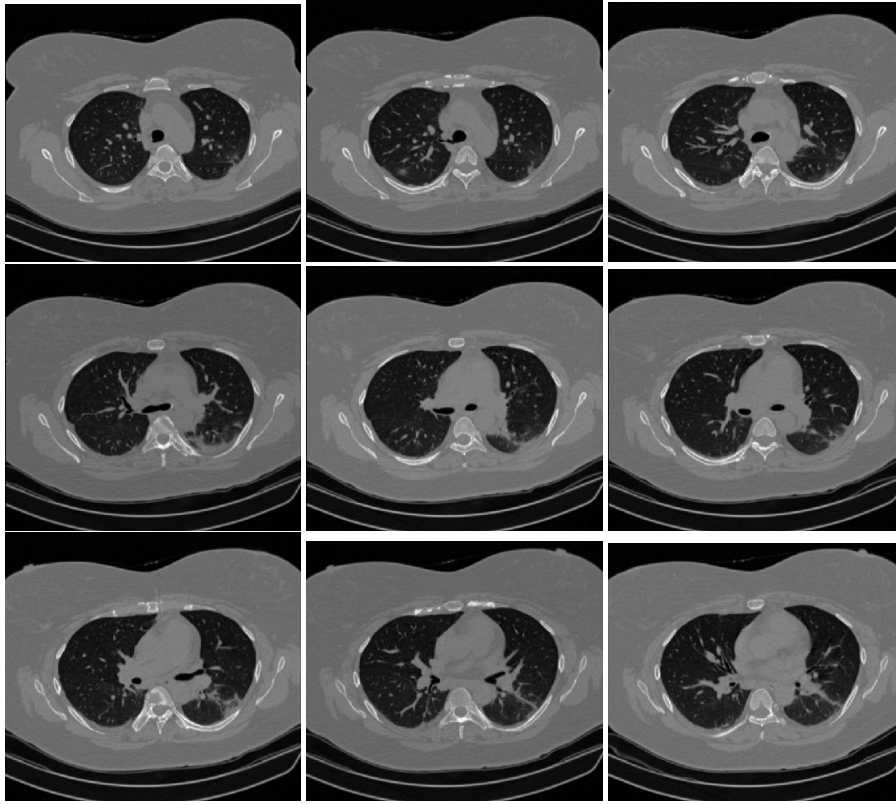
**Fig. 4.** 9 slices belonging to a CT-scan of the COVID-19 type, showing pneumonia bilateral thickening filtrates.

a short description of the best methods that the top three teams implemented and used. The index on the right of each team represents their official rank in the classification task.

Team **ACVLab** 1 proposed two spatiotemporal feature learning methodologies based on 2D-CNN + LSTM and 2D-CNN + Swin Transformer. At first, they trained the 2D CNN network which consisted of EfficientNet-b3a as the backbone and a fully connected layer with embedding size 224. Then they used the resulted embeddings from the 2D CNN network as input to the LSTM model which consisted of four Bi-LSTM layers with 128 hidden units. For the swin model they followed the same procedure and used the slice's embeddings from the 2D CNN model as input. In order to deal with the varying size of slices of the CT-scans the randomly selected 120 slices for the LSTM model and 224 slices for the swin model.

Team **FDVTS** 1 achieved the top performance in this task along with ACVLab team. In their approach, they employed a 3D Contrastive Mixup Classification network for COVID-19 diagnosis on chest CT images. They used the same network for both sub-challenges. Their network named CMC-COV19D employed the contrastive
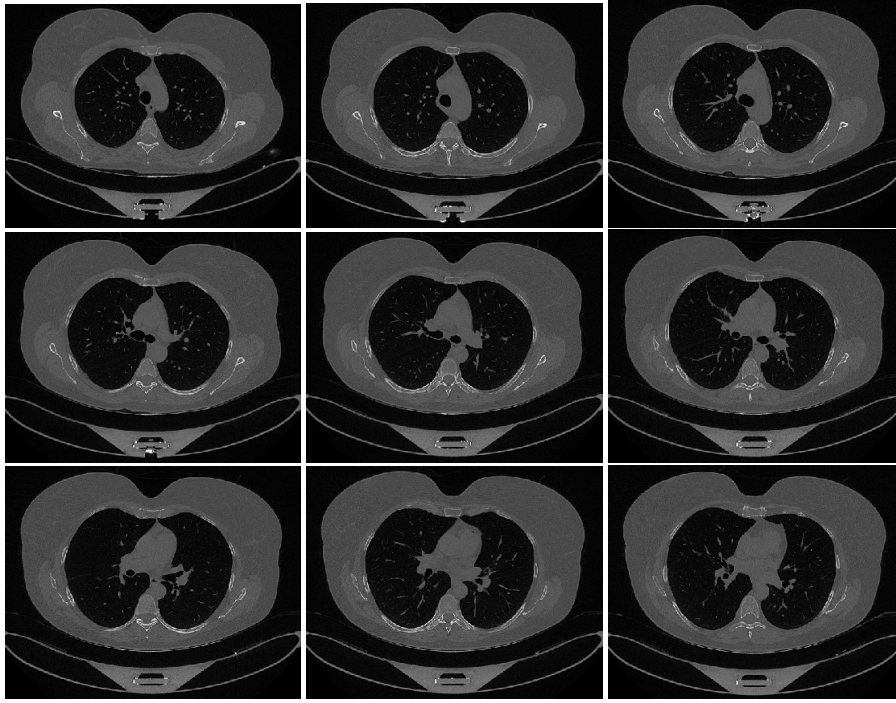
**Fig. 5.** 9 slices belonging to a CT scan of the non COVID-19 category

representation learning (CRL) as an auxiliary task to learn more discriminative representations of COVID-19. The CRL was comprised of the following components. 1) A stochastic data augmentation module, which transformed an input CT into a randomly augmented sample. They generated two augmented volumes from each input CT. 2) A base encoder, which mapped the augmented CT sample to a representation vector. 3) A projection network, which was used to map the representation vector to a relative low-dimension vector. 4) A classifier, which classified the vector to the final prediction.

Team **MDAP** 2 was the runner-up of the COVID19 Detection Challenge. Their proposed methodology named Cov3d, was a three dimensional convolutional neural network able to detect the presence and severity of COVID19 from chest CT scans. In order to cope with the varying input size problem they used 1D cubic interpolation to set the depth of the CT-scans to a fixed number. Cov3d used a neural network model analogous to a two dimensional ResNet-18 model but with 3D convolutional and pooling operations instead of their 2D equivalents. In particular, Cov3d used the ResNet-3D-18 model. This model has been pre-trained on the Kinetics-400 dataset which classifies short video clips.

**Table 4.** COVID19 Detection Challenge Results: F1 Score in %; the best performing submission is in bold

| Teams | Submission # | Macro F1 | F1 (NON-COVID) | F1 (COVID) | Github |
|---|---|---|---|---|---|
| ACVLab [7] | 1 | 86.8 | 96.66 | 76.95 | |
| | 2 | 85.75 | 96.26 | 75.22 | |
| | **3** | **89.11** | **97.45** | **80.78** | link |
| | 4 | 86.72 | 96.64 | 76.80 | |
| | 5 | 86.75 | 96.61 | 76.89 | |
| FDVTS [6] | 1 | 88.61 | 97.13 | 80.09 | |
| | **2** | **89.11** | **97.31** | **80.92** | |
| | 3 | 88.40 | 97.08 | 79.73 | link |
| | 4 | 88.65 | 97.14 | 80.15 | |
| | 5 | 88.70 | 97.15 | 80.24 | |
| MDAP [21] | **1** | **87.87** | **96.95** | **78.80** | |
| | 2 | 87.60 | 96.88 | 78.31 | |
| | 3 | 85.92 | 96.26 | 75.59 | link |
| | 4 | 87.12 | 96.76 | 77.47 | |
| | 5 | 85.97 | 96.32 | 75.61 | |
| Code 1055 [10] | 1 | 82.13 | 95.04 | 69.22 | |
| | 2 | 86.02 | 96.47 | 75.57 | |
| | 3 | 83.63 | 95.47 | 71.78 | link |
| | 4 | 83.93 | 95.57 | 72.29 | |
| | **5** | **86.18** | **96.37** | **76.00** | |
| CNR-IEMN [3] | 1 | 81.91 | 95.34 | 68.48 | |
| | 2 | 77.03 | 92.94 | 61.12 | |
| | 3 | 83.12 | 95.55 | 70.69 | link |
| | 4 | 84.12 | 95.90 | 72.34 | |
| | **5** | **84.37** | **95.98** | **72.76** | |
| Dslab [23] | 1 | 81.69 | 95.74 | 67.64 | |
| | 2 | 81.04 | 95.47 | 66.61 | |
| | 3 | 82.38 | 95.85 | 68.92 | link |
| | **4** | **83.78** | **96.22** | **71.33** | |
| | 5 | 83.53 | 96.19 | 70.87 | |
| Jovision-Deepcam [20] | 1 | 77.71 | 93.13 | 62.29 | |
| | 2 | 80.13 | 94.24 | 66.01 | |
| | **3** | **80.82** | **94.56** | **67.07** | link |
| | 4 | 77.66 | 93.01 | 62.31 | |
| | 5 | 75.51 | 92.36 | 58.65 | |
| ICL [8] | **1** | **79.55** | **93.77** | **65.34** | link |
| etro [2] | **1** | **78.72** | **93.48** | **63.95** | link |
| baseline | **1** | **69** | **83.62** | **54.38** | |

## 5.2   COVID19 Severity Detection Challenge Results

The winner of this Competition was **FDVTS** from Fudan University, China. The runner-up (with a slight difference from the winning teams) was **Code 1055** from Augsburg University, Germany. The team that ranked third was **CNR-IEMN**.

All results of the 6 teams are shown in Table 5; links are also provided to both, of the related Github codes and arXiv papers. In the following we provide a short description of the best methods that the top three teams implemented and used. The index on the right of each team represents their official rank in the classification task.

**Table 5.** COVID19 Severity Detection Challenge Results: F1 Score in %; the best performing submission is in bold

| Teams | Submission # | Macro F1 | F1 (MILD) | F1 (MODERATE) | F1 (SEVERE) | F1 (CRITICAL) | Github |
|---|---|---|---|---|---|---|---|
| FDVTS [6] | 1 | 46.21 | 53.90 | 35.66 | 56.57 | 38.71 | |
| | 2 | 41.15 | 62.50 | 35.30 | 51.81 | 15.00 | |
| | 3 | 47.14 | 62.94 | 35.21 | 60.61 | 29.79 | link |
| | 4 | 45.96 | 53.33 | 39.16 | 52.63 | 38.71 | |
| | **5** | **51.76** | **58.97** | **44.53** | **58.89** | **44.64** | |
| Code 1055 [10] | 1 | 46.67 | 58.16 | 40.85 | 55.24 | 32.43 | |
| | 2 | 49.36 | 56.95 | 30.14 | 52.48 | 57.90 | |
| | **3** | **51.48** | **61.14** | **34.06** | **61.91** | **48.83** | link |
| | 4 | 46.01 | 60.00 | 27.78 | 52.78 | 43.48 | |
| | 5 | 49.90 | 63.01 | 28.99 | 48.52 | 59.09 | |
| CNR-IEMN [3] | 1 | 39.89 | 34.86 | 35.44 | 58.00 | 31.25 | |
| | 2 | 43.04 | 53.09 | 18.52 | 63.52 | 37.04 | |
| | **3** | **47.11** | **55.67** | **37.88** | **55.46** | **39.46** | link |
| | 4 | 44.37 | 53.42 | 26.67 | 60.36 | 37.04 | |
| | 5 | 45.49 | 54.66 | 26.79 | 63.48 | 37.04 | |
| MDAP [21] | 1 | 37.57 | 62.25 | 36.64 | 51.38 | 0.00 | |
| | 2 | 33.94 | 55.48 | 17.65 | 55.51 | 7.14 | |
| | 3 | 39.84 | 64.12 | 40.56 | 50.24 | 4.45 | link |
| | 4 | 42.31 | 66.23 | 45.75 | 57.28 | 0.00 | |
| | **5** | **46.00** | **58.14** | **42.29** | **43.78** | **40.03** | |
| etro [2] | **1** | **44.92** | **57.69** | **29.57** | **62.85** | **29.57** | link |
| Jovision-Deepcam [20] | **1** | **41.49** | **62.20** | **40.24** | **38.59** | **24.93** | link |
| baseline | **1** | **40.3** | **61.10** | **39.47** | **37.07** | **23.54** | |

Team **FDVTS** 1 achieved the top performance in this task. They used the same methodology for both sub-challenges.

Team **Code 1055** 2 was the runner-up of COVID19 Severity Detection Challenge. The architecture they utilized was a three dimensional version of the recently published ConvNeXt architecture and was modified in order to be able to process the three-dimensional input tensors. This architecture was especially characterized by multiple alterations already proven to be useful in the context of Vision Transformers that are applied to the standard ResNet architecture. To make use of the ConvNeXt architecture, they replaced all 2D components with corresponding 3D components.

Team **CNR-IEMN** 3 developed an ensemble of convolutional layers with Inception models for COVID19 Severity Detection Challenge. In addition to the convolutional layers, three Inception variants were used, namely Inception-v3, Inception-v4 and Inception-Resnet. Regarding the pre-processing of the CT-scans they did the following steps: removed the non-lung slices and segmented the lung regions to remove unnecessary features, they resized the obtained slices into two volumes of size 224 × 224 × 32 and 224 × 224 × 16. The goal of the two volumes was to reduce the information lost due to resizing and to obtain two views of each CT scan. The two input volumes were fed into the convolution block.

## 6    Conclusions and Future Work

In this paper we have introduced a new large database of chest 3-D CT scans, obtained in various contexts and consisting of different numbers of CT scan slices. We have also developed a deep neural network, based on a CNN-RNN architecture and used it for COVID-19 diagnosis on this database.

The scope of the paper has been to present a baseline scheme regarding the performance that can be achieved based on analysis of the COV19-CT-DB database. This was used in the Competition part of the ECCV 2022 AIMIA Workshop, which included two Challenges: one targeting COVID-19 detection and another targeting COVID-19 severity detection.

The paper also presented the results of the methods developed in the framework of both Challenges and briefly described the main methodology followed by each of the three top ranking teams in the Challenges. By examining these results, it can be concluded that the detection of COVID-19 severity, based on a rather small of cases for each severity category, has been a harder problem to tackle. More data will be released in the future for each examined category, so that better performance can be reached.

The model presented as baseline in the paper will be the basis for future expansion towards more transparent modelling of COVID-19 diagnosis.

# References

1. Arsenos, A., Kollias, D., Kollias, S.: A large imaging database and novel deep neural architecture for covid-19 diagnosis. In: 2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP). pp. 1–5. IEEE (2022)

2. Berenguer, A.D., Mukherjee, T., Bossa, M., Deligiannis, N., Sahli, H.: Representation learning with information theory for covid-19 detection (2022). https://doi.org/10.48550/ARXIV.2207.01437, https://arxiv.org/abs/2207.01437

3. Bougourzi, F., Distante, C., Dornaika, F., Taleb-Ahmed, A.: Ensemble cnn models for covid-19 recognition and severity perdition from 3d ct-scan (2022). https://doi.org/10.48550/ARXIV.2206.15431, https://arxiv.org/abs/2206.15431

4. Caliva, F., De Ribeiro, F.S., Mylonakis, A., Demazi'ere, C., Vinai, P., Leontidis, G., Kollias, S.: A deep learning approach to anomaly detection in nuclear reactors. In: 2018 International joint conference on neural networks (IJCNN). pp. 1–8. IEEE (2018)

5. De Sousa Ribeiro, F., Leontidis, G., Kollias, S.: Introducing routing uncertainty in capsule networks. Advances in Neural Information Processing Systems **33**, 6490–6502 (2020)

6. Hou, J., Xu, J., Feng, R., Zhang, Y.: Fdvts's solution for 2nd cov19d competition on covid-19 detection and severity analysis (2022). https://doi.org/10.48550/ARXIV.2207.01758, https://arxiv.org/abs/2207.01758

7. Hsu, C.C., Tsai, C.H., Chen, G.L., Ma, S.D., Tai, S.C.: Spatiotemporal feature learning based on two-step lstm and transformer for ct scans (2022). https://doi.org/10.48550/ARXIV.2207.01579, https://arxiv.org/abs/2207.01579

8. Jung, O., Kang, D.U., Kim, G., Chun, S.Y.: Adaptive glcm sampling for transformer-based covid-19 detection on ct (2022). https://doi.org/10.48550/ARXIV.2207.01520, https://arxiv.org/abs/2207.01520

9. Khadidos, A., Khadidos, A.O., Kannan, S., Natarajan, Y., Mohanty, S.N., Tsaramirsis, G.: Analysis of covid-19 infections on a ct image using deepsense model. Frontiers in Public Health **8** (2020)

10. Kienzle, D., Lorenz, J., Schön, R., Ludwig, K., Lienhart, R.: Custom pretrainings and adapted 3d-convnext architecture for covid detection and severity prediction (2022). https://doi.org/10.48550/ARXIV.2206.15073, https://arxiv.org/abs/2206.15073

11. Kollias, D., Arsenos, A., Soukissian, L., Kollias, S.: Mia-cov19d: Covid-19 detection through 3-d chest ct image analysis. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 537–544 (2021)

12. Kollias, D., Bouas, N., Vlaxos, Y., Brillakis, V., Seferis, M., Kollia, I., Sukissian, L., Wingate, J., Kollias, S.: Deep transparent prediction through latent representation analysis. arXiv preprint arXiv:2009.07044 (2020)

13. Kollias, D., Tagaris, A., Stafylopatis, A., Kollias, S., Tagaris, G.: Deep neural architectures for prediction in healthcare. Complex & Intelligent Systems **4**(2), 119–131 (2018)

14. Kollias, D., Vlaxos, Y., Seferis, M., Kollia, I., Sukissian, L., Wingate, J., Kollias, S.D.: Transparent adaptation in deep medical image diagnosis. In: TAILOR. pp. 251–267 (2020)

15. Kollias, S., Bidaut, L., Wingate, J., Kollia, I., et al.: A unified deep learning approach for prediction of parkinson's disease. IET Image Processing (2020)

16. Li, Y., Xia, L.: Coronavirus disease 2019 (covid-19): role of chest ct in diagnosis and management. American Journal of Roentgenology **214**(6), 1280–1286 (2020)

17. Morozov, S.P., Andreychenko, A.E., Blokhin, I.A., Gelezhe, P.B., Gonchar, A.P., Nikolaev, A.E., Pavlov, N.A., Chernina, V.Y., Gombolevskiy, V.A.: Mosmeddata: data set of 1110 chest ct scans performed during the covid-19 epidemic. Digital Diagnostics **1**(1), 49–59 (2020)

18. Ribeiro, F.D.S., Leontidis, G., Kollias, S.: Capsule routing via variational bayes. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 34, pp. 3749–3756 (2020)

19. Tagaris, A., Kollias, D., Stafylopatis, A., Tagaris, G., Kollias, S.: Machine learning for neu-rodegenerative disorder diagnosis—survey of practices and launch of benchmark dataset. International Journal on Artificial Intelligence Tools **27**(03), 1850011 (2018)
20. Tan, W., Yao, Q., Liu, J.: Two-stage covid19 classification using bert features (2022). https://doi.org/10.48550/ARXIV.2206.14861, https://arxiv.org/abs/2206.14861
21. Turnbull, R.: Cov3d: Detection of the presence and severity of COVID-19 from CT scans using 3D ResNets [Preliminary Preprint] (7 2022). https://doi.org/10.26188/20226087.v1, https://melbourne.figshare.com/articles/preprint/Cov3d_Detection_of_the_presence_and_severity_of_COVID-19_from_CT_scans_using_3D_ResNets_Preliminary_Preprint_/20226087
22. Wang, X., Deng, X., Fu, Q., Zhou, Q., Feng, J., Ma, H., Liu, W., Zheng, C.: A weakly-supervised framework for covid-19 classification and lesion localization from chest ct. IEEE transactions on medical imaging **39**(8), 2615–2625 (2020)
23. Zheng, L., Fang, J., Tang, X., Li, H., Fan, J., Wang, T., Zhou, R., Yan, Z.: Pvt-cov19d: Pyramid vision transformer for covid-19 diagnosis (2022). https://doi.org/10.48550/ARXIV.2206.15069, https://arxiv.org/abs/2206.15069