

# Pictures of X-Rays Displayed in Monitors for Deep Learning-Based COVID-19 Screening: Implications for Mobile Application Development

**Pedro Silva**

Universidade Federal de Ouro Preto

**Eduardo Luz**

Universidade Federal de Ouro Preto

**Gladston Moreira**

Universidade Federal de Ouro Preto

**Caio Gomes**

Universidade Federal de Ouro Preto

**Larissa Viana**

Universidade Federal de Ouro Preto

**Rodrigo Silva** (✉ [rodrigo.silva@ufop.edu.br](mailto:rodrigo.silva@ufop.edu.br))

Universidade Federal de Ouro Preto <https://orcid.org/0000-0003-2547-3835>

---

## Research Article

**Keywords:** COVID-19, Deep Learning, EfficientNet, Pneumonia, Chest Radiography

**Posted Date:** June 30th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-37908/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Pictures of X-Rays Displayed in Monitors for Deep Learning-Based COVID-19 Screening: Implications for Mobile Application Development

Pedro Silva<sup>a,\*</sup>, Eduardo Luz<sup>b</sup>, Gladston Moreira<sup>b</sup>, Caio Gomes<sup>c</sup>, Larissa Viana<sup>c</sup>, Rodrigo Silva<sup>b</sup>

<sup>a</sup>Graduate Program in Computer Science, Universidade Federal de Ouro Preto (UFOP), MG, Brazil

<sup>b</sup>Department of Computer Science, Universidade Federal de Ouro Preto (UFOP), MG, Brazil

<sup>c</sup>Undergraduate Program in Control and Automation Engineering, Universidade Federal de Ouro Preto (UFOP), MG, Brazil

---

## Abstract

As the world faces the COVID-19 pandemic, Artificial Intelligence, in particular, Deep Learning (DL) have been called up for help. Several recent research papers have shown the usefulness of these techniques for COVID-19 screening in Chest X-Rays (CXRs). To make this technology accessible and easy to use for the healthcare workers a natural path is to embed it into a mobile app. In these cases, however, the DL models must be prepared to receive as inputs pictures taken with the smartphones.

Trying to raise awareness about the limitations of these models in a real-world setup, in this work, a dataset of CXR pictures taken of computer monitors with smartphones is built and DL models are evaluated on it. The results show that the current models are not able to correctly classify this kind of input. In the tested setup, augmenting the dataset with such pictures has shown to mitigate the problem, but it was not enough to raise accuracy to acceptable levels. As an alternative, this work shows that it is possible to build a model that discards pictures of monitors such that the COVID-19 screening module does not have to cope with them.

*Keywords:* COVID-19, Deep Learning, EfficientNet, Pneumonia, Chest Radiography.

---

\*Corresponding author

*Email address:* [pedroh21.silva@gmail.com](mailto:pedroh21.silva@gmail.com) (Pedro Silva)

*URL:* <http://www.decom.ufop.br/csilab/> (Pedro Silva)

## 1. Introduction

In the late months of 2019, a new coronavirus, named SARS-CoV-2, started affecting people in China. The virus quickly spread to other countries, and in a short time, it became a pandemic. In February of 2020, the World Health Organization (WHO) named the disease caused by SARS-CoV-2 as COVID-19.

The COVID-19 infection may manifest itself as a flu-like illness potentially progressing to an acute respiratory distress syndrome (Araujo-Filho et al., 2020). The disease severity resulted in global public health measures to contain person-to-person viral spread (Davarpanah et al., 2020). In most countries, these measures involve social distancing and large scale testing for early disease detection and isolation of sick patients.

The Reverse-Transcriptase Polymerase Chain Reaction (RT-PCR) is, currently, the gold standard for the diagnosis of COVID-19 (Araujo-Filho et al., 2020). However, effective exclusion of COVID-19 infection requires multiple negative tests (American College of Radiology, 2020) which became scarce in the world due to the pandemic.

The scarcity of test kits has started a race to search for alternative diagnostic methods and many researchers have turned to Artificial Intelligence (AI) for help. Given the success of AI, in particular, Deep Learning (DL), in tasks of pattern recognition, the application of these techniques to Chest X-Ray (CXR) based screening of COVID-19 has become very popular (Hemdan et al., 2020; Farooq and Hafeez, 2020; Li et al., 2020; Abbas et al., 2020; Wang and Wong, 2020; Luz et al., 2020).

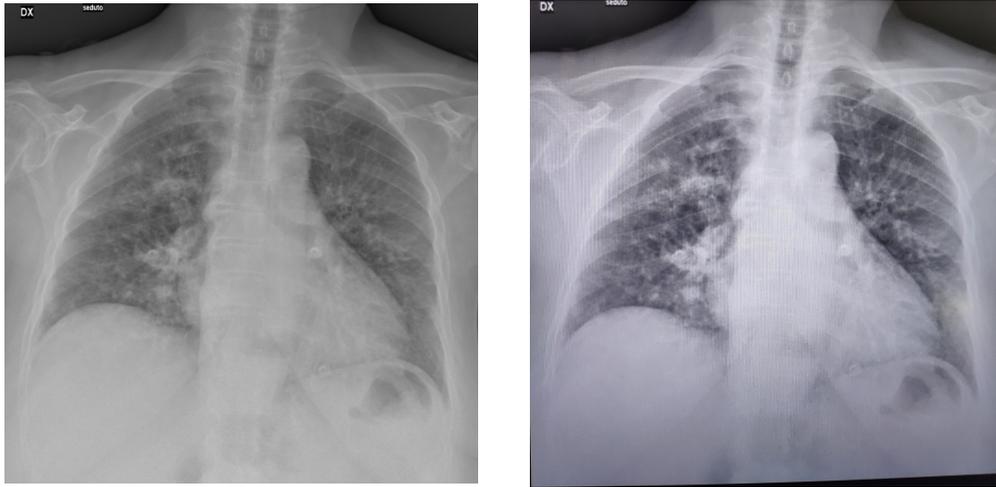
To fit these models, labeled training data is required. The main source of COVID-19 CXR images is the repository made available by Cohen et al. (Cohen et al., 2020) which was used, for instance, in (Hemdan et al., 2020; Farooq and Hafeez, 2020; Li et al., 2020; Abbas et al., 2020; Wang and Wong, 2020; Luz et al., 2020). So that models can learn to differentiate COVID-19 from healthy patients and other types of lung diseases, in (Wang and Wong, 2020), a protocol to build a more comprehensive dataset was proposed. This dataset, named COVID $x$ , merges five CXR databases and contains images from healthy patients as well as patients with multiple variants of pneumonia caused by different bacteria and viruses, including the SARS-CoV-2.

In COVID $x$ , the vast majority of CRXs belonging to non-COVID-19 patients were provided by the National Institute of Health (NIH) (Wang et al., 2017). All these images are

in DICOM<sup>1</sup> format and were collected by mining the NIH's own PACS <sup>2</sup>.

Despite being the most adequate file format, DICOM is not widely used by the general public. Hence, not by coincidence, the COVID-19 CXRs images come from much more heterogeneous sources composed by public sources as well as indirect collection from hospitals and physicians all over the world (Cohen et al., 2020; Wang and Wong, 2020).

In a time in which social distancing is recommended and mobility is restricted, mobile applications which embeds the DL models would be extremely useful to physicians and other healthcare practitioners. In this scenario, one may expect the input for the models to be a picture taken from the smartphone, in many cases, directly from a computer monitor attached to the PACS. As can be seen in Figure 1, this procedure may generate distortions and add noise to the images.



(a) COVID $x$  Image (Publicly available at <https://github.com/ieee8023/covid-chestxray-dataset/>) (b) A picture of the same image taken of a computer monitor

Figure 1: COVID $x$  X-ray and a monitor picture of the same image.

In this context, the question of whether these pictures can be used in combination with deep learning models to produce accurate COVID-19 diagnosis arises. To address this question, we split it in the following hypothesis:

---

<sup>1</sup>Digital Imaging and Communications in Medicine

<sup>2</sup>PACS (Picture Archiving and Communication System) is a medical imaging technology which provides economical storage, retrieval, management, distribution and presentation of medical images.

1. A DL model for CXR based diagnosis of COVID-19 misclassify smartphone pictures from computer monitors.
2. Augmenting the training dataset with smartphone pictures of displayed CXRs gives the model the ability to produce accurate results even with this noisy data input method.
3. Assuming that smartphone pictures of displayed CXRs do not make the cut for the diagnosis process, it is possible to build a model that distinguishes them from proper CXR images and discards them automatically before the diagnosis process.

At the time of the writing of this paper, the only free and publicly available models were the COVID-Net (Wang and Wong, 2020) and the EfficientNet-C19 (Luz et al., 2020). Since the difference in accuracy among these models is no more than 1.5%, the focus of our experiments will be the EfficientNet-C19 which is the most memory-efficient model of the two. Being memory efficient is critical when designing a web application.

Overall, the results show that testes models are far from being able to correctly classify most of the pictures taken of monitors. However, data augmentation with these pictures or automatically discarding them may be workarounds. The results reported here have important implications when considering turning the DL models in a web or mobile app. If not taken into account they may lead to very frustrating use of the technology in the real world.

The remainder of this paper is organized as follows: In Section 2, the EfficientNet-C19 is defined. In Section 3, the COVID $x$  dataset is described and the new dataset created for this work is presented. In Section 4, the quality metrics used to evaluate the DL models are defined. In Section 5, the computational experiments designed to test the hypothesis above are described and the results are reported. Finally, in Section 6 the conclusions and final remarks are presented.

## 2. EfficientNet-C19

The EfficientNet-C19 neural network (Luz et al., 2020) extends the the family of EfficientNets (Tan and Le, 2019) for the CXR based COVID-19 screening problem. The EfficientNet family has as its basic building block the architecture shown in Table 1. Its main component, the Mobile Inverted Bottleneck Conv Block, MBconv (Sandler et al., 2018), is depicted in FIGURE 2.

Table 1: EfficientNet baseline network : B0 architecture.

Stage	Operator	Resolution	#channels	#layers
1	Conv3x3	224x224	32	1
2	MBCConv1,k3x3	112x112	16	1
3	MBCConv6,k3x3	112x112	24	2
4	MBCConv6,k5x5	56x56	40	2
5	MBCConv6,k3x3	28x28	80	3
6	MBCConv6,k5x5	14x14	112	3
7	MBCConv6,k5x5	14x14	192	4
8	MBCConv6,k3x3	7x7	320	1
9	Conv1x1/Pooling/FC	7x7	1,280	1

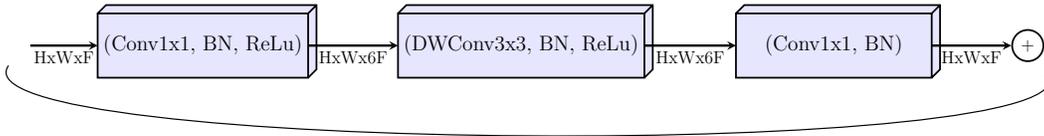


Figure 2: MBCConv Block (Sandler et al., 2018). DWCConv stands for depthwise conv, k3x3/k5x5 defines the kernel size, BN is batch normalization,  $HxWxF$  represents the tensor shape (height, width, depth), and x1/2/3/4 is the multiplier for number of repeated layers.

The idea behind the EfficientNet family is to start from the high quality yet compact baseline model presented in Table 1 and uniformly scale each of its dimensions systematically with a fixed set of scaling coefficients. The different scaling factors give rise to the different members of the family.

An EfficientNet is defined by three dimensions: (i) depth; (ii) width; and (iii) resolution as illustrated in FIGURE 3. Starting from the baseline model, called B0, in Table 1 each dimension is scaled by the parameter  $\phi$  according to Equation 1 where  $\alpha = 1.2$   $\beta = 1.1$  and  $\gamma = 1.1$  are constants obtained experimentally using a grid search. By varying  $\phi$ , one obtains new architectures of the EfficientNet family. For instance,  $\phi = 1$  gives rise to the EfficientNet B1,  $\phi = 2$  gives rise to the EfficientNet B2, and so on.

According to (Tan and Le, 2019), Eq. 1 provides a nice balance between performance and computational cost.

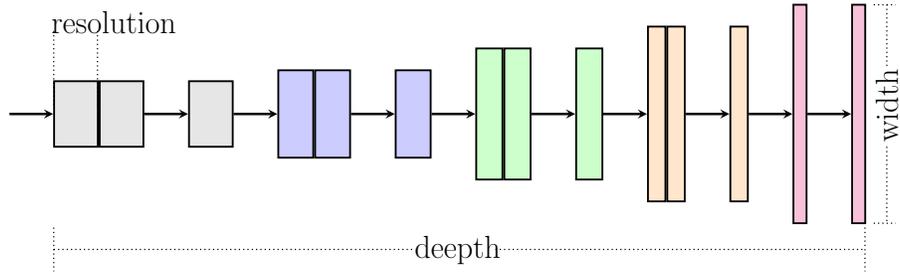


Figure 3: Efficient net compound scaling on three parameters. (Adapted from (Tan and Le, 2019))

$$\begin{aligned}
 \text{depth} &= \alpha^\phi \\
 \text{width} &= \beta^\phi \\
 \text{resolution} &= \gamma^\phi \quad , \\
 \text{s.t. } \alpha \cdot \beta^2 \cdot \gamma^2 &\approx 2 \\
 \alpha \geq 1, \beta \geq 1, \gamma &\geq 1
 \end{aligned} \tag{1}$$

In (Luz et al., 2018), as shown in Table 2, the authors add four new blocks to the baseline model to improve the EfficientNet performance on the COVID-19 screening problem. Here, this model will be called EfficientNet-C19 B0. For the complete discussion on the rationale behind these new blocks, see (Luz et al., 2020).

Table 2: EfficientNet-C19 baseline model or EfficientNet-C19 B0. (NC = Number of Classes).

Stage	Operator	Resolution	#channels	#layers
1-9	EfficientNet B0	224x224	32	1
10	BN/Dropout	7x7	1280	1
11	FC/BN/Swich/Dropout	1	512	1
12	FC/BN/Swich	1	128	1
13	FC/Softmax	1	NC	1

To obtain other models from the EfficientNet-C19 family, one changes the EfficientNet layers. That is, in Table 2, if instead of the EfficientNet B0, one uses the EfficientNet B3 in the first stages, the model becomes the EfficientNet-C19 B3.

Since the main question addressed in this paper has arisen from a smartphone application, it makes sense to test a model that would fit well in a mobile app. For reference,

the COVID-Net presented in (Wang and Wong, 2020) requires 2.1GB of memory while the EfficientNet-C19 B3 requires only 134MB.

### 3. Data set augmentation with pictures of CXRs displayed in computer monitors

To investigate whether smartphone pictures of CXRs displayed in monitors are suitable for COVID-19 identification in CXRs, a new dataset of such pictures must be built. Thus, in section 3.1 we describe the COVID $x$  dataset which is the biggest CXR dataset for the COVID-19 screening problem. Then, in Section 3.2, we describe the procedure used to build a picture dataset from COVID $x$ .

#### 3.1. The COVID $x$ dataset

The COVID $x$  dataset combines five data repositories and has been built to leverage the following types of patient cases from each of the data repositories (Wang and Wong, 2020):

- COVID-19 cases from the COVID-19 Image Data Collection (Cohen et al., 2020), the COVID-19 Chest X-ray Dataset Initiative <sup>3</sup>, the ActualMed COVID-19 Chest X-ray Dataset Initiative <sup>4</sup>, and COVID-19 radiography database (Chowdhury et al., 2020).
- Patient cases who have no pneumonia (i.e., normal) from the RSNA Pneumonia Detection Challenge dataset (RSNA) and the COVID-19 Image Data Collection (Cohen et al., 2020).
- Non-COVID19 pneumonia patient cases from RSNA Pneumonia Detection Challenge dataset (Cohen et al., 2020).

The COVID $x$  was designed to represent a classification with three classes:

- *normal* - for healthy patients.
- *COVID-19* - for patients with COVID-19.
- *pneumonia* - for patients with non-COVID-19 pneumonias.

---

<sup>3</sup><https://github.com/agchung/Figure1-COVID-chestxray-dataset>

<sup>4</sup><https://github.com/agchung/Actualmed-COVID-chestxray-dataset>

The majority of the instances which belong to the normal and pneumonia classes come from the RNSA dataset (RSNA). The COVID-19 class instances come from multiple sources. The COVID $x$  has a total of 13,800 images from 13,645 individuals and is split into two partitions, one for training purposes and one for testing (model evaluation). The distribution of images between the partitions is shown in Table 3. The source code to reproduce the dataset is publicly available<sup>5</sup>.

Table 3: COVID $x$  Images distribution among classes and partitions.

Type	Normal	Pneumonia	COVID-19	Total
Training	7966	5421	152	13569
Test	100	100	31	231

### 3.2. The COVID-19 dataset of CRXs displayed in monitors (C19-CRX-M) dataset

As shown in Table 3, the number of COVID-19 images for training is 152 images. The normal and pneumonia classes have a huge number of samples when compared to the COVID-19. It is impractical to take photos of all images from pneumonia and normal classes (13,387 images). In that sense, we only select 152 random images of normal samples and 152 random images of pneumonia samples from the COVID $x$  dataset.

To generate the dataset of screen pictures, the selected images from the COVID $x$  dataset were displayed in different computer screens of which pictures were taken with different smartphones. The selection process can be summarized as follows:

- For C19-CRX-M training set:
  - 152 Normal images randomly selected from the COVID $x$  training set.
  - 152 Pneumonia images randomly selected from the COVID $x$  training set.
  - All the 152 COVID-19 images available in the COVID $x$  training set.
- For C19-CRX-M test set:
  - All the images available in the COVID $x$  test set.

<sup>5</sup><https://github.com/lindawangg/COVID-Net>

The C19-CRX-M class distribution is shown in Table 4. The detailed information about the selected images and used devices is available at <https://github.com/ufopcsilab/C19-CRX-M>.

Table 4: C19-CRX-M Images distribution among classes and partitions.

Type	Normal	Pneumonia	COVID-19	Total
Training	152	152	152	456
Test	100	100	31	231

After describing the datasets, in the next section, we define the metrics used to assess the quality of the models.

#### 4. Evaluation metrics

Following the methodology in (Luz et al., 2020) and (Wang and Wong, 2020), in this work, three metrics are used to evaluate models: accuracy ( $Acc$ ), COVID-19 sensitivity ( $Se_C$ ), and COVID-19 positive prediction ( $+P_C$ ), i.e.,

$$Acc = \frac{TP_N + TP_P + TP_C}{\#samples}$$

$$Se_C = \frac{TP_C}{TP_C + FN_C} \tag{2}$$

$$+P_C = \frac{TP_C}{TP_C + FP_C}$$

where:

- $TP_N$  is the number of normal samples correctly classified
- $TP_P$  is the number of non-COVID-19 samples correctly classified
- $TP_C$  is the number of COVID-19 samples correctly classified,
- $FN_C$  is the number of COVID-19 samples classified as normal or non-COVID-19
- $FP_C$  is the number of non-COVID-19 and normal samples classified as COVID-19.

## 5. Computational Experiments

In this section, the experiments implemented to test each of the hypotheses raised in Section 1 are presented along with results and implications.

### 5.1. Misclassification of Pictures of Displayed CXRs

This experiment tackles the the following hypothesis:

- A DL model for CXR based diagnosis of COVID-19 misclassify smartphone pictures of Displayed CXRs.

#### 5.1.1. Setup

The COVIDNet Large of (Wang and Wong, 2020) and the EfficientNet-C19 B3 of (Luz et al., 2020) were used to classify the test images from the C19-CRX-M dataset. Then, accuracy, positive prediction, and sensitivity metrics as defined in Section 4 were computed.

#### 5.1.2. Results

Table 5, presents the results.

Table 5: Evaluation of deep learning model on the C19-CRX-M dataset

Approach	$Acc$	$Se_C$	$+P_C$
COVIDNet - Wang <i>et al.</i> (Wang and Wong, 2020)	16.88%	100.00%	13.90%
EfficientNet-C19 - Luz <i>et al.</i> (Luz et al., 2020)	16.02%	100.00%	15.27%

It is possible to observe that both approaches fail to appropriately classify the CXR pictures. As can be seen in the confusion matrices in Figure 4, the majority of the images are classified as COVID-19. This explains the high sensitivity with a low positive prediction for the COVID-19 class.

Since both models were trained in the same dataset, we conjecture that these results may have occurred because of the way the COVID $x$  is built. In COVID $x$  the majority of pneumonia and normal images come from the same source, that is, the NIH Clinic Center PACS system (RSNA). Meanwhile, the COVID-19 images come from very heterogeneous sources including donations from physicians. Hence, the COVID-19 set is more subject to external noise than their pneumonia and normal counterparts. This might have induced the DL models to “think” that the screen pictures belonged to the COVID-19 class. FIGURE 4 presents the confusion matrix for both models which supports the above claims.

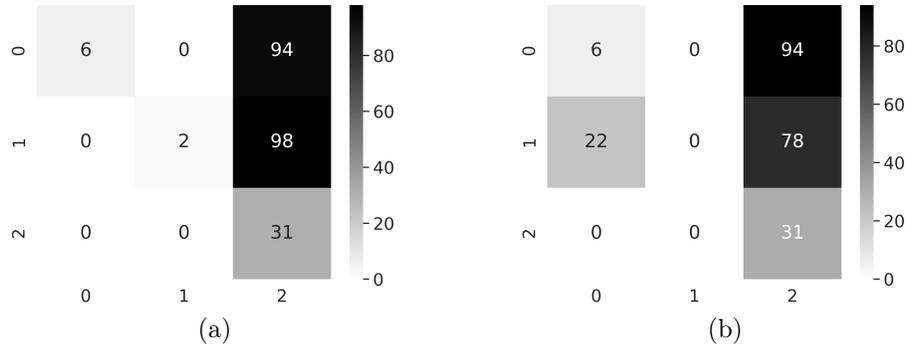


Figure 4: Confusion Matrices: (a) EfficientNet-C19, (b) COVIDNet.

## 5.2. Effect of dataset augmentation

Once one is aware of the problem with the CXR pictures, a natural step is to add data like this to the training dataset, as a data augmentation technique. Thus, in this section, the following hypothesis is tested:

- Augmenting the training dataset with smartphone pictures of displayed CXRs gives the model the ability to produce accurate results even with this noisy data input method.

### 5.2.1. Setup

Due to its effectiveness and suitability for smartphone applications, in this experiment, the focus is on two variants of the EfficientNet-C19 (Luz et al., 2020), the B0 and B3.

Both neural networks are trained with a combination between the C19-CRX-M dataset and the COVID $x$  images used to generate the C19-CRX-M. Thus, the models see the original and screen versions of every image in the training and dataset.

As shown in Table 6, the images used by both architectures consist of the C19-CRX-M dataset plus the images from the COVID $x$  dataset used to produce the C19-CRX-M. Furthermore, we also vary the resolution of the input images which can be  $224 \times 224$  (EfficientNet B0 default input size (Tan and Le, 2019)),  $300 \times 300$  (EfficientNet B3 default input size (Tan and Le, 2019)), and  $448 \times 448$ .

### 5.2.2. Results

Table 7 presents the results for the three quality metrics defined in Section 4.

Table 6: Distribution of images from the combination of the original images and the C19-CRX-M dataset

Type	Source	Normal	Pneumonia	COVID-19	Total
Training	Screen Picture	152	152	152	456
	Original Image	152	152	152	456
Test	Screen Picture	100	100	31	231
	Original Image	100	100	31	231

Table 7: . (EF = EfficientNet; ORI = COVID $x$  Original Images; CAM = C19-CRX-M Images)

Model	Test Data	$Acc$	$Se_C$	$+P_C$
EFB0 224x224	CAM	51.52%	83.87%	22.41%
	ORI	90.91%	90.32%	80.00%
	ORI + CAM	77.21%	87.10%	35.76%
EFB3 300x300	CAM	74.46%	87.10%	43.55%
	ORI	90.04%	87.10%	87.10%
	ORI + CAM	82.25%	87.10%	58.06%
EFB0 448x448	CAM	57.14%	96.77%	31.91%
	ORI	91.77%	96.77%	96.77%
	ORI + CAM	74.46%	96.77%	48.00%
EFB3 448x448	CAM	65.80%	90.32%	37.33%
	ORI	91.77%	100.00%	91.18%
	ORI + CAM	78.79%	95.16%	54.13%

The first point to highlight is the increase in both accuracy and positive prediction of COVID-19 screen pictures when compared with the results in Table 5. Nevertheless, this improvement remains far from the results obtained for the COVID $x$  images.

Increasing the input image size from 224x224 to 448x448 was beneficial for the EfficientNet-C19 B0 but it was not beneficial EfficientNet-C19 B3. We conjecture that the B3 version is more sensitive to noise than the B0 version. Thus, for B3, having a higher rate of compression in the input decreases this effect. On the other hand, since B0 is already robust to noise, having a larger input image gave it more information to exploit.

Figure 5 shows the confusion matrices for the four tested models. It is possible to see that, with the augmented dataset, the model mistakes get more diversified and the bias towards COVID-19 observed in the previous experiment (Figure 4) is not as extreme anymore.

Overall, the augmented dataset did significantly improve the accuracy of the models. Even so, the models still struggle to correctly classify the smartphone pictures of the displayed CXRs and the overall accuracy remains unacceptable for the intended application.

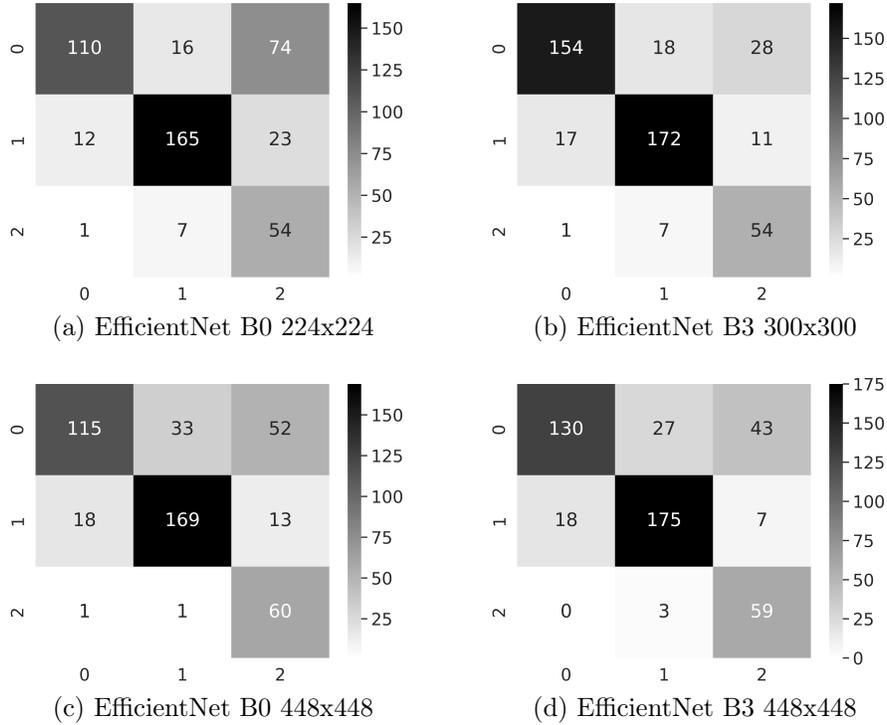


Figure 5: Confusion matrices for the four models tested on C19-CRX-M dataset. 0 is normal samples, 1, pneumonia and 2, COVID-19.

Thus, for the presented setting, the augmented dataset did not solve the problem satisfactorily. Nevertheless, given how the augmentation changed the distribution of the errors, a bigger and more varied dataset may improve the results even further.

### 5.3. Discarding Pictures of Displayed CXRs Automatically

Having verified the inability of the models in coping with the CXR screen pictures, in this section, we test the following hypothesis:

- It is possible to build a model that distinguishes pictures of displayed CXRs from proper CXR images and discards them automatically before the diagnosis process.

#### 5.3.1. Setup

In this experiment, we train the EfficientNet-C19 B0 on the dataset presented in Table 6 to verify if it can distinguish between the CXR pictures and proper CXR images. Hence,

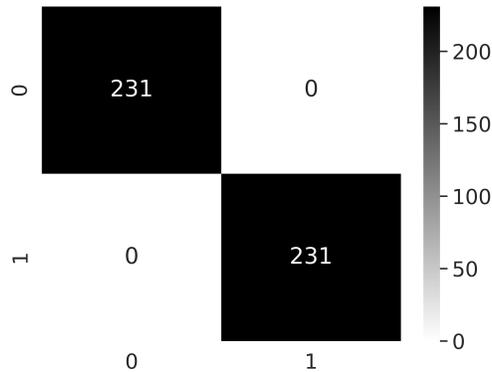


Figure 6: Confusion matrix for discarding CXRs screen pictures automatically.

the task becomes a two-class classification problem. One class represents the proper images obtained from the COVID $x$  dataset and the other class represents the pictures in the C19-CRX-M dataset.

As shown in Table 6, the training set is balanced. Thus, it has been decided not to perform any data augmentation and use only the data available.

### 5.3.2. Results

The confusion matrix obtained with the proposed setup is depicted in Figure 6. It can be seen that the proposed approach reached an accuracy of 100%. Thus, this methodology may be used as step previous to COVID-19 classification in which the signals are classified as good (proper images) or not (pictures of displayed CXRs).

## 6. Conclusion

Deep Learning applied to CXR based COVID-19 screening has shown a lot of promise, in particular, in a scenario of worldwide test kit shortage. Thus, having a mobile app embedding such technology would be an important step in order to make the technology really useful and accessible. In such an application, a natural input would be a picture taken from the smartphone itself often from the monitor attached to the PACS. In order to study the effect of such inputs in the DL models, the C19-CRX-M dataset was created by taking pictures of CXR images displayed in computer monitors.

The results show that the models trained on COVID $x$  fail to classify the CXR pictures and tend to classify the majority of them as COVID-19. We conjecture that this phenomenon might be due to the nature of the COVID $x$  dataset, in which COVID-19 images are from a different distribution when compared to the normal and pneumonia ones which are mostly DICOM files from NIH PACS.

We have also evaluated a data augmentation strategy in which the CXR pictures are merged with the original images. In this scenario, an overall reduction is observed in the accuracy, COVID-19 positive prediction, and COVID-19 sensitivity when compared with a scenario with the original images only. However, there is an improvement in the screen picture classification, from 16.88% and 13.90% to 88.25% and 58.06% of accuracy and COVID-19 positive prediction, respectively.

Despite the improvement, a 58.06% COVID-19 positive prediction remains too low for a real-world application. Because of that, an approach to filter the CXR pictures out of the COVID-19 classification pipeline has also been proposed. Since the noise is visually quite significant, the proposed method had no difficulty in filtering the pictures. The proposed methodology has presented 100% effectiveness which makes the system as whole much more reliable.

Overall, the results show that the tested models are far from being able to correctly classify most of the pictures taken of monitors. Thus, researchers and practitioners must be aware of these limitations, especially, when planning to put these models to production and embed them into a web service or mobile app. On the other hand, these same neural network architectures found it easy to distinguish these pictures from regular images. Having a model like that to discard screen pictures might be a workaround while the datasets are not big and comprehensive enough to allow for the construction of a reliable solution.

## References

- Abbas, A., Abdelsamea, M., Gaber, M., 2020. Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network URL: <http://dx.doi.org/10.1101/2020.03.30.20047456>, doi:10.1101/2020.03.30.20047456.
- American College of Radiology, 2020. ACR recommendations for the use of chest radiography and computed tomography (ct) for suspected covid-19 infection. ACR website.

- Araujo-Filho, J.d.A.B., Sawamura, M.V.Y., Costa, A.N., Cerri, G.G., Nomura, C.H., 2020. COVID-19 pneumonia: what is the role of imaging in diagnosis? *Jornal Brasileiro de Pneumologia* 46.
- Chowdhury, M.E.H., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M.A., Mahbub, Z.B., Islam, K.R., Khan, M.S., Iqbal, A., Al-Emadi, N., Reaz, M.B.I., 2020. Can ai help in screening viral and covid-19 pneumonia? [arXiv:2003.13145](https://arxiv.org/abs/2003.13145).
- Cohen, J.P., Morrison, P., Dao, L., 2020. Covid-19 image data collection. [arXiv preprint arXiv:2003.11597](https://arxiv.org/abs/2003.11597) .
- Davarpanah, A.H., Mahdavi, A., Sabri, A., Langroudi, T.F., Kahkouee, S., Haseli, S., Kazemi, M.A., Mehrian, P., Mahdavi, A., Falahati, F., Tuchayi, A.M., Bakhshayeshkaram, M., Taheri, M.S., 2020. Novel screening and triage strategy in iran during deadly coronavirus disease 2019 (covid-19) epidemic: Value of humanitarian tele-consultation service. *Journal of the American College of Radiology* , [1doi:10.1016/j.jacr.2020.03.015](https://doi.org/10.1016/j.jacr.2020.03.015).
- Farooq, M., Hafeez, A., 2020. Covid-resnet: A deep learning framework for screening of covid19 from radiographs. [arXiv preprint arXiv:2003.14395](https://arxiv.org/abs/2003.14395) .
- Hemdan, E.E.D., Shouman, M.A., Karar, M.E., 2020. Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images. [arXiv preprint arXiv:2003.11055](https://arxiv.org/abs/2003.11055) .
- Li, T., Han, Z., Wei, B., Zheng, Y., Hong, Y., Cong, J., 2020. Robust screening of covid-19 from chest x-ray via discriminative cost-sensitive learning. [ArXiv abs/2004.12592](https://arxiv.org/abs/2004.12592).
- Luz, E., Moreira, G., Junior, L.A.Z., Menotti, D., 2018. Deep periocular representation aiming video surveillance. *Pattern Recognition Letters* 114, 2–12.
- Luz, E., Silva, P.L., Silva, R., Silva, L., Moreira, G., Menotti, D., 2020. Towards an effective and efficient deep learning model for covid-19 patterns detection in x-ray images. [arXiv:2004.05717](https://arxiv.org/abs/2004.05717).
- RSNA, . Radiological Society of North America. RSNA pneumonia detection challenge. <https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data>. Accessed: 2020-04-01.

- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C., 2018. Mobilenetv2: Inverted residuals and linear bottlenecks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4510–4520.
- Tan, M., Le, Q.V., 2019. Efficientnet: Rethinking model scaling for convolutional neural networks. arXiv preprint arXiv:1905.11946 .
- Wang, L., Wong, A., 2020. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest radiography images. arXiv preprint arXiv:2003.09871 .
- Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., Summers, R.M., 2017. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) URL: <http://dx.doi.org/10.1109/CVPR.2017.369>, doi:10.1109/cvpr.2017.369.

**Declarations:**

Competing interests: The authors declare no competing interests.

# Figures



Figure 1

COVIDx X-ray and a monitor picture of the same image.

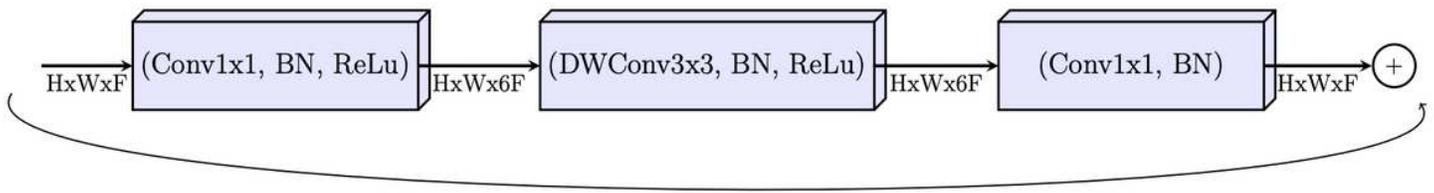
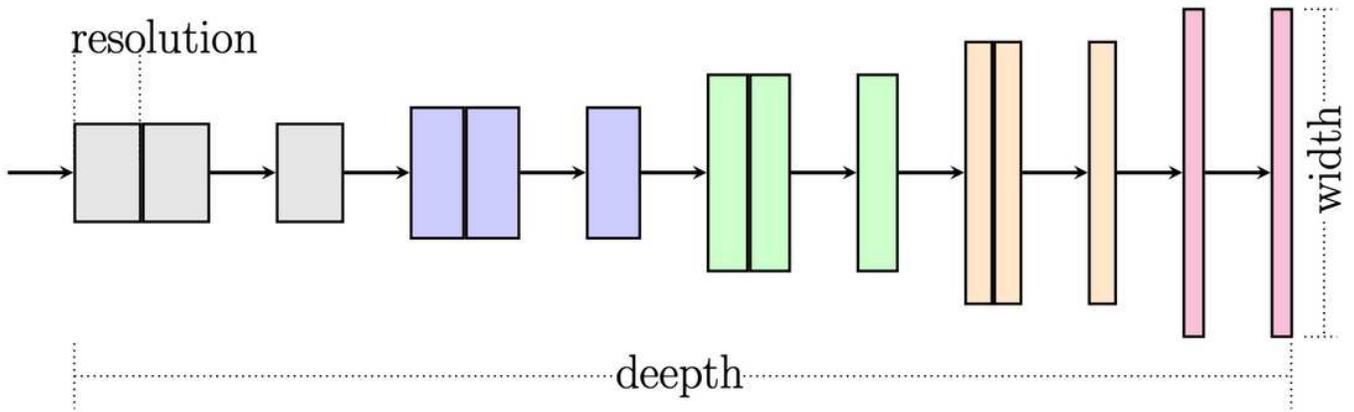


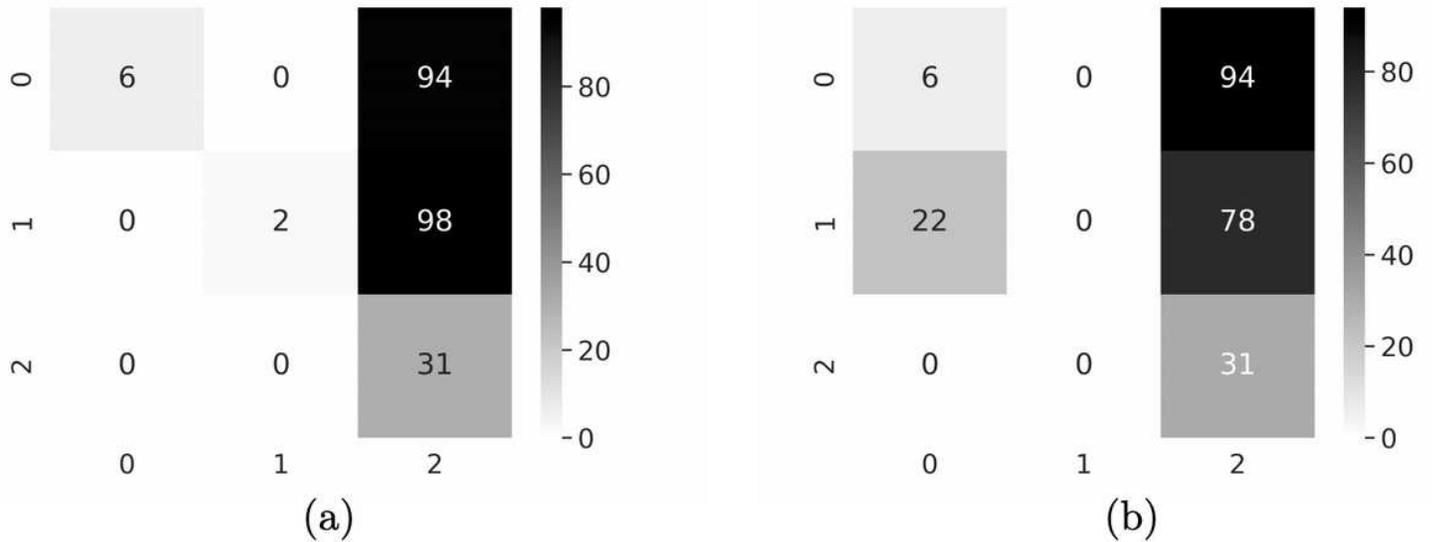
Figure 2

MBConv Block (Sandler et al., 2018). DWConv stands for depthwise conv,  $k3 \times 3/k5 \times 5$  defines the kernel size, BN is batch normalization,  $H \times W \times F$  represents the tensor shape (height, width, depth), and  $x1/2/3/4$  is the multiplier for number of repeated layers



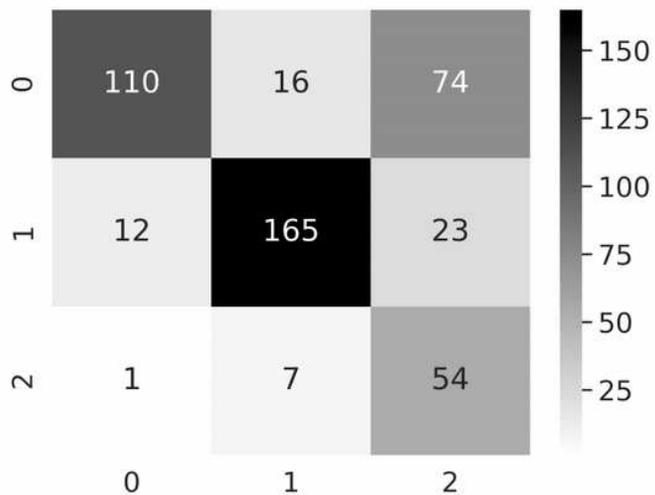
**Figure 3**

Efficient net compound scaling on three parameters. (Adapted from (Tan and Le, 2019))

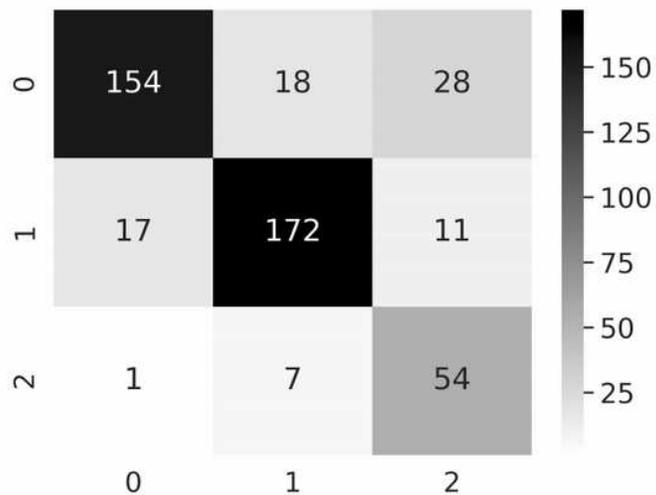


**Figure 4**

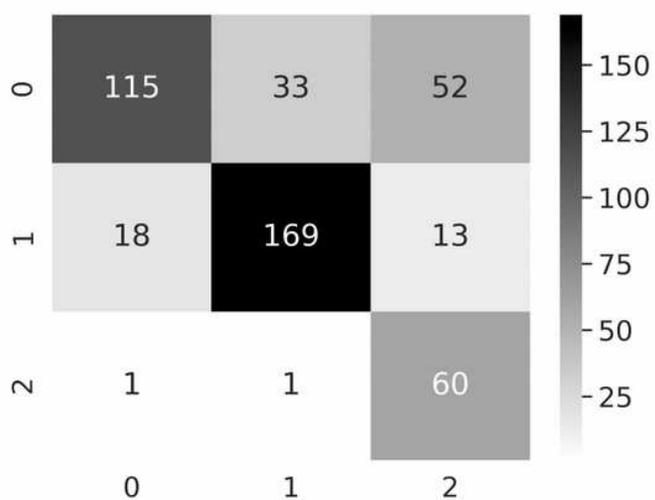
Confusion Matrices: (a) EfficientNet-C19, (b) COVIDNet.



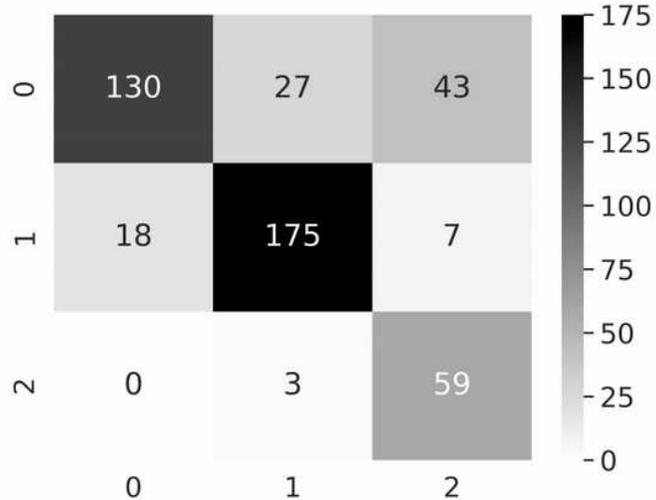
(a) EfficientNet B0 224x224



(b) EfficientNet B3 300x300



(c) EfficientNet B0 448x448



(d) EfficientNet B3 448x448

**Figure 5**

Confusion matrices for the four models tested on C19-CRX-M dataset. 0 is normal samples, 1, pneumonia and 2, COVID-19.

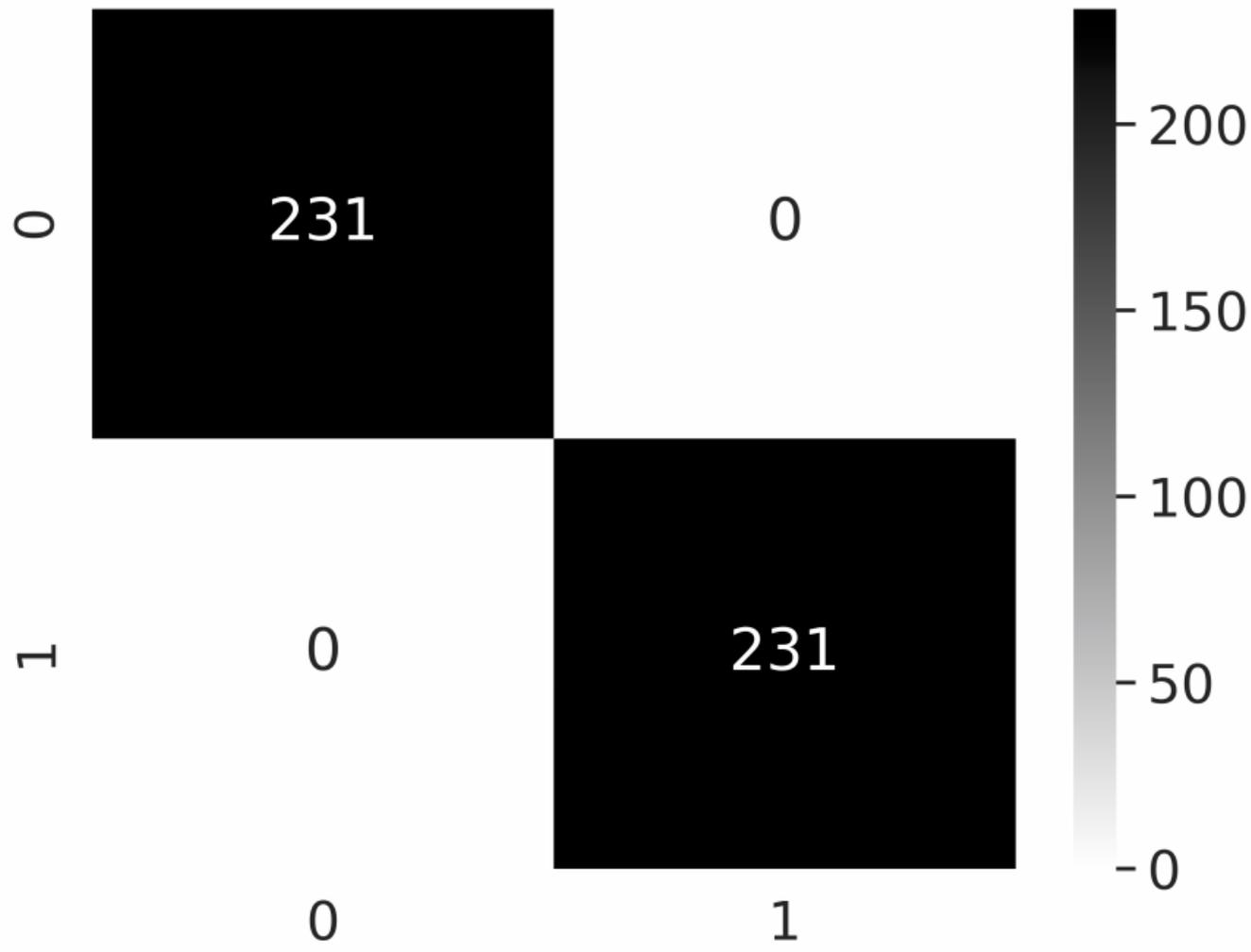


Figure 6

Confusion matrix for discarding CXRs screen pictures automatically