SYSTEM FOR ASSISTANCE IN DIAGNOSIS OF DISEASES PULMONARY

Gustavo Chichanoski and Maria Bernadete de Morais França

Department of Electronic Engineering,
State University of Londrina, Londrina, Paraná, Brazil

ABSTRACT

Covid-19 is caused by the SARS-COV2 virus, where most people experience a mild to moderate respiratory crisis. To assist in diagnosing and triaging patients, this work developed a Covid-19 classification system through chest radiology images. For this purpose, the neural network models ResNet50V2, ResNet101V2, DenseNet121, DenseNet169, DenseNet201, InceptionResnetV2, VGG-16, and VGG-19 were used, comparing their precision, accuracy, recall, and specificity. For this, the images were segmented by a U-Net network, and packets of the lung image were generated, which served as input for the different classification models. Finally, the probabilistic Grad-CAM was generated to assist in the interpretation of the results of the neural networks. The segmentation obtained a Jaccard similarity of 94.30%, while for the classification the parameters of precision, specificity, accuracy, and revocation were evaluated, compared with the reference literature. Where DenseNet121 obtained an accuracy of 99.28%, while ResNet50V2 presented a specificity of 99.72%, both for Covid-19.

KEYWORDS

Deep Learning, U-Net, Covid-19, Segmentation & Machine Learning

1. INTRODUCTION

Terms related to artificial intelligence have gained evidence in the media due to their presence in devices such as smartphones, virtual assistants, and autonomous cars, where as a general concept, researchers seek to automate decisions executed by humans. As with everyday tools, medicine, and more specifically the area of radiology, has undergone major technological changes in recent years.

As described in [3] and [4], typical symptoms of COVID-19 are characterized by fever, cough, shortness of breath, muscle aches, mental confusion, headaches, sore throat, rhinorrhea, pain in the chest, diarrhea, nausea, vomiting, pneumonia, decrease in white cell count or decrease in lymphocytes.

The symptoms that can be visualized in radiology are presented in three stages, initial, progression, and strong. In the early stage of the disease, the lungs show multiple small irregular shadows and interstitial changes, most apparent in peripheral zones of the lungs. In the progression stage, there are multiple opaque ground glass and infiltration in both lungs, and in the strong stage, there is a consolidation of the lungs.

Among the works carried out on the diagnosis of Covid-19 using CT images of the chest, the work [11] stands out, whose objective was to produce a design of deep convolutional neural

networks called COVID-Net, obtaining an accuracy of 93,3%, using 11.75 million parameters and with a sensitivity to COVID-19 of 91%, cited by [7] as one of the most successful methods in diagnosis. [11] used the dataset COVIDx, divided into three categories, No Disease, Pneumonia, and COVID-19, the work was made available at https://github.com/lindawangg/COVID-Net.

According to [12], in the 101 x-ray images collected at 4 institutions in Hunan, China, of patients between 21 and 50 years old, divided into non-emergency (87), mild or common cases, and emergencies (14), severe or fatal cases.

Most patients had GGO (Ground Glass Opacity), 87 (86.1% of the total cases), or a mixture of GGO with lung consolidation, 65 (64.4% of the total cases). However, when looking exclusively at emergency patients (14 cases), the number of patients with GGO reached 100% (14 cases), consolidation rose from 41.4% in non-emergency cases to 57.1% in emergency cases, while the mixture of both remained constant in 64.4% of both types of cases.

The article [7] proposes a new diagnostic approach, using an (FC)-DenseNet103 to perform lung segmentation and the Resnet-18 network to classify thorax radiology, available at https://github.com/jongcye/Deep-Learning-COVID-19-on-CXR-using-Limited-Training-Data-Sets.

The purpose of [7] is to train the weights of the ResNet-18 network using only a random real image package selected from the thorax radiology image, and this package cannot be completely zero or black. While the test is done by removing non-zero K slices from the image, passing each package through the trained network image, and defining the winner based on each class's highest number of votes. Resulting in a sensitivity of 100% for COVID-19 and an accuracy of 76.9%.

However, in [13], using the method suggested in [7] and in [14], he can create a new manual method to auxilian in the diagnosis of Covid-19, through RALE the lung was classified in a grade between 0 and 4 to each lung, where 0 is normal, 1 to 2 is wild, 3 to 5 is moderate and 6 to 8 is severe.

This article aims to improve the performance of [7] in segmentation by changing the error function to Log-Cosh Dice-Loss and using other models of neural networks. To make it possible was used a U-Net to separate lungs of background in thorax radiology, evaluating the model using F1 Score and Jaccard Score parameters. On other hand, for classification was used the models ResNet50V2, ResNet101V2, DenseNet121, DenseNet169, DenseNet201, VGG-16, VGG-19, and InceptionResnetV2, using as parameters, precision, accuracy, recall and specify.

2. REVIEW OF LITERATURE

2.1. Segmentation

To reduce the complexity of the chest x-ray image, the segmentation models isolate the lungs from the background. One of these models is U-Net, suggested in [8], shown in Figure 1.
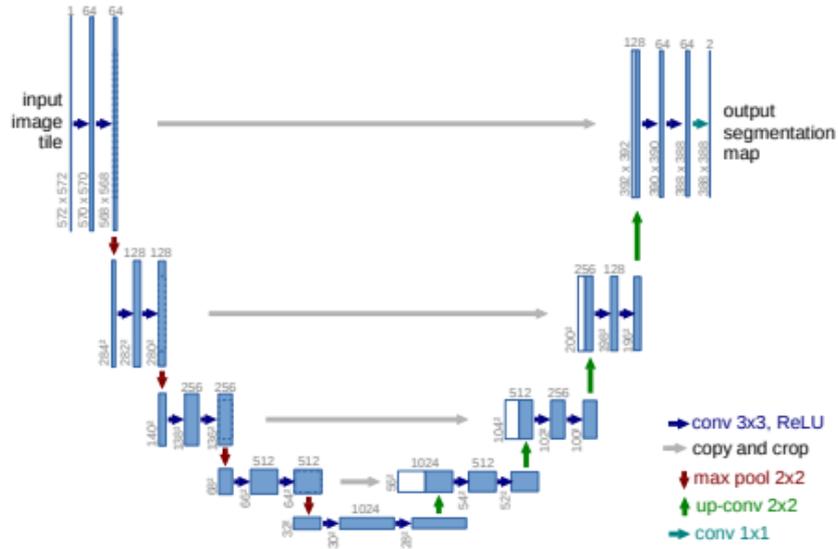
Figure 1. Architecture of U-Net model described in [8].

To reduce the complexity of the chest x-ray image, the segmentation models isolate the lungs from the background.

The U-Net model work in two steps: compression and expansion. Each level is composed of 3x3 convolution layers, ReLU activation, and a max-pooling layer in the compression step.

In the last level, the expansion step begins, consisting of up-sampling layers, concatenation layers, joining the output of the up-sampling layer with the inputs of the same level in the compression step, and then going through another convolution layer and activation. But at the last level of the expansion, add a convolution layer with a 1x1 filter and a Sigmoid activation layer.

To increase the performance of the segmentation model, the Sorensen-Dice coefficient was used to calculate the correlation between two images, according to the equation:

$$D = \frac{2 \cdot TP}{2 \cdot TP + FP + TN},$$

(1)

whose abbreviations can be described as TP (true negatives), FP (false positives), FN (False negatives), and D (Sorensen-Dice coefficient). However, this equation would not serve to be applied as a cost function, for that [7] describes a new equation, shown in:

$$L(y, \hat{p}) = 1 - \frac{2 y \hat{p} + 1}{y + \hat{p} + 1},$$

(2)

where y represents the real values and $\hat{p}$ the values predicted by the U-Net model, the addition of 1 in the numerator and denominator makes the function finite when y and $\hat{p}$ are 0. Thus, the closer the prediction to the true value, the smaller the value of L. However, this function does not behave like a convex function, decreasing the learning efficiency of the model, to correct this [8] suggests the correction for (2) adding the Lovsz extension:

$$LD(y,\hat{p}) = \log\left(\cosh\left(1 - \frac{2y\hat{p}+1}{y+\hat{p}+1}\right)\right), \tag{3}$$

whose characteristic curve is shown in Figure 2, transforming the function into convex, keeping the finite function in the range from -1 to 1.
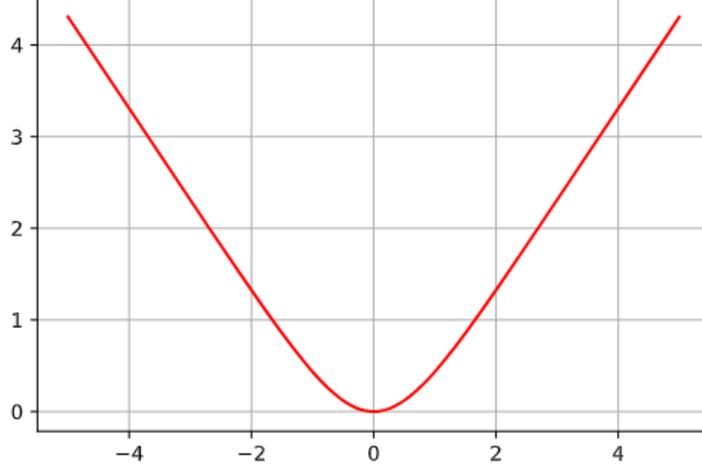

Figure 2. Characteristic curve of the Log-Cosh Dice error function.

2.2. Grad-CAM Probabilistic

To assist visualization in the interpretation of results of this new method of classification a new Grad-CAM was proposed, the Probabilistic Grad-CAM. It is a variation of the Grad-CAM, based on the equation proposed in [8], the Grad-CAM is calculated for only a single slice:

$$L^c = \text{Re}\,LU\left(\sum_k \alpha_k^c \cdot A^k\right), \tag{4}$$

Where α_k^c represents a partial linearization of the deep network through A, the importance of feature map k of target class c, calculated with the equation:

$$\alpha_k^c = \frac{1}{Z}\sum_i^u \sum_j^v \frac{\partial y^c}{\partial A_{i,j}^k}, \tag{5}$$

so, to calculate the Probabilistic Grad-CAM of the original input image, use the equation:

$$\left[l_{prob}^c\right] = \frac{1}{K_i}\left[\sum_{k=1}^K r^c(x_k)Q_k\left(l^c(x_k)\right)\right]_i, \tag{6}$$

where c is the class to be analyzed, x_k is the input slice of the model, l^c is the heatmap generated by Grad-CAM, Q_k is the operator that places the slice in the original position of the image, r^c is the class probability for the slice, K is the number of slices that use pixel, i is the pixel to be analyzed.

3. METHODOLOGY

3.1. Segmentation

The segmentation used two public datasets, made available by [10]. They are provided by the Department of Health and Human Services of Montgomery County (MC), Maryland, the United States, and the Shenzhen People's Hospital 3 of Guangdong Medical University in Shenzhen (Shenzhen), China. Containing a total of 800 chest radiographs. For the masks, it used the dataset provided by [10], which was generated based on [6], described in Table 1.

Table 1. Scattering of the segmentation dataset between Normal and Tuberculosis.

Dataset	No Disease	Tuberculosis	Total
Montgomery County	80	58	138
Shenzhen	326	336	662
Total	406	394	800

Both datasets form a set of X-ray images and their masks, which were also divided into training, validation, and testing. The proportion of division was 72%, 18%, and 10% respectively, whose objective was to reduce the bias. To reduce the impact of the reduced number of images in the dataset, the segmentation training dataset increases the number of artificial images to 10,000 using the data augmentation, described in [1], applying gamma correction, gaussian noise, elastic transformation, grid distortion and flip horizontal.

The images are of different sizes and types, making it necessary to normalize them to improve network performance. In this way, transform the image from uint8 to float32, resize it to 256 x 256 pixels, and normalized it from 0 to 1. In addition, a Gamma correction of 0.5 was applied. The weights were initialized using the Glorot distribution, being trained using the Log-Cosh Dice Loss error function, with the Adamax optimizer and a learning rate of 0.001, reducing according to the validation error and with a patience of 10 epochs.

3.2. Classification

For the classification of X-ray images, the dataset [8] was used, available free of charge, containing 7,103 images, of which 853 are X-ray images of patients with Covid-19, 1,887 of patients No Disease, and 4,363 of patients with Pneumonia.

The classification dataset used was publicly available in [8]. More than 7,103 images, 853 infected with Covid-19, 1,887 No Diseases, and 4,363 with Pneumonia were separated into training, validation, and testing, 61%, 15%, and 24%, shown in Table 2.

Table 2. Dispersion of images from the classification dataset between COVID-19, Normal, and Pneumonia, and the separation between training, validation, and testing in the number of images.

Labels	Train	Validation	Test	Total
Covid-19	457	114	282	853
No Disease	1,069	267	551	1,887
Pneumonia	2,790	698	875	4,363
Total	4,316	1,079	1,708	7,103

The pre-processing of the images consists of normalizing the values from 0 to 1, resized to 1024x1024 pixels, applying the Gamma correction of 0.55, and performing the image

segmentation with the mask generated by the segmentation model. Next, the image is cut into several random packets of size 224x224 pixels, avoiding null areas, as shown in Figure 3. This process aims to produce the inputs for the local classification model, from which the predictions for each packet are made.
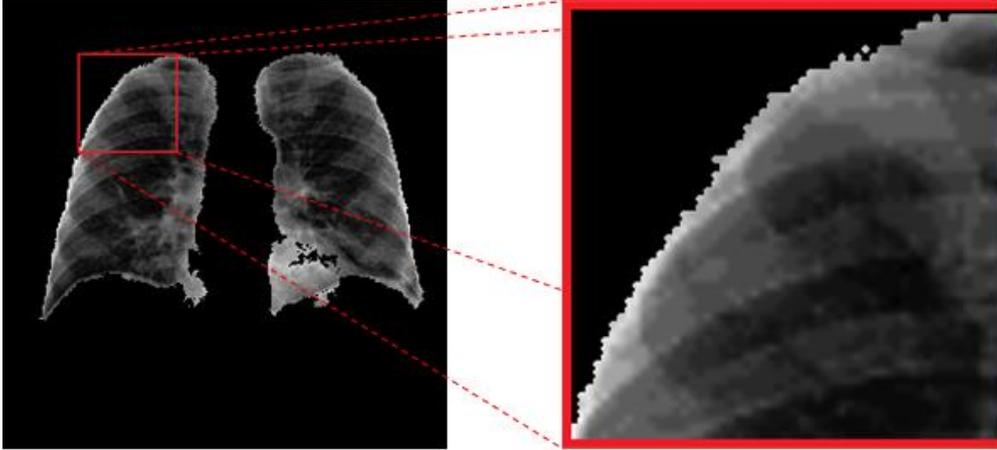

Figure 3. Slice generated by the classification model.

The model for the local approach is composed of the input layer, a convolution layer, a convolution layer, a base layer, and a dense layer. This architecture allows the exchange between the models to be tested, without interfering with the results. The input layer receives the image and converts it to TensorFlow tensors. The convolution layer converts the single-channel image to a three-channel image, with three 3x3 size filters, with a ReLU activation layer. The base layer varies depending on the model type selected: ResNet50V2, InceptionResnetV2, DenseNet121, and VGG-19. The dense layer calculates the predictions for the image, and the SoftMax activation layer filters the results.

3.3. Grad-CAM Probabilistic

Using the heatmaps generated by each package in the classification model, removing the negative parts, the heatmap of the original image is created. Finally, the value of each pixel is divided by the number of packets that contained its position. The number of packets can vary as needed, generating sharper images as the number of packets increases. In the analysis of the results, 400 packages were used.

4. RESULTS

The U-Net segmentation model was evaluated using a test dataset with 80 images through the parameters of the F1 Score and Jaccard Score, to compare the segmentation results with the reference mask. Using the F1 Score criterion, the five worst similarity results in the entire dataset were respectively 88.41%, 89.23%, 89.50%, and 89.65%, while the five best results were 98.48%, 98.51%, 98.53% 98.56%, and 98.57%. The average F1 Score performance was 96.37% similar, with a standard deviation of 2.26%.

The five worst results obtained by the Jaccard Score criterion were 79.23%, 80.56%, 80.99%, 81.25%, and 82.48%, on the other hand, the five best results were 97.01%, 97.06%, 97.11%, 97.16%. The average performance obtained by the Jaccard Score was 93.09% with a standard

deviation of 4. 07%. Figure 4 shows the histogram for the F1 Score criterion from the segmentation test dataset and Figure 5 shows the histogram for the Jaccard Score criterion.
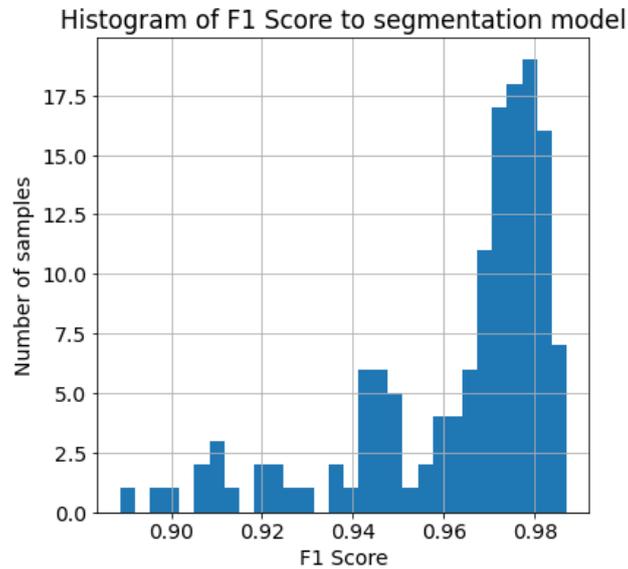

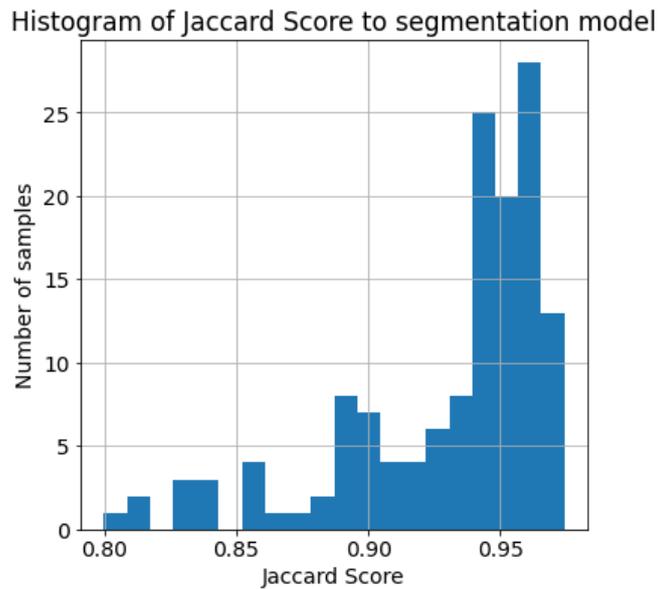Figure 4. Histogram of segmentation model to F1 Score.

Figure 5. Histogram of segmentation model to Jaccard Score.

Figure 6 (a) shows a pre-processed image of a chest X-ray contained in the test dataset. Figure 6 (b) shows the reference mask provided by the dataset and Figure 6 (c) shows the mask generated by the segmentation model.
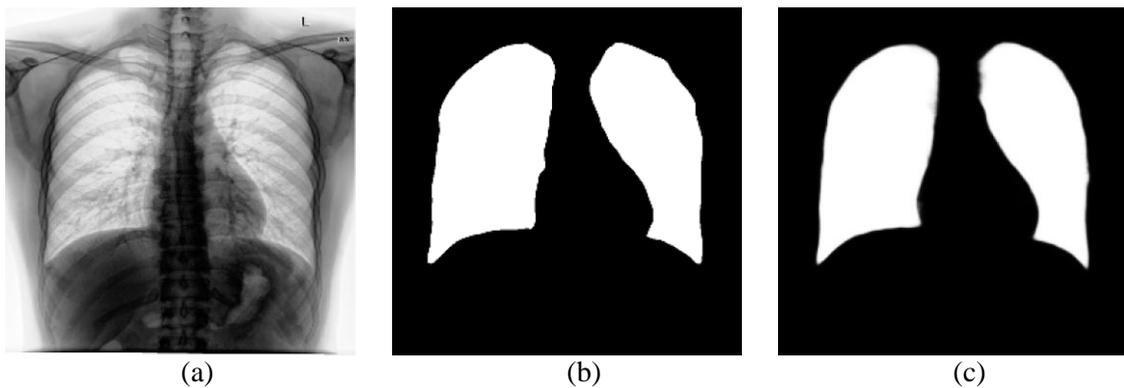
Figure 6. Example of mask generated by the model.
(a) Original image. (b) Original mask. (c) Predict mask.

The classification, in global approximation, was evaluated for precision, accuracy, specificity, and recall. Table 3 shows the results, in which the best performance for each of the criteria is highlighted.

Table 3. Results for Covid-19 using 100 slices of each image to test the dataset.

Models	Precision	Accuracy	Recall	Specify	Weights
ResNet50V2	98.23	**99.47**	**98.58**	**99.72**	23,570,980
DenseNet121	**99.28**	99.41	97.16	99.44	**7,040,613**
InceptionResNetV2	98.57	99.41	97.87	99.58	54,341,381
VGG-19	98.86	98.54	92.20	98.48	20,025,957
DenseNet169	99.21	98.13	89.36	97.94	12,647,909
DenseNet201	96.60	97.95	90.78	98.20	18,327,781
ResNet101V2	98.17	98.89	95.04	99.02	42,632,741
VGG-16	97.55	97.13	84.75	97.06	14,716,261

As the objective of this work is to develop a Medical Diagnostic System, the most important parameters of the networks are precision and specificity. The neural network with better precision was DenseNet121 and the network with better specificity was ResNet50V2. Analyzing the other parameters, we observed that, in general, ResNet50V2 presented the best performance.

Next, the performance of the networks was compared with the literature 7. Table 4 compares the precision results and Table 5 compares the specificity results. It is noted that all models presented a more accurate result for the classification of Covid19 and Pneumonia, however, for cases of No disease, the literature was better. The literature obtained a result with a better specificity for the classification of Covid-19, however, the results of the neural networks used in this work reached better values of specificity in the classification of Non-disease and Pneumonia.

Table 4. Precision of classification models for classes.

Models	Covid-19	No Disease	Pneumonia
[7]	90,30	95,70	90,30
ResNet50V2	98,23	94,32	95,90
DenseNet121	99,28	95,51	94,65
InceptionResNetV2	98,57	94,98	95,05
VGG-19	98,86	94,55	94,10
DenseNet169	99.21	88.05	93.08
DenseNet201	96.60	91.65	92.03
ResNet101v2	98.17	95.55	94.88
VGG-16	97.55	91.65	89.82

However, when looking at specificity, shown in Table 5, the literature [7] obtained the best result, 2.81% higher than the article, while for the other labels, higher values were obtained in all models, reaching 8.13% greater than [7] in best model.

Table 5. Specificity of classification models for classes.

Models	Covid-19	No Disease	Pneumonia
[7]	**100**	90.00	93.00
ResNet50V2	99.72	96.98	96.02
DenseNet121	99.30	96.59	**96.91**
InceptionResNetV2	99.58	96.58	96.45
VGG-19	98.48	**97.32**	96.19
DenseNet169	97.94	96.88	92.02
DenseNet201	98.20	95.12	94.65
ResNet101v2	99.02	96.88	95.17
VGG-16	97.06	95.12	94.26

Finally, through the Probabilistic Grad-CAM method, the heatmaps of the models in the article were generated using 400 image packages, to obtain a clearer image, as shown in Figure 7. In red are the locations that most influenced the decision and in blue the least influential, where you can see that the regions of interest are in the left lung. DenseNet121 shows less dispersed results.
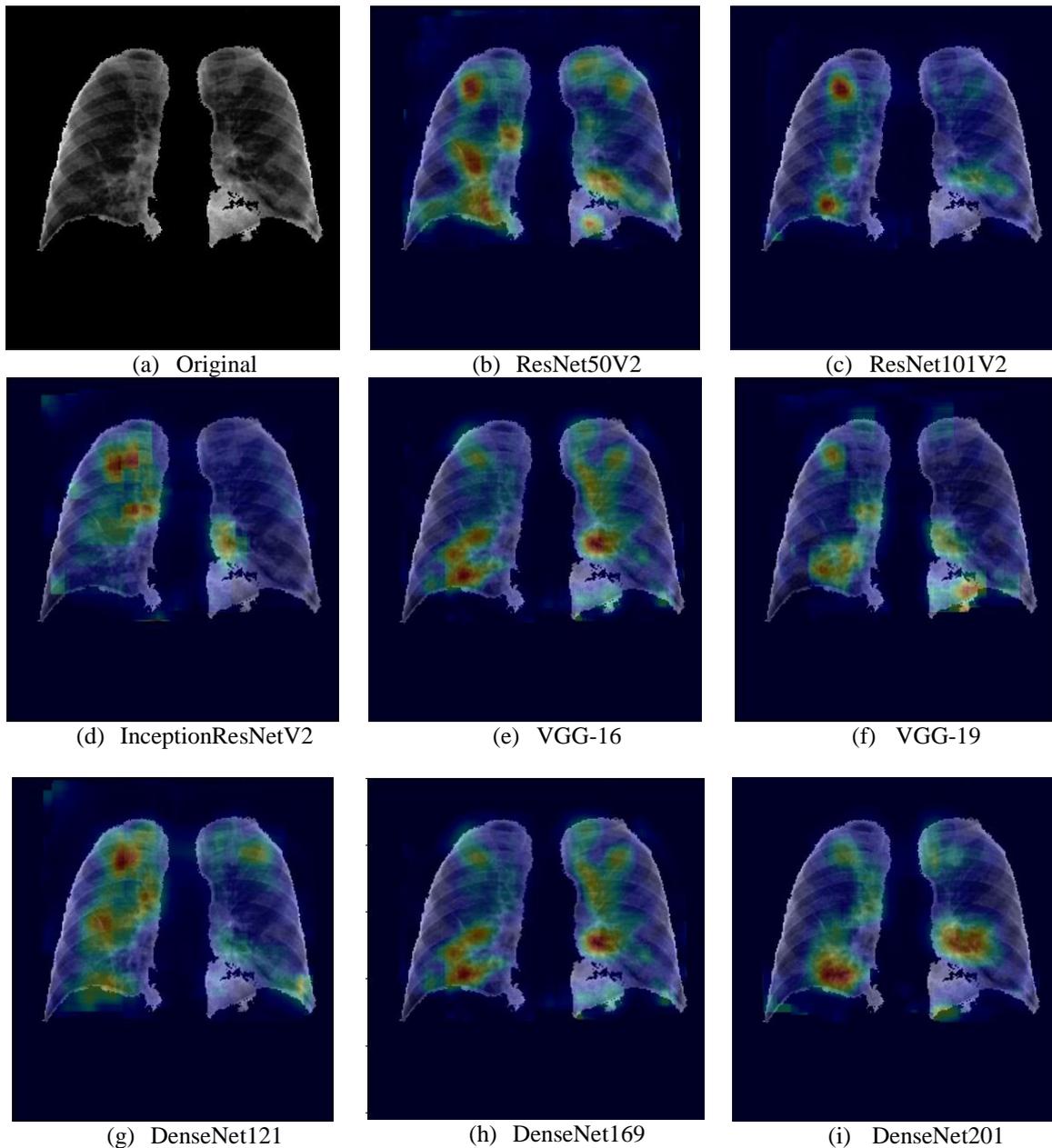
Figure 7. Grad-CAM Probabilistic generated by classifications models to Covid-19 using 400 slices.

5. CONCLUSIONS

With the current pandemic situation in the world, new diagnostic methods for these new diseases are increasingly necessary. This article improves the performance of the model suggested by [7]. The DenseNet121 network obtained an improvement in the precision of 9.94 when compared to the literature, however, in the specificity the performance was 0.7 worse. ResNet50V2 obtained an 8.78 higher precision than the literature and a 0.28 worse performance in specificity. Despite the worsening of specificity seen in ResNet50V2, this value was small compared to the high gain that this network achieved in precision.

Finally, when analyzing the images generated by the Probabilistic Grad-CAM, it can be seen that the regions of interest are in similar locations, however, the DesNet121 models obtained less diffuse regions.

ACKNOWLEDGMENTS

REFERENCES

[1]   Buslaev, A., Iglovikov, V.I., Khvedchenya, E., Parinov, A., Druzhinin, M., Kalinin, A.A.: Albumentations: Fast and flexible image augmentations. Information 11(2) (2020). https://doi.org/10.3390/info11020125, https://www.mdpi.com/2078-2489/11/2/125

[2]   Cohen, J.P., Morrison, P., Dao, L.: Covid-19 image data collection. arXiv 2003.11597 (2020), https://github.com/ieee8023/covid-chestxray-dataset.

[3]   De Oliveira Lima, C.M.A.: Information about the new coronavirus disease (COVID-19). Radiologia Brasileira 53(2), v–vi (2020). https://doi.org/10.1590/0100-3984.2020.53.2e1.

[4]   Dong, D., Tang, Z., Wang, S., Hui, H., Gong, L., Lu, Y., Xue, Z., Liao, H., Chen, F., Yang, F., Jin, R., Wang, K., Liu, Z., Wei, J., Mu, W., Zhang, H., Jiang, J., Tian, J., Li, H.: The Role of Imaging in the Detection and Management of COVID-19: A Review. IEEE Reviews in Biomedical Engineering 14(c), 16–29 (2021). https://doi.org/10.1109/RBME.2020.2990959.

[5]   Jadon, S.: A survey of loss functions for semantic segmentation. 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, CIBCB 2020 (2020). https://doi.org/10.1109/CIBCB48159.2020.9277638.

[6]   Jaeger, S., Candemir, S., Antani, S., Wáng, Y.X.J., Lu, P.X., Thoma, G.: Two public chest x-ray datasets for computer-aided screening of pulmonary diseases. Quantitative imaging in medicine and surgery 4(6), 475–477 (2014). https://doi.org/10.3978/j.issn.2223-4292.2014.11.20.

[7]   Oh, Y., Park, S., Ye, J.C.: Deep learning covid-19 features on cxr using limited training data asets. IEEE Transactions on Medical Imaging 39(8), 2688–2700 (2020).https://doi.org/10.1109/TMI.2020.2993291.

[8]   Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 9351, 234–241 (May 2015). https://doi.org/10.1007/978-3-319-24574-428, https ://arxiv.org/pdf /1505.04597.pdf.

[9]   Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.: Grad-cam: Visual explanations from deep networks via gradient-based localization. International Journal of Computer Vision 128(2), 336–359 (2020). https://doi.org/10.1007/s11263-019-01228-7, https://arxiv.org/pdf/1610.02391.pdf.

[10] Stirenko, S., Kochura, Y., Alienin, O., Rokovyi, O., Gordienko, Y., Gang, P., Zeng, W.: Chest x-ray analysis of tuberculosis by deep learning with segmentation and augmentation. arXiv (2018).

[11] Wang, L., Lin, Z.Q., Wong, A.: Covid-net: a tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. Scientific Reports 10(1), 19549 (Nov 2020). https://doi.org/10.1038/s41598-020-76550-z, https://doi.org/10.1038/s41598-020-76550-z.

[12] Zhao, W., Zhong, Z., Xie, X., Yu, Q., Liu, J.: Relation between chest CT findings and clinical conditions of coronavirus disease (covid-19) pneumonia: A multicenter study. American Journal of Roentgenology 214(5), 1072–1077 (2020). https://doi.org/10.2214/AJR.20.22976.

[13] TABIK, S. et al. COVIDGR Dataset and COVID-SDNet Methodology for Predicting COVID-19 Based on Chest X-Ray Images. IEEE Journal of Biomedical and Health Informatics, v. 24, n. 12, p. 3595–3605, 2020. ISSN 21682208.

[14] ZIMATORE, C. et al. Accuracy of the Radiographic Assessment of Lung Edema Score for the Diagnosis of ARDS. Frontiers in Physiology, v. 12, n. June 2021, 2021. ISSN 1664042X.

AUTHORS

Gustavo Chichanoski, Bachelor in Electronic Engineering, work in Department of Electronic Engineering, State University of Londrina, Londrina, Paraná

Maria Bernadete de Morais França, Doctor in Electronic Engineering work in Department of Electronic Engineering, State University of Londrina, Londrina, Paraná