

Statistical Analysis of X-ray Images and Laboratory Studies of COVID-19 Patients

Análisis estadístico de imágenes de rayos X y estudios de laboratorio de pacientes con COVID-19

SÁNCHEZ-TORRECITAS, Enrique†*, YAÑEZ-VARGAS, Juan Israel, QUINTANILLA-DOMÍNGUEZ, Joel and AGUILERA-GONZÁLEZ, José Gabriel

Universidad Politécnica de Juventino Rosas, Master of Engineering, Intelligent Systems. Mexico.

ID 1st Author: *Enrique, Sánchez-Torrecitas* / ORC ID: 0000-0002-5395-6749, CVU CONACYT ID: 1012800

ID 1st Co-author: *Juan Israel, Yañez-Vargas* / ORC ID: 0000-0001-5749-8442, CVU CONACYT ID: 295711

ID 2nd Co-author: *Joel, Quintanilla Domínguez* / ORC ID: 0000-0003-2442-2032, CVU CONACYT ID: 165237

ID 3rd Co-author: *José Gabriel, Aguilera-González* / ORC ID: 0000-0002-4160-448X, CVU CONACYT- ID: 329447

DOI: 10.35429/JSTA.2021.20.7.10.16

Received July 15, 2021; Acceptance December 30, 2021

Abstract

The large increase in COVID-19 cases that continue occurring in Mexico has generated the need to learn more about the damage caused by this virus, for which it has been proposed to carry out a statistical study of the 164 ray images database X and 29 laboratory studies provided by the Salamanca State Center for Critical Care (CECCS) of patients diagnosed with COVID-19 and not COVID-19, the purpose is to analyze if there is any correlation between X-rays and laboratory studies. First, the entire database was organized by patient, both for X-ray and laboratory studies. Afterwards, the statistical process of the laboratory studies was carried out. Subsequently, we worked on the selection of patients with a greater number of X-rays, segment of these the area of the lungs, since it was our area of interest. In these radiographs, the statistical process of entropy and box-and-whisker plot is performed. Finally, the results of some laboratory studies are correlated with the entropy and box-and-whisker plot results.

COVID-19, X-rays, Analysis of laboratory studies

Resumen

El gran aumento de casos de COVID-19 que se siguen presentando en México ha generado la necesidad de conocer más a fondo los estragos que genera este virus, por lo cual se ha propuesto realizar un estudio estadístico de la base de datos 164 imágenes de rayos X y 29 estudios de laboratorio proporcionados por el Centro Estatal de Cuidados Críticos de Salamanca (CECCS) de pacientes diagnosticados con COVID-19 y no COVID-19, cuyo propósito es analizar si existe alguna correlación entre los rayos X y los estudios de laboratorio. Primero, se organizó toda la base de datos por paciente, tanto de rayos X como de estudios de laboratorio. Después, se realizó el proceso estadístico de los estudios de laboratorio. Posteriormente, se trabajó en la selección de los pacientes con una mayor cantidad de rayos X, segmento de estas el área de los pulmones, ya que era nuestra zona de interés. En estas radiografías, se realiza el proceso estadístico de entropía y diagrama de caja y bigotes. Finalmente se hace una correlación de los resultados de algunos estudios de laboratorio con los resultados de entropía y diagrama de caja y bigotes.

COVID-19, Rayos X, Análisis de estudios de laboratorio

Citation: SÁNCHEZ-TORRECITAS, Enrique, YAÑEZ-VARGAS, Juan Israel, QUINTANILLA-DOMÍNGUEZ, Joel and AGUILERA-GONZÁLEZ, José Gabriel. Statistical Analysis of X-ray Images and Laboratory Studies of COVID-19 Patients. Journal of Scientific and Technical Applications. 2021. 7-20:10-16.

* Correspondence to Author (Email: m19030002@upjr.edu.mx)

† Researcher contributing as first author.

Introduction

The birth of the new coronavirus SARS-CoV-2 (COVID-19) as of 2019, has generated millions of infections and thousands of deaths throughout the world to date, creating a great challenge in the scientific community in the creation of vaccines and drugs to combat the virus, at the same time it begins with large studies for its timely detection and complete analysis.

Given the results published to date, it is observed that the behavior of COVID-19 varies in each country or region, and this depends on multiple factors such as: predisposition to diseases, the most common diseases in the country, age, climate, low defenses and more. Unfortunately, Mexico has a high vulnerability to COVID-19, due to the Mexican's health history (Ministry of Health | Government | gob.mx, s. F.).

According to the Secretary of Health of the Government of Mexico through CONACYT, they report that at the national level as of April 14, 2021 there are 2,228,133 confirmed, negative 3,579,530, suspect 439,531, deaths 210,294, recovered 1,815,237 and active 24,939, of which In the state of Guanajuato, as of the indicated date, there are 129,138 confirmed, 146,629 negative, 16,395 suspects, 10,215 deaths, 106,161 recovered, and 749 active (COVID-19 Tablero México, nd).

Due to the aforementioned, it is necessary to know the evolution of the virus in our region, for which it is intended to carry out a statistical study of the X-ray images and clinical data provided by the State Center for Critical Care of the City of Salamanca (CECCS), Guanajuato. These images and clinical laboratories belong to patients who tested positive for COVID-19 and who were admitted to CECCS.

This analysis seeks to make a statistical correlation between clinical data and X-ray images of patients with COVID-19 to understand the behavior of the virus both in the lungs and in the human body over time.

In the study by Jacobi et al. (2020, 38 p.) Shows a pictorial review in which they describe the most common manifestations and patterns of pulmonary abnormalities on a portable chest radiograph (CXR) in COVID-19.

In their article Liu et al. (2020, p. 370) report on epidemiological characteristics, laboratory and radiological studies, as well as possible biomarkers to predict the severity of the disease in patients infected with the new coronavirus (2019-nCoV) in the Shenzhen sub-province of China. The study was conducted on 12 patients who were infected with COVID-19 in the city of Wuhan, where the outbreak of this pandemic began. They mention the symptoms that patients present at the beginning, the resulting clinical characteristics and blood biochemical indices obtained.

In their research work, Zhang et al. (2020b, p. 1738) present the analysis carried out on 140 patients infected with SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) in Wuhan, China, who underwent a demographic investigation, computed tomography studies, analysis of comorbidity, different laboratory tests, presenting a statistical correlation between lymphocytes and eosinophils through the Spearman correlation coefficient.

Yun *et al.*, (2020, 95 p.) Show the statistical analysis of clinical data obtained from 2510 patients, who underwent the nucleic acid test for COVID-19, influenza type A and influenza type B, presenting the results obtained for each one of viruses, as well as differences in hematological indices (lymphocytes, platelets, white blood cells, monocytes, red blood cells, hematocrit, hemoglobin, basophils).

K. Liu *et al.*, (2020, p. 1027) present their report of a study carried out on 137 patients admitted to hospitals in the Hubei province in China, who underwent laboratory tests, computed tomography (CT) of the lungs in which They presented abnormal characteristics and some treatments that were carried out without any success.

In their article Ng et al. (2020) present a study carried out on 21 patients infected with COVID-19, cases that occurred in Shenzhen and Hong Kong, China. His research focuses on the analysis of computed tomography and radiographic follow-up, in addition to presenting the statistical analysis of some characteristics of medical history and hematological and biochemical findings.

According to the information available worldwide, there is still great uncertainty about the behavior of this virus, which is why all the possible information that can be provided for the benefit of the human being is of great relevance. In this research, a correlation is made between some laboratory studies and the X-ray images provided by the CECCS of the state of Guanajuato, where a similarity can be seen in the information obtained by the X-ray images and the results of the studies. from laboratory.

Material and methods

For the development of this research, the CECCS had the support, which provided approximately 164 X-ray images, of which 11 presented errors, since they did not correspond to chest images and only 153 images were used, which were taken with a mobile X-ray machine model Primax Cybermovil TS30. Information was obtained from 29 laboratory studies of 109 patients of which 86 tested positive for COVID-19 in the PCR Polymerase Chain Reaction test (Dra. Celia M. Alpuche Aranda, 2020), 18 tested negative for COVID-19 and 5 were under examination. Information collected from March 24 to October 20, 2020.

COVID-19 patient clinical laboratories

Laboratory clinical analysis data

For the analysis of clinical data, there are 29 laboratory studies that were performed on patients with COVID-19 upon entering the CECCS and during their stay, being the following: Glucose, Urea, Ureic Nitrogen, Creatinine, Uric Acid, Total Proteins, Albumin, Globulin, A / G Ratio, Alanine Aminotransferase, Aspartate Aminotransferase, Total Bilirubin, Direct Bilirubin, Indirect Bilirubin, Lactic Dehydrogenase, Creatinine-Phosphokinase, Chlorine, Potassium, Sodium, Hematology Sedimentation Rate, Globular Cartometry Time of Prothrombin, Fibrinogen, Procalcitonin, Calcium, Phosphorus, Alkaline Phosphatase, Magnesium, Triglycerides and C-Reactive Protein (Pagana & Pagana, 2015).

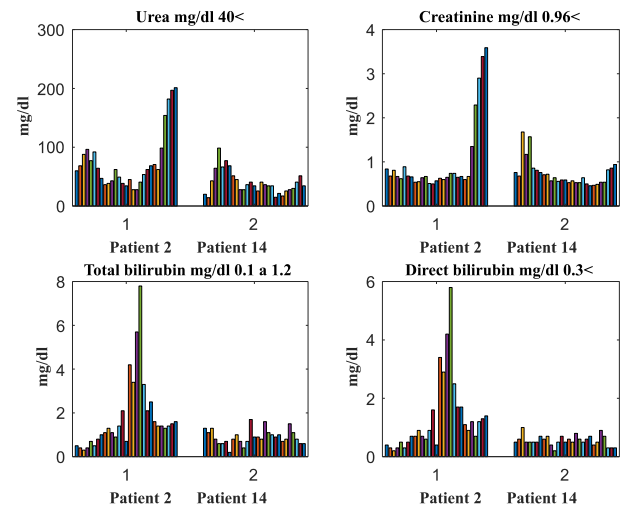


Figure 1 Charts of clinical laboratories. Relationship of the results of patient 2 and patient 14 where the x-axis corresponds to 29 days divided into two patients, while the y-axis represents the values of the units in which it is measured

Source: Own [Matlab]

With the results of the laboratory studies, an analysis of their evolution over time is carried out, as well as a comparison of the studies between different patients from admission to discharge or death.

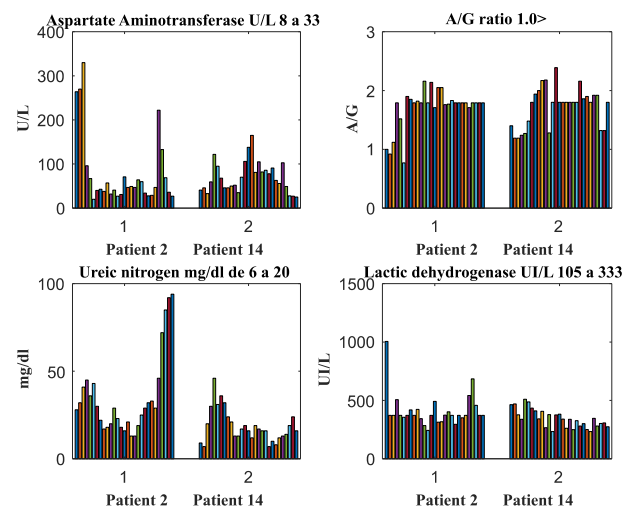


Figure 2 Charts of clinical laboratories. Relationship of the results of patient 2 and patient 14 where the x-axis corresponds to 29 days divided into two patients, while the y-axis represents the values of the units in which it is measured

Source: Own [Matlab]

The graphs in Figure 1 and Figure 2 show a comparison of some clinical studies performed on patient 2 (death) and patient 14 (discharge) during their stay at CECCS, such as Urea, Creatinine, Total Bilirubin, Direct Bilirubin, Aspartate Aminotransferase, Albumin / Globulin Ratio, Ureic Nitrogen and Lactic Dehydrogenase, where the x-axis represents the 29 days divided into two patients, while the y-axis represents the values of the laboratory study in the determined units. The graphs show the variations of the normal values of patient 2 in relation to patient 14.

In patient 2, Figure 1, normal values show a greater increase compared to patient 14, where it can be observed in the study of Urea and Creatinine that in the last results there is a very high increase. In the case of the Total Bilirubin and Direct Bilirubin graphs, there are very marked increases from day 16 to day 22 where the increases in the values are well above normal.

Figure 2 shows patient 2 in his first tests for Aspartate Aminotransferase, the results are high and over time there is a tendency to normalize, but in the latest studies, he again shows an increase. In patient 14, the opposite happens, since in the first results, a normal trend is appreciated, but later increases are generated to finally show a tendency to normalize again. In the graphs of Ureic Nitrogen and Lactic Dehydrogenase, it is observed how in patient 2 from day 25 to day 29 there is a very pronounced increase and as for patient 14 it is appreciated that there is a tendency to normalize.

X-ray images of COVID-19 patients

In the study of X-ray images, they were first organized by patient and by date of taking the X-ray. Subsequently, those patients with a greater amount of X-rays were selected, such as the example shown in Figure 3, since this will allow us to know what happens in the lungs over time. After the selection, we worked on the segmentation of the rectangular area of the lungs as can be seen in Figure 4 being the area of interest.

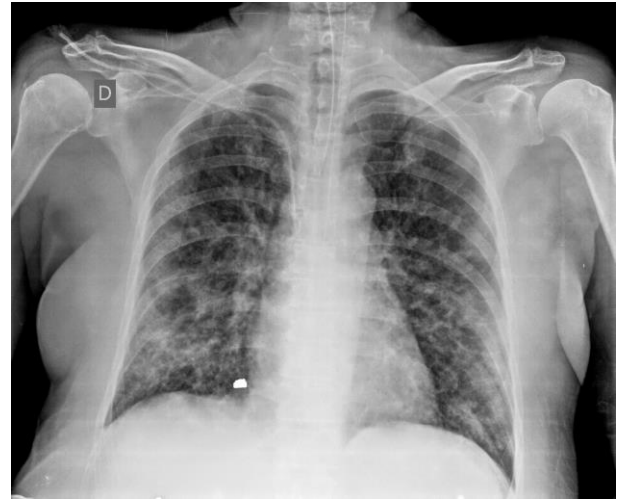


Figure 3 X-ray image of patient 14 provided by the CECCS

Source: Own [Matlab]

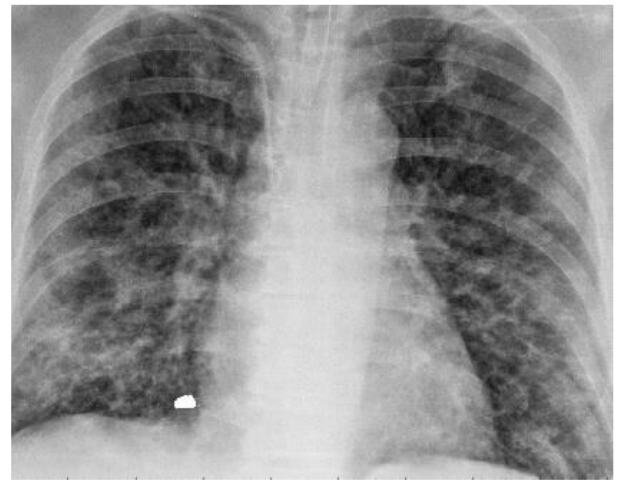


Figure 4 Patient 14 X-ray image provided by the CECCS. Segmentation of the lung area

Source: Own [Matlab]

The following graphs (Figure 5, Figure 7 and Figure 9) show the comparison of the Shannon (S) entropy statistical process performed to a sequence of X-ray images from the database provided by CECCS, images that were taken on consecutive days and that present a greater appreciation of the lungs, which are called Healthy lung (Figure 5), Patient 2 (Figure 7), Patient 14 (Figure 9), where the x-axis corresponds to the X-ray images X taken per day and the y-axis corresponds to the normalized values of entropy. Figure 5 shows the entropy results performed on a healthy lung and Figure 7 and Figure 9 show the entropy results performed on two COVID-19 positive patients named as Patient 2 and Patient 14.

Figure 7 shows the relationship that corresponds to the entropy obtained from 14 radiographs performed on a female patient who died, and Figure 9 shows the results of the entropy performed on 9 radiographs which correspond to a female patient. female who successfully graduated from COVID-19.

$$S = - \sum_{i=1}^k (p_i \log_2 p_i) \quad (1)$$

Equation (1) Shannon entropy (Gonzalez *et al.*, 2003), where S is the average output information of the source that is obtained through the weighted sum of the clusters or groups of pixels that goes from $i = 1$ to k states, where p_i represents the probability of each of the states and the \log_2 represents the information in binary form.

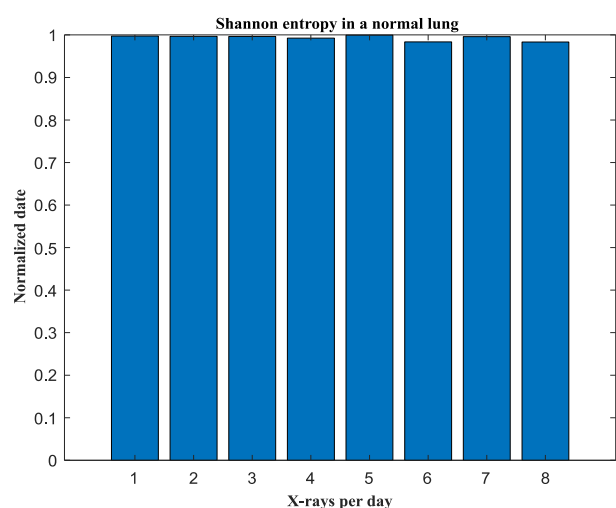


Figure 5 Shannon Entropy Plot of Healthy Lung. The x-axis represents the weighted summation of the pixels in each graphed x-ray image, while the y-axis represents the normalized entropy values
Source: Own [Matlab]

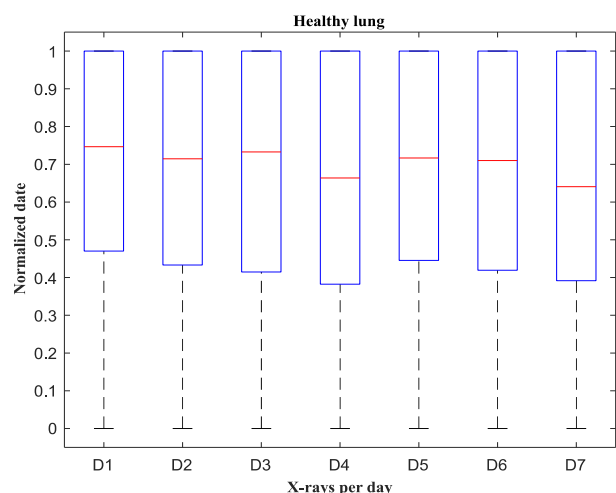


Figure 6 Normal lung box-and-whisker diagram. The x-axis represents the vector of the values of each radiograph taken in a sequence of days (D), while the y-axis represents the values of the normalized pixels
Source: Own [Matlab]

Results

The entropy plot of Figure 5 is statistically related to the box-and-whisker plot of Figure 6 where the x-axis represents x-ray images per day, while the y-axis represents my normalized data, through which it provides information on the distribution of the pixels, since, when relating Figure 5 of the healthy lung with Figure 6, a very high entropy can be seen which indicates that there is a very similar relationship between the values of the pixels and on the side. From the box-and-whisker plot, it can be seen that the median is between the values of 0.65 to 0.75.

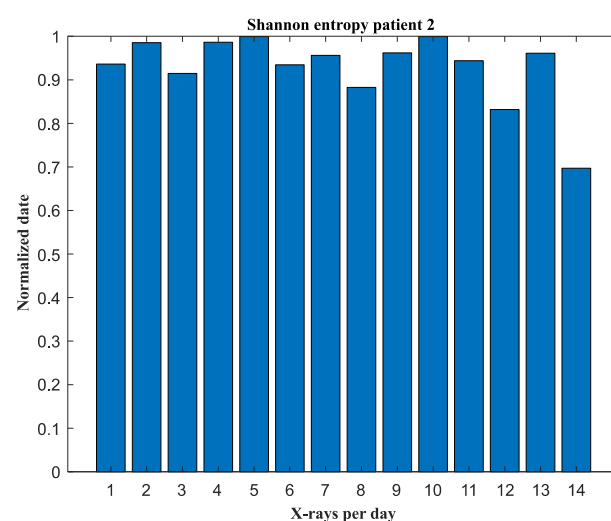


Figure 7 Shannon's Entropy Plot of the lung of patient 2 (deceased). The x-axis represents the weighted summation of the pixels of each graphed X-ray image per day, while the y-axis represents the normalized entropy values
Source: Own [Matlab]

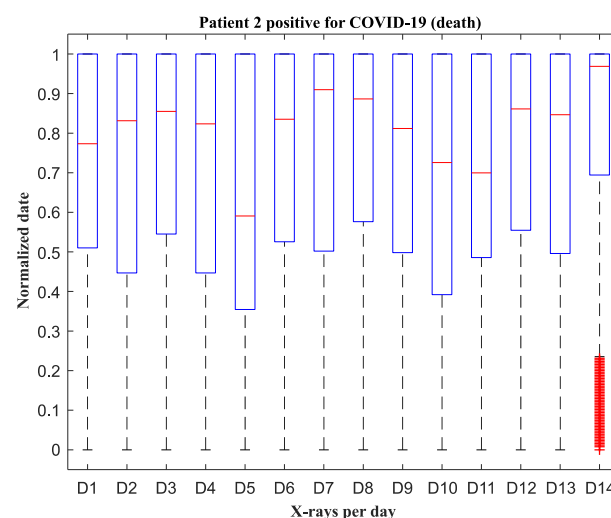


Figure 8 Box-and-whisker plot, patient 2. The x-axis represents the vector of the values of each radiograph taken in a sequence of days (D), while the y-axis represents the normalized pixel values
Source: Own [Matlab]

In the graphs of Figure 7 and Figure 8, the results of patient 2 who died from COVID-19 are presented, where, if the graph of the healthy lung is compared with the previous ones, the variability that occurs through the weather. In the entropy graph Figure 7, it can be seen how in the last X-ray image there a large change in entropy is, which statistically indicates that the pixel values underwent large changes, interpreting this as a greater invasion of the virus in the lungs. On the side of the box-and-whisker graph Figure 8, the change in the median values is statistically observed, since relating the last X-ray image, it is observed how the median is well above the value of 0.9 interpreting with this is that internally in the lungs very drastic changes were generated.

The graphs in Figures 9 and 10 correspond to patient 14 who was reported as discharge. Performing an analysis of the data obtained in entropy, it can be observed that statistically there are changes in the X-rays over time, but they were not very noticeable like the changes presented in patient 2. This can also be seen in the graph of boxes and whiskers where in the last X-ray it is observed that there is a very similar relationship with that of healthy lungs, since the median value is approximately 0.64.

Conclusions

In this article, a statistical analysis of the patients of the CECCS of the state of Guanajuato was carried out, which includes the X-ray history and laboratory studies of patients with COVID-19 and not COVID-19, where by relating the X-ray graphs of a normal lung with the results of the graphs of a lung of a patient infected with COVID-19, the drastic changes that occur in the infected lung are appreciated.

When making a correlation of the results of the X-ray entropy, the box and whisker diagrams with the graphs of some laboratory studies, it can be observed, according to what is presented, how there are very similar trends between the statistical results of the X-ray images and laboratory studies, not knowing for the moment if there is any antecedent that generates this relationship between them.

Acknowledgments

The authors thank the State Center for Critical Care of Salamanca (CECCS) and the Polytechnic University of Juventino Rosas (UPJR).

References

Jacobi, A., Chung, M., Bernheim, A., & Eber, C. (2020). Portable chest X-ray in coronavirus disease-19 (COVID-19): A pictorial review. *Clinical Imaging*, 64, 35–42. <https://doi.org/10.1016/j.clinimag.2020.04.001>

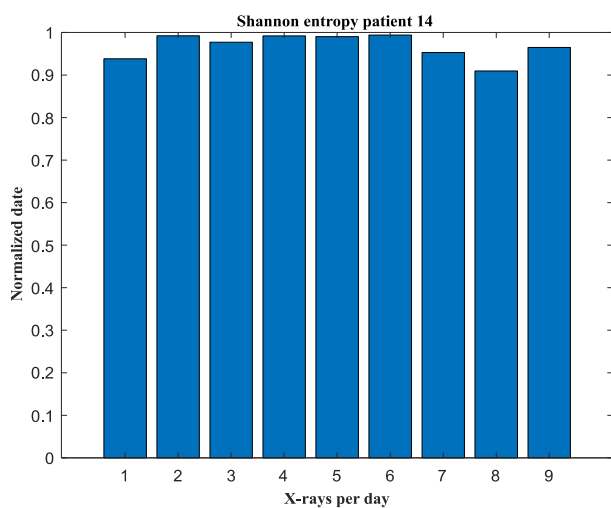


Figure 9 Shannon's Entropy Plot of the lung of patient 14 (discharge). The x-axis represents the weighted summation of the pixels of each graphed x-ray image per day, while the y-axis represents the normalized entropy values
Source: Own [Matlab]

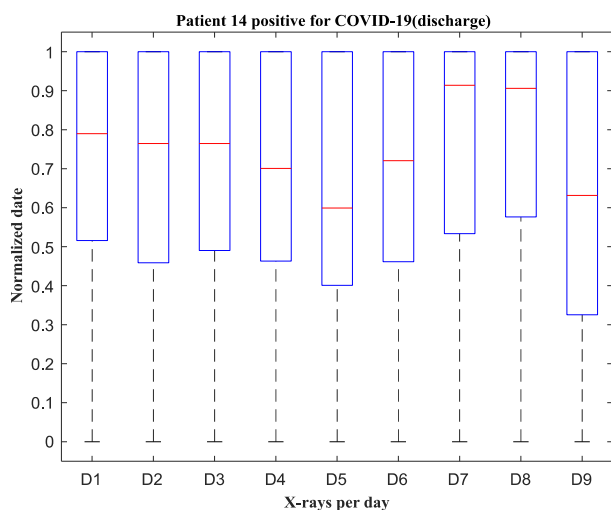


Figure 10 Box-and-whisker plot patient 14. The x-axis represents the vector of the values of each radiograph taken in a sequence of days (D), while the y-axis represents the normalized pixel values
Source: Own [Matlab]

Liu, Y., Yang, Y., Zhang, C., Huang, F., Wang, F., Yuan, J., Wang, Z., Li, J., Li, J., Feng, C., Zhang, Z., Wang, L., Peng, L., Chen, L., Qin, Y., Zhao, D., Tan, S., Yin, L., Xu, J., . . . Liu, L. (2020). Clinical and biochemical indexes from 2019-nCoV infected patients linked to viral loads and lung injury. *Science China Life Sciences*, 63(3), 364–374. <https://doi.org/10.1007/s11427-020-1643-8>

Yun, H., Sun, Z., Wu, J., Tang, A., Hu, M., & Xiang, Z. (2020). Laboratory data analysis of novel coronavirus (COVID-19) screening in 2510 patients. *Clinica Chimica Acta*, 507, 94–97. <https://doi.org/10.1016/j.cca.2020.04.018>

Zhang, J. J., Dong, X., Cao, Y. Y., Yuan, Y. D., Yang, Y. B., Yan, Y. Q., Akdis, C. A., & Gao, Y. D. (2020b). Clinical characteristics of 140 patients infected with SARS-CoV-2 in Wuhan, China. *Allergy*, 75(7), 1730–1741. <https://doi.org/10.1111/all.14238>

Liu, K., Fang, Y. Y., Deng, Y., Liu, W., Wang, M. F., Ma, J. P., Xiao, W., Wang, Y. N., Zhong, M. H., Li, C. H., Li, G. C., & Liu, H. G. (2020). Clinical characteristics of novel coronavirus cases in tertiary hospitals in Hubei Province. *Chinese Medical Journal*, 133(9), 1025–1031. <https://doi.org/10.1097/cm9.0000000000000744>

Ng, M. Y., Lee, E. Y. P., Yang, J., Yang, F., Li, X., Wang, H., Lui, M. M. S., Lo, C. S. Y., Leung, B., Khong, P. L., Hui, C. K. M., Yuen, K. Y., & Kuo, M. D. (2020). Imaging Profile of the COVID-19 Infection: Radiologic Findings and Literature Review. *Radiology: Cardiothoracic Imaging*, 2(1), e200034. <https://doi.org/10.1148/ryct.2020200034>

Gonzalez, R. C., Woods, R. E., & Eddins, S. L. (2003). *Digital Image Processing Using Matlab*. Prentice Hall.
 Secretaría de Salud | Gobierno | gob.mx. (s. f.). Secretaria de Salud. Retrieved May 9, 2021, de <https://www.gob.mx/salud>

COVID-19 Tablero México. (s. f.). COVID - 19 Tablero México. <https://datos.covid-19.conacyt.mx/>

Dra. Celia M. Alpuche Aranda. (2020, junio). Informe Técnico Reunión De Expertos Sobre Uso De Pruebas De Laboratorio Para Identificar SARS-COV-2 (N.o 1). Subsecretaría De Prevención Y Promoción De La Salud Dirección General De Epidemiología. https://coronavirus.gob.mx/wpcontent/uploads/2020/06/Docto_InfTec-SARSCOV2_26jun2020.pdf

Pagana, K., & Pagana, T. (2015). Laboratorio clínico: Indicaciones e interpretación de resultados (1.a ed.). Editorial El Manual Moderno.