EAI/Springer Innovations in Communication and Computing

Prashant Johri · Mario José Diván
Ruqaiya Khanam · Marcelo Marciszack
Adrián Will *Editors*

Trends and Advancements of Image Processing and Its Applications

EAI RESEARCH MEETS INNOVATION

Springer

EAI/Springer Innovations in Communication and Computing

Series editor
Imrich Chlamtac, European Alliance for Innovation, Ghent, Belgium

Editor's Note

The impact of information technologies is creating a new world yet not fully understood. The extent and speed of economic, life style and social changes already perceived in everyday life is hard to estimate without understanding the technological driving forces behind it. This series presents contributed volumes featuring the latest research and development in the various information engineering technologies that play a key role in this process.

The range of topics, focusing primarily on communications and computing engineering include, but are not limited to, wireless networks; mobile communication; design and learning; gaming; interaction; e-health and pervasive healthcare; energy management; smart grids; internet of things; cognitive radio networks; computation; cloud computing; ubiquitous connectivity, and in mode general smart living, smart cities, Internet of Things and more. The series publishes a combination of expanded papers selected from hosted and sponsored European Alliance for Innovation (EAI) conferences that present cutting edge, global research as well as provide new perspectives on traditional related engineering fields. This content, complemented with open calls for contribution of book titles and individual chapters, together maintain Springer's and EAI's high standards of academic excellence. The audience for the books consists of researchers, industry professionals, advanced level students as well as practitioners in related fields of activity include information and communication specialists, security experts, economists, urban planners, doctors, and in general representatives in all those walks of life affected ad contributing to the information revolution.

Indexing: This series is indexed in Scopus, Ei Compendex, and zbMATH.

About EAI

EAI is a grassroots member organization initiated through cooperation between businesses, public, private and government organizations to address the global challenges of Europe's future competitiveness and link the European Research community with its counterparts around the globe. EAI reaches out to hundreds of thousands of individual subscribers on all continents and collaborates with an institutional member base including Fortune 500 companies, government organizations, and educational institutions, provide a free research and innovation platform.

Through its open free membership model EAI promotes a new research and innovation culture based on collaboration, connectivity and recognition of excellence by community.

More information about this series at http://www.springer.com/series/15427

Prashant Johri • Mario José Diván
Ruqaiya Khanam • Marcelo Marciszack
Adrián Will
Editors

Trends and Advancements of Image Processing and its Applications

Springer

EAI
RESEARCH MEETS INNOVATION

Editors
Prashant Johri
Plot No 2, Sector 17 A, Yamuna Expressway
School of Computing Science &
Engineering, Galgotias University
G.B.Nagar, Greater Noida
Uttar Pradesh, India

Ruqaiya Khanam
Computer Science and Engineering
Department, SET
Sharda University
Greater Noida, Uttar Pradesh, India

Adrián Will
Information System Department
National Technological University -
Tucumán Regional Faculty
Tucumán, Tucumán, Argentina

Mario José Diván
Data Science Research Group at
Economy School
National University of La Pampa
Santa Rosa, Argentina

Marcelo Marciszack
Information System Department, Center for
Research, Development, and Transfer of
Information Systems (CIDS)
National Technological University -
Córdoba Regional Faculty
Córdoba, Argentina

ISSN 2522-8595 ISSN 2522-8609 (electronic)
EAI/Springer Innovations in Communication and Computing
ISBN 978-3-030-75944-5 ISBN 978-3-030-75945-2 (eBook)
https://doi.org/10.1007/978-3-030-75945-2

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To our parents:
Sh. A. N. Johri and Mithlesh Johri
Anita Koller
Bilqis Hamid
Gladys Beatriz Bussolini and Enrique Otto Marciszack

and to our families:
Methily Johri, Aniket Johri, and Mamta Saxena
José Ignacio, Santiago Agustín, and María Laura
Kashif A. Ahmad, Maaz, Almaas, and Fawaz
Erika, Federica, and María Alejandra
Amelia, Iván, and Jeremías

Preface

Image processing is a great tool for many applications in our day-to-day life. Anyone could be captivated by the variety of forms, colors, textures, and motion among other perceptible aspects in the world. Human beings have a great ability to acquire, integrate, and interpret the information surrounding them. It is a challenging task to impart these capabilities to a machine in a way to interpret the visual information deeply embedded in a still image, moving image, video, and graphics in the sensory world. As a result, there is a lot of application areas of digital image processing, such as multimedia computing, biomedical imaging, biometrics, security image data communication, pattern recognition, texture understanding, remote sensing, and image compression content-based image retrieval.

This book is intended to cover current technological innovations and applications in the emerging field of image processing and analysis techniques with application in the clinical industry, remote sensing, forensics, astronomy, manufacturing, defense, and many more that depend on image storage, display, and information disclosure about the world around us. The book presents the concepts and techniques of remote sensing, such as image mining and geographical and agricultural resources, medical image detection and diagnoses, image recognition and analysis, artificial intelligence, computer vision, machine learning, deep learning-based image analysis, pattern recognition and capsule networks in object tracking, remote sensing, moving object tracking, and wavelet transformation.

The book targets undergraduate, graduate, and post-graduate students, researchers, academicians, policymakers, government officials, academicians, technocrats, industry research professionals who are currently working in the field of academic research, and the research industry to improve the lifespan of the general public.

The book is organized into 16 chapters. Chapter 1 describes the application of convolutional neural network (CNN) for classifying COVID-19 based on computerized tomography scans. Chapter 2 outlines challenges related to image processing on mobile devices. Chapter 3 synthesizes a proposal applied to Smart City, using Internet of Things (IoT) and image processing. Chapter 4 proposes an application of artificial intelligence (AI) for dental image analysis. Chapter 5 analyzes the feature extraction techniques towards image processing of the plant extracts. Chapter 6 proposes a median filter based on entropy to remove noise in images. Chapters 7 and 8 describe the application of deep learning models for predicting COVID-19 and chronic myeloid leukemia, respectively. Chapter 9 addresses an automatic bean classification system based on visual features. Chapter 10 describes a supervised model to classify human sperm head. Chapter 11 proposes the future contribution of computational vision in the detection of maturity states of medicinal cannabis in Colombia. Chapter 12 describes the detection of brain tumor region based on magnetic resonance images (MRI) using clustering algorithms. Chapter 13 proposes the use of CNN for estimating human posture through pictures. Chapter 14 addresses a human skin detection technique supported by color models. Chapter 15 proposes a study of improved methods on image inpainting.

We have tried to gather a broad field of current and representative applications related to different perspectives of image processing. In such a sense, we want to express our gratitude to all of our contributors and friends, who brought their opinion and researches with different viewpoints, enriching the subject treatment.

Greater Noida, Uttar Pradesh, India — Prashant Johri
Santa Rosa, Argentina — Mario José Diván
Greater Noida, Uttar Pradesh, India — Ruqaiya Khanam
Córdoba, Argentina — Marcelo Marciszack
Tucumán, Argentina — Adrián Will

Acknowledgments

Our special thanks to God, our families, and friends for supporting, enlightening, and helping us every day.

This work has been partially supported by the CD 312/18 project at the National University of La Pampa (Argentina) and Grant UTN CCUTITU0007896TC from the National Technological University of Argentina.

Our gratitude to the Center for Research, Development, and Transfer of Information Systems (CIDS) of the Córdoba Regional Faculty of the National Technological University (Argentina).

Contents

Part I
Recent Trends and Advancements of Image Processing and its Applications

Using Convolutional Neural Networks for Classifying COVID-19 in Computerized Tomography Scans

Lúcio Flávio de Jesus Silva, Elilson dos Santos, and Omar Andres Carmona Cortes

1 Introduction

Computer-aided diagnosis (CAD) systems, according to Esteves et al. [1], "generally involve a classification step, which determines, for example, the presence or absence of a disease of interest." Thus, early detection of a disease can provide sufficient time for successful treatment or action.

CAD systems fit into a field called Medicine 4.0 [2]. One of the applications includes the combination of innovative artificial intelligence technologies, including machine learning algorithms, to develop support for clinical decisions. These decision support systems are related to procedures to improve medical decisions, providing evidence-based information at the time of contact between doctor and patient or even when deciding on treatment [3].

Machine learning (ML), according to Monard and Baranauskas [4], "is an Artificial Intelligence research field that studies the development of methods capable of extracting concepts (knowledge) from data samples." In general, ML algorithms are used in order to generate classifiers for a set of examples. Classification means the process of assigning, to a given data, the label of one (or more) class to which it belongs. In this sense, ML techniques are used to induce a classifier that, based on examples, is able to predict the class of new data of the amount in which it was trained.

L. F. de Jesus Silva · E. dos Santos
Universidade Estadual do Maranhão (UEMA), Programa de Pós-Graduação em Engenharia de Computação e Sistemas (PECS), Av. Cidade Universitária Paulo VI, Av. Lourenço Vieira da Silva, São Luís, MA, Brazil

O. A. C. Cortes (✉)
Instituto Federal de Educação, Ciência e Tecnologia do Maranhão (IFMA), Computer Department (DComp), São Luis, MA, Brazil
e-mail: omar@ifma.edu.br

P. Johri et al. (eds.), *Trends and Advancements of Image Processing and its Applications*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-75945-2_1

Machine learning approaches have evolved into deep learning approaches, which are a more robust and efficient way of dealing with the huge amounts of data generated from modern discovery approaches [5]. However, deep neural networks, especially convolutional neural networks (CNNs) that work with images, have many layers and connections, making them computationally complex for the complete training of the network. However, a solution called transfer learning can be applied to help with this problem. The idea is to use prior knowledge to solve similar problems, as humans do. Thus, transfer learning is a method of reusing a model or knowledge for another related task [6].

1.1 Problem Definition

The outbreak of coronavirus 2 2019 (COVID-19) of severe acute respiratory syndrome (SARS-CoV-2) has led to millions of people being infected and thousands of deaths worldwide. The World Health Organization (WHO) released the Situation Report reporting 41,570,883 cases and 1,134,940 deaths worldwide by October 23, 2020 [7]. The outbreak caused the health systems to collapse and questioned society's preparedness in the face of unknown virus pandemics, where immediate diagnosis and isolation could prevent a mass spread.

With the development of Computer-Aided Disease Diagnostic Systems, the possibility arose to determine, for example, the presence or absence of a disease of interest. Deep learning algorithms, mainly convolutional neural networks (CNNs), have been used successfully in the processing and analysis of digital images, including medical images [8]. The study on the performance of several CNN architectures in the analysis and classification of medical images can assist professionals and researchers in the development of these systems, applying the architecture that achieved the best performance in solving a certain problem.

In this work, 11 deep learning architectures will be presented in the task of computational analysis of medical images in order to compare the classifiers in the diagnostic task of COVID-19 analyzing computed tomography images. The architectures used are DenseNet121, DenseNet169, DenseNet201, Resnet50, Resnet50 v2, VGG16, VGG19, MobileNet, MobileNetV2, Xception, and Inception v3. To carry out the experiments, a database composed of 2477 computed tomography images was used, which is used to detect lung diseases, 1250 of which tested positive for COVID-19 and 1227 had a negative diagnosis.

1.2 Formulation of Hypotheses

In view of technological advances in the field of medicine, the deep learning algorithms have gradually helped professionals in the area in the analysis of clinical images; however, there is a lack of studies that identify which architectures are most suitable for analyzing these images. Therefore, which deep learning architectures performed better in the computational analysis of medical images?

1.3 Objective

Analyze and compare the performance of different deep learning architectures using transfer learning in the diagnosis of different pathologies in different databases, initially from COVID-19.

Specific Objective

The specific objectives of this work are:

- Carry out a bibliographic survey on different architectures of deep learning and transfer learning.
- Implement means for the pre-processing of computer tomography (CT) data.
- Determine techniques that assist in training the models.
- Compare the performance of the different deep learning architectures in the diagnosis of pathologies.

1.4 Methodology

To carry out this work, it was divided into five major activities: bibliographic survey, collection and analysis of sources, specification of requirements for the architectures used, architectural training, and performance analysis between different architectures.

The search for this content was carried out in books by renowned authors and cited in the context of research related to machine learning and deep learning; in addition to scientific articles with publications in magazines and university sites, search engines such as Google Scholar were also used. Moreover, specific words were used as a way to specify the search for this research material, such as machine learning, deep learning, artificial neural networks, image classification, and tomographic images. These descriptors generated significant assistance in the development of scientific research.

The main difficulty encountered was obtaining a robust dataset containing tomographic images of patients diagnosed with a specific pathology. However, there is a dataset called the SARS-COV-2 Ct-Scan Dataset [9], which united researchers with the study data, and this dataset was used for the first stage of this project.

A team of researchers from Qatar University, Doha, Qatar, and the University of Dhaka, Bangladesh, together with their collaborators from Pakistan and Malaysia in collaboration with doctors, have also created a database of chest X-ray images called COVID -19 Radiography Database containing positive cases of COVID-19 together with standard cases and images of viral pneumonia [10].

2 Literature Review

2.1 Machine Learning

According to Mitchell [11] "Machine Learning research seeks to develop techniques capable of extracting knowledge from examples of data." Machine learning algorithms are identified, the purpose of which is to enable a machine to be able to interpret new information and classify it appropriately after a certain training with a certain cluster of data by which the instances have family classification, from a generalization of what was shown previously. According to Libralão et al. [12], there are three paradigms that suggest how to learn the machine learning algorithm:

- Supervised: needs supervisors to achieve the best model intended in the training phase.
- Unsupervised: it allows the AM algorithm to learn to group the inputs taking into account a measure of similarity between them.
- For reinforcement: learning happens through bonuses, according to the performance achieved.

Also, according to Libralão et al. [12] most of the ML algorithms are inspired by biological systems (artificial neural networks (ANNs) and genetic algorithms), symbolic learning (decision trees), cognitive processes (case-based reasoning), and theory statistics (SVMs).

Artificial Neural Networks

ANN is an area of research that works with the simulation of human cognitive skills. Machines are developed that are capable of intelligent behavior, as if they were human behaviors. Human intelligence is the most developed among other creatures and the region responsible for this intelligence in humans is the brain [13].

Still according to Rauber [13], the essential elements are neurons, interconnected in networks, which allows the sharing of information between them, forming biological intelligence. An apparent interest that arises from these facts is the effort to copy the way the brain works in a computational environment. This means that the research tries to understand the functioning of the intelligence inhibited in neurons and diagram it in an artificial structure, such as a junction of hardware and software, thus converting biological neural networks into artificial neural networks.

Neuroscience Inspiration

What are the particularities of the human brain that enable it to behave intelligently? The following topics reflect the most significant characteristics that are mainly attractive to be used in an artificial neural network, according to Rauber [13]:

- Robustness and fault tolerance: The exclusion of neurons does not affect general functionality.
- Learning capacity: The brain has the ability to learn new tasks that have never been done before.
- Uncertain information processing: Even if the knowledge provided is incomplete, affected by noise, or relatively contradictory, precise reasoning is still possible.
- Parallelism: A huge number of neurons are active simultaneously. There is no restriction on a processor that must work with information after another.

Figure 1 shows the summarized model of a real neuron. The neuron is a cell that has a nucleus and a body where chemical and electrical behaviors demonstrate the treatment of information. The sum information is outputted by electrical impulses that diffuse through the axon. At the end of the axon are numerous branches that share information for neighboring neurons. The junction with other neurons is done through synapses that are linked to a dendrite of the receptor neuron. The synapse sends a chemical when it is alerted by the axon pulse. The substance spreads between synapse and dendrite, making the coupling between two neighboring neurons. According to the alerts (or inhibitions) that neighboring cells send to the cell under consideration, it treats the information once again and transmits it through its axon [13].

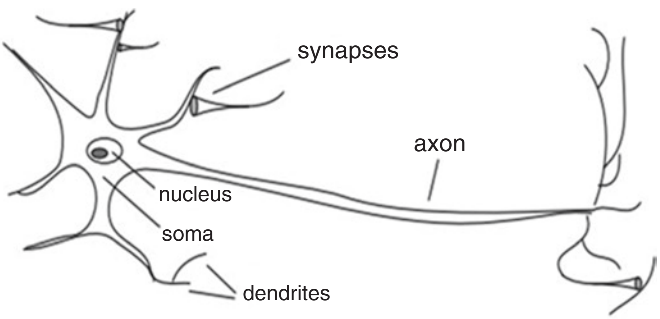

Fig. 1 Biological neuron, Rauber [13]

Brief History of Neurocomputing

An important point in the history of ANNs was the demonstration of a model of an artificial neuron by McCulloch and Pitts [14]. The works in this line of research ended in perceptron's conception by Rosenblatt [15] and, in a similar model, adaline by Widrow and Hoff [16]. Perceptron has the ability to classify between classes that are linearly separable. It was used to recognize characters, for example. This application was executed on a machine called MARK I PERCEPTRON, and it caused a great euphoria certainly exorbitant in relation to the imagination of the abilities of future intelligent robots.

Fundamentals

According to Prampero [17] "Artificial Neural Networks (ANNs) apply a mathematical model influenced by the neural architecture of living beings, contracting knowledge through experience."

Tatibana and Kaetsu [18] declare that "Artificial Neural Networks are formed by a grouping of nodes, which imitate the role of neurons, linked by a propagation principle. Each node acquires its inputs with the related weights, coming from other nodes, or from an external interaction. The input layer has a specific node, called a bias, which helps to increase the degrees of freedom, enabling a better adaptation of the network."

Also, on these inputs, an activation function is used, which uses a weighted sum as an argument in the network inputs. "The activation point of a node is defined by the activation function, usually a sigmoidal function or a step function" [19]. A layer that acquires signals from the outside environment is called an input layer and a layer that expresses signals to the environment is called an output layer. All others are defined as hidden layers and do not connect directly with the environment [20].

Learning Paradigms

According to Rauber [13], once the neural network has been determined, it must be trained. This means that the degrees of autonomy that the network has to solve the task under consideration must be adapted in the best way. Usually, this means that weights must be changed between neuron i and neuron j, according to an algorithm. A finite set T of n training samples is available to adapt the weights during the training of the network. A fundamental distinction in relation to the learning paradigm that is relevant for all types of systems with adaptation techniques is supervised learning and unsupervised learning.

In supervised learning, each training sample is accompanied by a cost that is the desired cost. This means that training set T is composed of n pairs of samples (xp yp) where T = {(xp yp)} np = 1. The dimension of the input vector is D, i.e., the input variables are associated with one multidimensional value, usually part of the

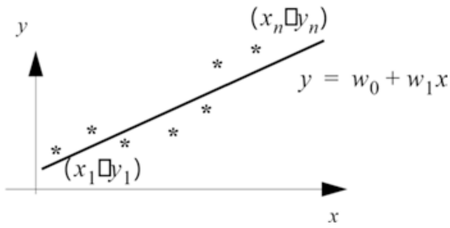

Fig. 2 Linear regression, Rauber [13]

real numbers: x = (x1… xj… xD)T, xj R. (The transposition (.)T is used to save space by registering a vector on a line.) The output variables are gathered in an output vector y = (y1… yi… yc) T [13].

According to Rauber [13], a demonstration of a supervised learning task is linear regression. The one-dimensional case is used to simplify the illustration, as shown in Fig. 2. In this problem the training set consists of pairs of real numbers (xp yp). The learning objective is the delimitation of coefficients w0 and w1 of the line y = w0 + w1x. The learning algorithm promotes to minimize the discrepancy between the desired value yp and the value that is the answer y′ = w0 + w1xp of the system, and this on average for each sample (xp yp).

In unsupervised learning, when there is only one information available, and that information is the values (xp) the learning task is to find correlations between the training examples T = {(xp)} np = 1. The number of groups or classes is not defined a priori. This means that the network has to find considerable statistical attributes, and it has to develop its own interpretation of the stimuli that enter the network. A synonym for unsupervised learning is clustering [13].

An example from the medical field is the detection of diseases using images, such as X-ray images. There are places within the image that can be attributed to the same material, such as bone. The number of materials (from agglomerations) is initially unknown. The objective of the system is to find the number of different materials and at the same time group each point of the image for the respective material. The entrance to the network could be the points of the image, for example, a small interval of 5 by 5 points. The expected response from the network would be the material to which this region of the image belongs [13].

2.2 Deep Learning

Deep learning allows computational models composed of several layers of processing to learn representations of data with varying levels of abstraction. According to Lecun [21] these methods have drastically improved the state of the art in speech recognition, object recognition, object detection, and many other domains. Deep learning discovers the intricate structure in large datasets using the backpropagation algorithm to indicate how a machine should change its internal parameters, which are used to calculate the representation in each layer from the representation in the previous layer. Deep convolutional networks brought advances in image, video, voice, and audio processing, while recurring networks illuminated sequential data, such as text and speech.

Convolutional Neural Networks

Convolutional neural networks are a variation of multilayer perceptron (MLP). By biological inspiration, CNNs emulate the basic mechanism of the animal's visual cortex. Neurons in CNNs share weights unlike MLPs, where each neuron has a separate weight vector [22].

Using the weight-sharing strategy, neurons are able to perform convolutions on the data with the convolution filter formed by the weights. This is followed by a pooling operation that is a form of nonlinear subsampling that progressively reduces the spatial size of the representation, thus reducing the number of parameters and computation on the network [22].

After several layers of convolution and pooling, the size of the input matrix (size of the characteristics map) is reduced and more complex characteristics are extracted. Eventually, with a small enough feature map, the content is compressed into a one-dimensional vector and fed into a fully connected MLP for processing [22].

Normally, between the convolution layer and the pooling layer, a ReLu layer is applied to which an unsaturated activation function on one side is applied element by element, such as $f(x) = \max(0, x)$, limiting the lowest layer output values to zero [22].

The last layer of the fully connected MLP, seen as the output, is a loss layer that is used to specify how training on the network penalizes the deviation between the predicted and true labels [22].

Architecture of Convolutional Neural Networks

A CNN architecture is built on three components: convolutional layer, pooling layer, and a fully connected one. Convolutional layers use filters that examine a part of the image and extract resources from it. These resources are usually colors, shapes, and borders that define a specific image [23]. CNNs can have as many convolutional layers as needed. The more convolutional layers, the more resources are extracted.

One-dimensional convolution is an operation between a vector of weights $m \in R$ ^ m and a vector of inputs seen as a sequence $s \in R$ ^ s. The vector m is the convolution filter. Concretely, we think of s as the input sentence and $si \in R$ is a unique characteristic value associated with the i-th word in the sentence. The idea behind the one-dimensional convolution is to take the scalar product of the vector m with each m-gram in the sentence to obtain another sequence c as shown in Eq. 1.

$$cj = m^t sj - m + 1 : j \tag{1}$$

After the convolutional layers, we add pooling layers that are responsible for grouping the resource maps in an image [23], providing a reduction in dimensionality, which is produced by applying a single maximum or average of the values within a box produced by the convolutional layer. Thus, the fully connected layers, which are normal MLP networks, receive a smaller image than the one presented at CNN's entrance.

2.3 Image Classification

Image classification is mainly the method of isolating images in previously defined groups, where images related to the same group have significant affinities, as well as having significant distinctions with images from the other groups. The definition of image classification rules consists of extracting their characteristics.

For the construction of an automatic classifier to become viable, [24], p. 3 declare that "the rules for the classification of images must be expressed numerically, thus offering a technique developed or adapted to recognize patterns within these rules." Promptly for image processing, rules based on pixel intensity are selected to classify objects or regions of images.

Tomography Scans

Computed tomography (CT) is a diagnostic imaging exam consisting of an image representing a section or "slice" of the body. It is obtained through computer processing collected after exposing the body to a succession of X-rays. Its main method is to analyze an X-ray beam's attenuation as it travels through a segment of the body [25]. Among the characteristics of tomographic images, pixels, matrix, field of view, grayscale, and windows stand out. The pixel is the smallest point that can be obtained in an image. The greater the number of pixels in a matrix, the better its spatial resolution, which allows for better spatial differentiation between structures.

Functional Images

Functional imaging is the state of the art in medical diagnostic imaging. Thus, it was mentioned that "Magnetic Resonance has made a huge advance in this direction through ultra-fast obtaining techniques, which have made it possible to measure changes in the level of oxygen utilization resulting from the BOLD effect" [26], p. 1. There are two broad imaging methods in this line that are becoming evident: dynamic PET (positron emission tomography) and fMRI (functional magnetic resonance imaging). The first comes from nuclear medicine and the second from magnetic resonance. For the first, a device was developed to obtain images and, for the second, hardware and software techniques, taking advantage of the device's same base [27]. Functional images have been used mainly in research on brain functioning, being even able to provide information for surgical planning.

PET and Dynamic PET

According to Queirós [27] PET (positron emission tomography) is a diagnostic method developed with the purpose of schematizing the use of tissue glucose, transforming it into an efficient instrument for the detection of tumors. The isotope usually applied is fluoride linked to deoxyglucose, called FDG. Fluorine 18 is a positron emitter, a particle that rapidly interacts with the electrons in the medium (called annihilation), forming a pair of gamma photons, which pass from the point of annihilation in the same path, but in opposite directions. According to Phelps [28], "FDG is a matter metabolized by the cell indifferently, as well as glucose, due to its similarity. The main use is in the search for metabolic variations that indicate cancer or metastasis."

fMRI

Functional magnetic resonance imaging (fMRI) is a method of the special use of magnetic resonance imaging (MRI) qualified to detect changes in blood flow in reaction to neurological activity, a phenomenon called the BOLD effect [29]. This BOLD effect is based on the magnetic state of hemoglobin; that is, hemoglobin has the ability to present different magnetic states according to its oxygenation state. Thus, according to Pauling and Coryell [30], deoxyhemoglobin is paramagnetic (highlighted), that is, it magnetizes itself in the direction of the magnetic field to which it is presented, and oxyhemoglobin is diamagnetic (indented), and these magnetic properties have a direct effect on the strength of the signal detected in the neural regions that are active. It is possible to ascertain that an increase in the oxyhemoglobin agglomeration in the blood flow will generate an increase in the strength of the signal captured and that in an inverse situation, that is, in the presence of a greater agglomeration of deoxyhemoglobin, there will be a decrease in local strength due to the realignment of T2 and T2 *. This is because the events that start with the increase in electrical activity and articulate the neurovascular response change the magnetic resonance signal over time and generate the hemodynamic response function [31].

To study fMRI, it is essential to obtain one or more time series of functional information, acquired during the production of sensory or motor stimuli or during the production of paradigms, which are collections of cognitive tasks, and to obtain anatomical data that contain the areas of interest that exercise the function of structural reference for the visual concession of active functional areas. After this concession, the location and particularization of the brain regions activated by the stimuli is made. To this end, it is essential to perform image processing phases since this whole process is subject to the interference of types of artifacts that can modify the images obtained [32]. fMRI has been widely used in scientific research and much less in clinical follow-up. According to Langleben and Moriarty [33], "in many cases, fMRI is related to other non-invasive techniques such as electroencephalography (EEG) and near-infrared spectroscopy (NIRS)."

2.4 Tools and Techniques

This section shows the tools, frameworks, and techniques that were used during the development of the experiments.

TensorFlow

TensorFlow is a machine learning system that operates on a large scale and in heterogeneous environments. TensorFlow uses data flow graphs to represent computing, shared state, and operations that change that state. It maps the nodes of a data flow graph on many machines in a cluster and, within one machine, on various computing devices, including multicore CPUs, general-purpose GPUs, and custom ASICs, known as Tensor Processing Units (TPUs). This architecture offers application developer flexibility: while in previous "parameter server" designs, shared state management is integrated into the system, TensorFlow allows developers to experiment with new optimizations and training algorithms. TensorFlow is compatible with a variety of applications, with a focus on training and inference in deep neural networks [34].

Keras

It is the high-level API of TensorFlow for creating and training deep learning models. It is used for rapid prototyping, cutting-edge research, and production, with the following advantages: It has a simple and consistent interface optimized for common use cases, and Keras modular and composite models are made by connecting configurable elements, with few restrictions [35].

NVIDIA CuDNN

It is a GPU-accelerated library of primitives for deep neural networks. CuDNN provides highly tuned implementations for standard routines, such as forward and backward convolution layers, pooling, normalization, and activation. CuDNN accelerates the widely used deep learning frameworks, including Caffe2, Keras, PyTorch, and TensorFlow [36].

Transfer Learning

It is the reuse of a pre-trained model in a new problem. The models used were trained with weights from ImageNet, which has millions of natural images in several categories. Training such a large mass of data might be very expensive, taking many days in advanced GPUs, then expending too much energy, consequently, money.

The purpose of transfer learning is to take advantage of all this knowledge that was generated using these natural images, whether in the most initial layers, such as the recognition of edges, curves, and colors, or in the innermost convolutional layers where they learn the texture, smooth. The more specialized layers are not interesting for learning, so only the initial layers are used that recognize basic details that theoretically every image has.

Callbacks

A callback is an object that can perform actions at various stages of the training (e.g., at the beginning or at the end of a season, before or after a single batch, etc.) [34]. Callbacks can be used to:

- Record TensorBoard records after each training batch to monitor your metrics
- Periodically save your model to disk
- Stop early
- Get a view of a model's internal states and statistics during the training

3 Experiments

This section details how the experiments were carried out, what equipment was used during the process, and how the data were obtained and treated for the training of the neural networks, in addition to the metrics used to evaluate the architectures.

3.1 Environment

All experiments were performed using a machine with the following configurations:

- CPU Intel Core i7 3770 @ 3.40GHz
- Memory 8GB Dual-Channel DDR3 @ 798MHz (10-10-10-30)
- GPU NVIDIA GeForce GTX 960 4GB (ZOTAC International)
- SSD 240GB KINGSTON (SATA-3)

3.2 Dataset

A Kaggle dataset called the SARS-COV-2 Ct-Scan Dataset provided by Soares et al. [9] was used, which united the researchers to the study data, and this dataset was used for the first stage of this project. The dataset consists of 2477 computed tomography images used to detect lung diseases, 1250 of which tested positive for COVID-19 and 1227 had a negative diagnosis. Figures 3 and 4 show a sample of the dataset with images of people with positive and negative diagnoses, respectively.

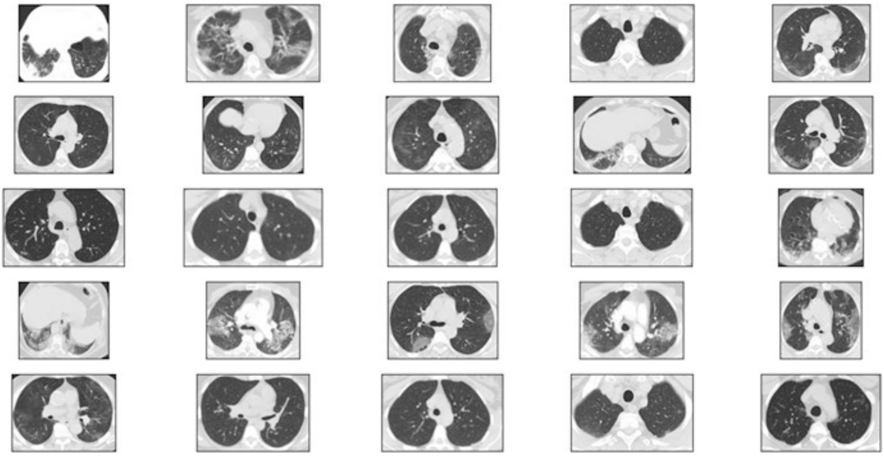

Fig. 3 Standardized COVID-19 images, [9]

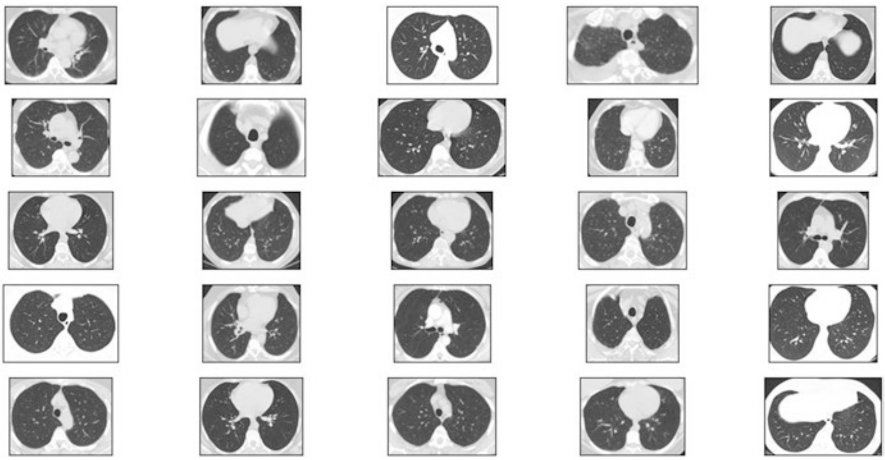

Fig. 4 Standardized non-COVID-19 images, [9]

The dataset has images with different resolutions; due to this, all images were converted to 128 × 128 resolution. In addition, the images were standardized and the validation sets were separated using ImageDataGenerator; this technique also makes image augmentation which generates new images while separating and normalizing. The following operations were also performed in order to avoid overfitting the neural network: rotation range = 10, width shift range = 0.2, height shift ranged = 0.2, zoom range = 0.1, horizontal flip = True, and vertical flip = True.

3.3 Definition of Models

The model was created using "include_top = false" to remove the specific piece of natural images when starting with the weights of ImageNet. Soon after, a pooling is added to reduce the size, and some dense layers are added. In dense layers, "relu" was used as activation and "he_normal" as kernel weight. Dropout with 0.3 was also used in each dense layer, and lastly, softmax to be able to separate the classes.

Adam was used with a learning rate of 0.002, the loss function as categorical_crossentropy, and accuracy as a metric. Soon after, the model is executed using cross-validation with k = 5 and fit_generator with 50 periods for each of the architectures.

The following callbacks were also used to perform actions at various training stages:

- ModelCheckPoint: to save the model that has the best loss during the training
- EarlyStop to stop training if the network stops learning
- ReduceLROnPlateau to decrease the learning rate if the loss of validation does not change

3.4 Evaluation Metrics

Initially, the meaning of true positives, true negatives, false positives, and false negatives was defined. True positives and true negatives are correct classifications, that is, classifying COVID-19 and non-COVID-19 correctly. In contrast, false positives and false negatives are wrong classifications. Thus, metrics were defined.

Equation 2 presents the first metric, called precision. This metric is the ratio between true positives and true positive plus false positives. Thus, a low precision indicates that the number of correct classifications is too low, or the number of false positives too is high.

$$P = \frac{TP}{TP + FP} \tag{2}$$

The next metric is the accuracy, as presented in Eq. 3, which is the percentage of correct classifications. A low accuracy could indicate that the number of wrong classifications (false positives and false negatives) is high.

$$A = \frac{TP + TN}{TP + FP + TN + FN} \tag{3}$$

The recall presented in Eq. 4 is the ratio between the true positives and true positives plus false positives. This metric indicates that the algorithm is performing well in classifying true positives. However, if this metric is low, it can mean that a high number of misclassifications is going on. Thus, this metric is essential to minimize the number of false negatives, which can produce the patient's worst scenario.

$$R = \frac{TP}{TP + FN} \tag{4}$$

Finally, the F1 score presented in Eq. 5 is the harmonic mean between precision and recall. In this context, F1 ends up being a big picture of the performance because precision takes into account false positives, and recall takes into account false negatives. Thus, F1 score gives an idea of whether the classifying algorithm is providing too many incorrect classifications.

$$\text{F1score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{5}$$

4 Results

Table 1 shows the best results for the accuracy and loss of the 11 CNNs according to the metrics presented above. In the same table, the average of all metrics used during training is also presented. As can be seen, DenseNet169 achieves the best precision, recall, F1_score, and accuracy of 96.3%, 96.4%, 96.4%, and 96.4%, respectively. In addition, DenseNet169 presented the best fold accuracy with 96.9%. MobileNetV2 had the worst results.

Considering that DenseNet169, VGG16, Xception, and VGG19 obtained the best results, these will be detailed below. DenseNet169 presented the best result. Figure 5 shows its average confusion matrix, in which we can observe that DenseNet169 was incorrect in only 18 cases; however, 12 cases gave false positives for COVID-19. On the other hand, the notable fact is that only six cases of COVID-19 were erroneously classified as non-COVID-19. Below, we detail the results of DenseNet169, showing the confusion matrices and training curves.

Figures 6, 7, 8, 9, and 10 show the accuracy and loss fold. As can be seen, both curves demonstrate the expected behavior, which is to achieve a value of one (100%)

Table 1 Result of 11 CNNs: accuracy, precision, recall, F1_score, accuracy fold, and loss fold

Model	Best fold accuracy	Best fold loss	Average precision	Average recall	Average F1_score	Average accuracy	Average loss
DenseNet169	*0.969*	*0.097*	*0.963*	*0.964*	*0.964*	*0.964*	*0.097*
VGG16	0.957	0.108	0.952	0.951	0.951	0.951	0.134
Xception	0.965	0.126	0.951	0.951	0.951	0.951	0.150
VGG19	0.967	0.109	0.947	0.946	0.946	0.946	0.134
DenseNet201	0.961	0.119	0.949	0.948	0.948	0.948	0.138
DenseNet121	0.957	0.153	0.946	0.947	0.947	0.947	0.155
ResNet50	0.959	0.108	0.943	0.942	0.942	0.943	0.149
ResNet50V2	0.949	0.134	0.940	0.940	0.939	0.939	0.163
InceptionV3	0.957	0.140	0.940	0.938	0.939	0.939	0.162
MobileNet	0.955	0.112	0.935	0.934	0.934	0.934	0.165
MobileNetV2	0.752	0.576	0.634	0.613	0.588	0.617	0.644

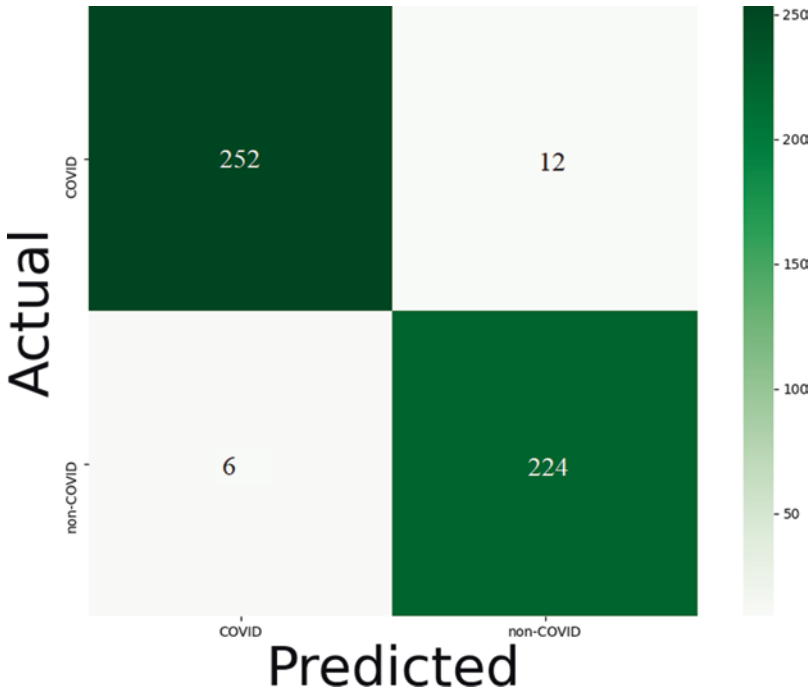

Fig. 5 Confusion matrix – DenseNet169

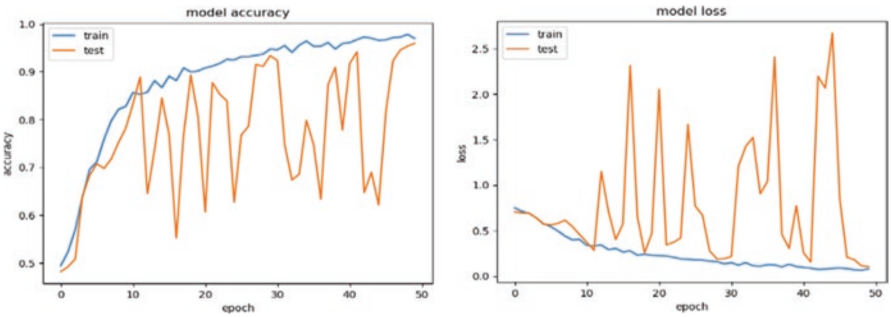

Fig. 6 Accuracy and loss fold 1 – DenseNet169

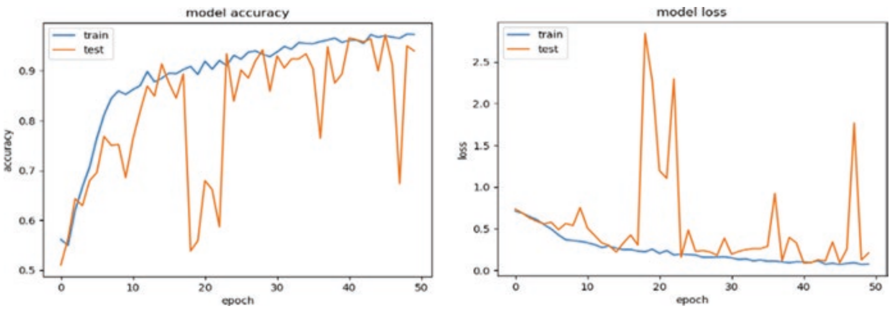

Fig. 7 Accuracy and loss fold 2 – DenseNet169

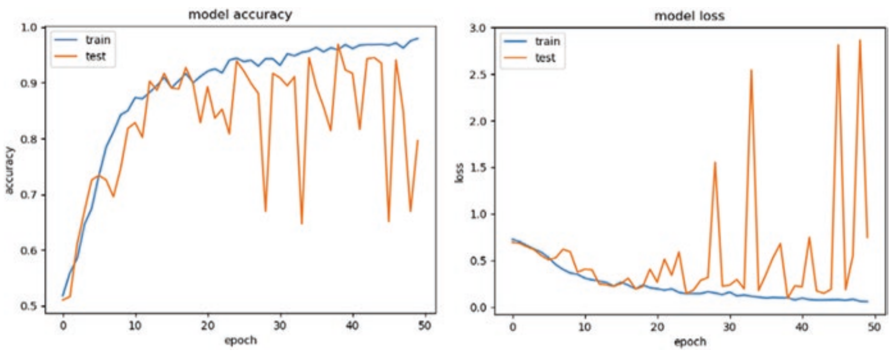

Fig. 8 Accuracy and loss fold 3 – DenseNet169

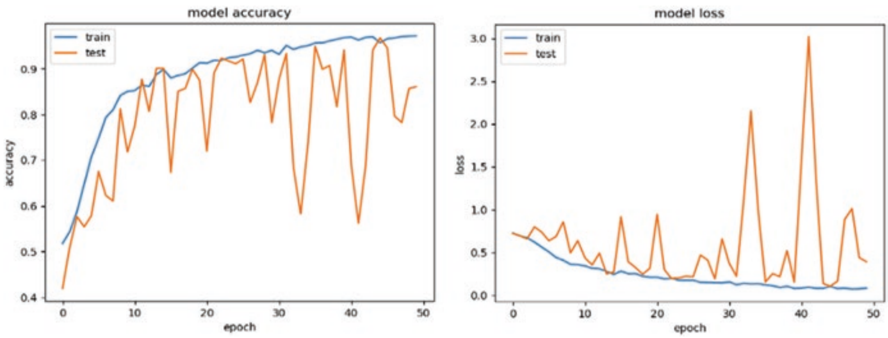

Fig. 9 Accuracy and loss fold 4 – DenseNet169

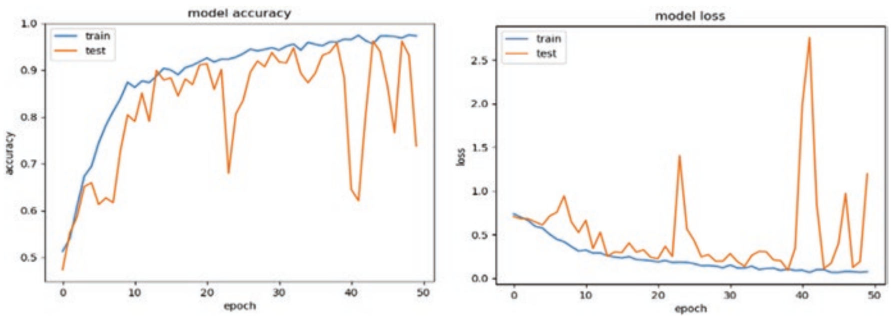

Fig. 10 Accuracy and loss fold 5 – DenseNet169

in precision and decrease to zero (0%) in the loss. In other words, the curves tend to converge.

VGG16 also achieved good results. Figure 11 shows its average confusion matrix, in which we can observe that VGG16 was incorrect in 24 cases; however, 12 cases gave false positive for COVID-19, and 12 cases of COVID-19 were wrongly classified as non-COVID-19. Below, we detail the results of VGG16, showing the confusion matrices and training curves.

Figures 12, 13, 14, 15, and 16 show the accuracy and loss fold achieved by VGG16. As can be seen, both curves demonstrate the expected behavior, which is to achieve a value of one (100%) in precision and decrease to zero (0%) in the loss. In other words, the curves tend to converge.

Figure 17 shows the average confusion matrix of the Xception, which also achieved good results, in which we can observe that Xception was incorrect in 24 cases; however, 14 cases gave false positives for COVID-19. On the other hand, only ten cases of COVID-19 were erroneously classified as non-COVID-19. Below,