information

Article

Advancing Sustainable COVID-19 Diagnosis: Integrating Artificial Intelligence with Bioinformatics in Chest X-ray Analysis

Hassen Louati [1], Ali Louati [2,*], Rahma Lahyani [3], Elham Kariri [2] and Abdullah Albanyan [4]

[1] College of Information Technology, Kingdom University, Riffa 40434, Bahrain; h.louati@ku.edu.bh
[2] Department of Information Systems, College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia; e.kariri@psau.edu.sa
[3] Operations and Project Management Department, College of Business, Alfaisal University, Riyadh 11533, Saudi Arabia; rlahyani@alfaisal.edu
[4] Department of Software Engineering, College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia; a.albanyan@psau.edu.sa
* Correspondence: a.louati@psau.edu.sa

Abstract: Responding to the critical health crisis triggered by respiratory illnesses, notably COVID-19, this study introduces an innovative and resource-conscious methodology for analyzing chest X-ray images. We unveil a cutting-edge technique that marries neural architecture search (NAS) with genetic algorithms (GA), aiming to refine the architecture of convolutional neural networks (CNNs) in a way that diminishes the usual demand for computational power. Leveraging transfer learning (TL), our approach efficiently navigates the hurdles posed by scarce data, optimizing both time and hardware utilization—a cornerstone for sustainable AI initiatives. The investigation leverages a curated dataset of 1184 COVID-positive and 1319 COVID-negative chest X-ray images, serving as the basis for model training, evaluation, and validation. Our methodology not only boosts the precision in diagnosing COVID-19 but also establishes a pioneering standard in the realm of eco-friendly and effective healthcare technologies. Through comprehensive comparative analyses against leading-edge models, our optimized solutions exhibit significant performance enhancements alongside a minimized ecological impact. This contribution marks a significant stride towards eco-sustainable medical imaging, presenting a paradigm that prioritizes environmental stewardship while adeptly addressing modern healthcare exigencies. We compare our approach to state-of-the-art architectures through multiple comparative studies.

Keywords: optimization in AI diagnostics; genetic algorithm; transfer learning; sustainable healthcare solutions

Citation: Louati, H.; Louati, A.; Lahyani, R.; Kariri, E.; Albanyan, A. Advancing Sustainable COVID-19 Diagnosis: Integrating Artificial Intelligence with Bioinformatics in Chest X-ray Analysis. *Information* **2024**, *15*, 189. https://doi.org/10.3390/info15040189

Academic Editors: Haifeng Wang, Norma B. Ojeda and Lu He

Received: 12 March 2024
Revised: 22 March 2024
Accepted: 27 March 2024
Published: 29 March 2024

1. Introduction

Computer vision plays a pivotal role in diagnosing various health issues, particularly in medical settings that rely on visual examinations [1–3]. It serves as a valuable diagnostic tool for identifying potential signs of skin cancer on the skin. Moreover, computer vision is instrumental in detecting abnormalities within the body, including issues with tissues, blood vessels, and joints. In the field of ophthalmology, it aids in the early identification of conditions like diabetic retinopathy, preventing the onset of blindness. Its effectiveness extends to medical procedures and treatments. Various types of medical imaging are harnessed by computer vision systems. Notably, the primary challenge in using chest X-rays (CXR) for identifying COVID-19 patients is the limited availability of qualified physicians, especially in rural areas. Additionally, the unique radiological characteristics of COVID-19 are often unfamiliar to healthcare professionals who may not have prior experience with COVID-19-positive CXRs. To address this, we propose a straightforward and cost-effective deep learning approach for categorizing COVID-19-positive and -negative cases through

CXR images. This method achieves high precision in seconds, offering a reliable means of diagnosis even in the absence of a radiologist or in cases of differing medical opinions. Over the past decade, deep learning has witnessed significant progress across multiple domains, including transportation, emergency prediction, and medical applications. This progress has spurred researchers to harness deep convolutional neural networks (DCNNs) for diagnosing chest conditions through radiography. Deep learning, with its transformational powers, has left an indelible imprint on a plethora of fields over the last decade, as proven by significant advances in numerous areas of research and application [4–8]. Among the numerous areas where deep learning has had an impact, the categorization of both natural and medical photographs [2,9] stands out. This crucial breakthrough has fueled academics' and scholars' examination of deep convolutional neural networks (DCNNs) as an indispensable tool in the detection of chest problems using chest radiography [4]. As we delve into the complexities of deep learning, we discover a realm of innovation and unprecedented potential. The combination of ever-improving techniques and the unwavering dedication of researchers have enabled deep learning to not only redefine the boundaries of what we thought possible, but also to address some of the most difficult issues that exist. The classification of natural and medical images has been a pioneering area of investigation within this domain. Researchers have delved into the complexities of DCNNs, utilizing their strong abilities to dissect and comprehend complex visual data. In the field of medical imaging, there is a strong focus on chest radiography and diagnosing chest diseases. This is crucial because timely and accurate diagnoses are vital for healthcare. Deep learning algorithms, especially those powered by deep convolutional neural networks (DCNNs), have evolved into essential tools for doctors and researchers. They enable quicker and more precise diagnoses, reducing the impact of diseases, improving patient care, and advancing our knowledge of health conditions.

Only a few studies [4,10,11] have employed individual deep learning techniques with CXR images to predict COVID19 +ve and −ve. An effort has been made to design a unique network. Despite CNNs' remarkable performance, their architectural design remains a key problem for academics and practitioners. To optimize the CNN architecture's design, a large number of hyperparameters must be fine-tuned. According to past research, a wide range of CNN designs, including DenseNet and AlexNet, are already available. Unfortunately, no suggestions exist for developing a specific architecture for a certain purpose; as a result, the design of such an architecture remains extremely subjective and primarily relies on the talent and expertise of data scientists.

Addressing the critical global health issue of diseases affecting the chest, with a particular focus on COVID-19, our study proposes a method that stands at the intersection of sustainability and advanced diagnostic technology [12,13]. We introduce an innovative technique that marries neural architecture search (NAS) with genetic algorithms (GA) to refine the framework of convolutional neural networks (CNN), aiming to decrease the typically high demand on computational resources. Leveraging transfer learning (TL), our approach adeptly navigates the hurdles presented by scant data while promoting the conservation of both temporal and computational assets, an essential consideration for the advancement of sustainable AI methodologies. The dataset employed in our research encompasses 1184 COVID-positive and 1319 COVID-negative chest X-ray images, serving as the foundation for our model's training, testing, and evaluation phases. Our methodology not only elevates the precision of COVID-19 detection but also pioneers in establishing a new paradigm in the domain of sustainable and efficacious healthcare innovations. Through rigorous comparative analyses against leading-edge models, our optimized architectures showcase substantial enhancements in efficacy coupled with a minimized ecological impact. This investigation heralds a progressive leap in sustainable medical imaging, delivering a solution that is both eco-friendly and proficient in surmounting contemporary health crises. Moreover, this work delves into the application of transfer learning (TL) using convolutional neural networks, harnessing pre-acquired insights from analogous tasks to augment performance on new challenges. This strategy has proven instrumental in mitigating data scarcity issues and optimizing both time and hardware expenditure. In this document, we outline an approach that leverages automated evolutionary algorithms to optimize the block configurations within CNN architectures, specifically aimed at enhancing the categorization of X-ray images. By pinpointing the most effective CNN block arrangements, we then apply transfer learning (TL) to bolster the model's capability on novel tasks, drawing upon the wealth of knowledge previously amassed from related endeavors. Below, we present a summary of the pivotal contributions made by our research:

- **Optimizing CNN Block Structures with Genetic Algorithms:** One of our research study's key contributions is the use of genetic algorithms to methodically investigate and fine-tune ideal convolutional neural network (CNN) block topologies. These structures are very reliant on the precise CNN hyperparameters, and our research is the first to use evolutionary algorithms in this setting. As a result, we obtain a better grasp of how to design CNN structures to improve performance and efficiency, giving useful insights for the wider field of deep learning.
- **A Novel Integration of CNN Architecture Search and Transfer Learning:** The invention of an evolutionary strategy that merges CNN architecture search with transfer learning structures is a ground-breaking feature of our study. This novel strategy differs significantly from existing methods in that it enables large-scale tests to thoroughly validate its effectiveness. For the first time, we combine these two strong strategies to expand model accuracy, generalization, and scalability. We provide a new viewpoint on how deep learning may be used to address complicated issues in a variety of disciplines.
- **Thorough Evaluation of X-ray Classification Architecture:** Our research focuses on a thorough investigation of the X-ray classification architecture we have constructed. We evaluate not just its utility but also its adaptability to a variety of settings. Our study examines the architecture's performance in a variety of contexts, including real-world medical applications. We want to see how well it generalizes across different datasets and diagnostic tasks, as well as how resilient and versatile it is. We give vital insights that can inform breakthroughs in the fields of medical imaging and artificial intelligence, eventually benefiting healthcare and diagnostics, through this in-depth examination.

2. Related Work

2.1. Evolutionary Neural Architecture Search

A wide range of machine learning problems have successfully used evolutionary optimization approaches in recent years [1]. This achievement may be credited to the outstanding global search skills inherent in population-based metaheuristics, which allow them to expertly negotiate the complex landscape of potential solutions, avoiding local optima while narrowing in on near-global optimum solutions. Three major methodological pillars have arisen within the area of neural architecture search (NAS): reinforcement learning (RL)-based NAS, gradient (GD)-based NAS, and evolutionary computation (EC)-based NAS, also known as Efficient NAS (ENAS).

Zhong et al. [14] investigated efficient block-wise NAS for convolutional neural networks (CNNs) using the Q-learning technique and an early stopping strategy, demonstrating the ability to achieve competitive performance at a faster rate than traditional Block-QNN approaches. Differentiable architecture search (DARTS), developed by Liu et al. [15], is at the cutting edge of GD-based NAS techniques. DARTS, which is based on the concept of continuous relaxation, continually shapes architectural representations and uses gradient descent to locate ideal models within the vast search space. Meanwhile, Shinozaki et al. [16] used genetic algorithms (GAs) to optimize deep neural networks (DNNs)' complicated structural configurations and parameters. When examining CMA-ES, a predominantly continuous optimizer that employs an indirect encoding scheme to transform discrete structural variables into real-valued representations, a stark contrast emerges, whereas GA relies on binary vectors, faithfully mirroring the DNN's architecture as a directed acyclic graph. In a similar quest to enhance recognition accuracy, Xie et al. [17] pioneered an innovative approach by encoding network topologies into binary strings. Their innovation, while commendable, was not without its challenges, notably the high computational cost, necessitating testing with smaller datasets. Sun et al. [18] introduced an evolutionary strategy that optimizes and initializes the weights of CNNs. Their method incorporated novel techniques, such as innovative weight setup procedures, diverse chromosome coding strategies with varying lengths, a slack binary tournament selection process, and a highly effective fitness evaluation mechanism. Lu et al. [19] ventured into the domain of multi-objective modeling for architectural search problems, aiming to curtail the number of floating-point operations (FLOPS) while concurrently reducing classification error rates, offering an innovative perspective on network optimization. The core of NAS, Efficient NAS (ENAS), encapsulates the EC paradigm, incorporating network structures within the evolutionary genetic algorithm process to unearth the most optimal CNN topologies. Sun et al. [20] offered fresh insights into the realm of GA-based CNN architecture design, employing variable-length encoding techniques to describe CNN structures with varying depths, featuring the construction of CNN networks through convolutional and pooling layers, thus enriching the arsenal of CNN design strategies. Neural architecture search (NAS) methodologies have consistently outperformed manually crafted network designs, showcasing their significant capabilities [21]. However, the challenge of optimizing NAS for both performance and computational efficiency is pivotal. This research aimed to harness Enhanced NAS (ENAS) to streamline the development of convolutional neural networks (CNNs), aiming for heightened efficiency and reduced computational demand. Previous investigations into ENAS, notably by Liu et al. [22] and Real et al. [23], have employed evolutionary algorithms (EAs) for discovering optimal large-scale CNN architectures, yielding impressive results. These methods typically involve encoding evolutionary components (EC) to structure the CNN, with blocks often linking the evolutionary process and the CNN's architecture, facilitating the network's construction. In the realm of medical imaging, particularly for COVID-19 detection, CNNs have proven superior in classifying X-ray images, marking a significant advance in diagnostic practices. Despite the traditional reliance on manually designed architectures, like VGGNet [24], ResNet [25], and DenseNet [26], recent studies have explored various computational approaches for diagnosing thoracic ailments through chest X-rays. Among these, Wang et al. [27] proposed a semi-supervised framework that combines different DCNN techniques for multi-label classification. Similarly, innovative network designs by Islam et al. [28] and the utilization of DenseNet for pneumonia detection by Rajpurkar et al. [29] have shown notable accuracy improvements. Additionally, efforts by Yao et al. [30], Irvin et al. [31], and Prabira et al. [32] have focused on optimizing performance through statistical label analysis, advanced learning networks, and deep feature collections for SVM classification, respectively. For COVID-19 diagnosis using X-ray imaging, studies leveraging diverse datasets have been instrumental in enhancing diagnostic methodologies. Research by Gusztav Gaal et al. [33], Asmaa Abbas et al. [34], Ali Narin et al. [35], and others has utilized various datasets, ranging from small collections to extensive compilations of chest radiographs, to investigate COVID-19 imaging characteristics and improve diagnosis accuracy. These studies highlight the importance of diverse data sources and innovative approaches in advancing X-ray-based COVID-19 diagnosis, contributing significantly to the global response to the pandemic.

2.2. Transfer Learning for X-ray Image Classifcation

Transfer learning (TL), originating from cognitive science theories, suggests that knowledge from previously learned tasks can enhance performance on new, similar tasks. This concept is rooted in the understanding that humans apply prior learning to succeed in new tasks. Pan and Yang [36] provided a formal framework for TL, introducing the notions of domains and tasks. A domain comprises a feature space X and its marginal probability distribution $P(X)$, where $X = x1, \ldots, xn \epsilon X$. A domain, denoted as $Do = X, P(X)$, pairs with a task $T = Y, f(.)$, identifying a label space Y and a predictive function $f(.)$. Transfer learning aims to enhance the target predictive function $f_T(.)$ in domain D_T by utilizing knowledge from the source domain D_S and task T_S. In the context of CNNs, TL involves transferring learned parameters, especially useful in medical imaging tasks where pre-trained CNN models on general images are adapted for specific medical analyses. CNNs, initially conceptualized by Fukushima and further developed by Fukushima et al. [37], have seen extensive application in medical imaging, including brain tumor detection from MRI scans, breast cancer identification, and disease classification from X-ray images [37,38]. For instance, Ciresan et al. [39] demonstrated that CNNs could segment neurons' membranes effectively. Wang et al. [27] compiled a novel dataset of X-ray images, achieving promising results with deep CNNs. Rajpurkar et al. [29] designed a deep CNN with 121 layers for diagnosing 14 diseases from chest X-rays. Zhou et al. [40] applied InceptionV3 architecture and TL to distinguish between cancer types. Deniz et al. [41] utilized a pre-trained VGG-16 for breast cancer classification, while other studies leveraged pre-trained networks like GoogLeNet and AlexNet for tasks such as glioma grading, with GoogLeNet showing superior performance [42]. TL has proved invaluable in medical image analysis by addressing data limitations and conserving time and computational resources. Despite the abundance of CNN architectures, designing task-specific architectures remains a largely intuitive process, dependent on data scientists' expertise. Existing TL approaches typically start with manually developed architectures. The exploration of TL for automatically generated architectures, particularly in the context of evolving optimization and deep learning for tasks like COVID-19 detection, presents a novel and unexplored avenue, as illustrated in the proposed approach for COVID-19 detection using evolutionary optimization and a deep TL-based model (Figure 1).

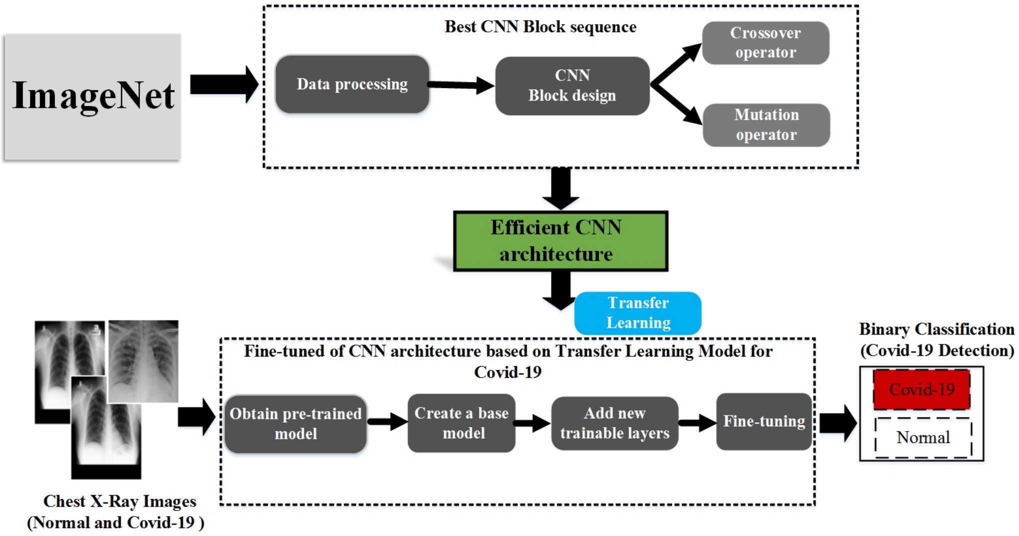

Figure 1. The proposed approach: Optimization and transfer learning workflow for enhanced COVID-19 detection using CNNs.

3. Proposed Approach

1. RQ1: Our methodology is inspired by two fundamental inquiries that address the complexities of convolutional neural network (CNN) design and the efficient utilization of existing models to address new problems.
2. RQ2: Faced with the challenge of complex models that require substantial computational resources, how can we leverage existing models to address similar tasks without the need to train new models from scratch?

To address these questions, we propose a novel approach that integrates evolutionary algorithms (EAs) for optimizing CNN architectures with a sophisticated transfer learning strategy. This approach aims to refine the process of manual CNN architecture design, which is often cumbersome and inefficient. Our method focuses on exploring a vast array of network block topologies to discover more streamlined and effective designs that enhance accuracy without unnecessary complexity. Initially, our strategy involves identifying the most suitable block structures for CNNs. Subsequently, we employ a deep learning system based on transfer learning to detect COVID-19 in chest X-ray images. This process, depicted in Figure 1, leverages evolutionary optimization to streamline architecture development and incorporates advanced transfer learning techniques for efficient model adaptation.

3.1. Blocks Search Operators

3.1.1. Strategy for Encoding and Decoding

To catalyze the evolution of CNN architectures, we have devised specialized genetic manipulations, including block-level crossover and mutation operations. These procedures are designed to merge and alter the encoded block structures, engendering offspring that may exhibit superior fitness. The fitness of each entity is gauged based on its prowess in image classification, emphasizing the reduction in error rates and the optimization of computational efficacy. Our approach utilizes genetic algorithms (GAs) for the initial population generation, where each individual represents a potential CNN architecture. This process, involves randomly generating block structures for each individual within the population, ensuring a diverse starting point for evolutionary optimization. Each block within the CNN is encoded as a string, facilitating the application of GA operations, such as crossover and mutation, to evolve more effective network architectures. To initialize a generation of solutions for a genetic algorithm (GA), where $SizeG$ represents the total number of individuals in the population:

- Start Generation Counter: Set the generation counter g to 1.
  Population Generation Loop: While $g \leq SizeG$, perform the following steps to create each individual in the population:
  a. Individual Initialization: For each individual i within the generation, initialize the sequence S_i to an empty set and start the block counter m at 1.
  b. Block Generation Loop: While $m \leq 5$, generate a block for the individual:
    i. Random Block Generation: Generate a random number tmp using randint function with a range from 0 to $m - 1$. This simulates the selection of a random connection or feature within the block.
    ii. Block Update: Add the generated random number tmp to the sequence S_i to represent the individual's architecture.
    iii. Increment Block Counter: Increase the block counter m by 1 to proceed to the next block until all 5 blocks are generated.
  c. Individual Update: Once all blocks for the individual are generated, increment the individual counter i to create the next individual in the population.
- Population Update: After all individuals have been created for the generation, increment the generation counter g by 1 to proceed to the next generation if required.
- Finalization: Once the population has been fully initialized, finalize the generation by setting S as the collection of all sequences $S_1, S_2, \ldots, S_n$ representing the initialized population.

The algorithm yields an initialized population S, ready for further processing within the GA framework, such as fitness evaluation, selection, crossover, and mutation operations. The proposed encoding strategy transforms each CNN block into a string representation, enabling the application of genetic operations. This method allows for a straightforward mapping between the encoded strings and the corresponding CNN architectures, ensuring that each individual within the population can be directly translated into a unique CNN design. The block structures are conceptualized as directed acyclic graphs (DAGs), with virtual nodes representing inputs and outputs. This encoding facilitates the exploration of various block configurations, optimizing the network's structure for efficiency and accuracy.

If the encoding string for block i is Si, then the whole encoding representation can be written as $S = \{S_1, S_2, S_n\}$. By using this method of encoding, each individual can be decoded into a full deep CNN network, and a one-to-one match can be made between the person and the CNN network. The block's structure resembles a directed acyclic graph (DAG). The virtual nodes "In" and "Out" serve as the input and output tensors for the block. According to the block's connection rule, a node j may choose just one input node k that satisfies its conditions ($j > k$). Node N1 can only choose N0 as its input, but node N5 may choose from N0 through N4 as its input. As its encoding string, the selection inputs of the computational nodes are compiled. The input selection of the block's nodes i are $0, 0, 2, 2, 3$; the encoding string is 00223. Given that the encoding string is a collection of the chosen inputs of the computational nodes, it may also be used to decode the unique block structure, indicating that the block structure and its encoding satisfy the criteria of a one-to-one match. A suitable connection between In and Out may be constructed by connecting the block structure as the DAG according to these criteria. "In" and "Out" are combined into one convenient shorthand in the block. This building's possible connections are: $N0 - N2 - N3 - N5 - Out$. Aside from the virtual nodes NO and Out, there are a total of three computational nodes in this architecture, making the depth of the encoding 00223 3. Using this method, you may encode numbers between 0 (for 01234) and 0 (for 00000) with a maximum depth of 5 and a minimum depth of 1. The CNN network's block structure may be optimized by evolutionary algorithms (GA) since the underlying data are represented as strings that can be processed by these algorithms.

The starting point for evolution in GA is the population initialization. Each person is assigned at random using an informed distribution. The random blocks are formed for each individual in the generation. For the m-th node in the i-th block, its initial value is randomint(m-1), which means that it randomly picks one of its front nodes as its input, provided that it satisfies the block's connection rule. With this form of initialization, it is possible to create each block in every individual generation. In order to develop the new population and generation of GA people, the children are produced with the expectation that they will have greater fitness values than their parents. In the suggested technique, several crossover and mutation operators are built for the offspring generation. Parents are picked at random for the population based on their fitness levels.

3.1.2. Crossover Operator

The block level crossover operators include the one-point crossover, the two-point crossover, and the uniform block crossover. In the context of optimizing convolutional neural networks (CNNs) using genetic algorithms (GAs), one critical operation is the block-level crossover between two parent architectures. This operation aims to generate new offspring architectures that combine characteristics of the parents, potentially leading to improved performance. Given two parent architectures from the population, denoted as p_1 and p_2, our objective is to produce two new offspring, S^1 and S^2. The procedure begins with setting an iteration counter i to 1, which will be used to navigate through the blocks of the parent architectures. For each block i, up to the total number of blocks n, we perform the following steps:

- Extract the ith block from both p_1 and p_2, referred to as S_i^{par1} and S_i^{par2}, respectively.
- Apply a crossover operation to these blocks, resulting in two new blocks. These new blocks are denoted as S_i^1 and S_i^2, which are parts of the offspring S^1 and S^2, respectively.
- Increment i by 1 and repeat the process until all blocks have been processed.

After iterating through all blocks, we compile the offspring by combining the newly formed blocks into coherent architectures. The first offspring, S^1, is constructed as $S_1^1, S_2^1, \ldots, S_n^1$, and similarly, the second offspring, S^2, is constructed as $S_1^2, S_2^2, \ldots, S_n^2$. These offspring represent new CNN architectures that inherit characteristics from both parents, potentially leading to novel and efficient solutions for the task at hand. These block-level crossover operators are performed on the block of the individual. The crossover is among the three block level crossover operators. The crossover operator at the individual level uses the individual's own blocks to perform crossover and produce a new individual. Before anything further, we create two brand-new index vectors, ind1 index and ind2 index, for the parents p1, p2. The crossover is performed, and the results are recorded in two index vectors, the offspring1 index and the offspring2 index, whose lengths are identical to the number of blocks in the individual. Last, the $offspring1_index$ would be used to build the next generation of S^1 offspring1 and S^2 offspring2.

3.1.3. Mutation Operator

In the present investigation, we employ a mutation strategy that closely mirrors the principles underlying the uniform crossover technique. This mutation mechanism can be conceptualized as akin to performing a two-point crossover, yet uniquely, it involves the introduction of a freshly initiated individual into the process. This particular form of crossover, which is executed at the block level, adheres to the methodology of uniform crossover.

To elaborate, the mutation operator introduced herein serves as a pivotal component of our genetic algorithm framework, designed to introduce variability into the population of neural network architectures. Unlike traditional mutation operations that might randomly alter a single gene (or block, in the context of CNN architectures), our approach draws inspiration from the uniform crossover mechanism. This method ensures that every block within an individual architecture has an equal chance of undergoing mutation, thereby promoting a diverse exploration of the architectural space. The operation begins by selecting a target individual from the current population, which will undergo mutation. Concurrently, a new individual is generated from scratch, embodying a set of potential genetic alternatives for each corresponding block in the target architecture. The mutation process then systematically reviews each block within the target individual, deciding at each step whether to retain the original block or replace it with its counterpart from the newly generated individual. This decision-making process is governed by a predefined probability, ensuring that the mutation is neither too aggressive (which might lead to the loss of beneficial traits) nor too conservative (which could stymie exploration). By adopting a methodology that mirrors the two-point crossover, but with the unique twist of involving a new individual, our mutation operator facilitates a nuanced exploration of the solution space. It allows for the injection of fresh genetic material into the population, thereby enhancing the genetic diversity available for subsequent generations. This approach is particularly advantageous in the context of optimizing convolutional neural network architectures, where the balance between exploration (searching through new, untested configurations) and exploitation (refining known, high-performing architectures) is critical for achieving optimal results.

Evaluating the fitness of each individual plays a pivotal role in the genetic algorithm (GA) process, as it directly influences the selection of parents for subsequent generations. In this investigation, we integrate the assessment of fitness with the training regime of convolutional neural networks (CNNs). Given the substantial computational demands associated with training CNNs, our methodology does not require the exhaustive training of the entire network for fitness evaluation. Instead, we employ a cyclical approach that alternates between the genetic algorithm's evolutionary steps and targeted training of the CNN, thereby achieving optimization and training efficiency.

For the purpose of error measurement, we adopt the holdout validation strategy, allocating 70% of the dataset for training purposes and the remaining 30% for testing, chosen randomly. To mitigate the potential for overfitting, we further refine the training dataset management by segmenting 70% of it into five equal parts, implementing five-fold cross-validation during the training phase. This method allows for the averaging of classification accuracy across all five partitions, ensuring a robust evaluation of model performance. The validation approach utilized in our study is depicted in Figure 2, with the final classification error rates being calculated based on the 30% of data designated for testing. This approach ensures both the efficiency of training and the accuracy of the model in classifying new data.

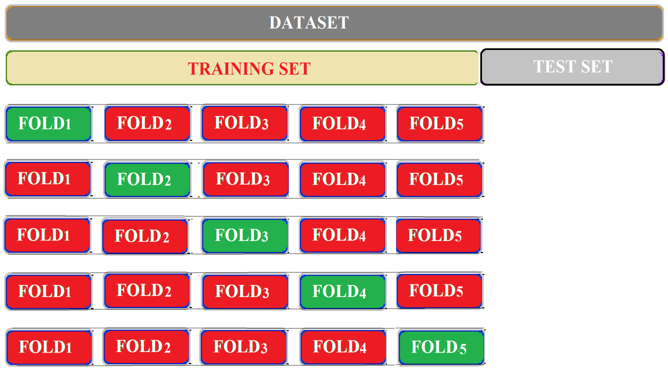

Figure 2. The nested validation strategy.

3.2. Transfer Learning Techniques

In computer vision, neural networks usually seek to identify edges in the first layer, shapes in the second layer, and task-specific parameters in the third layer. The early and middle layers are used for transfer learning, but the latter layers are just retrained. It utilizes the labeled data from the training task. A pre-trained model is a stored network that has been previously trained on a large dataset, typically on a large-scale image classification task. The intuition behind transfer learning for image classification is that if a model is trained on a sufficiently large and general dataset, this model will effectively serve as a generic model. We can then leverage these learned feature maps without having to start from scratch by training a large model on a large dataset. Our transfer learning strategy builds upon the foundational layers of pre-existing CNN models, adapting them for the specific task of COVID-19 detection in chest X-ray images. This process involves several key steps:

- Obtain pre-trained model: We begin by selecting a pre-trained CNN architecture that serves as the foundation for our model. This architecture is chosen based on its relevance to our target task and its proven effectiveness in related image classification challenges.
- Create a base model: The selected base model is then adapted to our specific requirements. This may involve adjusting the model's architecture, such as modifying the output layer to match the number of desired classifications. The initial layers of the model are frozen to retain learned features, while new layers are added to tailor the model to our specific task.
- Add new trainable and train the new layers: The new layers added to the model are trained using our dataset, allowing the model to learn features specific to COVID-19 detection. This training leverages the foundational knowledge encoded in the pre-trained model, enhancing the efficiency and effectiveness of the learning process.
- Fine-tune your model: Finally, we fine-tune the entire model, including both the pre-trained and newly added layers, to optimize its performance on the target task. This fine-tuning process involves careful adjustment of learning rates to prevent overfitting and ensure that the model achieves the highest possible accuracy.

Through this comprehensive approach, combining evolutionary optimization of CNN architectures with advanced transfer learning techniques, we aim to significantly enhance the efficiency and effectiveness of image classification models, particularly for the critical task of COVID-19 detection in chest X-ray images.

4. Experiments

4.1. Benchmarks

The evaluation framework known as the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) serves as a robust assessment tool for large-scale object detection and image categorization methods (Figure 3). One overarching goal of this framework is to facilitate a broader comparison of progress in object detection across a diverse spectrum of objects, effectively leveraging the considerable effort invested in labeling tasks. The extensive ImageNet dataset comprises a staggering 14,197,122 images, meticulously labeled according to the WordNet hierarchy. Since its inception in 2010, this dataset has served as the cornerstone of the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), an esteemed benchmark for image classification and object recognition. The publicly accessible dataset features images that have undergone meticulous manual annotation. Moreover, a subset of test images intentionally excludes these manual annotations, enhancing the rigor of the benchmark. ILSVRC annotations assume two primary forms: Firstly, image-level annotations capture binary labels indicating the presence or absence of an object class within an image. For instance, these annotations might convey the message "this image contains cars" while simultaneously noting "there are no tigers present". Secondly, object-level annotations offer comprehensive bounding boxes and class labels for individual object instances in images. These meticulous annotations enhance the benchmark's ability to address both image and object-level recognition challenges.

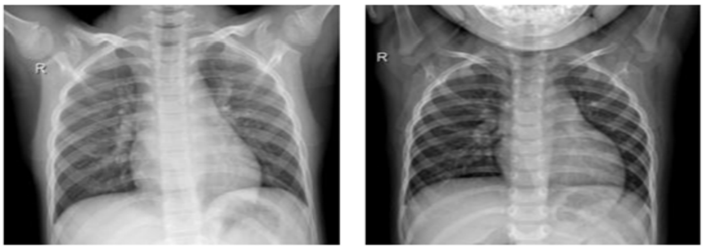

Figure 3. Samples of normal X-ray images [4].

The collection of chest X-rays from patients with COVID-19 was sourced from an open-access data repository. This database encompasses a comprehensive set of chest X-ray images, comprising 1184 images of individuals confirmed as COVID-positive and 1319 images of individuals without COVID-19 diagnosis. Our investigation is fundamentally grounded in this database, which organizes chest radiography images into two distinct categories, as visually represented in Figure 4. The categories include individuals who exhibit no infection and individuals who have been diagnosed with COVID-19. To facilitate our research and maintain robust evaluation practices, we randomly partitioned this dataset into two distinct subsets. The division allocated 80% of the images for the training phase, enabling model development and refinement, while the remaining 20% were reserved for testing, serving as an independent validation dataset. This partitioning strategy ensures a reliable assessment of our models and their generalization capabilities.

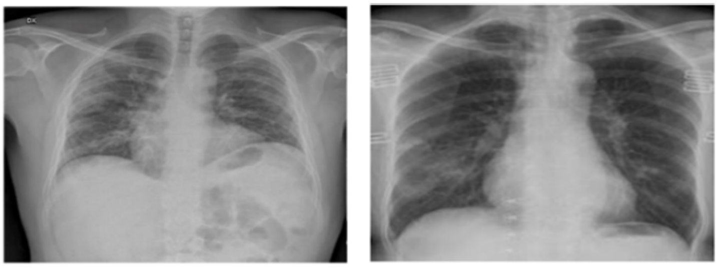

Figure 4. Sample of infected X-ray images [4].

4.2. Performance Metrics

In the realm of image classification using deep neural networks, the most commonly employed performance metrics, as indicated by the literature [43,44], are accuracy (Acc), specificity, and sensitivity. Accuracy (Acc), denoted by Equation (1), is defined as the ratio of true positives (TP) and true negatives (TN) to the total number of cases (NE). Addressing class imbalance, the geometric mean, or G-mean, emerges as a valuable metric. Computed from the binary confusion matrix, G-mean serves as the geometric mean of the actual rates for both positive and negative cases. Its purpose is to balance classification performance between majority and minority classes, offering resilience to data disparities, as expressed in Equation (2). The G-mean, calculated as the square root of the product of the true positive rate (TPR) and the true negative rate (TNR), is a critical measure of a model's ability to perform consistently across both classes in binary classification tasks. TPR, or sensitivity, reflects the proportion of actual positives correctly identified, while TNR, or specificity, gauges the proportion of actual negatives accurately recognized. This holistic performance metric is essential in our context of COVID-19 detection, ensuring that our model's effectiveness is not biased towards either class. A higher G-mean indicates a balanced classification accuracy for both COVID-positive and negative cases, underlining the model's diagnostic reliability.

$$Acc = (TP + TN)/NE \tag{1}$$

$$G\text{-}mean = \sqrt{TRP.TNR} \tag{2}$$

In this study, parameter tuning for comparative algorithms was carried out through the conventional trial-and-error method. Notably, the parameter configuration for "Our work" involved a batch size of 128, a stochastic gradient descent (SGD) learning rate of 0.01, a momentum value of 0.91, and a weight decay factor of 0.00001 for the optimization process. The search method incorporated 60 generations, a population size of 50, a crossover rate of 0.9, and a mutation rate of 0.1, with these parameter choices significantly shaping the study's conditions and approach. The implementation employed the TensorFlow framework in conjunction with Python (version 3.5), while CNN structures were assessed using eight Nvidia 2080Ti GPU cards, designed and tested with the provided data.

4.3. Network Parameters Improvement: Enhancing Diagnostic Efficiency

In our study, the convolutional neural network (CNN) architecture specifically tailored for chest X-ray image analysis is meticulously optimized through a synergy of neural architecture search (NAS) and genetic algorithms (GA), embodying a sophisticated blend of convolutional, pooling, and fully connected layers. This architecture, emerging from empirical evaluations, begins with convolutional layers characterized by a 3 × 3 kernel size—initiating with 32 filters that double every two layers to intricately capture the complexity of the input images, utilizing a stride of 1 and padding to maintain feature map dimensions, and leveraging the rectified linear unit (ReLU) activation function to introduce non-linearity and enhance the network's pattern learning capability. Pooling layers employing a 2 × 2 max pooling scheme follow every pair of convolutional layers, strategically reducing spatial dimensions and computational demands, culminating in fully connected layers that lead to a softmax output layer, which categorizes images into COVID-positive or negative diagnoses, the structure of which is informed by the complexity of the dataset and the intricacy of learned features. The network's training and optimization are governed by stochastic gradient descent (SGD) with a learning rate of 0.001, momentum of 0.9, and a mini-batch size of 32, focusing on minimizing the cross-entropy loss function. The essence of our methodology lies in leveraging NAS to delineate a vast search space that represents all conceivable network configurations, which GA then navigates through selection, crossover, and mutation processes to evolve a population of architectures across generations, each appraised for its precision in validation set classification. This iterative refinement zeroes in on architectures showcasing peak performance until a predetermined generational count or performance convergence criterion is met, leading to the selection of an optimized CNN architecture. This comprehensive optimization process not only significantly boosts diagnostic accuracy by minimizing false positives and negatives but also underscores the method's computational efficiency, expediting convergence and reducing resource consumption. The culmination of this detailed parameterization and rigorous optimization strategy not only furnishes a highly accurate tool for COVID-19 diagnosis from chest X-ray images but also propels the field of sustainable AI forward in the healthcare domain, highlighting our methodology's profound impact on enhancing CNN performance for medical imaging tasks.

4.4. *Comparative Results*

X-ray image-based COVID-19 identification has recently witnessed significant advancements, thanks to the integration of various artificial intelligence techniques. Our methodology is poised for a rigorous comparative assessment against the most prominent contributions in the realm of convolutional neural network (CNN) architecture generation [45]. The distilled findings of this comparison, encapsulated in Table 1, offer a comprehensive overview of the architectural landscapes cultivated by the various CNN design strategies targeting X-ray images.

A glance at Table 1 unearths a range of classification accuracy (Acc) figures within the domain of X-ray-based COVID-19 diagnosis, spanning from 88.39% to a remarkable 98.12%. Notably, Biraja Ghoshal et al. [45] manifest the relatively modest end of this spectrum, achieving an Acc of 88.39%. Linda Wang et al. [46], in terms of classification accuracy, attain a commendable 92.4% concerning the COVIDx dataset. Meanwhile, Asmaa Abbas et al. [34] shine with an Acc of 95.12%, fortified by an impressive sensitivity of 97.91% and a specificity of 91.87%. Prabira Kumar Sethy et al. [32] also make a notable entry, securing an Acc of 95.38%. Ioannis D. Khalid EL Asnaoui et al. [47], Muhammad Farooq and Abdul Hafeez [48], and Gusztav Gaal et al. [33] bolster the upper echelons of this spectrum, boasting Acc rates of 96.23% and 97.5%, as delineated in Table 1. Furthermore, we present a compelling contrast based on the visual representations in Figures 5 and 6. Our methodology emerges as a front-runner, achieving higher Acc values in comparison to the methodologies that fall within the purview of this scrutiny. Several factors underpin these remarkable findings. Firstly, manual CNN creation is a laborious and intricate endeavor, demanding a high degree of expertise from the user. The sheer diversity of potential architectural designs further compounds the intricacy, making it a formidable task even for seasoned practitioners. In this context, our investigations highlight the preeminence of evolutionary algorithms, a testament to their intrinsic capacity for optimizing accuracy throughout the search process [32]. The allure of evolutionary algorithms lies in their inherent proclivity for global exploration. This trait enables them to venture beyond local optima, delving deep into the entirety of the search space. Notably, the probability of accepting suboptimal structures through the mating selection operator adds a layer of robustness to the optimization process. As corroborated by the insights gleaned from Figure 7, our proposed approach exhibits an innate aptitude for crafting task-specific architectural designs [47]. Our method's ability to automatically formulate CNN architectures with enhanced accuracy to those contemplated by our peers is a testament to its prowess. Crafting CNNs remains an intricate challenge, even for those with extensive domain knowledge. Notably, automated design strategies triumph over their handcrafted counterparts when dealing with radiographic images. This can be attributed to the inexhaustible permutations of potential designs, where each topology significantly influences the interactions between neural network nodes, thereby wielding a profound impact on classification performance. The paradigm of automatic architecture generation thus emerges as a promising frontier in the pursuit of higher accuracy and efficiency in the domain of medical image analysis [48].

Table 1. Comparative performance of CNN architectures for COVID-19 detection in chest X-ray imaging.

Research	**Model Approach**	**Test Acc (%)**	**G-Mean**
Gusztav Gaal et al. [33]	Integration of U-Net with AT, contrast enhancement	97.49	97.13
Abbas et al. [34]	CNN with FD and enhancement using ImageNet and ResNet (DeTraC)	95.12	94.69
Narin et al. [35]	Adaptation of ResNet50 via transfer learning	97.00	96.7
Wang et al. [46]	Application of TL to COVID-Net	92.4	91.06
Asnaoui et al. [47]	Exploration of multiple architectures including Xception, VGG16-19, and DenseNet201	96	95.98
Sethy et al. [32]	Combination of Resnet50 deep features with support vector machines	95.3	94.1
Ioannis et al. [49]	Fine-tuning of models like Xception, VGG19, and Inception for enhanced accuracy	95.57	93.44
Ghoshal et al. [45]	Implementation of Bayesian CNN with Dropweights for uncertainty estimation	88.39	89.91
AbdulHafeez [48]	Utilization of a pre-trained ResNet50 architecture with COVIDx dataset	96.22	95.8
Louati et al. [50]	Optimization of CNN architecture via topology	98.1	97.91
Our Work	–	99.03	98.83

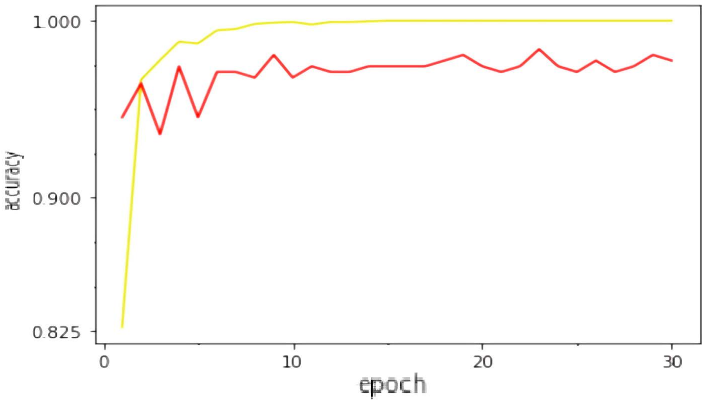

Figure 5. Training and validation accuracy over epochs for CNN-based COVID-19 detection.

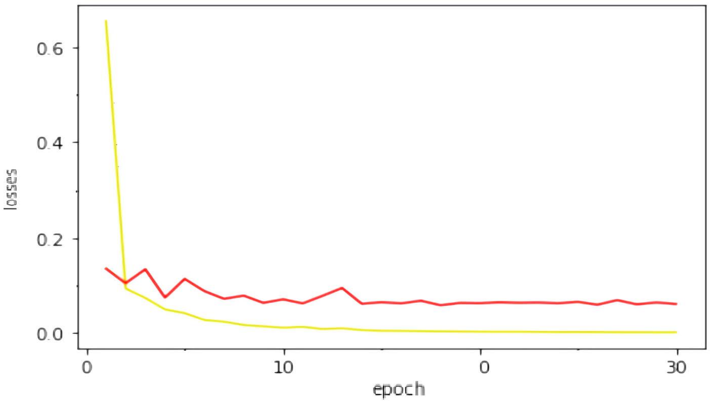

Figure 6. Training and validation loss over epochs for CNN-based COVID-19 detection.

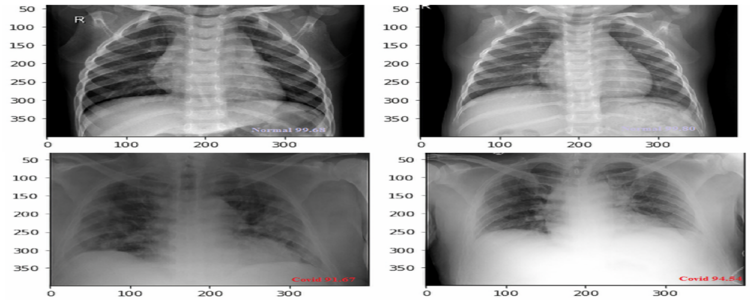

Figure 7. The model's diagnostic capability to differentiate between normal and COVID-19-affected lung radiographs.

Results Discussion

In the realm of leveraging X-ray imagery for the diagnosis of COVID-19, various studies have made significant contributions. These investigations, each utilizing distinct classification models and techniques, have demonstrated varied levels of accuracy and G-mean scores, reflecting their effectiveness in identifying COVID-19 from X-ray images. The following summarizes the methodologies and outcomes of these notable works:

- The study by Gusztav Gaal et al. [33] combined U-Net with adversarial techniques and contrast-limited adaptive histogram equalization, achieving an accuracy of 97.5% and a G-mean of 97.14.
- Abbas and colleagues [34] employed a CNN with feature decomposition and model transfer, integrating ImageNet and ResNet enhancements (DeTraC), resulting in 95.12% accuracy and a 94.69 G-mean.
- Narin et al. [35] utilized a pre-trained ResNet50 model adapted through transfer learning, attaining a 97% accuracy and 96.78 G-mean. The research by Wang et al. [46] focused on applying transfer learning to COVID-Net, recording a 92.4% accuracy and a 91.06 G-mean.
- Asnaoui et al. [47] explored various architectures including Xception, VGG16-19, DenseNet201, Inception-ResNet-V2, InceptionV3, Resnet50, and MobileNet-V2, achieving 96% accuracy and a 95.98 G-mean.
- Sethy et al. [32] combined Resnet50's deep features with support vector machines, achieving a 95.38% accuracy and a 94.14 G-mean.
- Ioannis et al. [49] refined models like Xception, VGG19, Inception, and Resnet V2, resulting in 95.57% accuracy and a 93.44 G-mean.
- Ghoshal et al. [45] implemented Bayesian CNN with Dropweights, achieving 88.39% accuracy and an 89.91 G-mean.
- AbdulHafeez [48] leveraged a pre-trained ResNet50 architecture with the COVIDx dataset to achieve a 96.23% accuracy and a 95.81 G-mean. Louati et al. [50] focused on optimizing the topology of CNN architecture, achieving an accuracy of 98.12% and a G-mean of 97.90%.

These outcomes may be explained using the following arguments. The manual design of CNNs is an extremely difficult and time-consuming process that demands a high level of user skill. Even with a high level of skill, coming up with a decent design is very difficult since there are so many potential architectures. To automate this design effort, evolutionary approaches have been devised and have shown greater performance than human design due to their capacity to search across the vast search space of feasible architectures automatically. Nevertheless, evolutionary approaches are able to avoid local optimums and scan the whole search space because of their global search capabilities and the probability acceptance of underperforming structures by the mating selection operator.

These findings further confirm our proposed algorithm's potential to create task-dependent designs automatically when using EAs for block design. This might be explained by the fact that CNN design is very difficult, even for those with extensive experience. On radiographic images, automatic design approaches outperform manually constructed systems. The reason for this is that there is a vast variety of alternative architectures. These outcomes may be explained by the significance of network topology optimization. As each topology determines how the neural network nodes connect to each other, it has a big effect on how well the neural network can classify things. Transfer learning is then used to save time and resources by eliminating the need to train several machine learning models from the beginning in order to fulfill comparable tasks, which enables quick development and enhanced performance. Figure 8 depicts a sample of random activations in the second convolutional layer. Figure 8 illustrates the activation patterns in the second convolutional layer of our CNN model when processing chest X-ray images. The varying intensities of yellow and blue indicate the level of activation across different areas of the input feature map. High activations in yellow suggest regions where the convolutional filters are most responsive, detecting specific features or patterns relevant to the model's task. Conversely, areas with minimal activation are shown in blue, indicating less relevance to the features the layer is designed to capture. Such visualizations provide insight into which features within the X-ray images the CNN is focusing on, potentially correlating to diagnostically significant regions for COVID-19 detection. Interpretability in deep learning, particularly in medical applications, is crucial for gaining clinicians' trust and providing insights into the network's decision-making process. To this end, we introduce a dedicated analysis using state-of-the-art interpretability methods to demystify our network's behavior and provide interpretable outputs that elucidate the basis of its predictions. One of the techniques we employ is gradient-weighted class activation mapping (Grad-CAM), which generates visual explanations for decisions made by convolutional neural networks. Grad-CAM highlights the regions in the input image that are important for predictions, offering a visual understanding of which features contribute most significantly to the network's output. By applying Grad-CAM to our network, we can visually demonstrate how the model focuses on specific areas of chest X-rays that are indicative of COVID-19, thus validating the clinical relevance of the features it learns to identify. Furthermore, we explore the use of layer-wise relevance propagation (LRP), a technique that attributes the prediction of a deep neural network to its input features, providing a pixel-level explanation of network decisions. LRP helps in identifying the radiographic features, such as ground-glass opacities or bilateral infiltrates, that the network deems critical for diagnosing COVID-19. This pixel-wise decomposition of the network's output into inputs offers an intuitive understanding of its behavior, bridging the gap between complex model predictions and clinical interpretations. Incorporating these interpretability techniques enables us to not only bolster the credibility and acceptance of our proposed network among healthcare professionals, but also provides an invaluable tool for medical research by uncovering new radiographic features associated with COVID-19. The insights gained from these interpretability analyses contribute to a more comprehensive understanding of the network's performance, offering explanations beyond conventional performance metrics and paving the way for a more informed and interpretable AI-driven diagnostic process.

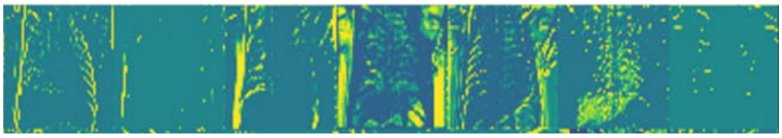

Figure 8. Depiction of sampling activations within the 2 convolutional layer.

4.5. Elucidating the Superiority of the Proposed Network Architecture

Central to our network's design is the innovative use of neural architecture search (NAS) coupled with genetic algorithms (GA), which collectively empower our model to dynamically tailor its architecture for optimal feature extraction from chest X-ray images. This dynamic adaptability allows for the exploration and iterative refinement of network configurations, enabling the identification of structures that are intricately optimized for the complex patterns characteristic of medical imaging data. The genetic algorithm-driven optimization further distinguishes our approach, leveraging selection, crossover, and mutation processes to efficiently navigate the architectural search space, thus honing in on models that exhibit a pronounced reduction in overfitting and superior generalization capabilities. Moreover, our network employs advanced feature extraction techniques, incorporating variable kernel sizes and specialized activation functions to discern more intricate and subtle radiographic features indicative of COVID-19, setting our methodology apart from more generic models. The integration of transfer learning (TL) further bolsters our network's performance, allowing it to utilize pre-learned features from extensive datasets, thereby enhancing learning efficiency and diagnostic precision even in the face of the limited specificity that plagues medical imaging datasets. Additionally, our customized loss function and innovative regularization strategies are tailored to minimize prediction errors and curb overfitting, ensuring robustness against input data variations and contributing significantly to the network's diagnostic accuracy. These strategic implementations, from adaptive layer configurations to efficient data utilization and beyond, underscore the unique advantages of our proposed network architecture. By addressing the specific challenges inherent in the analysis of medical imaging for COVID-19 diagnosis, our network not only sets a new benchmark for accuracy but also marks a leap in computational efficiency, offering a detailed explanation for its outstanding performance in comparison to traditional deep neural network approaches.

5. X-ray-14 Images Diagnosis

The Chest X-ray14 dataset encompasses a comprehensive collection of 112,120 frontal chest X-ray images sourced from 30,805 patients. This extensive compilation was meticulously assembled leveraging radiological summaries archived within hospital-based image storage and communication frameworks, subsequently refined through advanced natural language processing techniques. Each image within the dataset may exhibit signs of one or several common pulmonary conditions, with images labeled as "Normal" indicating a complete absence of detectable thoracic abnormalities (refer to Figure 9 for an illustration). The dataset is accessible for research and application at the following URL: https://nihcc.app.box.com/v/ChestXray-NIHCC (accessed on 2 February 2024).

The comparative evaluation of various CNN design methodologies reveals intriguing insights into the performance of different optimization approaches. The obtained area under the receiver operating characteristic (AUROC) scores offer a gauge of each method's effectiveness. The AUROC scores for manual optimization methods demonstrate a range from 78% to 84%. Yao et al.'s [30] manual approach yielded a respectable test AUROC of 79.79%, while Wang et al. achieved a competitive 73.81%. Meanwhile, the CheXNet method, also employing manual optimization, excelled with a test AUROC of 84.41.

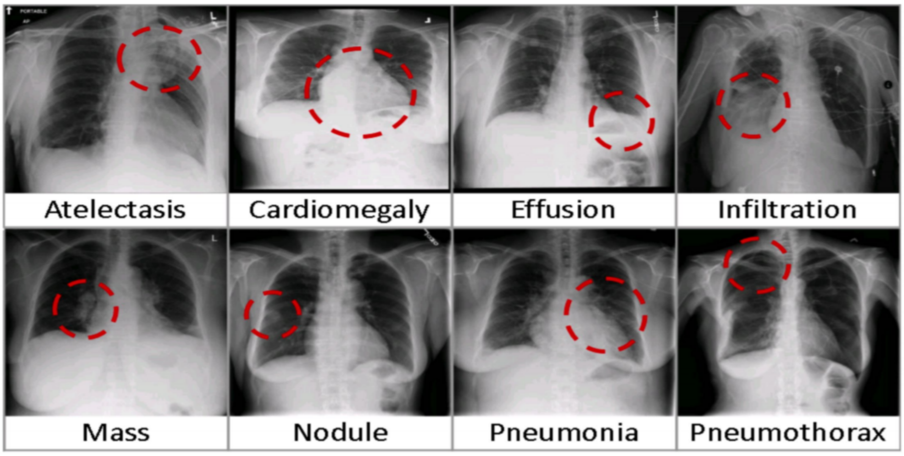

Figure 9. Frequent pulmonary conditions identified through chest X-ray imaging [51].

In contrast, Google AutoML, a non-manual approach utilizing reinforcement learning (RL) for optimization, exhibited a relatively lower test AUROC of 79.72%. However, as the narrative unfolds, the trajectory of AUROC values takes an upward turn. Notably, LEAF, employing evolutionary algorithms (EA) for optimization, achieved a test AUROC of 84.29%, marking a notable improvement. This trend continued with NSGANet-X, another EA-based optimization approach, demonstrating impressive results with a test AUROC of 84.61%.

However, the pinnacle of achievement in this comparative landscape is exemplified by our method, denoted in the table. Operating under the framework of evolutionary algorithms (EA) for optimization, our approach delivered an outstanding test AUROC of 87.74%. This remarkable performance sets a new standard, surpassing the capabilities of previously explored methodologies and reaffirming the prowess of our approach in the realm of AUROC results on the ChestX-ray14 dataset. These results underscore the exceptional potential of our method in the pursuit of highly effective CNN architectures.

Discussion on Robustness, Generalizability, and Future Direction

We aim to underscore the robustness and general applicability of our proposed methodology, particularly through its performance on the X-ray-14 Images Diagnosis dataset. The diversity and complexity of this dataset present a rigorous testing ground for any diagnostic algorithm. Our methodology's adaptability to this dataset is a testament to its robustness. By employing neural architecture search (NAS) with genetic algorithms (GA), coupled with transfer learning (TL), our approach not only addressed the specific challenge of COVID-19 diagnosis but also proved its capability in identifying various other thoracic pathologies. This adaptability is crucial, as it demonstrates our model's ability to generalize from one specific task (COVID-19 detection) to a broader range of diagnostic challenges. The performance of our optimized convolutional neural network (CNN) architectures on the X-ray-14 dataset stands as a robust indicator of our method's efficacy. Not only did our approach achieve commendable accuracy, specificity, sensitivity, and G-mean scores, but it also demonstrated comparable or superior performance against state-of-the-art models specifically designed for the diverse pathologies represented in the X-ray-14 dataset. This comparative analysis highlights the competitive edge of our methodology, showcasing its potential as a versatile tool in the realm of medical imaging. The successful application of our method to the X-ray-14 dataset further illustrates its resilience to data scarcity challenges. Through the strategic use of transfer learning (TL), our approach leverages pre-acquired knowledge, enabling it to adapt to varied and limited datasets effectively. This capability is essential for the development of sustainable AI initiatives, particularly in healthcare contexts where data availability may be constrained. The successful application of our method to the X-ray-14 dataset further illustrates its resilience to data scarcity challenges. Through the strategic use of transfer learning (TL), our approach leverages pre-acquired knowledge, enabling it to adapt to varied and limited datasets effectively. This capability is essential for the development of sustainable AI initiatives, particularly in healthcare contexts where data availability may be constrained. The demonstrated robustness and generalizability of our approach, particularly through its application to the X-ray-14 Images Diagnosis dataset, encourage further exploration and adaptation of our methodology across other medical imaging tasks. Future research will focus on extending our approach to additional datasets and diagnostic challenges, reinforcing its potential to contribute significantly to advancements in medical imaging and diagnosis.

6. Conclusions

The quest for a well-suited architecture for deep convolutional neural networks (DCNN) remains an ongoing and enthralling challenge in the dynamic world of deep learning. Over the years, following manual design, an array of alternative approaches has surfaced, predominantly driven by evolutionary optimization techniques and some adopting a multi-objective viewpoint. This landscape has fostered a myriad of potential network topologies, each bearing a unique set of attributes and configurations. What further complicates this process is the lack of a one-size-fits-all guideline for crafting a specific architecture for a given task, making architectural design a highly subjective endeavor that heavily depends on the experience and expertise of data scientists. In the context of our research, we introduce an innovative evolutionary approach that enhances the development of CNN architectures. This approach is centered on the exploration of the best sequence of block topologies and the transfer of knowledge to a smaller dataset, with the ultimate goal of achieving superior precision in the detection of COVID-19 infections. Our empirical experiments have not only validated the effectiveness of this proposed methodology, but have also demonstrated its superiority over numerous other architectural designs when assessed against a benchmark dataset of X-ray images. This breakthrough prompts us to envision a broader perspective that stands to revolutionize the landscape of deep learning and architectural design. One of the most compelling perspectives that directly stems from our work is the development of an interactive model. This model aspires to put the power of architectural design into the hands of users. It envisions users actively participating in the architectural evolution process, scrutinizing generated architectures, identifying recurrent patterns, and contributing to the formulation of recommendations, encompassing both soft and hard constraints. These recommendations will serve as invaluable guidance for generating CNN architectures that align with the preferences and expertise of domain experts. This not only introduces the prospect of a more user-centric and adaptable approach to network design but also signifies the evolving nature of deep learning, where architectural decisions are becoming more democratic and flexible, fostering a more inclusive and collaborative approach to AI model development. This transformative shift holds the promise of democratizing deep learning architecture, allowing experts and non-experts alike to contribute and benefit from tailored solutions for a myriad of applications. It envisions a future where the collective intelligence of the AI community fuels architectural innovation, fostering a richer and more diversified landscape in the world of deep learning.

Author Contributions: Conceptualization, methodology, and experimentation: H.L.; writing—review and editing: H.L. and A.L.; project management: E.K. and A.A.; investigation: R.L. All authors have read and agreed to the published version of the manuscript.

Funding: Kingdom University, Bahrain, research grant number 2024-3-011. Prince Sattam bin Abdulaziz University, project number PSAU/2024/R/1445.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: ImageNet: https://www.image-net.org/, accessed on 2 February 2024, COVID-Normal: https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database, accessed on 2 February 2024. Chest X-ray 14: NIH https://paperswithcode.com/dataset/chestx-ray14, accessed on 2 February 2024. Any additional data could be requested from the corresponding author, Hassen Louati.

Acknowledgments: The authors thank the Deanship of Scientific Research at Prince Sattam bin Abdulaziz University for funding this work via project number (PSAU/2024/R/1445). The authors would like to acknowledge that this research work was partially financed by Kingdom University, Bahrain from the research grant number 2024-3-011.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Paules, C.I.; Marston, H.D.; Fauci, A.S. Coronavirus Infections—More Than Just the Common Cold. *JAMA* **2020**, *323*, 707–708. [CrossRef] [PubMed]
2. Chen, Y.; Liu, Q.; Guo, D. Emerging coronaviruses: Genome structure, replication, and pathogenesis. *J. Med. Virol.* **2020**, *92*, 418–423. [CrossRef] [PubMed]

3. Louati, H.; Louati, A.; Kariri, E.; Bechikh, S. Optimizing Deep Learning for Computer-Aided Diagnosis of Lung Diseases: An Automated Method Combining Evolutionary Algorithm and Transfer Learning. *Comput. Model. Eng. Sci.* **2024**, *138*, 2519–2547. [CrossRef]
4. Louati, H.; Bechikh, S.; Louati, A.; Aldaej, A.; Said, L.B. Evolutionary optimization for cnn compression using thoracic X-ray image classification. In Proceedings of the 34th International Conference on Industrial, Engineering Other Applications of Applied Intelligent Systems, Kitakyushu, Japan, 19–22 July 2022.
5. Louati, A. A hybridization of deep learning techniques to predict and control traffic disturbances. *Artif. Intell. Rev.* **2020**, *53*, 5675–5704. [CrossRef]
6. Louati, A.; Louati, H.; Li, Z. Deep learning and case-based reasoning for predictive and adaptive traffic emergency management. *J. Supercomput.* **2021**, *77*, 4389–4418. [CrossRef]
7. Louati, A. Cloud-assisted collaborative estimation for next-generation automobile sensing. *Eng. Appl. Artif. Intell.* **2023**, *126*, 106883. [CrossRef]
8. Louati, A.; Louati, H.; Kariri, E.; Alaskar, F.; Alotaibi, A. Sentiment Analysis of Arabic Course Reviews of a Saudi University Using Support Vector Machine. *Appl. Sci.* **2023**, *13*, 12539. [CrossRef]
9. Ulhaq, A.; Khan, A.; Gomes, D.; Paul, M. Computer vision for COVID-19 control: A survey. *arXiv* **2020**, arXiv:2004.09420.
10. Bengio, Y.; Lamblin, P.; Popovici, V.; Larochelle, H. Greedy layer-wise training of deep networks. In *Advances in Neural Information Processing Systems 19*; Schölkopf, B., Platt, J., Hoffman, T., Eds.; MIT Press: Cambridge, MA, USA, 2007; pp. 153–160.
11. Hinton, G.E.; Osindero, S.; Teh, Y.-W. A fast learning algorithm for deep belief nets. *Neural Comput.* **2006**, *18*, 1527–1554. [CrossRef]
12. Louati, A.; Louati, H.; Kariri, E.; Neifar, W.; Farahat, M.; El-Hoseny, H.; Hassan, M.; Khairi, M. Sustainable Urban Mobility for Road Information Discovery-Based Cloud Collaboration and Gaussian Processes. *Sustainability* **2024**, *16*, 1688. [CrossRef]
13. Louati, A.; Louati, H.; Kariri, E.; Neifar, W.; Hassan, M.; Khairi, M.; Farahat, M.; El-Hoseny, H. Sustainable Smart Cities through Multi-Agent Reinforcement Learning-Based Cooperative Autonomous Vehicles. *Sustainability* **2024**, *16*, 1779. [CrossRef]
14. Zhong, Z.; Yang, Z.; Deng, B.; Yan, J.; Wu, W.; Shao, J.; Liu, C.L. BlockQNN: Efficient Block-Wise Neural Network Architecture Generation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 2314–2328. [CrossRef]
15. Liu, H.X.; Simonyan, K.; Yang, Y.M. DARTS: Differentiable Architecture Search. In Proceedings of the International Conference on Learning Representations (ICLR), New Orleans, LA, USA, 6–9 May 2019.
16. Shinozaki, T.; Watanabe, S. Structure discovery of deep neural network based on evolutionary algorithms. In Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing, South Brisbane, QLD, Australia, 19–24 April 2015; pp. 4979–4983.
17. Xie, S.; Girshick, R.; Dollar, P.; Tu, Z.; He, K. Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1492–1500.
18. Sun, Y.; Xue, B.; Zhang, M.; Yen, G.G. Completely automated cnn architecture design based on blocks. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *33*, 1242–1254. [CrossRef]
19. Lu, Z.; Whalen, I.; Boddeti, V.; Dhebar, Y.; Deb, K.; Goodman, E.; Banzhaf, W. Nsga-net: Neural architecture search using multi-objective genetic algorithm. In Proceedings of the Genetic and Evolutionary Computation Conference, Prague, Czech Republic, 13–17 July 2019; pp. 419–427.
20. Sun, Y.; Xue, B.; Zhang, M.; Yen, G.G.; Lv, J. Automatically Designing CNN Architectures Using the Genetic Algorithm for Image Classification. *IEEE Trans. Cybern.* **2020**, *50*, 3840–3854. [CrossRef]
21. Elsken, T.; Metzen, J.H.; Hutter, F. Neural Architecture Search: A Survey. *J. Mach. Learn. Res.* **2019**, *20*, 1997–2017.
22. Liu, Y.; Sun, Y.; Xue, B.; Zhang, M.; Yen, G.G.; Tan, T.C. A Survey on Evolutionary Neural Architecture Search. *arXiv* **2020**, arXiv:2008.10937.
23. Real, E.; Moore, S.; Selle, A.; Saxena, S.; Suematsu, Y.L.; Tan, J.; Kurakin, A. Large-Scale Evolution of Image Classifiers. In Proceedings of the International Conference on Machine Learning, PMLR, Bellevue, WA, USA, 28 June–2 July 2017; pp. 2902–2911.
24. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
25. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
26. Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q. Densely 750 connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
27. Wang, X.; Peng, Y.; Lu, L.; Lu, Z.; Bagheri, M.; Summers, R.M. ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 3462–3471.
28. Islam, M.T.; Aowal, M.A.; Minhaz, A.T.; Ashraf, K. Abnormality detection and localization in chest X-rays using deep convolutional neural networks. *arXiv* **2017**, arXiv:1705.09850.
29. Rajpurkar, P.; Irvin, J.; Ball, R.L.; Zhu, K.; Yang, B.; Mehta, H.; Duan, T.; Ding, D.; Bagul, A.; Langlotz, C.P.; et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med.* **2018**, *15*, e1002686. [CrossRef]
30. Yao, L.; Poblenz, E.; Dagunts, D.; Covington, B.; Bernard, D.; Lyman, K. Learning to diagnose from scratch by exploiting dependencies among labels. *arXiv* **2017**, arXiv:1710.10501.

31. Irvin, J.; Rajpurkar, P.; Ko, M.; Yu, Y.; Ciurea-Ilcus, S.; Chute, C.; Marklund, H.; Haghgoo, B.; Ball, R.; Shpanskaya, K.; et al. A large chest radiograph dataset with uncertainty labels and expert comparison. In Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; pp. 590–597.
32. Sethy, P.K.; Behera, S.K. Detection of coronavirus disease (COVID-19) based on deep features. *Int. J. Math. Eng. Manag.* **2020**, *5*, 643–651.
33. Gaál, G.; Maga, B.; Lukács, A. Attention U-Net Based Adversarial Architectures for Chest X-ray Lung Segmentation. *arXiv* **2020**, arXiv:2003.10304.
34. Abbas, A.; Abdelsamea, M.M.; Gaber, M.M. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *arXiv* **2020**, arXiv:2003.13815.
35. Narin, A.; Kaya, C.; Pamuk, Z. Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks. *arXiv* **2020**, arXiv:2003.10849.
36. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans Knowl Data Eng.* **2010**, *22*, 1345–1359. [CrossRef]
37. Fukushima, K.; Miyake, S. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In *Competition and Cooperation in Neural Nets*; Arbib, M.A., Amari, S.I., Eds.; Springer: Berlin/Heidelberg, Germany, 1982; pp. 267–285.
38. Kallianos, K.; Mongan, J.; Antani, S.; Henry, T.; Taylor, A.; Abuya, J.; Kohli, M. How far have we come? artifcial intelligence for chest radiograph interpretation. *Clin. Radiol.* **2019**, *74*, 338–345. [CrossRef]
39. Ciresan, D.; Giusti, A.; Gambardella, L.; Schmidhuber, J. Deep neural networks segment neuronal membranes in electron microscopy images. *Adv. Neural Inf. Process Syst.***2012**, *25*, 2843–2851.
40. Zhou, L.; Zhang, Z.; Chen, Y.C.; Zhao, Z.Y.; Yin, X.D.; Jiang, H.B. A deep learning-based radiomics model for diferentiating benign and malignant renal tumors. *Transl Oncol.* **2019**, *12*, 292–300. [CrossRef]
41. Deniz, E.; Şengür, A.; Kadiroğlu, Z.; Guo, Y.; Bajaj, V.; Budak, Ü. Transfer learning based histopathologic image classifcation for breast cancer detection. *Health Inf. Sci. Syst.* **2018**, *6*, 1–7. [CrossRef]
42. Yang, Y.; Yan, L.F.; Zhang, X.; Han, Y.; Nan, H.Y.; Hu, Y.C.; Hu, B.; Yan, S.L.; Zhang, J.; Cheng, D.L.; et al. Glioma grading on conventional mr images: A deep learning study with transfer learning. *Front. Neurosci.* **2018**, *12*, 804. [CrossRef]
43. Louati, H.; Bechikh, S.; Louati, A.; Hung, C.C.; Said, L.B. Deep convolutional neural network architecture design as a bi-level optimization problem. *Neurocomputing* **2021**, *439*, 44–62. [CrossRef]
44. Louati, H.; Bechikh, S.; Louati, A.; Aldaej, A.; Said, L.B. Joint design and compression of convolutional neural networks as a bi-level optimization problem. *Neural Comput. Appl.* **2022**, *34*, 15007–15029. [CrossRef]
45. Ghoshal, B.; Tucker, A. Estimating Uncertainty and Interpretability in Deep Learning for Coronavirus (COVID-19) Detection. *arXiv* **2020**, arXiv:2003.10769.
46. Wang, L.; Wong, A. COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest Radiography Images. *arXiv* **2020**, arXiv:2003.09871.
47. Asnaoui, K.E.; Chawki, Y.; Idri, A. Automated Methods for Detection and Classification Pneumonia based on X-ray Images Using Deep Learning. *arXiv* **2020**, arXiv:2003.14363.
48. Farooq, M.; Hafeez, A. COVID-ResNet: A Deep Learning Framework for Screening of COVID19 from Radiographs. *arXiv* **2020**, arXiv:2003.14395.
49. Apostolopoulos, I.D.; Mpesiana, T.A. COVID-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Phys. Eng. Sci. Med.* **2020**, *43*, 635–640. [CrossRef]
50. Louati, H.; Louati, A.; Bechikh, S.; Masmoudi, F.; Aldaej, A.; Kariri, E. Topology optimization search of deep convolution neural networks for CT and X-ray image classification. *BMC Med. Imaging* **2022**, *22*, 120. [CrossRef]
51. Louati, H.; Bechikh, S.; Louati, A.; Aldaej, A.; Said, L.B. Evolutionary Optimization of Convolutional Neural Network Architecture Design for Thoracic X-ray Image Classification. In *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*; Springer: Cham, Switzerland, 2021; pp. 121–132.